Dipartimento di Economia e Finanza

# Metodi e analisi statistiche

2018

Università degli Studi di Bari Aldo Moro

Dipartimento di Economia e Finanza

# METODI E ANALISI STATISTICHE

2018

Università degli Studi di Bari Aldo Moro

Tutti i diritti di traduzione, riproduzione e adattamento,
totale o parziale, con qualsiasi mezzo (comprese
le copie fotostatiche e i microfilm) sono riservati

© Copyright 2018 by Università degli Studi di Bari Aldo Moro
www.uniba.it

Prima edizione: dicembre 2018

ISBN 978-88-6629-048-3

*Comitato scientifico:*
Ernesto Toma
Francesco D. d'Ovidio
Vittorio Nicolardi
Alessio Pollice
Nunziata Ribecco

Gli articoli qui presentati sono stati oggetto, oltre che di valutazione interna, anche di revisione anonima (in "doppio cieco").

*Editing finale*: F. D. d'Ovidio

# Sommario

Ernesto Toma
***Presentazione*** ..... pag. 5

Maria Carella, Frank Heins
***Crisi economica nel Sud d'Europa e transizione allo stato adulto: un focus sulle regioni italiane***..... « 7

Federico Matteo, Nunziata Ribecco
***On the Improbability Principle in Long-Run Upward Stock Trends*** ..... « 23

Crescenzio Gallo, Alessandro Rinaldi
***The What-If Analysis: an Application of the* Sensitivity Analysis *to Financial Derivatives***..... « 35

Francesco D. d'Ovidio, Najada Firza, Dante Mazzitelli
***Customer Relationship Management: il caso di BPPB***..... « 45

Caterina Marini, Vittorio Nicolardi
***Giovani lavoratori e contratto di apprendistato. Evidenza italiana***..... « 57

Antonella Bianchino, Agata Maria Madia Carucci, Giovanni Vannella
***Il nuovo sistema statistico "a misura di comune": una applicazione sul territorio lucano*** ..... « 83

Monica Carbonara
***Le tecniche di autocorrelazione spaziale per l'analisi della povertà in Puglia***..... « 109

Angela Maria D'Uggento, Mara Turturo
***Studio pilota su una coorte di pazienti con Lupus eritematoso sistemico*** ..... « 117

Laura Antonucci, Corrado Crocetta, Francesco D. d'Ovidio, Najada Firza, Ernesto Toma, Domenico Viola
***Indicatori per l'analisi della customer satisfaction in aeroporto*** ..... « 129

Fabio Manca, Claudia Marin
***Metodologie statistiche nei processi di sintesi dei nanomateriali*** ..... « 147

Thaís García-Pereiro, Roberta Pace
***L'analisi delle biografie come chiave di lettura dell'evoluzione delle convivenze di primo ordine. Il caso italiano*** ..... « 155

Mauro Bisceglia
***Su qualche equazione alle differenze*** ..... « 179

Vito Ricci
***L'apprezzo di Molfetta del 1417: analisi statistiche*** ..... « 209

Laura Antonucci, Gianluca Bruno, Corrado Crocetta, Leonardo Di Gioia, Massimo Russo
***Analisi della customer satisfaction per i servizi erogati da Sanitaservice ASL FG s.r.l.*** ..... « 245

Progetto Nazionale Lauree Scientifiche
***Questionario Indagine sul Cyber-bullismo*** ..... « 261

Floriana Rinaldi
***Cyberbullismo@Euclide: i risultati di un'indagine tra gli studenti*** ..... « 265

Giuliana Gatti, Cosimina Ligorio
***Non solo cyberbulli a spasso in città*** ..... « 277

Antonia Maria de Candia, Carmela Visaggi
***Atteggiamenti psicosociali sugli usi dei mezzi tecnologici*** ..... « 297

# Presentazione

Fra le collane di "working paper" curate da molti anni dal Dipartimento di Economia e Finanza (ex Dipartimento di Scienze Economiche e Metodi Matematici), sia su temi dichiaratamente economici (SERIES - *Southern Europe Research in Economic Studies*) che su altri temi di interesse delle diverse sezioni dipartimentali (la collana *Geografia Economica*, i *Saggi di Storia Economica* e infine gli *Studi di Diritto Pubblico*), l'ultima nata, per così dire, è la collana di cui fa parte il presente volume collettaneo: *Metodi e applicazioni statistiche*, che raccoglie vari studi matematico-statistici di fonte interna ma anche esterna al Dipartimento.

Si è detto trattarsi dell'ultima nata, ma nella realtà essa eredita il prestigioso ruolo che, a partire dal 1927, pur con diverse interruzioni nel tempo, fu degli *Annali dell'Istituto di Statistica dell'Università degli studi di Bari*, poi divenuti (a seguito dell'estinzione degli Istituti universitari) *Annali del Dipartimento di Scienze Statistiche dell'Università degli studi di Bari*. In quei volumi apparvero nel corso degli anni gli scritti di molti valenti Autori, e anche i primi scritti del recentemente scomparso Giovanni Girone, che fu un riconosciuto luminare della Statistica italiana, Autore internazionalmente noto e autorevole Preside della Facoltà di Economia prima di diventare Rettore della nostra Università. Dopo la lunga parentesi ai vertici dell'Ateneo, egli riprese a coltivare la passione della propria vita, e i suoi ultimi scritti comparvero sul precedente volume di questa stessa collana, a cui egli non mancò mai di fornire il proprio supporto e contributo, anche in collaborazione con vari colleghi.

Contributo che purtroppo manca in questo volume, pubblicato, come i precedenti, esclusivamente in formato PDF *open access*, a seguito di un processo di revisione interna ed esterna (alcuni revisori afferiscono alla nostra università, altri a varie università e istituzioni italiane, qualcuno a università straniere) che ha interessato pressoché tutti i saggi pubblicati, alcuni dei quali possono essere considerati un lavoro compiuto e rifinito, mentre altri sono "work in progress", ma tutti con aspetti di adeguata innovatività scientifica.

Motivata eccezione alla procedura di *peer review* è costituita dai tre saggi meramente divulgativi che chiudono il volume; essi derivano dall'esperienza del Piano Nazionale Lauree Scientifiche, localmente coordinato dalla collega Nunziata Ribecco, che quest'anno ha affrontato con la collaborazione di diverse scuole del territorio provinciale l'attualissimo argomento del cyber-bullismo.

Sperando che il volume incontri il favore dei lettori, colgo l'occasione per ringraziare tutti coloro che vi hanno contribuito, sia Autori che *referees* esterni.

Bari, 14/12/2018

Il Direttore del Dipartimento di Economia e Finanza
*Ernesto Toma*

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018* ISBN: 978-88-6629-048-3
DIPARTIMENTO DI ECONOMIA E FINANZA pp. 7-22

# Crisi economica nel Sud d'Europa e transizione allo stato adulto: un focus sulle regioni italiane

**Maria Carella**[1*], **Frank Heins**[2]

[1]*Università degli studi di Bari "Aldo Moro",*
[2]*Istituto di Ricerche sulla Popolazione e le Politiche Sociali, Consiglio Nazionale delle Ricerche*

**Riassunto:** Il lavoro si propone di analizzare gli effetti della crisi economica sulla transizione allo stato adulto nelle regioni italiane. In una prima fase si valutano le ripercussioni della instabilità economica sui tassi di disoccupazione dei giovani adulti distinti per categorie socio-demografiche e per tipologie familiari di appartenza. Nella fase successiva si cerca di comprendere in che misura la crisi economica abbia influenzato la biografia e le scelte dei giovani nel passaggio alla vita adulta tenendo conto delle loro condizioni lavorative e della composizione dei nuclei familiari. Per perseguire questo obiettivo si effettuano analisi di carattere essenzialmente esplorativo che vengono condotte osservando le differenze territoriali e di genere.

**Keywords:** Crisi economica; Transizione allo stato adulto; Regioni italiane.

## 1. Introduzione

Nei paesi dell'Europa meridionale, la "Great Recession", iniziata nel 2007/2008, ha avuto notevoli ripercussioni sugli indicatori economici causando una forte riduzione dei livelli occupazionali dei giovani. Il progressivo degrado del mercato del lavoro ha danneggiato indistintamente tutte le economie del Sud Europa; tuttavia le modalità e i tempi di diffusione della crisi sono apparsi diversi, evidenziando rilevanti eterogeneità tra i paesi e tra le regioni dell'area. La recessione ha infatti colpito la Grecia e la Spagna con maggiore intensità, provocando una caduta eccezionа-

[*] Autore corrispondente: maria.carella1@uniba.it
Il lavoro qui descritto è frutto di un progetto comune: tuttavia M. Carella ha provveduto alla redazione del paragrafo 1, F. Heins ha redatto il paragrafo 6, mentre i due autori hanno scritto congiuntamente i paragrafi 2, 3, 4, 5.

le dell'occupazione: nel 2013, in Grecia e in Spagna un giovane su due di età compresa tra i 18 e i 24 anni era disoccupato[1], contro uno su tre in Italia e Portogallo (Cottini e Vitali, 2014; Golaz V. et al., 2018)

Alla luce di queste evidenze, nel presente contributo si osservano dapprima gli andamenti della produzione e dei tassi di disoccupazione che hanno caratterizzato i paesi del Sud Europa negli ultimi 15 anni: i trend vengono analizzati in un'ottica comparativa rispetto ai valori medi europei (dell'UE 28). In una fase successiva si cerca di valutare l'impatto della crisi economica sulla transizione allo stato adulto dei giovani italiani.

Lo studio si prefigge un duplice obiettivo. In una prospettiva macro, si esaminano le dinamiche sottese all'andamento dei tassi di disoccupazione dei giovani distinti per tipologia familiare di appartenenza. Nel contesto di un'analisi micro ci si propone di comprendere in che misura l'incertezza economica abbia inciso sulle loro scelte nella fase di passaggio alla vita adulta. Per perseguire questo intento si osservano contestualmente le condizioni lavorative dei giovani e il nucleo familiare al quale essi appartengono rivestendo il ruolo di: figli con genitori; partner in coppia; genitori con figli.

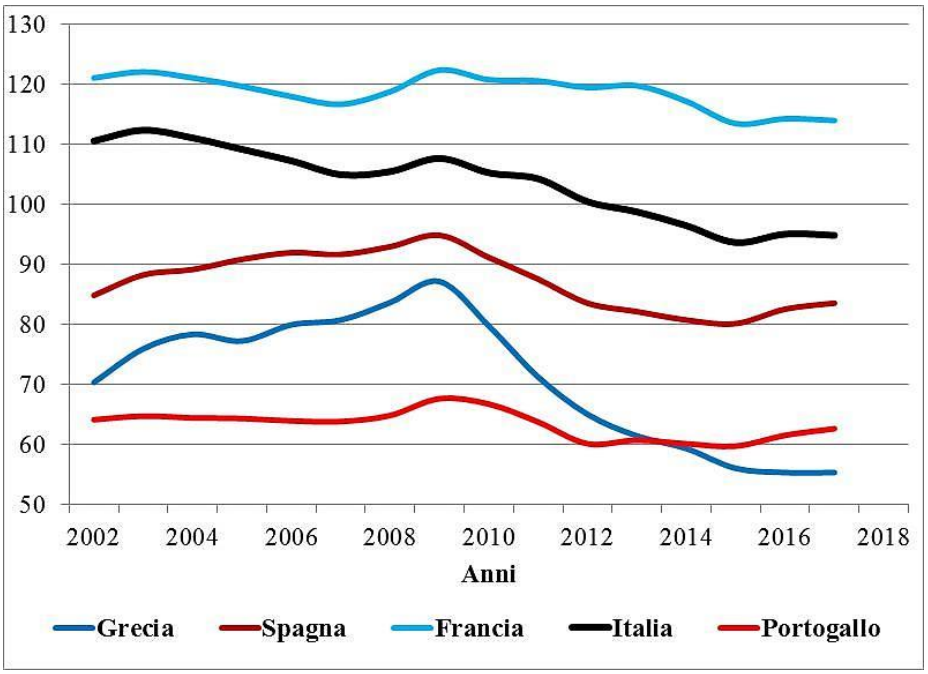
Lo studio si articola secondo due dimensioni trasversali: la prospettiva di genere e il contesto geografico. Relativamente alla prima si argomenta sull'assunto che la crisi economica colpisca uomini e donne secondo modalità e ritmi differenziati. Con riguardo al secondo aspetto, si conduce un'analisi comparativa sistematica a livello territoriale (per ripartizione e province) al fine di verificare se, esaminando i livelli di occupazione dei giovani e gli stili di vita per tipologie lavorative e familiari, si rilevano differenze di rilievo negli ambiti geografici di riferimento. In questo senso l'Italia rappresenta un privilegiato laboratorio per condurre questo tipo di studio: storicamente le differenze tra le regioni si rinvengono sia nei comportamenti demografici sia nelle condizioni socio-economiche degli italiani.

## 2. La 'great recession' nei paesi del sud d'Europa

La crisi finanziaria degli anni 2007-2008 ha avuto implicazioni importanti sull'economia reale di tutti i paesi europei. Tuttavia la recessione ha colpito particolarmente i paesi del Sud Europa dove dal 2008 al 2015 si è manifestata a ritmi differenziati incidendo soprattutto sulla produzione e sull'occupazione.

[1] Qui e nel seguito il termine "disoccupato" viene utilizzato, come ormai d'uso, nel più ampio senso di "non occupato", invece di indicare soltanto chi abbia perso un lavoro.

**Figura 1.** *Andamento del PIL 2002-2017, UE dei 28 = 100*




Fonte: Eurostat database [http://ec.europa.eu/eurostat/data/database]

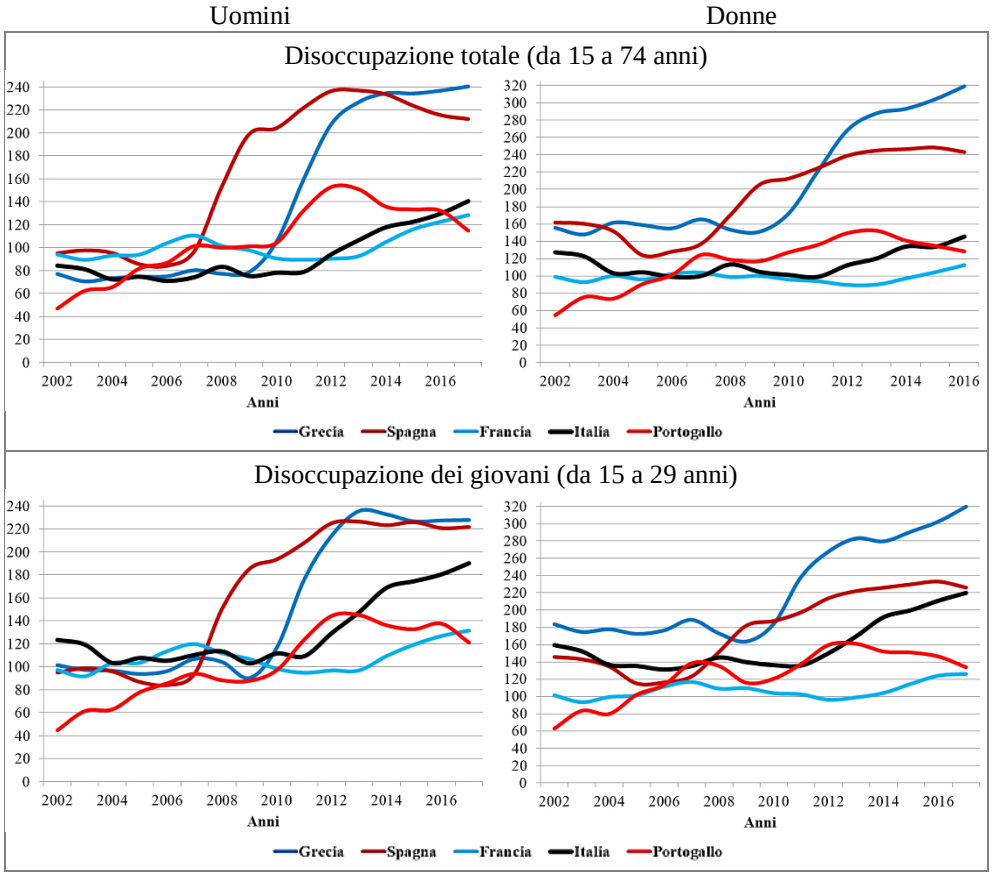
Esaminando l'andamento del PIL pro capite, fatta 100 la media europea (UE28), si nota che dal 2006 al 2016 solo in Francia l'indicatore economico si è sempre posizionato su livelli superiori rispetto al valore medio dell'UE, mentre i restanti paesi dell'area sono rimasti sempre al disotto del medesimo (Fig. 1).

In generale nell'Europa meridionale si intravedono due fasi recessive distinte: la prima relativa al periodo 2008-2010, durante la quale le dinamiche della produzione, seppur in calo, appaiono stabili e una seconda fase iniziata nel 2011 in cui, fatta eccezione per la Francia, la distanza in termini di variazioni del prodotto interno lordo dalla media dei paesi UE si è ampliata (Fig. 1). Dal 2010 la Grecia ha sofferto maggiormente lo shock recessivo rispetto ai suoi omologhi, mentre l'Italia e la Spagna hanno perso meno in termini di produzione. Tuttavia le modalità e i tempi di ripresa per questi ultimi paesi non coincidono: dal 2014 la crescita del PIL appare più sostenuta sul territorio spagnolo mentre in Italia, dove l'economia stenta a ripartire, le variazioni positive della produzione restano alquanto contenute.

Sul versante del lavoro, in tutti i paesi dell'Europa meridionale l'intensa contrazione degli indicatori economici ha prodotto un consistente calo dell'occupazione che è risultato particolarmente significativo in Spagna e in Grecia. L'eccezionale

emorragia di posti di lavoro che i due paesi hanno sperimentato soprattutto nel periodo 2009-2013 ha prodotto uno scollamento rilevante tra le traiettorie dei loro tassi di disoccupazione e quelli osservati nei restanti contesti dell'area (Fig. 2).

**Figura 2.** *Andamento dei tassi di disoccupazione 2002-2017, UE dei 28 = 100*




Fonte: Eurostat database [http://ec.europa.eu/eurostat/data/database]

In generale lo shock recessivo ha colpito particolarmente i giovani: il loro tasso di disoccupazione, osservato in relazione all'età compresa tra 15 e 29 anni, è aumentato in tutta l'area, malgrado si rilevino sostanziali differenze tra i paesi (Fig. 2).

Di fatto, mentre la Spagna e la Grecia hanno sofferto maggiormente la crisi economica fin dal biennio 2009-2010 sperimentando da subito un rapido e drammatico incremento della disoccupazione giovanile, in Italia e in Portogallo è soprattutto nel corso della seconda fase recessiva, cioè nel biennio 2012-13, che la distanza in termini di variazioni dei tassi di disoccupazione dalla media della zona

Euro si è dilatata. Rispetto al Portogallo, il ritardo dell'Italia non accenna a ridursi nemmeno con la ripresa economica che, iniziata tra il 2014 e 2015, è apparsa lenta e irregolare, evidenziando una differente capacità di reazione da parte del paese nel raggiungere i valori pre-crisi.

Nel contesto sud europeo un caso a parte rappresenta la Francia dove il tasso di disoccupazione dei più giovani (15-24 anni), dopo un'inziale aumento rilevato nel 2009, è rimasto stabile raggiungendo la soglia del 25% (Golaz V. et al., 2018).

Il peggioramento delle condizioni occupazionali ha influito su donne e uomini in modo diverso: gli uomini hanno subito una perdita di posti di lavoro superiore a quella delle donne nei paesi caratterizzati da una maggiore segregazione del mercato che ha protetto l'occupazione femminile. In generale "la segregazione determina una presenza massiccia delle donne nel settore dei servizi – compreso il pubblico impiego – e una sotto rappresentazione nei settori dell'industria manifatturiera, nell'edilizia e nei segmenti del settore finanziario dominati dalla presenza maschile" (Cirillo et al., 2016). Tuttavia, in periodo di crisi, l'andamento dei tassi di disoccupazione delle donne giovani (età 15-29) mostra un aumento più consistente rispetto a quello osservato per gli uomini.

In generale, negli anni dal 2008 al 2014 il divario di genere si è ridotto, sebbene il tasso di disoccupazione femminile sia rimasto elevatissimo in Grecia (30,2 %) e in Spagna (25,4 %) (ISTAT, 2015).

In uno scenario così delineato il deterioramento delle condizioni economiche e lavorative dei giovani ha avuto importanti implicazioni anche sui comportamenti demografici, dilatando i tempi nei percorsi di transizione allo stato adulto. Nei paesi dell'Europa mediterranea, storicamente accomunati da un ritardo nei tempi di "fare famiglia" (Assve et al. 2013; Castiglioni e Della Zuanna, 2017), la recessione non solo ha favorito la posticipazione degli eventi caratterizzanti il passaggio alla vita adulta ma in molti casi ha prodotto anche un ritorno presso le famiglie d'origine.

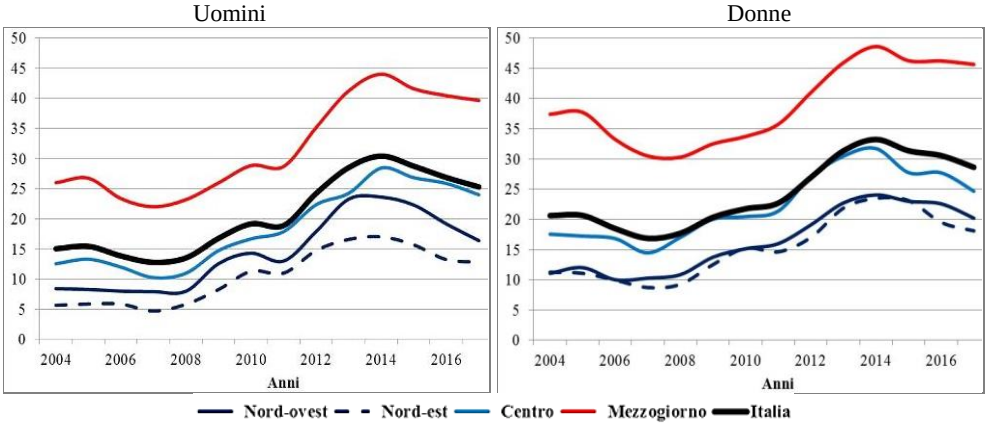
## 3. La 'great recession' in Italia

In Italia il crollo dell'occupazione dei giovani identifica una problematica di grande rilievo dagli inizi degli anni 2000. La crisi economica ha aggravato ulteriormente i livelli di disoccupazione rendendo più impervio l'ingresso al mercato del lavoro proprio alle generazioni che, essendo meno numerose e più istruite, avrebbero dovuto accedervi con maggiore facilità.

Nel 2014, anno ancora in piena crisi economica, il numero di disoccupati in età compresa tra i 15 ei 34 anni si aggirava intorno a circa un milione e 600 mila con un incremento di 700.000 unità rispetto al 2008; nel medesimo intervallo temporale (2008-2014) il tasso di disoccupazione dei giovani sotto i 30 anni si è raddoppiato passando dal 15,3 % al 31,6% (De Rose e Strozza, 2015; Istat, 2015).

Relativamente a questi ultimi si rinvengono interessanti differenze di genere: sembrerebbe che i giovani di sesso maschile (15-29 anni) siano stati più colpiti dalla crisi rispetto alle loro coetanee tradizionalmente sfavorite sul mercato del lavoro. Tra il 2008 e il 2014 il tasso di disoccupazione dei primi è aumentato di circa 17 punti percentuali superando il 30%, mentre quello delle donne è cresciuto meno rapidamente registrando una variazione positiva di 15 punti e raggiungendo quota 33,2% (Fig. 3). L'impatto della crisi sull'andamento dei tassi disoccupazione ha così ridotto il persistente divario tra i sessi nel contesto lavorativo incidendo più criticamente sui settori di attività economiche a maggiore presenza maschile.
A livello territoriale, l'incremento della disoccupazione giovanile è stato più consistente nel Mezzogiorno: in sei anni (dal 2008 al 2014) il tasso di disoccupazione per le donne e per gli uomini under 30 delle regioni meridionali e insulari è cresciuto rispettivamente di 18 e 11 punti percentuali superando a fine periodo in entrambi i casi il 40% (48,6% per gli uomini e 44,0% per le donne). Procedendo da sud a nord la situazione migliora, i tassi restano sotto la soglia del 30% (fatta eccezione per quelli delle donne in età 15-29 anni delle regioni centrali) e la variazione nella disoccupazione decresce (Fig. 3).

**Figura 3.** *Andamento della disoccupazione dei giovani adulti (15-29) nelle ripartizioni italiane, per sesso, 2004-2017*




Fonte: Istat - Rilevazione sulle forze di lavoro 2004-2017 [dati.istat.it]

Disaggregando i dati a livello provinciale, le differenze geografiche appaiono più evidenti e confermano lo storico e persistente divario esistente tra le condizioni occupazionali che si osservano nelle regioni meridionali e settentrionali del nostro paese. Tra i comuni capoluoghi di provincia, quelli del Mezzogiorno registrano i più elevati incrementi della disoccupazione giovanile dal 2008 al 2014. La figura 4 conferma quanto detto, evidenziando l'esistenza di una relazione lineare e crescente tra i tassi di inizio e fine periodo di crisi e le province osservate: salendo nel grafico dal Nord al Mezzogiorno d'Italia la disoccupazione aumenta, in particolate i giovani maggiormente penalizzati risiedono nelle province campani, sardi e calabresi.

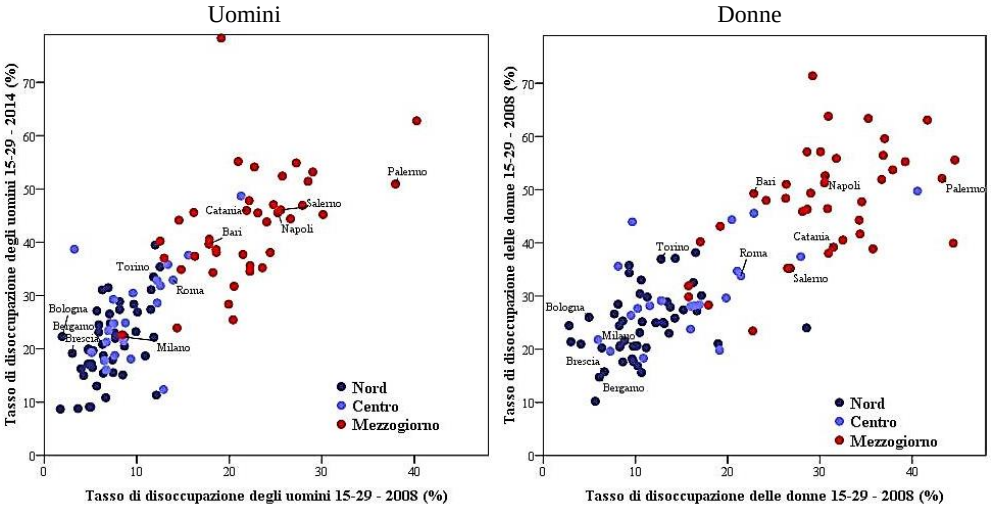
Questa relazione appare più evidente esaminando i dati degli uomini under 30 che risultano omogenei e concentrati. In sole due province i loro tassi di disoccupazione sono rimasti pressoché invariati (Sondrio e Prato); in altre del Nord hanno subito un aumento contenuto rimanendo su livelli uguali o inferiori al 10% (Verona, Vicenza, Cuneo, Padova, Bolzano). Nelle restanti province delle regioni settentrionali, l'incremento nella disoccupazione tra i due periodi è più apprezzabile ma i tassi non eccedono il 30% (fatta eccezione per Novara, Torino e per le tre liguri: Genova, La Spezia e Imperia). Una situazione molto diversa si osserva nelle regioni del Mezzogiorno, dove la disoccupazione giovanile per il sesso maschile ha raggiunto livelli record superando il 30% e il 50% in 5 province delle isole e della Calabria.

Per le donne under 30 si rileva invece una maggiore dispersione dei dati dovuta all'eterogeneità delle situazioni sul territorio nazionale dove le brusche variazioni dei tassi di disoccupazione hanno interessato non solo le province meridionali e insulari (in quattro comuni delle isole è stato superata quota 63% nel 2014) ma altresì quelle delle regioni centrali (+32 a Lucca e + 23 a Terni; a Massa-Carrara il tasso, già elevato in partenza, ha raggiunto 49%).

Circoscrivendo l'analisi ai giovani in età 20-34, al censimento del 2011 si osserva un andamento dei tassi di disoccupazione non dissimile da quello fin qui esaminato. Tali tassi appaiono allineati e crescenti man mano che si esaminano le province salendo da nord a sud.

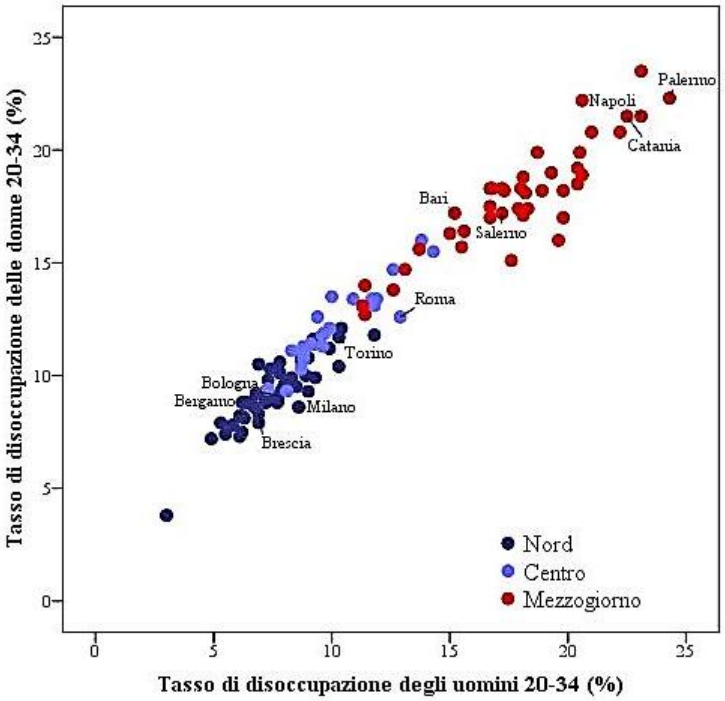
Questo dato, indagando su una sottoclasse di individui che include anche i 30-34enni, mostra quanto la crisi abbia inciso e reso ancora più vulnerabili i giovani di età più elevata che in Italia si avviano alle prime fasi della transizione allo stato adulto tradizionalmente in ritardo rispetto ai loro coetanei dei paesi del centro e nord Europa.

**Figura 4.** *Differenze territoriali della disoccupazione dei giovani adulti (15-29) nelle province italiane, per sesso, 2008 e 2014*




Fonte: Istat - Rilevazione sulle forze di lavoro 2008 e 2014 [dati.istat.it]

**Figura 5.** *Le differenze territoriali della disoccupazione dei giovani adulti (20-34) di ambo i sessi nelle province italiane, 2011*




Fonte: Istat - 15° censimento della popolazione e delle abitazioni 2011

## 4. Dati e metodi

Intento principale del presente lavoro è valutare le possibili conseguenze della instabilità occupazionale prodotta dalla recessione economica sulla transizione all'età adulta dei giovani italiani. Per perseguire questo obiettivo sono stati utilizzati i dati del Censimento della popolazione 2011 e quelli della Rilevazione continua sulle forze di lavoro (RFL) per gli anni 2004 a 2017. I primi consentono di osservare la distribuzione delle forme familiari a cui afferiscono i giovani distinti per sesso e per province all'apice della crisi. Si delinea un primo quadro descrittivo dal quale emergono interessanti eterogeneità a livello territoriale in merito al ruolo che i giovani rivestono nel nucleo familiare di appartenenza.

In un secondo step, si ricorre ai dati della RFL per analizzare le possibili interrelazioni esistenti tra status occupazionale e contesti familiari in cui i giovani vivono come: figli con i genitori; in coppia; genitori con figli. In generale, l'indagine sulle forze di lavoro si riferisce ad un campione rappresentativo della popolazione italiana e permette di effettuare analisi dettagliate sulla condizione occupazionale degli intervistati differenziandoli per area geografica di residenza. Osservando questa variabile e la posizione che gli individui ricoprono all'interno della famiglia nelle diverse classi di età si possono distinguere le fasi del ciclo di vita individuale. Di fatto esaminando congiuntamente due aspetti determinanti nel passaggio alla vita adulta, vale a dire lavoro e famiglia, si possano rinvenire e valutare le possibili implicazioni che la crisi ha prodotto su alcune sequenze del processo transizionale. In tal senso il nostro lavoro si propone un primo studio di carattere esplorativo.

Va aggiunto che per evidenziare opportunamente le differenze tra i comportamenti femminili e quelli maschili nella transizione allo stato adulto, le analisi sono state stratificate anche per genere.

Le statistiche ufficiali a livello internazionale includono nella classe giovanile gli individui in età compresa tra i 15 e 24 anni: tuttavia i giovani selezionati ai fini del nostro studio appartengono alla fascia di età 20-34. La scelta di spostare le soglie ad un'età più elevata rispetto a quella che tradizionalmente identifica la popolazione giovanile nasce dall'esigenza di indagare su un campione di giovani che verosimilmente potrebbero aver concluso gli studi e raggiunto l'indipendenza economica. In altri termini le analisi condotte sulla classe di età 20-34 permettono di tener conto del ritardo con cui si compie nel nostro paese la transizione allo stato adulto.

In Italia e in generale nei paesi del Sud Europa, prima dei 20 anni, i giovani sono per la maggior parte in istruzione e difficilmente lasciano l'abitazione dei geni-

tori; al contempo alcuni traguardi tipici dell'età adulta vengono raggiunti solo dopo i 30 anni (De Rose e Strozza, 2015).

## 5. La transizione allo stato adulto dei giovani in periodo di crisi economica

### *5.1 Giovani e contesti familiari di appartenenza*

La letteratura demografica evidenzia che i giovani italiani e i loro omologhi dei paesi sud- europei sperimentano difficoltà di vario tipo nell'avviare e completare gli stadi che definiscono il processo di transizione alla vita adulta (Aassve et al., 2013). L'acquisizione dell'indipendenza economica necessaria all'affrancamento dalla famiglia di origine è senz'altro una delle determinanti di maggior rilievo di tali difficoltà (Aassve et al., 2001).

In Italia gli esiti prolungati della crisi economica hanno inciso sulla dilatazione di questo processo, prorogando ulteriormente il compimento delle tappe transizionali. Recenti studi documentano che, negli anni della recessione, l'instabilità finanziaria e lavorativa ha indotto i giovani a rimandare alcune scelte di vita strettamente connesse alle sequenze tipiche del passaggio allo stato adulto: uscita dalla casa dei genitori, matrimonio , nascita del primo figlio (Sobotka et al., 2011; Becker et al., 2010; Caltabiano et al., 2017).

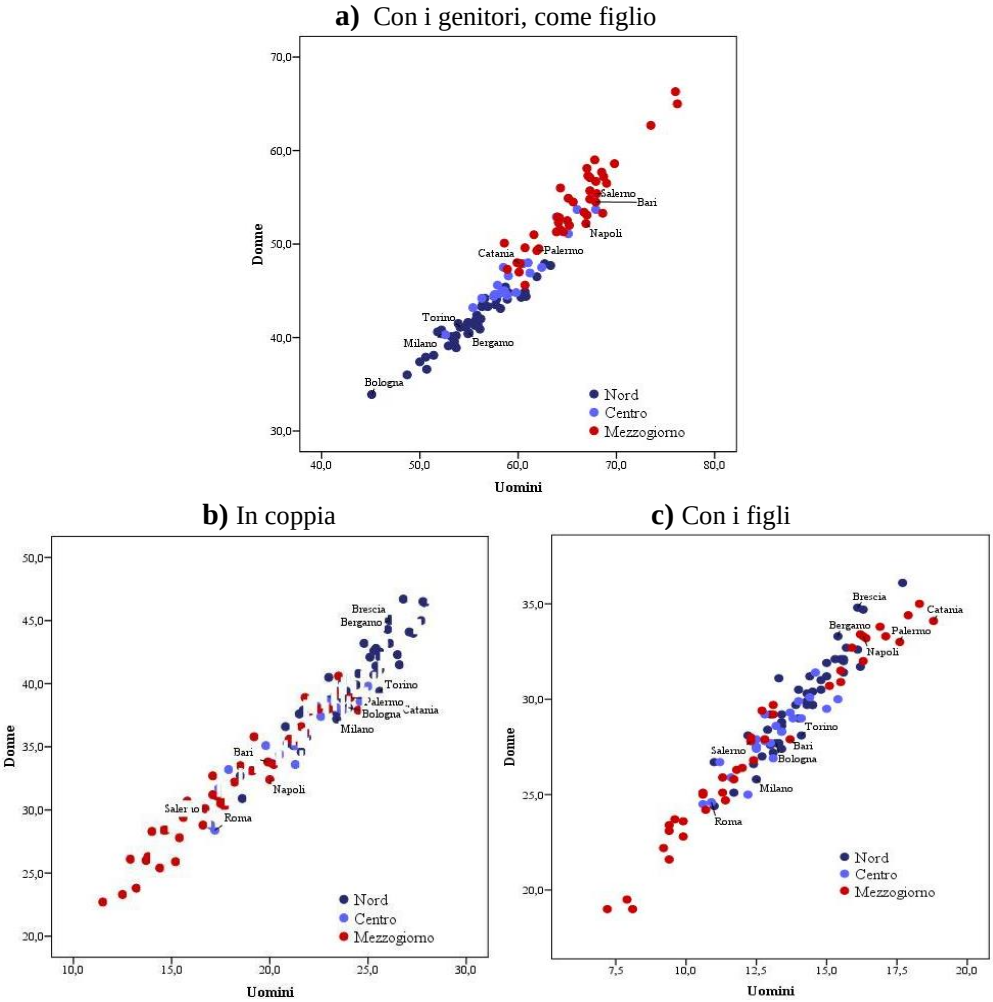
In particolare la permanenza prolungata dei giovani a casa dei genitori identifica una peculiarità del "familistic way of life" che da sempre caratterizza i paesi a forti legami familiari (Castiglioni e Della Zuanna, 2017). In Italia le determinanti di questo fenomeno si rinvengono sostanzialmente in tre tipologie di fattori: economici (difficoltà nell'accesso al mercato del lavoro e abitativo); culturali (abitudini e norme sociali incidono ancor oggi sui comportamenti demografici); socio-politici (il sistema italiano di welfare non offre un adeguato supporto alle giovani famiglie).

In sintesi, la partenza dalla famiglia d'origine risulta da sempre tardiva e per decenni è rimasta quasi esclusivamente legata al matrimonio (Billari et al., 2001). In periodo di crisi i tempi di uscita si sono ulteriormente allungati: tra il 2007 e il 2014, la percentuale di giovani in età 18-34 anni che risiede con i genitori è aumentata di 5 punti percentuali raggiungendo il 65,8% (Eurostat, 2015): nel 2014, quasi 7 milioni di giovani vivevano con le loro famiglie, due su tre contro uno su due in età compresa tra i 25 ei 34 anni (49,4%). Il fenomeno è più maschile che femminile (58% degli uomini e 41% delle donne tra i 25 e i 34 anni) mentre a livello territoriale riguarda indistintamente i giovani delle regioni settentrionali e meridionali.

Studenti e disoccupati permangono più a lungo con i genitori, tuttavia anche tra i giovani adulti (25-34 anni) che possiedono un' occupazione il 60% convive con il nucleo familiare d'origine (Carella e Heins, 2018). I giovani italiani lasciano la casa dei genitori 4 anni dopo (circa a 31 anni) rispetto alla media dei giovani europei.

Partendo dalle considerazioni sin qui espresse, si osservano alla data del censimento del 2011 i giovani in età 20-34 anni nel contesto familiare di appartenenza in base al ruolo da essi assunto come: figli con i genitori, partner in coppia, genitori con figli (Fig. 6.a, 6.b e 6.c).

**Figura 6.** *Le differenze territoriali delle forme familiari dei giovani adulti (20-34) per sesso nelle province italiane, 2011 .*




Fonte: Istat - 15° censimento della popolazione e delle abitazioni 2011

I dati disaggregati a livello provinciale mostrano alcune differenze sul territorio che rilevano ancora una volta la storica discrasia tra nord e mezzogiorno d'Italia per quanto attiene andamenti e comportamenti demografici. Due evidenze appaiono interessanti: salendo da nord a sud, la percentuale di giovani adulti che risiedono con i genitori aumenta (nei comuni del Mezzogiorno gli uomini superano quasi sempre il 60%) mentre quella di coloro che hanno lasciato la famiglia di origine per vivere in coppia decresce (in molte province del sud i giovani di sesso maschile non raggiungono il 20%).

Diverso il quadro che si delinea nel caso dei giovani che hanno sperimentato il ruolo di genitore e vivono con i figli: qui la situazione appare eterogenea e non sembra si possa individuare una possibile relazione tra contesto familiare e provincia di residenza.

### *5.2. I cambiamenti del ruolo dei giovani nel contesto familiare e lavorativo*

Le analisi che seguono mirano a ricostruire il profilo dei giovani adulti italiani e le fasi del loro ciclo di vita negli anni della crisi economica osservando contestualmente il ruolo che essi rivestono all'interno del contesto familiare e lavorativo.

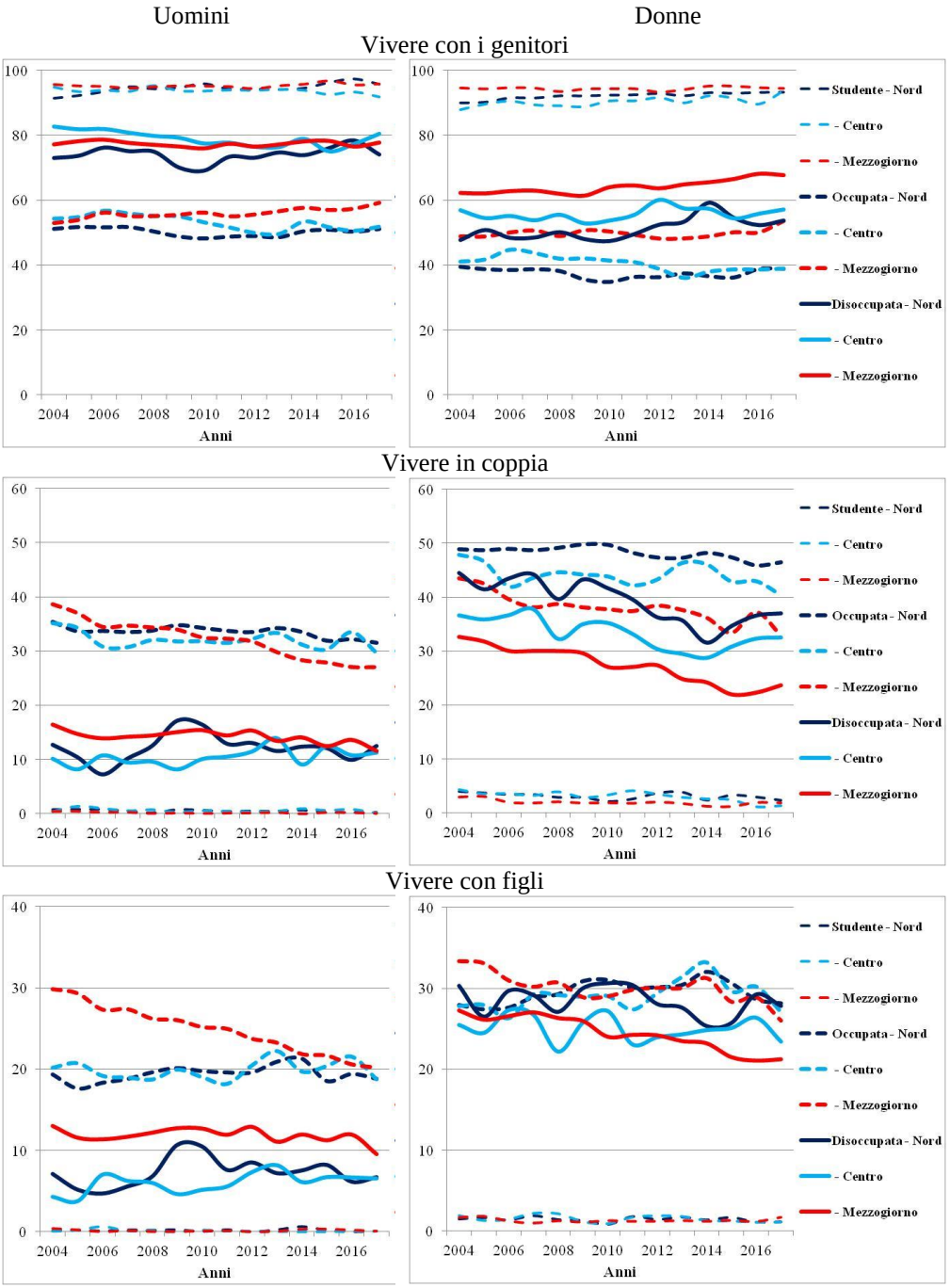
Sul versante familiare, il cambiamento del ruolo in famiglia segna le diverse fasi del ciclo di vita degli individui: il passaggio da figlio a persona in coppia, da quest'ultima a genitore in coppia o a genitore solo può identificare gli eventi che segnano le fasi della transizione allo stato adulto (quali l'uscita dalla famiglia di origine, la convivenza, il matrimonio, i figli avuti) (Istat, 2015).

Per effettuare questo esercizio si utilizzano i dati della RFL e si incrociano le variabili afferenti la condizione occupazionale dei giovani adulti (occupato, disoccupato, studente) e il contesto familiare e si stratificano i dati per genere e ripartizione territoriale.

In particolare, analizzando la composizione dei nuclei familiari attraverso il ruolo dei componenti si individuano tre modi di "vivere e fare famiglia" (Fig. 7):

1) Vivere con i genitori: questa tipologia interessa soprattutto studenti e disoccupati. In periodo di crisi (tra il 2008 e il 2014) la loro percentuale di sesso maschile è rimasta sempre elevata e pressoché invariata in tutte le ripartizioni geografiche superando rispettivamente il 90% e il 70%. Tra le donne, le giovani senza impiego che vivono con i genitori sono più presenti nelle regioni del Mezzogiorno dove la loro quota è aumentata (+9%) raggiungendo il 59%. Resta comunque non trascurabile la percentuale dei giovani per lo più uomini (superiore al 50 %) che pur possedendo un lavoro protraggono la convivenza con i genitori.

**Figura 7.** *Forme dei nuclei familiari dei giovani adulti (20-34) per sesso e per condizione occupazionale nelle 3 ripartizioni, 2006-2017*




Fonte: Istat - Rilevazione sulle forze di lavoro 2004 a 2017

2) Vivere in coppia: questo contesto familiare concerne per la maggior parte coloro che lavorano. Dal confronto territoriale emerge che le regioni del Nord registrano la percentuale più elevata di giovani occupati che hanno costituito un nuovo nucleo in coppia; relativamente al sesso, la quota delle donne che lavorano e vivono con il partner è più alta di quella degli uomini. In generale negli anni della crisi questa tipologia di giovani appare in calo nelle regioni del Mezzogiorno.
3) Vivere come genitore con figlio: questo modo di fare famiglia si caratterizza per diversità territoriali, traiettorie altalenanti nel periodo di crisi e comportamenti differenziati per sesso. Gli uomini interessati sono per la maggior parte coloro che possiedono un impiego e risiedono nel Mezzogiorno: tuttavia la loro percentuale ha subito una variazione negativa in periodo di crisi (-5%) raggiungendo il 21%. Le donne in nucleo con figli sperimentano più frequentemente questa forma familiare rispetto agli uomini e sono per la maggior donne impiegate che risiedono indifferentemente nelle tre ripartizioni geografiche superando, seppur con andamenti altalenanti, la quota del 30%.

## 6. Conclusioni

Le analisi condotte in questo lavoro mostrano che in Italia la crisi economica non solo ha allungato ulteriormente i tempi di transizione allo stato adulto, ma ha prodotto altresì un'implosione degli eventi che favoriscono il susseguirsi delle sequenze caratterizzanti questo processo.

La relazione tra condizione lavorativa e contesto familiare appare evidente: tra il 2008 e il 2014 i giovani adulti disoccupati o in formazione hanno prolungato la loro permanenza a casa dei genitori mantenendo inalterato un trend storicamente peculiare nel nostro paese. Interessante appare il dato che attiene i giovani occupati: tra i medesimi la percentuale di coloro che scelgono di lasciare la famiglia di origine e di vivere con il proprio partner in un nuova abitazione risulta in calo soprattutto nelle regioni del Mezzogiorno. Si potrebbe ipotizzare che l'incertezza economica abbia indotto soprattutto i giovani collocati in un contesto lavorativo precario a rinviare la convivenza in coppia.

Vivere da genitore con figli è invece una scelta che in tempo di crisi i giovani seguono con cliché differenziati: a riguardo le dinamiche non sono omogenee e cambiano in base al sesso e al territorio. Le donne sembrerebbe non intendano ri-

nunciare a questo contesto familiare (fatta eccezione per le giovani del Mezzogiorno che diminuiscono) . Diversa la posizione degli uomini delle regioni meridionali e insulari dove la percentuale di coloro che sono occupati e vivono in un nucleo con figli appare in netto declino.

In sintesi, l'instabilità finanziaria e lavorativa dovuti alla recessione hanno avuto evidenti implicazioni sul processo di formazione della famiglia rinviando le tappe. Tuttavia le modifiche intervenute nei calendari degli eventi sono il risultato di cambiamenti che si sono via via diffusi tra le generazioni. Inoltre le analisi mostrano che, anche in periodo di crisi, le esperienze nei percorsi di transizione allo stato adulto variano secondo il genere e il territorio di residenza, come, ovviamente, la crisi economica. Tra aspettative e realtà con la quale i giovani adulti si sono confrontati in questo ultimo decennio, si potrebbe ipotizzare che la forte crisi economica attuale potrebbe ancora rendere manifeste ulteriori conseguenze nelle fasi finali della transizione alla vita adulta dei giovani italiani.

## Riferimenti bibliografici

Aassve, A.; Cottini, E.; Fraboni, R.; Vitali, A. (2015). Giovani e formazione delle unioni. In : De Rose, A.; Strozza, S. (ed.) *Rapporto sulla popolazione. L'Italia nella crisi economica*. Bologna, Il Mulino: 33-61.

Aassve, A.; Cottini, E.; Vitali, A. (2013). Youth prospects in a time of economic recession. *Demographic Research*, 29: 949-962.

Aassve, A.; Arpino, B.; Billari, F. C. (2013). Age norms on leaving home: multilevel evidence from the European Social Survey. *Environment and Planning A*, 45, 2:383-401.

Becker, S.; Bentolila, S.; Fernandes, A.; Ichino, A. (2010). Youth Emancipation and perceived job insecurity of parents and children. *Journal of Population Economics*, 23: 1175-1199.

Bell, D. N.; Blanchflower, D. G. (2011). Young People and the Great Recession. *Oxford Review of Economy Policy*, 27, 2: 241-267.

Billari, F.; Philipov, D.; Baizán, P. (2001). Leaving home in Europe. The experience of cohorts born around 1960. *International Journal of Population Geography*, 7, 5: 339-356.

Billari, F. C.; Liefbroer, A. C. (2010). Towards a new pattern of transition to adulthood? *Advances in Life Course Research*, 15, 2-3: 59-75.

Bory, S. (2009). Trajectoires réversibles: le cas des jeunes adultes italiens. *Informations sociales*, 156: 132-141.

Caltabiano, M.; Comolli, C. L.; Rosina, A. (2017). The effect of the Great Recession on permanent childlessness in Italy. *Demographic Research*, 37, 20: 635-668.

Carella, M.; Garcia-Pereiro, T.; Pace, R. (2016). Les unions consensuelles en Espagne et en Italie: profils de femmes. In: Girard, C.; Pennec, S.; Sanderson, J. P. (ed.) *Trajectoires et âges de la vie*, Bari, Italie. AIDELF, Paris. erudit.org: 1-16.

Carella, M.; Heins, F. (2018). Les conséquences de la crise économique sur les comportements démographiques en Italie. In : Golaz, V.; Lefèvre, C., Veron, J. (ed.) *La crise dix ans après. Quels effets sur la conjoncture démographique des pays d'Europe du Sud?* Document de Travail, 239. INED, Paris: 63-86.

Castagnaro, C.; Guarneri, A.; Prati, S.; Rinesi, F. (2015). Comportamento riproduttivo. De Rose, A.; Strozza, S. (ed.) *Rapporto sulla popolazione. L'Italia nella crisi economica.* Il Mulino, Bologna: 63-83.

Castiglioni, M.; Dalla Zuanna, G. (2017). *La famiglia è in crisi. Falso!* Laterza, Bari.

Cicchelli, V. (2001). La dépendance familiale des jeunes adultes en France et en Italie. *Recherches et prévisions*, 65: 31-37.

Cirillo, V.; Corsi, M.; D'Ippoliti, C. (2016). Gender, Class and the Crisis. Fadda, S.; Tridico, P. *Varieties of Economic Inequality*, Routledge, Abingdon Oxon: 48-72.

Cottini, E.; Vitali, A. (2014). I giovani e la Grande Recessione. Cappellari, L.; Lucifora, C. (ed.) *Economia della popolazione: disuguaglianza, famiglia e migrazioni*. Vita e Pensiero, Milano: 67-86.

De Rose, A.; Strozza, S. (ed.) (2015). *Rapporto sulla Popolazione. L'Italia nella crisi economica*. Il Mulino, Bologna.

ISTAT (2015). *Rapporto annuale. La situazione del Paese nel 2014*. Istat, Roma.

ISTAT (2014). *Rapporto annuale. La situazione del Paese nel 2013*. Istat, Roma.

ISTAT (2009). *Rapporto annuale. La situazione del Paese nel 2008*. Istat, Roma.

Rimoldi, S. (2015). Famiglie in crisi, famiglie nella crisi. *Rivista Italiana di Economia, Demografia e Statistica*, LXIX, 2: 17-36.

Rosina, A.; De Rose, A. (2015). Crisi e welfare inadeguato affossano le nascite. *Welfare Oggi*, 1.

Salvini, S. (2015). Living in couple. Marriage and cohabitation in a changing Italy. *Rivista Italiana di Economia, Demografia e Statistica*, LXIX, 2: 37-56.

Salvini, S.; Vignoli, D. (2014). *Convivere o sposarsi?* Il Mulino, Bologna.

Sobotka, T.; Skirbekk, V.; Philipov, D. (2011). Economic recession and fertility in the developed world. *Population and Development Review*, 37, 2: 267-306.

Testa, M. R.; Basten, S. (2014). Certainty of meeting fertility intentions declines in Europe during the "Great Recession". *Demographic Research*, 31, 23: 687-734.

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018*
DIPARTIMENTO DI ECONOMIA E FINANZA
ISBN: 978-88-6629-048-3
pp. 23-34

# On the Improbability Principle in Long-Run Upward Stock Trends

**Federico Matteo*, Nunziata Ribecco**

*Dipartimento di Economia e Finanza, Università degli Studi di Bari Aldo Moro*

**Abstract:** In global stock markets, it rarely is possible to see a long-run upward - or downward - trend in price movement for an asset, people explain this fact with different interpretations. Using the Improbability Principle as instrument, we investigate if relative to the average this kind of trend can be considered as a random fluctuation. In order to assess the pattern in the market securities we consider a Random Walk, but with a GARCH component to model the variance, with a generalised hyperbolic distribution as conditional distribution to capture the volatility clustering and the heterogeneity of the residual densities. Thus, we show that in a large portfolio it is possible to see an upward movement due to the improbability principle. We estimate the price of the end of the period starting from the year before, with a Monte Carlo simulation based on the estimated residuals of the whole five years. So, we consider the ratio between the number of times the simulated price is greater than the real one and the number of simulations as the probability of having an over-appreciated stock price due to a false movement. In our analysis, only one stock has resulted spurious: the ZAGG company.

**Keywords:** The Improbability Principle, Monte Carlo Simulation, GARCH, Generalised Hyperbolic Distribution, Market Efficiency Theory.

## 1. Introduction

We have selected a data frame of 100 daily stock prices of the last five years, the period considered starts in 01 January 2013 and finishes in 01 January 2018. Data is provided by a Bloomberg terminal and stock prices belong to the firms of the US

* Corresponding Author: matteo.federico96@gmail.com

technology sector, we have randomly chosen them from a list of all equities exchanged in the NASDAQ (National Association of Securities Dealers Automated Quotation). To estimate the patterns and the chaotic motion of the stock prices we consider a theory on which the efficient market hypothesis is based (EMH), i.e. current stock prices fully reflect available information about the value of the firm. So, since new information are unpredictable by nature, the price of tomorrow will be equal to today's price plus a random variable, whose value cannot be predicted.

In financial literature this assessment about security markets is known as Random Walk, since there is no reason to think that prices are too high or too low, they adjusted to their average value with time. We consider the logarithm of the price instead of the simple price because we are interested in the change of the price (the daily logarithmic return) and not in the latter as a whole number. Initially, we use the following model, let be $j = 1, ..., 100$:

$$\log(P_t^j) = \varphi_1 \log(P_{t-1}^j) + \varepsilon_t^j$$

What we are supposing with our analysis is that even if an upward movement is really uncommon in a random walk, when we consider the random chance aspect, maybe it can be actually typical of the random walk itself. The idea derives from the laws 2 and 3 of the Improbability Principle (Hand, 2014). Before going on, the reader is supposed to know the laws of the improbability principle, so herein I want to provide a list of the set of laws and their definitions, for some examples see http://improbability-principle.com.

**Table 1.** *The five laws of the Improbability Principle*.

| |
|---|
| The Law of Inevitability |
| The Law of Truly Large Numbers |
| The Law of Selection |
| The Law of the Probability Lever |
| The Law of Near Enough |

It is important to remember that Professor David Hand has not demonstrated these laws in his book, since it is not thought to be an academic book. In this sense, another important purpose of this work is chalking up a mathematical and statistical demonstration for the set of laws, but bear in mind that the methodology used and described in this report is not a general demonstration, but instead a statistical and technical application of improbability principle in finance. As a matter of fact, even if in another sector and for other events the improbability

principle is still applicable, the techniques used must change in order to apply the five laws in different situations: there are different ways, herein we will propose one of these, adapted to our contest. Let's consider now the definitions of each law in the order of the Table 1, we will use them during the article: something must happen; with a large enough number of opportunities, any outrageous thing is likely to happen; you can make things as likely as you want if you choose after the event; slight changes can make highly improbable events almost certain; events which are sufficiently similar are regarded as identical. Let's consider this example: if I toss a coin ten times, the probability of obtaining ten heads in a row is really small, but if we put one hundred people in a single room this event is more likely to occur than before. Definitely it is possible that no one tosses ten heads in a row, but if we let more and more people enter in the room, at a certain point the possibility to look at that event becomes quite large, thanks to the laws 1 and 4. We want to demonstrate the same concept with stock prices. So, if we consider a large portfolio, like the one of one hundred stocks, can we look at an uncommon upward movement due to chance? Please note we do not consider the downward movement, for this reason in the data frame there are not bankrupted firms, we are only interested in the upward trend.

## 2. Material and methods

We make some assumptions about the random variable, i.e. all the residuals are independent, uncorrelated and normally distributed. The investigation of the assumptions is made with the autocorrelation function and with a qq-plot. We verify also with the Dickey Fuller standard test if $\varphi_1 \simeq 1$ and with the Jarque-Bera test the random variable normality, i.e. $\varepsilon_t \sim IIDN(0, \sigma^2)$. Considering the period under examination, we estimate the stock price at the time *n* using a Monte Carlo simulation based on past values to randomly generate the distribution of the last price at the time *n* using a sample of residuals and then find out where the real last price is positioned in the simulated distribution. The Monte Carlo simulation uses residuals deriving from an AR(1) based on the real data, but it could not be good to use old data in order to generate new prices, because structures of different years could be different and not necessary they need to reoccur in future. In the next chapter we will introduce the GARCH estimation with the volatility modelled by a conditional density, in order to solve the problem of past years. This new approach will be more accurate and more realistic.

A Monte Carlo simulation is useful to achieve an approximation of the real distribution. As a matter of fact, the empirical probability density function is approximated using a sum, let be $P^{[1]}$, $P^{[2]}$,..., $P^{[N]}$ a sequence of evenly distributed prices on the interval $]-\infty, \infty[$, so

$$\int_{-\infty}^{\infty} f(P_n) dP_n \cong \lim_{n\to\infty} \sum_{i=0}^{N} f\left(P_n^{[i]}\right),$$

where:

$$P_n^{[i]} = P_s + \sum_{x=0}^{N} \varepsilon_{x,t} \ .$$

with the value of $\varepsilon$ randomly sampled with replacement from $t \in [0,1306]$, and with $i = 0,..,N$, $s = 1047$, $x = 0,...,\tau$ , $\tau = 259$, $N = 20000$. This methodology used for one equity, is then repeated for the one hundred stocks in the data frame. From now, I explain the procedure and the symbols used in a simple way in order to drive the methodology home. We consider a sample of the residuals $\varepsilon_t$ of the whole five years, we have also tried to use only one-year residuals, but in this case the sample was too small and the simulated density ill-suited, since the larger the sample the greater the simulated density. Then we ask ourselves if when we are at the date 2 January 2017 ($s = 1047$), considering the sample of the whole five years residuals, ($t_{max} = 1306$ *days*), is it possible to look at a marked upward movement until the 1-th January 2018 and considering this as a result of a chance process? (So, in a period of $\tau = 259$ days). The Monte Carlo simulation considered is formed of twenty thousand simulations ($N = 20000$) to generate twenty thousand log prices at the time $n = 1306$ = 1/Jan./18 starting from a year before. Once we have got those simulated prices for all stocks, we are interested in learning where the real price of each stock $P_n^*$ is located in the one hundred simulated distribution. So, the output of the algorithm considered is a p-value for each equity (one hundred p-values).

Please, note that the following is a statistical test that we personally developed to capture the effect of the Improbability Principle in this context, in addition, from now it will be generally called algorithm. Firstly, let's consider $v_j$ vectors, with $j = 1, ..., m$ and $m = 100$:

$$v_j = \begin{pmatrix} v_1 \\ \vdots \\ v_i \\ \vdots \\ v_n \end{pmatrix}$$

whose $v_i$ elements, with $i = 1, ..., N$ , are:

$$v_i := \begin{cases} 1 \; if \; P_n^{[i]} > P_n^* \\ 0 \quad otherwise \end{cases} .$$

After that, we consider the quadratic norm of the *j-th* vector, called $\varphi$ for convenience:

$$\varphi_j := \|v_j\|^2 = \sum_{i=o}^{N} v_i^2 .$$

Thus, the p-value is equal to $\varphi$ divided by the number of simulations *N*. Or in other words, how many times the simulated price is greater than the real verified price, divided by the number of simulations:

$$p_{(j)} = \frac{\varphi_j}{N} .$$

When the p-value related to a stock is lower than 0,01, in the period 1/01/17 - 1/01/18 (259 days) we have watched an upward movement which is really rare and uncommon for the random walk. Note that in the case of a downward movement we should use a "<" in stead of ">" to calculate the $v_i$ element. But remember that we are not interested in downward trending price, so we avoid a selection bias due to the fact that bad performing companies may have been removed from the data: in the data-frame there are not bankrupted firms, as said before. Note also that with this formula $p_{(j)} \in [0{,}00001, 1]$, while if $p = 0$, it means that no one of the $P^{[i]}$ generated from the simulation is greater than the real value, i.e. the real price is anomalous and the stock trend spurious. The application of this algorithm to the stocks of the technology sector has produced several $p_{(j)} < 0{,}05$, and only two smaller than 0,01. However, we have to consider that the procedure used is a multiple comparison, so before evaluating a p-value significant, we adopt a conservative approach using the false discovery rate (FDR) procedure by Benjamini-Hochberg (1995), in order to conceptualise the rate of type I errors in null hypothesis. The BH procedure is very practical, because it simply requires that we are able to compute p-values for each of the individual tests and this permits the procedure to be defined and to distinguish genuine from spurious upward trends. We choose an $\alpha = 0{,}05$, and then define *k* to be the largest *j* for which:

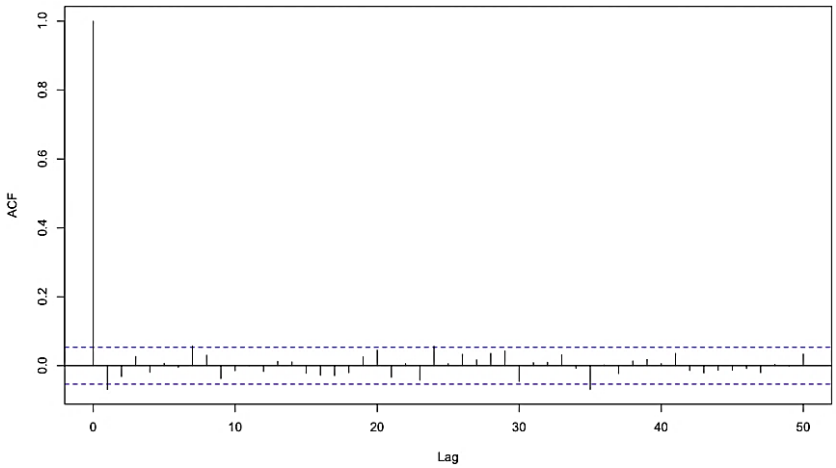
$$p_{(k)} < \frac{j}{m} \cdot \alpha .$$

The procedure is rejecting tests with p-values smaller or equal to $p_{(k)}$, given $p_{(1)}$, $p_{(2)},\ldots,p_{(j)},\ldots,p_{(m)}$ p-values, generated from $H_1,H_2,\ldots,H_j,\ldots,H_m$ null hypotheses. Results are shown in the next paragraph.

## 3. Results

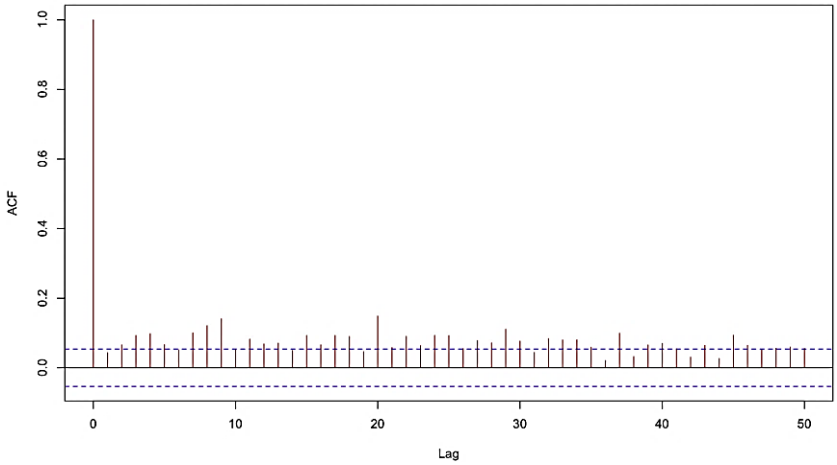
Firstly, we consider the ACF (Autocorrelation Function) only with a graphical analysis. What we obtain are plots similar of the one in Figure 1 for the one hundred stocks.

**Figure 1.** *The ACF of the residuals.*




In t = 0 the ACF is equal to one (because it is the correlation between the value in time zero and itself), after that, all the vertical lines are quite near to zero, so they are negligible, we get the same plots for the other stocks and this means that residuals are uncorrelated. The problem is when we calculate some non-linear transformations, such as the absolute value of residuals, as shown in Figure 2.

**Figure 2.** *The ACF of absolute value of the residuals.*



Now the vertical lines go over the dotted horizontal lines, but only in few cases the ACF values are no longer negligible, this means that residuals can still be considered dependent and as a consequence we can sum them. However, there is the presence of a systematic behaviour: lower residuals are followed by lower ones, higher residuals are followed by higher ones. In literature, if a problem of that kind occurs, it is known as volatility clustering. An observation of this type goes against simple random walk and have led to the use of GARCH models in financial time series analysis. In addition, there are also been some cases of non-stationary time series with a $\varphi > 1$, in this case we force R to use the maximum likelihood estimation in order to estimate the parameters of the series and get $\varphi \approx 1$. Then, we test if returns follow a normal distribution as supposed before, I use a graphical approach as well as the Jarque-Bera test, but the p-value generated from the latter is really small, a number lower than $2e^{-16}$, this implies the rejection of the normality hypothesis.

The histogram and a qq-plot of a typical stock are shown in Figure 3 and in Figure 4, they are also similar for the remaining stocks.

It is easy to see that the histogram has a kurtosis higher than K = 3 expected in a normal distribution, while in other stocks there is also a marked asymmetry. Moreover, to clarify the difference between a stock price distribution and a normal distribution I have also inserted in the plot the Kernel density of the residuals (in black) and the Normal density (in red).

**Figure 3.** *Histogram of residuals.*

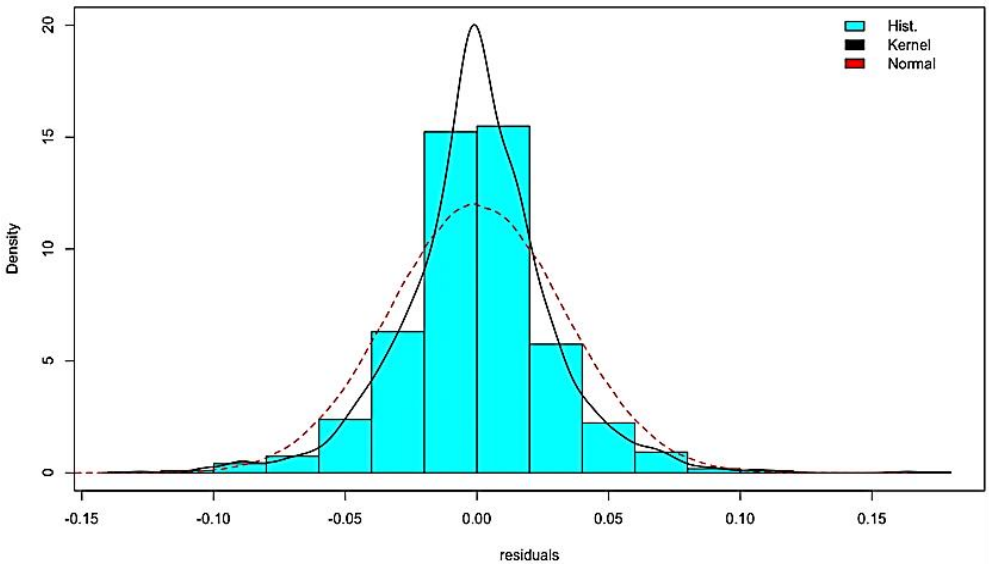
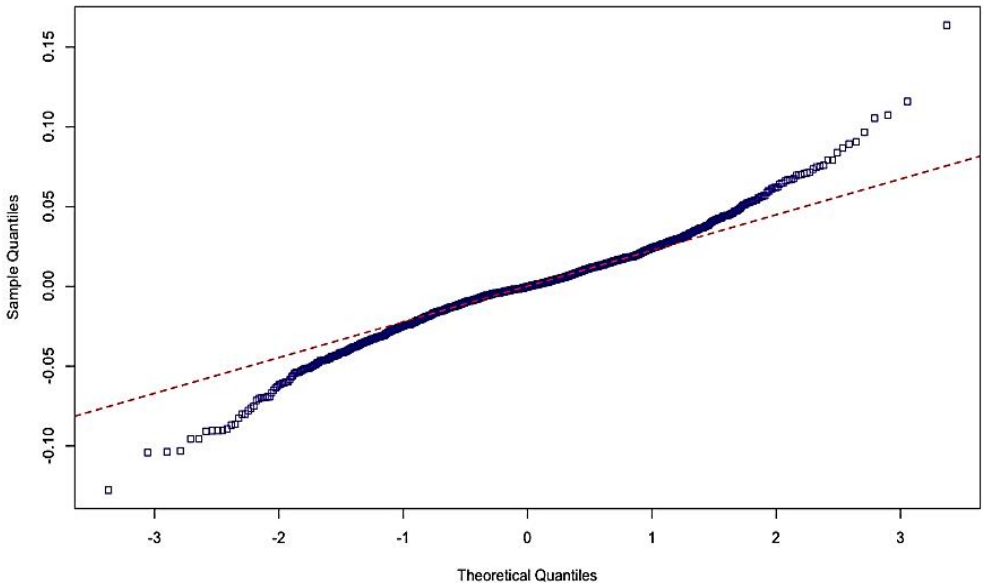


**Figure 4.** *Quantile-quantile plot of the residuals.*

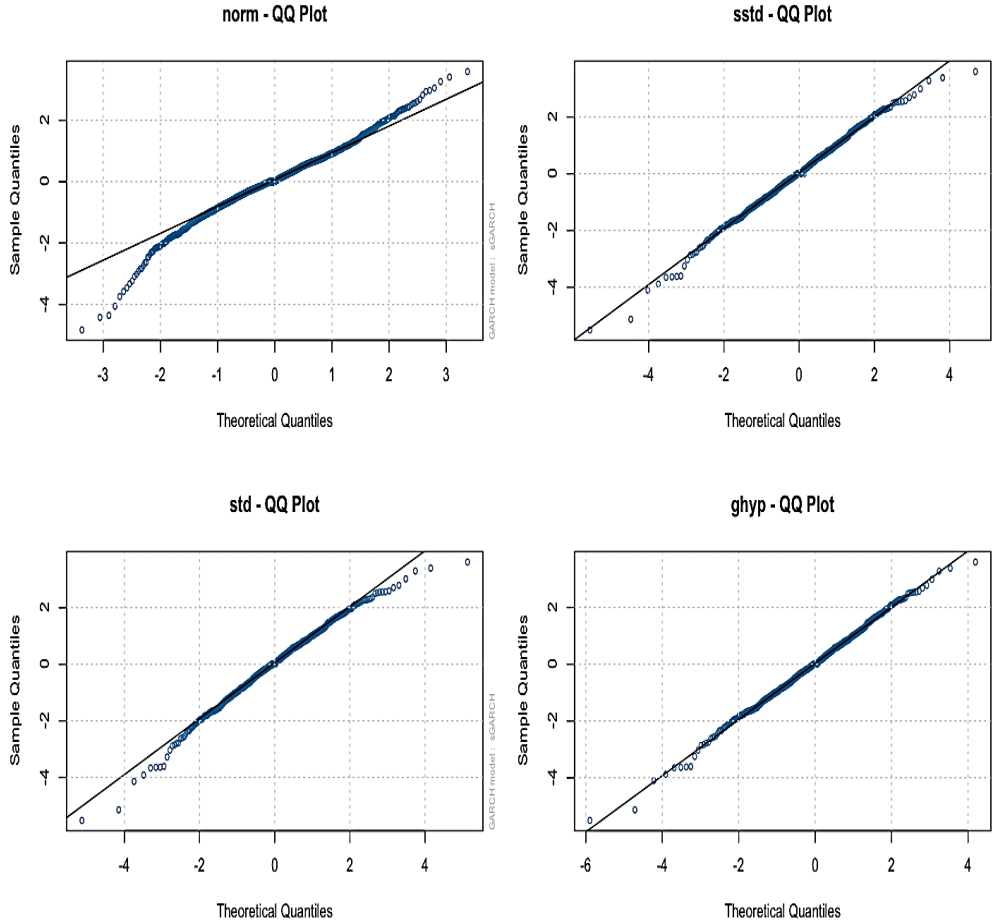



In the Normal qq-plot it is evident that the points fall along a line in the middle of the graph, but curve off in the extremities. Normal qq-plots that exhibit this behaviour usually mean data have more extreme values than would be expected if they truly came from a normal distribution. It is evident that the normality assumption is not suitable. We need to figure out a model that solve these problems. We use a GARCH(1,1) in order to solve the volatility clustering problem and overcome the simple random walk:

$$\log(P_t) = \varphi \log(P_{t-1}) + \varepsilon_t$$
$$\sigma^2 = \alpha + \alpha\log(P_{t-1}^2 + \beta\sigma_{t-1}^2$$
$$\varepsilon_t \sim IIDN(0, \sigma_t^2)$$

We verify with the Dickey Fuller standard test if $\varphi_1 \simeq 1$ and it results significant. Then, thanks to the study of the QQ-plots (but also with the Kolmogorov-Smirnov test) we choose the best probability density function fitting to data. The rugarch package in R has been really helpful to do this. The options considered are: normal, t-student (std), asymmetric t-student (sstd) and generalised hyperbolic distribution (ghyp). In Figure 5 it is possible to see how the generalised hyperbolic distribution captures every estimated point, in particular the extreme values in the tails, as well as the little asymmetry and the elevated kurtosis that caused us problems.

**Figure 5.** *The distributions under investigation.*



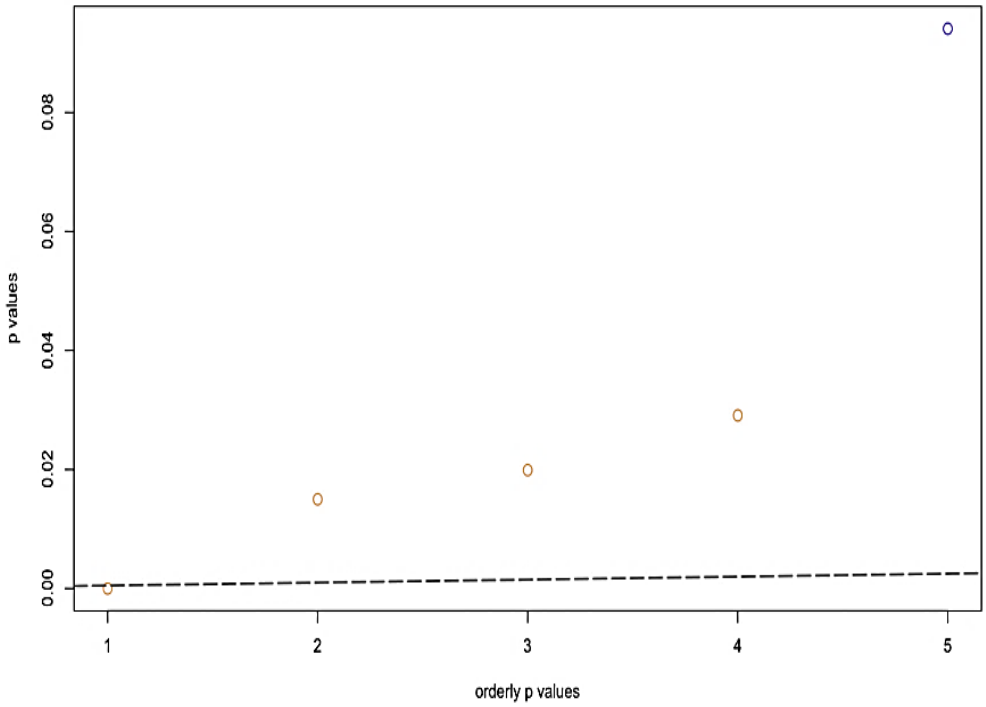

Therefore, we replace $\varepsilon_t \sim IIDN(0, \sigma_t^2)$ with Z~ghyp(ζ,ρ). Where z is the standardised random variable of $\varepsilon_t$, while the parameters ζ and ρ capture the symmetry and the kurtosis of the empirical PDF. Then, we run the BH procedure, we see that no one of the one hundred p-values considered is significant, neither the two smaller than 0,01. At this moment we need to increase the number of the stocks in the analysis, in particular using the second law of the improbability principle, the law of truly large numbers. If we increase the number of possibilities, we will expect to see a significant upward movement due to the random chance. You could consider the previous example, where a group of people tosses coins, now it is the moment to let other people enter in the room and allow them to play.

In order to realise this, we have considered other four sectors of the NASDAQ, which are: Basic industries, Consumer service, Energy and Healthcare. Each data

frame is newly composed of one hundred stocks belonging to the US markets. Now, the procedure previously described is repeated for the four new sectors.

We consider the FDR procedure for each sector separately, because we suppose that stocks from different sectors have different patterns and characteristics. The final result is that of 500 stocks, only two p-values are significant according to the FDR, they belong to the Energy and to the Consumer Service sectors. In Figure 6 is shown the FDR with a close up on the points that cross the line. Of course, the plot is similar for the second sector.

**Figure 6.** *A close-up on the FDR procedure of the Energy sector*.

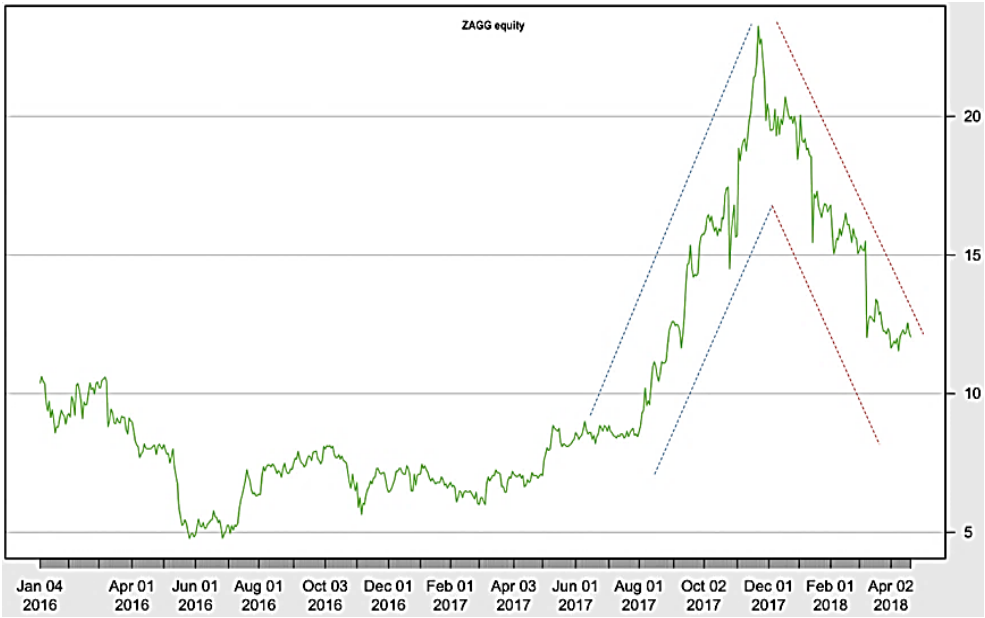



It is evident that from 2017, those stocks are exposed to a long-run upward trend. These are ZAGG and IO stocks and in our model they have produced p-values equal to 0,00035 and 0. A p-value equal to 0.0002 is considered significant with the B-H procedure, while a 0 value means that none of the $P_n^{[i]}$ simulated is bigger than the real price, so we can consider it as a real uncommon and rare upward movement, but we study the trends and we affirm that it is more probably to be an error of our simulation, due to the fact that the majority of past log returns are negative. So, a model based on past data do not fit well for this equity.

## 4. Conclusions

At the beginning of our study, we theorise the market efficiency theory true, in order to explain the randomness of the stock price movements. Well, if it is really true, once we demonstrate with our procedure that the ZAGG equity upward movement is false and due to random, at one point also the market will realise the spuriousness of this stock movement and the price will come back to its average value. Actually, this is exactly what happened the months after our study. In Figure 8, the ZAGG in the period Jan 2016 – Apr 2018.

**Figure 7.** *The ZAGG stock chart*




We believe this particular stock movement is spurious (i.e. it deviates significantly from the random walk) and that it looks like a price bubble (under behavioural finance theory) followed by a correction.

Our main conclusion is that, taking the improbability principle into account, almost all upward price trends could be accounted for within an efficient market theory. Moreover, this technique can also be used to assess the value of a stock price.

We hope that this work could open the way to lots of further very interesting work analysing the effect of the improbability principle in the market place and market response.

## Acknowledgments

This research has been partially financed from the Università degli Studi di Bari Aldo Moro. We would like to thank David Hand, Mikko Pakkanen and in particular Anthony Bellotti from the Imperial College London, for giving us important guidance throughout the project.

## Bibiography

Ghalanos, A. (2018). *Introduction to the rugarch package*. CRAN library, February, 4.

Chatfield, C. (2003). *The Analysis of Time Series: An Introduction*. Sixth Edition. Chapman and Hall/CRC.

Hand, D. (2014). *The improbability principle*. Bantam Press.

Fama, E. F. (1963). Mandelbrot and the Stable Paretian Hypothesis. *Journal of Business*. Vol. 36, No. 4, pp. 420-429.

Fama, E. F. (1965). Random Walks in Stock Market Prices. *Financial Analysts Journal*. Vol. 21, No. 5:55-59.

Pakkanen, M. (2017). *Lecture notes*.

Tsay, R. S. (2010). *Analysis of Financial Time Series*. Third edition. John Wiley & Sons, Inc.

Bollerslev, T. (1986). Generalized Autoregressive conditional heteroskedasticity. *Journal of Econometrics*. Volume 31, Issue 3, April, pp. 307-327.

Benjamini, Y.; Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*. Vol. 57, No. 1 (1995), pp. 289-300.

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018* ISBN: 978-88-6629-048-3
DIPARTIMENTO DI ECONOMIA E FINANZA pp. 35-44

# The What-If Analysis: an Application of the *Sensitivity Analysis* to Financial Derivatives

**Crescenzio Gallo[1*], Alessandro Rinaldi[2]**

[1]*University of Foggia, Italy,* [2]*University of Bari "Aldo Moro", Italy*

**Abstract:** For some time now, the issue of cost-benefit analysis has been one of the major challenges facing researchers and scientists around the world. In the current economic context of growing competitiveness and dynamism, it is essential to know the impact of financial policies or programmes that may be difficult to measure or predict. Sensitivity Analysis is a set of techniques that can be used to examine the degree of uncertainty in cost-benefit analysis (CBA) and how the latter affects the results of a study. This technique is based on the evaluation of a series of possible scenarios and the verification of the final results obtained with the variation of the individual variables. In this way, it is possible to show how these solutions are reactive and at the same time sensitive to the variations of the values of specific model's input variables.
This paper presents the *"What ... If"* analysis applied to the financial field, in particular to Credit Derivatives.

**Keywords:** Sensitivity Analysis, Credit Derivatives, First-to-Default Basket.

## 1. Introduction

Sensitivity Analysis (Saltelli *et al*., 2004; Saltelli *et al*., 2008) enhances simple evaluation since it provides a range of values rather than specific data. This allows a more comprehensive analysis, examining both positive and negative scenarios. It concerns the study of variations in the company's economic value as one or more of the parameters that determine it change. Generally, the parameters considered are those used to define prospective financial statements, such as, for example:

[*] Corresponding author: crescenzio.gallo@unifg.it

- the turnover rate of the invested capital;
- the depreciation rate;
- the cost of debt;
- the margin on sales;
- the tax rate.

In order to evaluate a correct sensitivity analysis, it is advisable to define, in addition to the base scenario, at least a couple of cases of improvement and a couple of cases of worsening. One or more Sensitivity Analyses allow to identify a range of values instead of having only one number available, to determine the decisions and operational choices. Sensitivity Analysis then allows the negative scenarios to be considered so that it is possible to estimate what the safety margins are in the event of a worsening of the market context. Finally, using this technique, it is possible to identify the most important drivers of value on which attention must be focused once the investment has been made.

In the following sections the Scenario Analysis is presented in detail, applying it to Credit Derivatives and credit risk with some explanatory examples and conclu–sions.

## 2. The Scenario Analysis: a support for investment decisions

The Scenario Analysis (Gao *et al.*, 2016) can be explained by means of a logical-mathematical model that the analyst, manager, scientist or engineer develops in order to predict the behavior of reality, predict the trend of a market or make important decisions regarding an investment. The Scenario Analysis allows you to answer the question: "What happens to the final output when one or more input parameters change?". It is therefore a set of mathematical methodologies (Saltelli *et al*., 2008; Winston *et al*., 2004) that allow, through the answer to questions such as *"What ... If"* or *"If ... Then ... Else"*, to open the black box (Fig. 1) and see what happens inside it so as to understand the underlying algorithmic procedure through which a certain final result is obtained.

The Scenario Analysis allows us to describe the possible evolution of economic-financial phenomena by hypothesizing future alternative trends (the so-called *scenarios*) of certain guiding variables, basing the forecast on prepositions such as: "*If ... Then*" and "*What ... If* ". A scenario represents an initial state and a sequence of events that describes the key to the driving forces and the changes that lead to an image of the future.

**Figure 1.** *The Black Box.*

This methodology not only supports but at the same time integrates the various data simulation techniques, meaning by simulation that reality model that allows to evaluate, describe and predict the dynamism of a sequence of events or processes that follow the application of certain constraints and conditions set by the analyst or the user. To this end, it is necessary and important that certain conditions, such as the right to use this methodology, are met in order to ensure its proper use:

- Diversification of scenarios corresponding to different degrees of trend of the economic state being studied by the expert;
- Presence of a not high number of input variables compared to the number of outputs generated during the various simulations;
- Detailed representation of the various future economic and financial situations (the "scenarios") by the input variables used in all their possible and multiple combinations.

The estimation of the multifaceted and multiple future realizations of the input variables on the basis of alternative scenarios allows to reach a better understan–ding of the various effects between the outputs obtained from the various simu–lations and those that define the scenarios themselves. Generally, each scenario is associated with a probability of occurrence that reflects the greater or lesser credibility that the contingencies and or assumptions formulated "ex ante" actually occur.

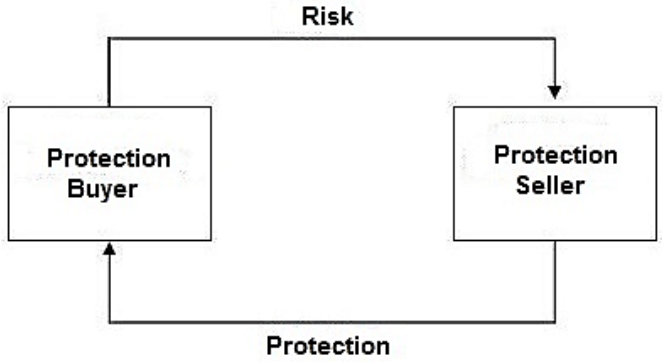
## 3. Credit derivatives and credit risk

In the financial field, as already mentioned in the previous section, there are Financial Derivatives, i.e. financial instruments whose value derives from that of an underlying asset. Im other words, the price of a derivative depends on the value of a certain asset that you do not own in such a way as to be able to negotiate the risk, without buying it directly on the market.

Underlying assets may be of a financial (such as equities, interest and exchange rates, indexes) or a real (such as coffee, cocoa, gold, oil, etc.) nature. Derivatives are basically nothing more than real bets made by investors (banks, companies and

countries) in order to guarantee themselves mainly against risks deriving from the volatility of market variables (the so-called *financial risks*) considering them separately from the operation that originally generated them (Fig. 2).

**Figure 2**. *Insurance purpose of financial derivatives*.




Among the various types of risk against which financial derivatives are issued, credit risk is distinguished (Saunders *et al*., 1997), which gives these financial instruments the name of "Credit Derivatives" (Schönbucher, 2003; Chacko *et al.*, 2015). Credit risk reflects the possibility that an unexpected change in the creditworthiness of a counterpart, to which an exposure exists, generates a corresponding unexpected change in the value of the credit position market.

In addition to the financial scandals that have tormented the national and international economic-financial panorama (Markham, 2015; MacDonald, 2017), in recent years the derivatives market has been characterized by astonishing exponential growth, reaching enormous figures (Bomfim, 2015).

In fact, derivatives have become the main elements for an efficient allocation of capital and for the redistribution of risks in a situation of great uncertainty. In a financial system such as the current one, derivatives help to increase efficiency, contributing to reduce transaction costs and to form prices that readily incorporate information about economic trends (Rösch *et al*., 2013). In such a system, the effects of derivatives are not defined only in the financial sphere; in fact, they are able to influence the investment and financing choices of agents and therefore capable of producing effects for the real economy.

Over the years, economic operators such as banks and industrial and commercial companies, having realized the enormous potential for profits from financial derivatives, have started to invest on all fronts. For banks, derivatives are a fundamental tool for managing credit risk efficiently and effectively, enabling

them to deal with risk positions that are difficult to manage and, at the same time, reducing exposure to parties with whom they have reached too large a size. Since 2008, with the economic crisis, many economists and scholars have pointed the finger at derivative financial instruments, considering them the cause of all evils; in reality, the underlying problem does not lie in the instrument as an end in itself, but in the reckless and unscrupulous use that has been made of it over the last few years (Dickinson, 1997).

## 4. Sensitivity Analysis applied to credit derivatives

Within the credit derivatives there are various subtypes of derivative contracts on credits; but specifically the efficient border will be analyzed only on a particular category, that is: "*First to default Basket Product*" (Lichters *et al*., 2015). To this end, several scenarios will be examined, showing the variation of the efficient border as the variables (Hazard rate *h*, maturity date *T*, risk-free interest rate *r*, number of assets *n* and Recovery Rate *RR*) determining its final premium change. The latter represents the expected return, which is achieved by means of the formula:

$$P = L \cdot (1 - RR) \cdot \frac{H}{r + H} \cdot \left[1 - e^{-T(r+H)}\right]$$

where $L$ is the nominal value of the credit, $(1-RR)$ indicates the loss in the event of default, $H/(r+H)$ is the discounting factor and $1-e^{-T(r+H)}$ highlights the total risk rate (default).

Suppose we are faced with a series of scenarios concerning "Credit Default Products". As shown by the various graphs relating to the various scenarios (see Figures 3–7), it is possible to observe how, when considering the variables *h*, *T* and *n* in the medium/long term and keeping the others constant, the protection buyer has convenience in investing in derivatives, as he sees his premium increase to 1711.604 in the case of a change in the Hazard Rate, 1371.011 in the case of a change in the time factor and finally 934.70 in the case of a change in the number of underlying assets.

Conversely, in the event of an increase in the risk-free interest rate and the Recovery Rate over time, the investor is faced with two negative scenarios, which see the premium fall to 62.37 in the first case and 569.78 in the second, leading the same to desist on a possible investment in shares that would see its return fall with a much higher risk than other types of scenarios.

**Figure 3**. *Scenario 1 — Sensitivity Analysis with respect to the Hazard rate.*

CREDIT DEFAULT PRODUCT: "WHAT IF ANALYSIS" COMPARED TO THE *HAZARD RATE*

| | | BASIC HYPOTHESIS | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOMINAL VALUE OF CREDITS | L | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 |
| RECOVERY RATE | RR | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% |
| RATE OF INTEREST RISK FREE | r | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% |
| NUMBER OF UNDERLYING ACTIVITIES | n | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| HAZART RATE | h | 0,100 | 0,010 | 0,020 | 0,030 | 0,040 | 0,050 | 0,060 | 0,070 | 0,080 | 0,090 | 0,100 |
| HAZARD CUMULATIVE RATE (n*h) | H | 0,700 | 0,070 | 0,140 | 0,210 | 0,280 | 0,350 | 0,420 | 0,490 | 0,560 | 0,630 | 0,700 |
| DEADLINE | T | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| PRIZE | | 1.711,604 | 365,4199 | 662,4932 | 904,0818 | 1100,621 | 1260,578 | 1390,824 | 1496,934 | 1583,434 | 1653,996 | 1711,604 |

| HAZARD RATE | PRIZE |
|---|---|
| 0,010 | 365,4199 |
| 0,020 | 662,4932 |
| 0,030 | 904,0818 |
| 0,040 | 1100,621 |
| 0,050 | 1260,578 |
| 0,060 | 1390,824 |
| 0,070 | 1496,934 |
| 0,080 | 1583,434 |
| 0,090 | 1653,996 |
| 0,100 | 1711,604 |

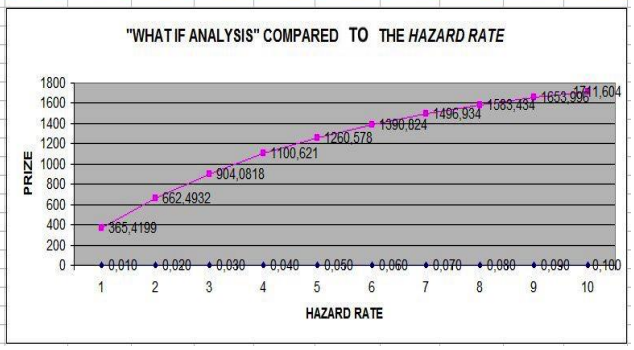



**Figure 4**. *Scenario 2 — Sensitivity Analysis with respect to the deadline.*

CREDIT DEFAULT PRODUCT: "WHAT IF ANALYSIS" RESPECT TO *DEADLINE*

| | | BASIC HYPOTHESIS | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOMINAL VALUE OF CREDITS | L | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 |
| RECOVERY RATE | RR | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% |
| RATE OF INTEREST RISK FREE | r | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% |
| NUMBER OF UNDERLYING ACTIVITIES | n | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| HAZART RATE | h | 0,100 | 0,010 | 0,020 | 0,030 | 0,040 | 0,050 | 0,060 | 0,070 | 0,080 | 0,090 | 0,100 |
| HAZARD CUMULATIVE RATE (n*h) | H | 0,700 | 0,070 | 0,140 | 0,210 | 0,280 | 0,350 | 0,420 | 0,490 | 0,560 | 0,630 | 0,700 |
| DEADLINE | T | 3 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| PRIZE | | 1.478,003 | 258,0952 | 476,9362 | 662,4932 | 819,8286 | 953,2346 | 1066,351 | 1162,263 | 1243,587 | 1312,543 | 1371,011 |

| DEADLINE | PRIZE |
|---|---|
| 1 | 258,0952 |
| 2 | 476,9362 |
| 3 | 662,4932 |
| 4 | 819,8286 |
| 5 | 953,2346 |
| 6 | 1066,351 |
| 7 | 1162,263 |
| 8 | 1243,587 |
| 9 | 1312,543 |
| 10 | 1371,011 |

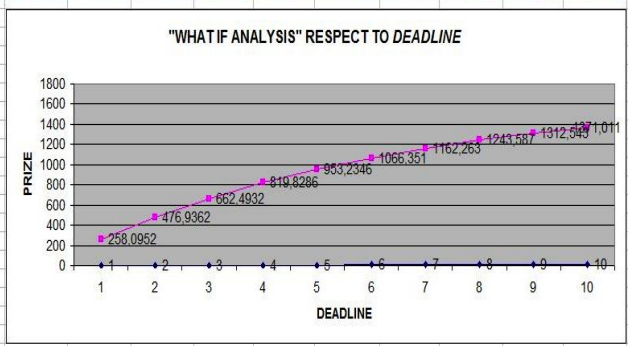


**Figure 5**. *Scenario 3 — Sensitivity Analysis with respect to the risk-free interest rate.*

CREDIT DEFAULT PRODUCT: "WHAT IF ANALYSIS" COMPARED TO THE *RATE RISK FREE*

| | | BASIC HYPOTHESIS | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOMINAL VALUE OF CREDITS | L | 5.000.000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 |
| RECOVERY RATE | RR | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% |
| RATE OF INTEREST RISK FREE | r | 2,50% | 2,00% | 2,50% | 3,00% | 3,50% | 4,00% | 4,50% | 5,00% | 5,50% | 6,00% | 6,50% |
| NUMBER OF UNDERLYING ACTIVITIES | n | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| HAZART RATE | h | 0,100 | 0,010 | 0,020 | 0,030 | 0,040 | 0,050 | 0,060 | 0,070 | 0,080 | 0,090 | 0,100 |
| HAZARD CUMULATIVE RATE (n*h) | H | 0,700 | 0,070 | 0,140 | 0,210 | 0,280 | 0,350 | 0,420 | 0,490 | 0,560 | 0,630 | 0,700 |
| DEADLINE | T | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| PRIZE | | 1.711,604 | 667,0752 | 662,4932 | 657,955 | 653,4599 | 649,0075 | 644,5974 | 640,229 | 635,902 | 631,6159 | 627,3702 |

| RATE RISK FREE | PRIZE |
|---|---|
| 2,00% | 667,0752 |
| 2,50% | 662,4932 |
| 3,00% | 657,955 |
| 3,50% | 653,4599 |
| 4,00% | 649,0075 |
| 4,50% | 644,5974 |
| 5,00% | 640,229 |
| 5,50% | 635,902 |
| 6,00% | 631,6159 |
| 6,50% | 627,3702 |

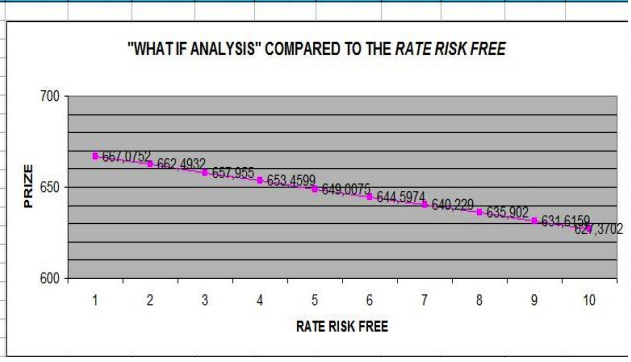



**Figure 6**. *Scenario 4 — Sensitivity Analysis with respect to the number of assets.*

CREDIT DEFAULT PRODUCT: "WHAT IF ANALYSIS" RESPECT TO THE *UNDELYING ACTIVITIES*

| | | BASIC HYPOTHESIS | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NOMINAL VALUE OF CREDITS | L | 5.000.000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 |
| RECOVERY RATE | RR | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% | 60% |
| RATE OF INTEREST RISK FREE | r | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% |
| NUMBER OF UNDERLYING ACTIVITIES | n | 5 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| HAZART RATE | h | 0,100 | 0,010 | 0,020 | 0,030 | 0,040 | 0,050 | 0,060 | 0,070 | 0,080 | 0,090 | 0,100 |
| HAZARD CUMULATIVE RATE (n*h) | H | 0,700 | 0,070 | 0,140 | 0,210 | 0,280 | 0,350 | 0,420 | 0,490 | 0,560 | 0,630 | 0,700 |
| DEADLINE | T | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| PRIZE | | 1.510,397 | 218,0292 | 317,376 | 411,714 | 500,2944 | 583,7895 | 662,4932 | 736,6827 | 806,6189 | 872,5475 | 934,7001 |

| NUMBER OF UNDERLYING ACTIVITIES | PRIZE |
|---|---|
| 2 | 218,0292 |
| 3 | 317,376 |
| 4 | 411,714 |
| 5 | 500,2944 |
| 6 | 583,7895 |
| 7 | 662,4932 |
| 8 | 736,6827 |
| 9 | 806,6189 |
| 10 | 872,5475 |
| 11 | 934,7001 |

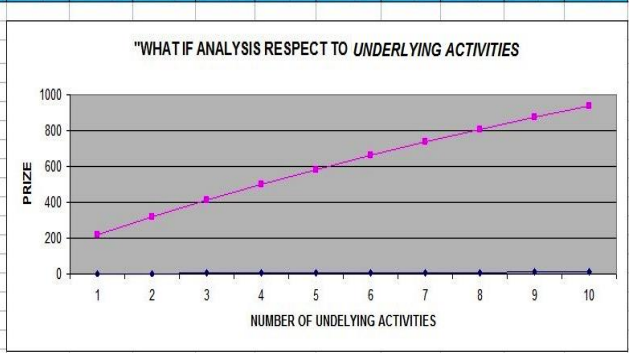


**Figure 7**. *Scenario 5 — Sensitivity Analysis with respect to Recovery Rate.*

| CREDIT DEFAULT PRODUCT: "WHAT IF ANALYSIS" RESPECT TO *RECOVERY RATE* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | BASIC HYPOTHESIS | | | | | | | | |
| NOMINAL VALUE OF CREDITS | L | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 | 5.000,000 |
| RECOVERY RATE | RR | 60,00% | 12,76% | 13,50% | 15,00% | 18,20% | 22,50% | 28,60% | 35,00% | 65,00% |
| RATE OF INTEREST RISK FREE | r | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% | 2,50% |
| NUMBER OF UNDERLYING ACTIVITIES | n | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| HAZART RATE | h | 0,100 | 0,010 | 0,020 | 0,030 | 0,040 | 0,050 | 0,060 | 0,070 | 0,080 |
| HAZARD CUMULATIVE RATE (n*h) | H | 0,700 | 0,070 | 0,140 | 0,210 | 0,280 | 0,350 | 0,420 | 0,490 | 0,560 |
| DEADLINE | T | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 |
| PRIZE | | 1.711,604 | 1444,898 | 1432,642 | 1407,798 | 1354,799 | 1283,581 | 1182,55 | 1076,552 | 579,682 |

| RECOVERY RATE | PRIZE |
|---|---|
| 12,76% | 1444,898 |
| 13,50% | 1432,642 |
| 15,00% | 1407,798 |
| 18,20% | 1354,799 |
| 22,50% | 1283,581 |
| 28,60% | 1182,550 |
| 35,00% | 1076,552 |
| 65,00% | 579,682 |

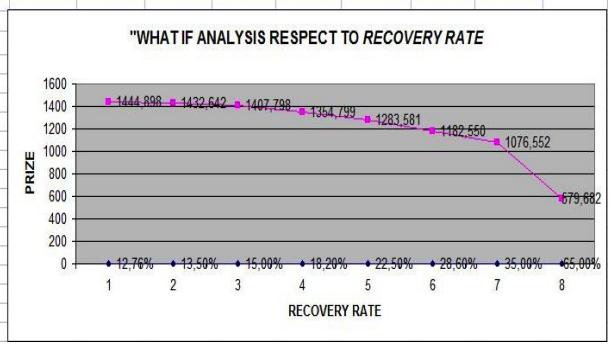



Table 1 shows the percentage variations of the variables and the incidence of the premium on each of them. Fig. 8 also shows the trend in the incidence of the premium *P* for each variable, which is also more significant and symptomatic for the variables *h*, *n*, *T* and less eloquent for the variable *r*, assuming values close to zero. In the case of the RR (Recovery Rate) variable, however, it can be seen that the curve's trend assumes a positive upsurge in the last phase after a sequence of negative values and this is therefore a sign of the strong volatility of this factor as regards the return on investment and the rate associated with it.

**Figure 8**. *Incidence of the "First to Default Basket Products" prize for each input variable.*

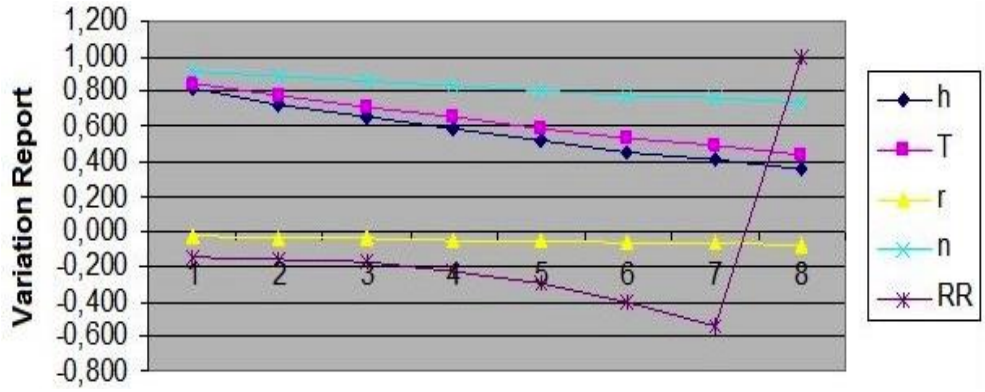


***Table 1***. *Variables' percentage variations and incidence of the premium on each of them.*

| CREDIT DEFAULT PRODUCT | *Statement of Variation* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| % variation of *h* | 100.00 | 50.00 | 33.33 | 25.00 | 20.00 | 16.67 | 14.29 | 12.50 |
| % variation of *P* | 81.30 | 36.47 | 21.74 | 14.53 | 10.33 | 7.63 | 5.78 | 4.46 |
| | | | | | | | | |
| % variation of *T* | 100.00 | 50.00 | 33.33 | 25.00 | 20.00 | 16.67 | 14.29 | 12.50 |
| % variation of *P* | 84.79 | 38.91 | 23.75 | 16.27 | 11.87 | 8.99 | 7.00 | 5.54 |
| | | | | | | | | |
| % variation of *r* | 25.00 | 20.00 | 16.67 | 14.29 | 12.50 | 11.11 | 10.00 | 9.09 |
| % variation of *P* | -0.69 | -0.69 | -0.68 | -0.68 | -0.68 | -0.68 | -0.68 | -0.67 |
| | | | | | | | | |
| % variation of *n* | 50.00 | 33.33 | 25.00 | 20.00 | 16.67 | 14.29 | 12.50 | 11.11 |
| % variation of *P* | 45.73 | 29.58 | 21.52 | 16.69 | 13.48 | 11.20 | 9.49 | 8.17 |
| | | | | | | | | |
| % variation of *RR* | 5.80 | 11.11 | 21.33 | 23.63 | 27.11 | 22.38 | 85.71 | -100.00 |
| % variation of *P* | -0.85 | -1.73 | -3.76 | -5.26 | -7.87 | -8.96 | -46.15 | -100.00 |
| | | | | | | | | |
| IMPACT OF *P* ON VARIABLES | | | | | | | | |
| ***H*** | 0.813 | 0.729 | 0.652 | 0.581 | 0.517 | 0.458 | 0.404 | 0.357 |
| ***T*** | 0.848 | 0.778 | 0.712 | 0.651 | 0.593 | 0.540 | 0.490 | 0.444 |
| ***R*** | -0.027 | -0.034 | -0.041 | -0.048 | -0.054 | -0.061 | -0.068 | -0.074 |
| ***N*** | 0.915 | 0.887 | 0.861 | 0.834 | 0.809 | 0.784 | 0.759 | 0.736 |
| ***RR*** | -0.146 | -0.156 | -0.176 | -0.222 | -0.290 | -0.401 | -0.538 | 1.000 |

## 5. Conclusion

The procedure illustrated is interesting in order to determine the best possible result on the basis of a series of scenarios, hypothesized as the individual variables characterizing financial derivatives change. In more detail, the objective was to apply the Sensitivity Analysis to a particular type of financial derivatives, the "*First to Default Basket Products*", so as to understand which was the most optimal scenario to focus on.

The problem has been tackled from a purely conceptual point of view, giving a theoretical contribution that can be the fruit and cue for further future developments.

## References

Bomfim, A. N. (2015). *Understanding credit derivatives and related instruments*. Academic Press.

Chacko, G.; Sjöman, A.; Motohashi, H.; Dessain, V. (2015). *Credit Derivatives, Revised Edition: A Primer on Credit Risk, Modeling, and Instruments*. FT Press.

Dickinson, A. (1997). Restitution and incapacity: A choice of law solution. *RLR*, 5:66.

Gao, L.; Bryan, B. A.; Nolan, M.; Connor, J. D.; Song, X.; Zhao, G. (2016). Robust global sensitivity analysis under deep uncertainty via scenario analysis. *Environmental Modelling & Software*, 76:154–166, 2016.

Lichters, R.; Stamm, R.; Gallagher, D. (2015). Pricing portfolio credit products. In: *Modern Derivatives Pricing and Credit Exposure Analysis*. Springer: 333–350.

MacDonald, S. B. (2017). *Separating fools from their money: a history of American financial scandals*. Routledge.

Markham, J. W. (2015). *A financial history of modern US corporate scandals: From Enron to reform*. Routledge.

Rösch, D.; Scheule, H. (2013). *Credit Securitisations and Derivatives: Challenges for the Global Markets*. John Wiley & Sons.

Schönbucher, P. J. (2003). *Credit derivatives pricing models: models, pricing and implementation*. John Wiley & Sons.

Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S. (2008). *Global sensitivity analysis: the primer*. John Wiley & Sons.

Saunders, A.; Thomas, H. (1997). *Financial institutions management*. Irwin.

Saltelli, A.; Tarantola, S.; Campolongo, F.; Ratto, M. (2004). *Sensitivity analysis in practice: a guide to assessing scientific models*. John Wiley & Sons.

Winston, W. L.; Goldberg, J. B. (2004). *Operations research: applications and algorithms, volume 3*. Thomson-Brooks/Cole, Belmont, CA.

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018* ISBN: 978-88-6629-048-3
DIPARTIMENTO DI ECONOMIA E FINANZA pp. 45-56

# Customer Relationship Management: il caso di BPPB

**Francesco D. d'Ovidio[1*], Najada Firza[2], Dante Mazzitelli[1]**

[1] *Università degli studi di Bari Aldo Moro (Italy),*
[2] *Università Cattolica "Zoja e Këshillit të Mirë", Tirana (AL)*

**Riassunto:** Il presente lavoro si pone alla base dei meccanismi decisionali della Cusomer Relationship Management bancario per la fidelizzazione della clientela, che implicano un focus sulle caratteristiche della clientela medesima. L'accurata compilazione del questionario MIFID da parte del cliente è il primo passo verso la raccolta dati nel processo di profilazione della clientela.
In questo studio sono presi in considerazione appunto i dati provenienti dai que–stionari MIFID compilati dalla clientela di un istituto bancario localmente molto importante. I dati sono stati dapprima analizzati tramite analisi statistiche descritti–ve e successivamente sono state avviate analisi con tecniche multivariate di esplo–razione dati. Nello specifico, utilizzando in modo esplorativo-confermativo la *clu–ster analysis*, innanzitutto si suddividono i clienti in gruppi con caratteristiche si–mili al loro interno e in un secondo momento se ne identificano le peculiarità.

**Keywords:** CRM; profilazione della clientela; MIFID; Cluster Analysis; k-means.

## 1. Introduzione

Oggigiorno le banche assomigliano sempre più a vere e proprie imprese, ed anche i prodotti e servizi che le banche offrono sono oculatamente diversificati e personalizzati: ciò perché nel corso del tempo è stato dimostrato che il profitto "aziendale" è vincolato alla soddisfazione del cliente. Per tale motivo, il rapporto che nasce tra la banca ed il proprio cliente deve essere costantemente alimentato cercando di raggiungere un elevato grado di soddisfazione della clientela.

---

[*] Autore corrispondente: francescodomenico.dovidio@uniba.it
Il lavoro qui descritto è frutto di un progetto comune, ma F.D. d'Ovidio ha provveduto alla redazione del § 3, N. Firza ha redatto i § 2 e 4, mentre D. Mazzitelli. ha redatto i § 1 e 5.

Il CRM (*customer relationship management*) rappresenta oggi la leva strategica tramite la quale arrivare all'ottimizzazione del profitto aziendale. Quando il cliente è soddisfatto dei servizi e dei prodotti che l'azienda (la banca) offre, difficilmente si rivolge a un'azienda (banca) concorrente.

Questa ricerca si propone di identificare il profilo della clientela allo scopo di evidenziare le diverse strategie del marketing bancario per la fidelizzazione ed il mantenimento del portafoglio clienti. Partendo dall'analisi dei dati pervenuti dalla compilazione del questionario MIFID[1] andremo ad identificare il profilo della clientela per meglio soddisfare le sue esigenze e nel contempo ridurre il rischio di gestione patrimoniale intrinseco della banca stessa.

La metodologia qui utilizzata è la *cluster analysis*[2], che in un primo momento viene adoperata in modalità esplorativa, con tecniche gerarchiche, per identificare il numero di cluster più adeguato per la partizione ottimale nella successiva analisi "confermativa" del processo con metodi non gerarchici (*k-means*); di seguito si procede quindi a elaborare i risultati definiti di *profilazione della clientela*.

## 2. Materiali e metodi

I dati a disposizione sono stati forniti dalla Banca Popolare di Puglia e Basilicata; essi trattano le risposte della clientela al questionario MIFID negli anni che vanno dal 1997 al 2010, nonché il loro livello di fidelizzazione negli anni successivi. Nello specifico, si tratta di 250 clienti campionati fra tutti coloro che hanno usufruito del servizio di gestione patrimoniale della banca.

Il questionario MIFID considerato è la versione che la Banca Popolare di Puglia e Basilicata ha revisionato nel 2007. Esso comprende cinque sezioni o dimensioni:

- L'esperienza finanziaria;
- L'obiettivo temporale;
- Gli obiettivi di investimento;
- La situazione finanziaria;
- La propensione al rischio.

[1] La Commissione Europea ha adottato il piano di azione per i servizi finanziari dal maggio 1999, comprendendo il MIFID che entrato in vigore nel novembre 2007. La direttiva si rivolge ai singoli investitori che vogliono investire in prodotti finanziari utilizzando i canali di intermediazione finanziaria. Lo scopo del MIFID è di proteggere e salvaguardare il rapporto tra la banca ed il cliente individuando un profilo di rischio del cliente.

[2] Il termine "cluster analysis" fu introdotto nel 1932 dallo statistico statunitense Robert Tryon.

**Tabella 1.** *Distribuzione della clientela bancaria per consistenza patrimoniale e classe di età.*

| Classe di età | Consistenza del patrimonio (in migliaia di €) | | | | | Totale |
|---|---|---|---|---|---|---|
| | meno di 50 | 50 -100 | 100-500 | 500-1000 | oltre 1000 | |
| < 35 | 1 | 2 | 2 | 4 | 0 | 9 |
| 35-50 | 3 | 5 | 39 | 16 | 14 | 77 |
| 50-65 | 2 | 9 | 41 | 22 | 19 | 93 |
| > 65 | - | 13 | 30 | 12 | 16 | 71 |
| Totale | 6 | 29 | 112 | 54 | 49 | 250 |

In un primo momento effettuiamo un'analisi descrittiva dei dati a disposizione. La Banca vuole adottare una strategia di fidelizzazione della clientela tramite il CMR[3] definendo la "clientela coltivabile", ossia i clienti più importanti, e la "clientela di primo piano" ossia i clienti che hanno più rapporti con la stessa banca.

Su 250 clienti, suddivisi per classe di età, osserviamo, come illustrato nella tabella 1, la distribuzione della consistenza patrimoniale della clientela.

In base a considerazioni sia di ordine aziendale (identificazione della migliore clientela) che di ordine statistico (riduzione degli *outliers*), si è quindi scelto di focalizzare il target di clientela bancaria in base alla consistenza patrimoniale, eliminando dal dataset i soggetti con una consistenza patrimoniale minore di 100mila euro.

Andiamo ora ad effettuare un'analisi descrittiva delle caratteristiche della *clientela target* rispetto alla consistenza patrimoniale.

Innanzitutto risulta rilevante la relazione tra la consistenza patrimoniale ed il livello di istruzione: la tabella 2 evidenzia, infatti, come il 43% dei clienti con una consistenza patrimoniale di oltre un milione di € è laureato, mentre il 46% dei clienti che hanno una consistenza patrimoniale tra 100mila e 500mila € ha un istruzione a livello di scuola media superiore, così come il 52% dei clienti con una consistenza patrimoniale tra 500mila e un milione di €.

Un altro dato interessante che si riscontra in tabella 2 è che oltre il 60% dei clienti ha un livello di istruzione pari o superiore alla secondaria di II grado (scuola media superiore) in tutti i gruppi distinti per consistenza patrimoniale, ma la massima quota, ben l'85%, compete al gruppo di consistenza patrimoniale da 500mila a un milione di euro.

[3] CRM, come si è accennato introduttivamente, è acronimo di *Customer Relationship Management*, ossia *gestione delle relazioni con la clientela*. Esso si pone quattro principali obiettivi (Duse, 2011):

- il mantenimento della clientela;
- l'incremento dei clienti più importanti;
- la fidelizzazione della clientela che ha maggiori rapporti con la banca:
- la trasformazione degli attuali clienti in veri e propri *promotori* della banca, incoraggiando nuovi clienti grazie all'espressione della propria soddisfazione.

**Tabella 2.** *Distribuzione della clientela bancaria secondo il livello di istruzione, per classe di consistenza patrimoniale.*

| Livello di istruzione | Consistenza del patrimonio in K € | | | Totale |
|---|---|---|---|---|
| | 100-500 | 500-1000 | oltre 1000 | |
| Nessuno | 1 (0,9%) | - | - | 1 |
| Licenza Elementare | 18 (16%) | 3 (6%) | 5 (10%) | 26 |
| Diploma Scuola Media Inferiore | 25 (22%) | 5 (9%) | 13 (27%) | 43 |
| Diploma di Scuola Media Superiore | 51 (45%) | 28 (52%) | 10 (20%) | 89 |
| Laurea | 17 (16%) | 18 (33%) | 21 (43%) | 56 |
| Totale | 112 (100%) | 54 (100%) | 49 (100%) | 215 |

**Tabella 3.** *Distribuzione della clientela bancaria secondo il livello di conoscenza degli strumenti finanziari, per classe di consistenza patrimoniale.*

| Livello di conoscenza degli strumenti finanziari | Consistenza del patrimonio in K € | | | Totale |
|---|---|---|---|---|
| | 100-500 | 500-1000 | oltre 1000 | |
| Basso | 21 (19%) | 2 (4%) | 6 (12%) | 29 |
| Medio | 57 (51%) | 30 (56%) | 25 (51%) | 112 |
| Alto | 34 (30%) | 22 (40%) | 18 (37%) | 74 |
| Totale | 112 (100%) | 54 (100%) | 49 (100%) | 215 |

La tabella 3 evidenzia, dal canto suo, uno degli elementi più rilevanti nel nostro studio: il livello di conoscenza degli strumenti finanziari dei 215 clienti in base alle dichiarazioni rese nel modello MIFID. Appare evidente che, come è logico, la maggior parte del collettivo dichiara un livello *medio* di conoscenza degli strumenti finanziari, ma si sottolinea che dichiarano un livello *alto* circa il 37% dei clienti con oltre un milione di euro di consistenza patrimoniale e ben il 40% dei clienti con una consistenza patrimoniale tra 500mila € e un milione.

È presumibile che per l'istituto bancario potrebbe essere interessante quantificare precisamente la consistenza patrimoniale di quei clienti (di cui addirittura 6, il 12%, in fascia patrimoniale superiore al milione di euro) che dichiarano una bassa conoscenza degli strumenti finanziari, allo scopo di implementare specifiche strategie informative e di marketing, in particolar modo qualora la somma in oggetto sia rilevante per la stessa banca.

Nella tabella 4, la clientela bancaria viene classificata in base alle operazioni finanziarie effettuate durante l'ultimo anno, evincendo che chi aveva una grande consistenza patrimoniale ma una bassa conoscenza di strumenti finanziari (appunto i sei clienti che avevano dichiarato un basso livello di conoscenza finanziaria in tabella 3) non ha effettuato alcuna operazione finanziaria nell'ultimo periodo. Questo comportamento valorizza ancora di più l'idea che avere una strategia di un compor-

tamento specifico da parte della banca nei confronti di questo gruppo di clienti sia una mossa vincente. La tabella 4 evidenzia, ancora, che il 66% all'interno del gruppo con una consistenza finanziaria maggiore ha effettuato più di 40 operazioni finanziarie, per un valore superiore a 50mila euro.

**Tabella 4.** *Distribuzione della clientela bancaria secondo il numero di operazioni finanziarie effettuate nell'ultimo anno, per classe di consistenza patrimoniale.*

| Operazioni Finanziarie Effettuate nell'ultimo anno | Consistenza patrimonio in K € | | | Totale |
|---|---|---|---|---|
| | 100-500 | 500-1000 | oltre 1000 | |
| Nessuno | 19 (17%) | 5 (9%) | 6 (12%) | 30 |
| Fino 5 operazioni, fino a 10 K€ | 17 (15%) | 8 (15%) | 2 (4%) | 27 |
| Da 10 a 20 operazioni, fino a 30 K€ | 16 (14%) | 10 (19%) | 6 (12%) | 32 |
| Da 20 a 40 operazioni, fino a 50 K€ | 23 (21%) | 5 (9%) | 3 (6%) | 31 |
| Oltre | 37 (33%) | 26 (48%) | 32 (66%) | 95 |
| Totale | 112 (100%) | 54 (100%) | 49 (100%) | 215 |

**Tabella 5.** *Distribuzione della clientela bancaria secondo gli obiettivi di investimento dichiarati, per classe di consistenza patrimoniale.*

| Obiettivi investimenti | Consistenza patrimonio in K € | | | Totale |
|---|---|---|---|---|
| | 100-500 | 500-1000 | oltre 1000 | |
| Conservare i risparmi | 9 (8%) | 3 (6%) | 1 (2%) | 13 |
| Rivalutare i risparmi | 17 (15%) | 7 (13%) | 12 (25%) | 36 |
| Incrementare capitale anche con perdite temporanee | 41 (37%) | 19 (35%) | 12 (25%) | 72 |
| Accrescere sensibilmente il capitale nel medio lungo termine | 30 (27%) | 11 (20%) | 14 (28%) | 55 |
| Incrementare il più possibile il capitale anche sopportando grosse perdite | 15 (13%) | 14 (26%) | 10 (20%) | 39 |
| Totale | 112 (100%) | 54 (100%) | 49 (100%) | 215 |

La tabella 5 tratta la suddivisione del collettivo in base agli obbiettivi di investimento. Si noti che solo una percentuale molto piccola della clientela dichiara obiettivi di semplice *conservazione dei risparmi* (8% del primo gruppo, 6% del secondo gruppo e addirittura il 2% del terzo gruppo, quello con la consistenza patrimoniale maggiore). Nella stessa tabella si evidenzia che, a prescindere dal gruppo di appartenenza, la maggior parte della clientela (oltre il 70%), ha obiettivi di incremento del capitale, come peraltro era prevedibile data popolazione studiata.

Inoltre il 26% dei clienti appartenenti al secondo gruppo (500-1000K €) ed il 20% dei clienti appartenenti al terzo gruppo (oltre 1000K €) si pone come obiettivo quello di incrementare il più possibile il capitale anche sopportando grosse perdite: dunque si tratta di gente disposta a rischiare molto.

La tabella 6 evidenzia, infine, la percentuale del proprio patrimonio che i clienti detengono e investono nella banca BPPB, rispetto al totale del capitale. Per l'Istituto di credito è interessante saper leggere ed interpretare questi dati al fine poter avere un piano di condotta per ogni gruppo della clientela.

In particolare, nella terza classe di consistenza patrimoniale (oltre un milione di €) si contano ben 18 correntisti, ossia il 37%, che investono meno del 25%. Essendo il terzo gruppo quello più abbiente, è bene che la Banca investighi per comprendere il motivo per cui essi investono meno di un quarto del proprio patrimonio e stabilire una adeguata strategia di proposte per incrementare il portafoglio dei prodotti finanziari di questo specifico target.

Solo tre clienti di tale classe di consistenza patrimoniale, peraltro, investono almeno il 75% del loro capitale nella banca, e appena 8 vi detengono tra metà e tre quarto del patrimonio: insomma, meno di un quarto dei clienti più solvibili affida almeno metà del proprio capitale a BPPB, mentre gli altri preferiscono differenziare gli utilizzi. Invece, i clienti con minore consistenza patrimoniale (100-500mila euro) sembra privilegiare l'Istituto in misura maggiore degli altri (16% vi investono almeno tre quarti del patrimonio, 35% fra metà e tre quarti, per un totale pari al 51%).

In generale, quasi il 40% dei clienti investe in BPPB tra un quarto e metà del proprio capitale, senza grandi differenze tra classi di consistenza patrimoniale.

**Tabella 6.** *Distribuzione della clientela bancaria secondo l'incidenza percentuale delle somme detenute in Banca Popolare di Puglia e Basilicata, sul totale del proprio patrimonio, per classe di consistenza patrimoniale.*

| Incidenza % delle somme detenute in BPPB sul totale del patrimonio | Consistenza patrimonio in K € | | | Totale |
|---|---|---|---|---|
| | 100-500 | 500-1000 | oltre 1000 | |
| Almeno 75% | 18 (16%) | 11 (20%) | 3 (6%) | 32 |
| Tra 50% e 75% | 39 (35%) | 7 (13%) | 8 (16%) | 54 |
| Tra 25% e 50% | 42 (38%) | 20 (37%) | 20 (41%) | 82 |
| Meno di 25% | 13 (11%) | 16 (30%) | 18 (37%) | 47 |
| Totale | 112 (100%) | 54 (100%) | 49 (100%) | 215 |

## 3. Cluster Analysis

A partire dai dati fin qui esplorati, è stata avviata una ulteriore esplorazione (multivariata) tramite tecniche gerarchiche di *cluster analysis*.

Com'è noto, la Cluster Analysis è una tecnica esplorativa *simmetrica* (in quanto non pone vincoli gerarchici tra le variabili osservate) il cui scopo è aggregare progressivamente gli $n$ elementi di un certo insieme in base alle loro caratteristiche in una "partizione ottimale" di $k$ gruppi (*cluster*) non predefiniti, il più possibile

omogenei al loro interno ed eterogenei rispetto agli altri: ossia mantenendo il vincolo della minima variabilità *within* delle variabili prese in considerazione e della massima variabilità *between* (cfr., ad es., Delvecchio 2010; Fabbris 1997).

Per la formazione dei cluster occorre individuare:

- le variabili (misurate in qualsiasi scala, anche nominale) che caratterizzano ciascun elemento;
- la "matrice di somiglianze" degli *n* elementi fra loro;
- la tecnica con la quale raggruppare gli elementi;
- il numero dei gruppi (con le tecniche gerarchiche questo elemento può essere valutato "a posteriori", mentre le tecniche gerarchiche richiedono che esso sia prestabilito).

Le variabili considerate, peraltro, hanno tutte il medesimo ruolo e la medesima importanza nella matrice di dati, senza alcun vincolo relazionale l'una dall'altra se non, all'interno dei gruppi finali, il concetto di "minima distanza". Dato che la distanza è uno dei modi con cui si può misurare la concordanza fra variabili, la partizione ottimale può essere interpretata come quella che massimizza la relazione di concordanza fra le variabili per le unità di ogni cluster e minimizza quella fra le determinazioni delle variabili osservate in un cluster e quelle osservate in un altro.

Al fine di avere risultati più attendibili nella suddivisione in gruppi dei nostri dati sono state utilizzate diverse tecniche di agglomerazione gerarchica (metodo di Ward, del legame medio, del legame singolo, del legame completo). Da un confronto incrociato dei risultati ottenuti con le varie tecniche, i risultati più attendibili sono stati riscontrati con il metodo di Ward (1963)[4].

Il metodo di Ward consente di identificare una partizione ottimale di *3 cluster*, sia tramite il *taglio* del dendrogramma (qui non riportato per l'eccessiva articolazione dei 215 "cluster iniziali", ossia i singoli soggetti) e sia con l'osservazione dei coefficienti di distanza calcolati dall'algoritmo di agglomerazione (Zani e Cerioli, 2007), mostrati in tabella 7, dove si nota che, mentre con tre cluster (212.mo stadio) il livello di distanza risulta non molto diverso dai precedenti, la successiva aggregazione avviene a un livello di distanza ben superiore: ciò evidenzia che, in termini di omogeneità interna dei gruppi, la partizione con soli due cluster risulta inadeguata (per non parlare della "partizione banale" composta da un unico cluster, la cui distanza di aggregazione è ancor maggiormente differenziata rispetto alla precedente).

[4] Il metodo di Ward (o *della minima devianza*) è un metodo gerarchico di *Cluster Analysis* che in genere fornisce risultati più attendibili poiché lo stesso considera la distanza minima tra gli elementi di uno stesso cluster (definito calcolando gli incrementi subiti dalla somma dei quadrati delle distanze in ogni aggregazione effettuata) prevalente sulla distanza massima tra i cluster ottenuti, comunque considerata nell'aggregazione.

**Tabella 7.** *Coefficienti di agglomerazione con metodo di Ward e identificazione della partizione ottimale.*

| Stadio | Combinazione Cluster | | Coefficiente di distanza | *k* |
|---|---|---|---|---|
| | Cluster *a* | Cluster *b* | | |
| 1 | 207 | 215 | 0,000 | 215 |
| 2 | 204 | 214 | 0,000 | 214 |
| 3 | 202 | 213 | 0,000 | 213 |
| … | … | ... | ... | ... |
| … | … | … | … | … |
| … | … | … | … | … |
| 210 | 2 | 5 | 160,316 | 5 |
| 211 | 2 | 12 | 180,343 | 4 |
| 212 | 13 | 20 | 214,168 | **3** |
| 213 | 1 | 2 | **324,173** | 2 |
| 214 | 1 | 13 | **440,716** | 1 |

Secondo la classificazione gerarchica di Ward, al *Cluster 1* risultano appartenere 38 unità statistiche (clienti), mentre al *Cluster 2* afferiscono 86 clienti e al terzo cluster 91.

I risultati ottenuti con il metodo di Ward sono stati poi validati con il metodo non gerarchico *k-means*, tramite il quale si ottengono in genere dei risultati più robusti. Il metodo di cluster analysis noto come *k-means* (MacQueen, 1967) è infatti un metodo non gerarchico molto efficiente: in altri termini, al contrario di quanto avviene con i metodi gerarchici, una unità "borderline" che in precedenza sia stata aggregata a un gruppo può essere *estratta* da tale gruppo e assegnata a un altro nel caso che la sua distanza dal centroide di quest'ultimo gruppo, a valle delle aggregazioni effettuate, sia divenuta minore della sua distanza dal centroide del gruppo di appartenenza[5]. Il vincolo di un metodo non gerarchico è, tuttavia, la necessità di prestabilire il numero di gruppi finali.

Anche secondo la tecnica *k-means*, al *Cluster 1* risultano appartenere 38 soggetti, 86 al *Cluster 2*, mentre i restanti 91 clienti afferiscono al *Cluster 3*.

## 4. Risultati

Per valutare il risultato pratico della Cluster Analysis, occorre innanzitutto contrassegnare (tramite il medesimo programma di aggregazione) l'appartenenza di ogni

[5] Come logica conseguenza, non è certo (anzi, è presumibile il contrario) che tutte le unità statistiche appartenenti a un cluster determinato con tecniche gerarchiche risultino appartenere al medesimo cluster determinato con tecniche non gerarchiche. Gli algoritmi di classificazione non gerarchica comunque assicurano, a parità di partizionamento, maggiore omogeneità entro i gruppi e maggiore eterogeneità tra i gruppi rispetto agli algoritmi non gerarchici (Forgy, 1965).

singolo soggetto a uno dei cluster identificati. Così facendo è possibile incrociare le variabili di profilatura della clientela (ossia le cinque dimensioni macro del questionario MIFID) con i gruppi ottenuti dall'analisi, allo scopo di definire le caratteristiche dei tre cluster, cioè come i soggetti appartenenti a tali gruppi si relazionano con le cinque dimensioni citate, e quale sia la modalità prevalente in ogni cluster.

Innanzitutto, la tabella 8 rileva le caratteristiche dei tre gruppi in base all'esperienza finanziaria: si noti che l'81% del primo cluster ed il 71% del secondo presentano una alta esperienza finanziaria, per cui è evidente che i clienti appartenenti a entrambi i primi due gruppi sono eccellenti conoscitori/utilizzatori dei prodotti finanziari, per cui possiamo definirli "esperti".

**Tabella 8.** *Distribuzione della clientela bancaria secondo l'esperienza finanziaria, per cluster di appartenenza.*

| Esperienza finanziaria | Partizioni | | | Totale |
|---|---|---|---|---|
| | I cluster | II cluster | III cluster | |
| Bassa | - | 1 (1%) | 18 (20%) | 19 |
| Media | 7 (18%) | 24 (28%) | 55 (60%) | 86 |
| Alta | 31 (82%) | 61 (71%) | 18 (20%) | 110 |
| Totale | 38 (100%) | 86 (100%) | 91 (100%) | 215 |

**Tabella 9.** *Distribuzione della clientela bancaria secondo l'obiettivo temporale, per cluster di appartenenza.*

| Obiettivo temporale | Partizioni | | | Totale |
|---|---|---|---|---|
| | I cluster | II cluster | III cluster | |
| Breve | - | 1 (1%) | 47 (52%) | 48 |
| Media | 26 (68%) | 62 (72%) | 41 (45%) | 129 |
| Lungo | 12 (32%) | 23 (27%) | 3 (3%) | 38 |
| Totale | 38 (100%) | 86 (100%) | 91 (100%) | 215 |

Nella tabella 9 si evidenziano invece gli obiettivi temporali della clientela, ed è immediato constatare che il primo ed il secondo cluster hanno prevalentemente obiettivi di medio termine (il 68% del primo gruppo ed il 72% del secondo gruppo), mentre il terzo gruppo si suddivide quasi equamente tra clienti che hanno obiettivi di breve periodo e clienti con obiettivi di medio periodo.

Per quel che concerne gli obiettivi di investimento, nella tabella 10 si osserva una netta differenza tra i gruppi: il primo gruppo ha nella sua totalità obiettivi *alti*) per quanto riguarda l'investimento (incremento di capitale), il secondo gruppo dichiara in misura preponderante obiettivi *bassi* (integrazione del reddito), mentre il terzo gruppo ha, con pochissime eccezioni, obiettivi che possono definirsi *medi* (gestione della liquidità).

**Tabella 10.** *Distribuzione della clientela bancaria secondo gli obiettivi di investimento, per cluster di appartenenza.*

| Obiettivi d'investimento | Partizioni | | | Totale |
|---|---|---|---|---|
| | I cluster | II cluster | III cluster | |
| Integrazione del reddito | - | 81 (94%) | 2 (3%) | 83 |
| Gestione della liquidità | - | 5 (6%) | 89 (97%) | 94 |
| Incremento del capitale | 38 (100%) | - | - | 38 |
| Totale | 38 (100%) | 86 (100%) | 91 (100%) | 215 |

**Tabella 11.** *Distribuzione della clientela bancaria secondo la situazione finanziaria, per cluster di appartenenza.*

| Situazione finanziaria | Partizioni | | | Totale |
|---|---|---|---|---|
| | I cluster | II cluster | III cluster | |
| Bassa | - | 2 (2%) | 9 (10%) | 11 |
| Media | 25 (66%) | 61 (71%) | 72 (79%) | 158 |
| Alta | 13 (34%) | 23 (27%) | 10 (11%) | 46 |
| Totale | 38 (100%) | 86 (100%) | 91 (100%) | 215 |

A riguardo della situazione finanziaria (legata alla consistenza patrimoniale ma non solo), la tabella 11 pone in luce che la maggior parte dei clienti appartenenti a tutti i tre cluster fanno rilevare un livello *medio*, ma il I cluster si differenzia non avendo nessun componente con *bassa* situazione finanziaria e il III cluster presenta il minimo numero di soggetti con situazione finanziaria *alta*.

Nella tabella 12 osserviamo infine la propensione al rischio che hanno i tre gruppi, osservando che la quasi totalità dei soggetti appartenenti al I cluster presenta una *alta* propensione al rischio finanziario. Notiamo invece che il secondo gruppo ha prevalentemente una propensione *media* al rischio finanziario (64%), mentre il restante 36% presenta una alta propensione. Il terzo gruppo si suddivide, infine, tra individui che hanno una *media* propensione al rischio (il 58%) ed individui che hanno una propensione al rischio *bassa* (42%).

**Tabella 12.** *Distribuzione della clientela bancaria secondo la propensione al rischio finanziario, per cluster di appartenenza.*

| Propensione al rischio finanziario | Partizioni | | | Totale |
|---|---|---|---|---|
| | I cluster | II cluster | III cluster | |
| Bassa | - | - | 38(42%) | 38 |
| Media | 1 (3%) | 55 (64%) | 53 (58%) | 109 |
| Alta | 37 (97%) | 31 (36%) | - | 68 |
| Totale | 38 (100%) | 86 (100%) | 91 (100%) | 215 |

## 5. Conclusioni

Il presente lavoro indaga sull'importanza dei rapporti con la clientela nell'ambito bancario, e su come tali rapporti debbano essere gestiti nel dettaglio partendo dalla profilatura (e in un certo senso dalla soddisfazione della clientela), tramite informazioni estratte dall'analisi dei dati.

I risultati ottenuti dalle analisi dei cluster sono esposti nel successivo prospetto sinottico, che sottolinea appunto i diversi profili della clientela. Tale quadro, essendo comunque sempre disponibili i riferimenti di appartenenza di ogni cliente ai cluster, fornisce utili informazioni al management della banca per definire le strategie più opportune e personalizzate.

*Profilatura sinottica della clientela, distinta per cluster di appartenenza.*

| *Profili MIFID* | Partizioni | | |
|---|---|---|---|
| | I cluster: Esperti aggressivi | II cluster: Esperti dinamici | III cluster: Prudenti |
| *Profilo di rischio* | Alto | Medio/Alto | Medio/Basso |
| *Esperienza finanziaria* | Alta | Alta | Media |
| *Obiettivo d'investimento* | Alto (Increm. Capitale) | Basso (Integraz. Reddito) | Medio (Gestione Liquidità) |
| *Obiettivo temporale* | Medio/Lungo | Medio/Lungo | Medio/Breve |
| *Situazione finanziaria* | Media | Media | Media |

Il *primo cluster* è formato da clienti esperti di strumenti finanziari non solo perché conoscono i prodotti finanziari ma soprattutto perché utilizzano tali prodotti in modo autonomo e consapevole per incrementare in modo "aggressivo" le loro entrate. Il loro profilo si può riepilogare quindi nel seguente modo: l'esperienza acquisita nel settore finanziario permette loro di avere un'alta propensione al rischio, infatti tali clienti non hanno interesse ad investire nel breve periodo, in cui il mercato è più volatile, ma investono nel medio/lungo periodo perché sanno che, nella grande maggioranza dei casi, col passare del tempo i rischi intrinsechi di un prodotto finanziario si abbassano (ragionamento calzante soprattutto per i prodotti ad alto rischio ed alto rendimento).

Il *secondo cluster* è composto sempre da clienti esperti nel settore ma che hanno un profilo meno propenso all'alto rischio finanziario. La propensione più avversa al rischio può essere un fattore caratteriale o economica del cliente, oppure una strategia di investimento (effettivamente orientata soprattutto all'integrazione del reddito) o un mix di questi fattori.

Il *terzo cluster*, infine, è formato dai cosiddetti "prudenti" i quali hanno media esperienza finanziaria, assumono un profilo di rischio medio/basso ed operano nel medio/breve periodo, prefiggendosi come obiettivo di investimento la gestione di liquidità detenuta. Sicuramente il profilo di rischio medio basso dipende dalla poco elevata conoscenza ed esperienza che questi clienti hanno dei prodotti finanziari.

Ognuno dei gruppi costituiti necessita, come si vede, di una diversa strategia relazionale e di gestione della clientela.

La metodologia utilizzata (cluster analysis) è tutt'altro che innovativa, ma funziona benissimo quando si tratta di variabili socio-economiche sulla cui gerarchia relazionale non vi siano informazioni. Inoltre, se ben interpretata, la *cluster analysis* permette di approfondire e sviscerare nel dettaglio dinamiche relazionali tra management e soddisfazione della clientela bancaria, dando la possibilità di far capire i meccanismi che fidelizzano la clientela in modo da operare le strategie giuste per intervenire sul processo di fedeltà del cliente stesso.

## Riferimenti bibliografici

Delvecchio, F. (2010). *Statistica per l'analisi di dati multidimensionali*, CLEUP, Padova.

Duse, M. (2011). *Il CRM strategico. Come migliorare la competitività aziendale fidelizzando e centralizzando il cliente*. Franco Angeli. Milano.

Fabbris, L. (1997) *Statistica multivariata. Analisi esplorativa dei dati*, McGraw-Hill, Milano.

Forgy, E. W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics*, 21: 768–769.

MacQueen, J. B. (1967). Some Methods for classification and Analysis of Multivariate Observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1:281-297

Ward, J. H., Jr. (1963). Hierarchical Grouping to Optimize an Objective Function, *Journal of the American Statistical Association*, 58, 236–244.

Zani, S.; Cerioli, A (2007). *Analisi dei dati e data mining per le decisioni aziendali*, Giuffrè, Milano.

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018* ISBN: 978-88-6629-048-3
DIPARTIMENTO DI ECONOMIA E FINANZA pp. 57-81

# Giovani lavoratori e contratto di apprendistato. Evidenza italiana

**Caterina Marini, Vittorio Nicolardi***

*Università degli Studi di Bari Aldo Moro*

**Riassunto:** Il presente lavoro analizza la dinamica del fenomeno occupazione in Italia limitatamente alle fasce più giovani della popolazione (15 – 29 anni) e ad una sola tipologia di rapporto di lavoro, il contratto di apprendistato. Alla stregua delle esperienze positive degli altri Paesi europei, il contratto di apprendistato è considerato dai governi italiani del XXI secolo una forma ottimale di ingresso nel mondo del lavoro principalmente per i più giovani non solo perché completa la formazione con il praticantato, ma anche perché costituisce una modalità di acquisizione di competenze ed esperienze lavorative da spendere nella professione. Il lavoro si delinea come un tentativo di fotografare un fenomeno ben circoscritto utilizzando il dato di fonte ISTAT riveniente dalla Rilevazione Continua delle Forze di Lavoro, per il periodo I trim. 2014 – II trim. 2018. Nonostante le note criticità del dato utilizzato, i risultati ottenuti sono soddisfacenti. Dall'analisi è emerso come il contratto di apprendistato venga sottoutilizzato in Italia per una serie di limitazioni che non solo sono propriamente ricollegabili alle caratteristiche del settore produttivo italiano, ma anche all'impianto della norma di riferimento. I risultati dell'analisi evidenziano caratteristiche di genere importanti sia per quel che riguarda l'inquadramento del lavoratore, sia per quel che riguarda il titolo di studio posseduto. Evidenze ricollegabili anche alle azioni del governo messe in atto su più fronti e dirette al mercato del lavoro emergono dall'analisi in modo evidente.

**Keywords:** Mercato del lavoro, RCFL, Contratto di apprendistato.

## 1. Introduzione

La crisi economico-finanziaria che ha travolto l'economia americana a partire dal 2007 ha profondamente modificato livelli e andamenti delle principali variabili

* Autore corrispondente: vittorio.nicolardi@uniba.it.

macroeconomiche per quasi tutte le economie avanzate, e indotto pesanti conseguenze economico-sociali laddove ha riguardato la condizione occupazionale e professionale dei lavoratori. In Italia, oramai da almeno un decennio una delle economie più fragili dell'eurozona, le conseguenze della crisi sono state gravi dal punto di vista economico e hanno avuto pesanti ripercussioni sulla crescita dell'economia e sull'occupazione. Infatti, nel 2009, mentre la crisi finanziaria interessava seriamente mercati e sistemi produttivi dell'intero pianeta, e causava una delle recessioni più dannose nella storia dell'economia mondiale (si ricordi per importanza storica, a titolo di esempio, anche solo la grande depressione che ebbe origine nel 1929 sempre negli Stati Uniti d'America, le cui cause ed effetti sono spesso messi a confronto con la crisi del 2009), il PIL italiano crollava drammaticamente del 5,5% rispetto all'anno precedente e il tasso di disoccupazione totale cominciava a crescere ininterrottamente sino a raggiungere il livello più alto degli ultimi 20 anni nel 2014 (12,7%). Che una crisi mondiale di entità così importante provochi una decelerazione dell'attività economica, una contrazione dell'occupazione a tutti i livelli, una riduzione degli investimenti determinata dall'incertezza globale dei mercati, una inversione della pressione inflazionistica causata dalla debolezza stessa dell'economia è argomentazione ampiamente discussa e prevista in letteratura. Ciò che rende possibile e prolungata la ripresa nei diversi Paesi, però, è non solo la solidità pre-crisi delle rispettive economie e della struttura finanziaria annessa, ma anche l'insieme di interventi post-crisi di politica economico-finanziaria e di natura giuslavoristica messi in atto dai corrispondenti governi al fine rilanciare il sistema economico attraverso il sostegno alla competitività e produttività dell'impresa da un lato, e l'occupazione dall'altro.

Il presente lavoro focalizza l'attenzione sul fenomeno dell'occupazione in Italia che, a fronte del generale aumento della disoccupazione che ha colpito drammaticamente soprattutto le fasce più giovani della popolazione per tutto il periodo 2008 – 2014, diventa oggetto di numerosi interventi di natura giuslavoristica ed economico-finanziaria finalizzati a invertire l'andamento del fenomeno. Difatti, similmente a quanto disposto da molti governi dei Paesi europei diversamente colpiti dalla crisi americana, anche l'Italia mette in atto una serie di interventi di varia natura volti a tamponare l'emorragia dei posti di lavoro e, pertanto, finalizzati al sostegno dell'occupazione e principalmente all'ingresso nel mercato del lavoro dei giovani.

L'analisi degli effetti di questi interventi di politica economica e giuslavoristica mostra, invero, una serie di difficoltà per ragioni di diversa natura. Innanzitutto, qualsiasi intervento di tipo legislativo mirato al mercato del lavoro non può prescindere dai paralleli interventi finalizzati alla ripresa economica generale. In se-

condo luogo, soprattutto per quel che riguarda le fasce di età più giovani dei lavoratori, a favore delle quali il governo italiano ha emanato anche interventi specifici per l'assunzione oltre quelli generali di lotta alla disoccupazione, tutti gli interventi di tipo legislativo sono normalmente di durata pluriennale e spesso si sovrappongono nei loro effetti. Infine, proprio l'analisi degli effetti degli interventi di politica economica sul mercato del lavoro è compromessa dalla disponibilità ridotta di informazioni a livello disaggregato che possano permettere un'analisi dinamica dell'andamento dell'occupazione per verificare quali siano non solo gli effetti reali dei singoli interventi, ma anche gli effetti congiunti data la sovrapposizione suddetta. Infatti, allo stato attuale la disponibilità di strutture di dati dinamiche relative al mercato del lavoro non sembra sufficiente per le finalità prepostesi. Soprattutto per quel che riguarda i rapporti di lavoro diversi dai più noti e diffusi contratto di lavoro subordinato a tempo indeterminato e contratto di lavoro a termine, oramai previsti dalla normativa italiana da più di vent'anni (L. n. 196/1997 e D. Lgs. n. 276/2003, per citarne alcuni) e principalmente rivolti all'inserimento nel mondo del lavoro dei giovani, le banche dati attualmente disponibili o non contemplano tutte le variabili necessarie ad un'analisi strutturale del fenomeno occupazionale giovanile a causa della natura stessa della rilevazione, oppure l'indisponibilità del dato è legata ai problemi di tutela della privacy che di fatto non consente di identificare il singolo individuo per seguirne i cambi di stato relativamente alla sua condizione lavorativa e/o professionale nel tempo. Nel primo caso, il riferimento è esplicitamente rivolto al questionario utilizzato per la Rilevazione Continua delle Forze di Lavoro dell'ISTAT (RCFL, d'ora innanzi) che per sua natura, essendo strutturato dal lato dell'offerta, consente di rilevare solo in parte il dato riguardante le forme contrattuali per prestazioni di lavoro temporaneo, anche se finalizzate alla formazione, così come previsto dalla normativa italiana. Ciò limita, di fatto, l'analisi dell'efficacia dei vari strumenti messi a disposizione dei datori di lavoro da parte del legislatore sulle modalità di ingresso dei giovani nel mercato del lavoro. Nel secondo caso, il riferimento è al dato amministrativo raccolto nelle banche dati del Ministero del Lavoro e delle Politiche Sociali e dell'Istituto Nazionale di Previdenza Sociale che, per sua natura, è rilevato indirettamente dal lato della domanda per il tramite delle Comunicazioni Obbligatorie. Pur nella sua relativa completezza per l'informazione raccolta, tale dato non è reso disponibile a causa di una mancata procedura di cifratura che ne consenta l'uso superando i limiti di tutela della privacy del lavoratore[1].

[1] V. Nicolardi, C. Marini (2017).

A ciò si aggiunga anche il fatto che le diverse banche dati disponibili non contemplano tutte le variabili necessarie ad un'analisi strutturale del mercato del lavoro, essendo ognuna costruita in funzione degli scopi istituzionali costitutivi degli Enti che le gestiscono.

Pertanto, la possibile e necessaria integrazione delle varie fonti di dati provenienti dai vari database oramai rappresenta l'unica strada da percorrere per riuscire a delineare un quadro completo dei più vari fenomeni economici e rappresenta già da anni materia di approfondimento e discussione da parte dell'ISTAT. E a tal proposito, la gestione centralizzata dei dati amministrativi costituisce da parte dell'ISTAT un passo importante verso l'efficienza della produzione e il contenuto informativo delle statistiche, e rientra tra gli obiettivi previsti dal Programma di modernizzazione dell'Istituto[2]. Risale infatti al recente passato, maggio 2016, la definizione del Sistema di Integrazione di Microdati (SIM), un "contenitore" di dati costruito mediante l'analisi del contenuto informativo e l'integrazione concettuale e fisica di microdati di fonte amministrativa e statistica di carattere censuario[3].

Il presente lavoro analizza la dinamica del fenomeno occupazione in Italia specificatamente limitata alle fasce più giovani della popolazione (15 – 29 anni) e ad una sola tipologia di rapporto di lavoro, il contratto di apprendistato, normato giuridicamente dal legislatore principalmente per tentare di risolvere il dramma della disoccupazione giovanile italiana che, seguendo un trend crescente per tutto il periodo 2008 – 2014, raggiunge il suo punto di massimo nel 2014 (31,6%), uno dei più alti nell'eurozona dopo Grecia, Spagna e Croazia.

Il dato utilizzato è di fonte ISTAT ed è quello riveniente dalla RCFL, per il periodo I trim. 2014 – II trim. 2018 (ultimo dato disponibile). La scelta di utilizzare come momento di inizio dell'analisi l'anno 2014 piuttosto che un anno precedente si basa sulla volontà da parte degli autori di studiare, in relazione alla tipologia di contratto prescelta, l'evoluzione dell'occupazione in Italia a seguito dell'ultima riforma dell'ordinamento giuridico per il mercato del lavoro, nota con l'acronimo JOBS ACT (JA, d'ora innanzi), e della contemporanea politica economico-finanziaria a sostegno del nuovo quadro normativo, non tralasciando però di valutare e commentare il quadro di riferimento giuridico vigente sino al 2013. Inoltre, anche la decisione di utilizzare una serie temporale non molto lunga si pone l'obiettivo di cogliere gli effetti immediati e prolungati nel medio e breve periodo di un complesso intreccio di interventi che nella maggior parte dei casi si sovrap-

[2] M. C. Runci, G. Di Bella, L. Galiè (2016), M. Calzaroni (2008)
[3] Atti Parlamentari, XVII LEGISLATURA, Doc. LXIX, n. 4 (2016)

pongono. Infatti, per far fronte all'emorragia di posti di lavoro gli interventi mirati del governo italiano per rendere "appetibile" l'assunzione di un giovane lavoratore in età 15 – 29 anni sono stati numerosi, assieme alle altre manovre di intervento per sostenere la disoccupazione in generale e di cui, in misura significativa, hanno beneficiato anche i giovani. A tal proposito, in Marini – Nicolardi (*in pubblicazione*) si dimostra come l'ultima riforma del mercato del lavoro JA, accompagnata dai sussidi all'occupazione, sia stata favorevolmente accolta dal tessuto imprenditoriale italiano anche per l'inserimento dei giovanissimi nel mercato del lavoro e, in particolare, il sussidio all'occupazione particolarmente "gradito" per l'assunzione con il nuovo contratto a tempo indeterminato a tutele crescenti dei giovani. Nello specifico del presente lavoro, la scelta è ricaduta sul rapporto di apprendistato perché, alla stregua delle esperienze positive degli altri Paesi europei, è considerato dai governi italiani del XXI secolo una forma ottimale di ingresso nel mondo del lavoro principalmente per i più giovani in quanto prevede un connubio ottimale tra formazione e praticantato, ma anche una modalità di acquisizione di competenze ed esperienze lavorative da spendere in tempi successivi sul percorso professionale. Riguardo all'utilità del contratto di apprendistato per agevolare il passaggio dal periodo di formazione prettamente scolastico all'ingresso nel mondo del lavoro dei più giovani, la discussione in ambito scientifico internazionale è ancora viva (Marsden e Ryan, 1990; Sako, 1991; Quintini e Manfredi, 2009). L'evidenza degli studi sinora condotti mostra come per quelle situazioni in cui il mercato del lavoro preveda l'esistenza di un sistema duale basato su apprendistato "di formazione", cosiddetto di primo livello, e apprendistato professionalizzante i risultati positivi dipendano fondamentalmente dalla elevata percezione che si ha delle potenzialità dello strumento normativo e dall'uso efficiente dello stesso al punto da consentire anche ai giovani che hanno conseguito da poco la qualifica professionale di poter trovare occupazione anche con poca o nessuna esperienza. Tale è il caso molto positivo di Paesi come la Germania, l'Olanda e la Svizzera, mentre risultati diametralmente opposti si hanno per Paesi come Grecia, Portogallo, Italia e Spagna dove non solo la percezione delle potenzialità del contratto di apprendistato è molto bassa, ma la presenza di un tessuto imprenditoriale costituito principalmente da piccole e medie imprese non ne rende facile l'utilizzo per via del basso numero di dipendenti (CEPR, 2015; Pacifico et al., 2018; OECD, 2017). Chiaramente, al fine di poter comparare strutturalmente e più precisamente dati di Paesi diversi superando la sintesi eccessiva riportata, è opportuna un'analisi del dato giuridico più dettagliata per distinguere ulteriormente le finalità formative dell'apprendistato. Pur in assenza di esplicite definizioni legislative in merito, infatti, i contratti di apprendistato si possono funzionalmente distin-

guere a seconda che puntino all'addestramento propriamente detto o alla formazione del lavoratore. Nel primo caso, in linea di massima, si possono definire come orientate all'addestramento pratico quelle forme di apprendistato in cui l'aspetto "formativo" privilegia l'acquisizione delle informazioni e delle competenze necessarie ad occupare un determinato ruolo in uno specifico contesto produttivo e, pertanto, addestramento che si caratterizza per essere "firm specific". Orientate alla formazione professionale sono, invece, quelle forme di apprendistato dirette allo sviluppo delle "attitudini" e all'acquisizione delle "capacità" e delle "competenze" necessarie per affrontare le variazioni sincroniche e diacroniche del lavoro, intendendo per competenza la "comprovata capacità di utilizzare, in situazioni di lavoro, di studio o nello sviluppo professionale e personale, un insieme strutturato di conoscenze e di abilità acquisite nei contesti di apprendimento formale, non formale o informale" (l'art. 2, c. 1, lett. e, d. lgs. n. 13 del 2013)[4].

L'analisi del fenomeno occupazione giovanile sviluppata nel presente lavoro è un primo approccio di analisi disaggregata, per quanto parziale a causa dell'incompletezza del dato disponibile, per mostrare come a partire dal 2014, a seguito delle ultime riforme approvate dal governo italiano, anche il rapporto di lavoro di apprendistato abbia contribuito, seppur in misura limitata e con caratteristiche territoriali marcate a livello nazionale, a consentire l'ingresso nel mondo del lavoro dei giovani, anche con limitata o nulla esperienza lavorativa.

Il lavoro è organizzato come segue: nella Sezione 2 si descrive il quadro normativo e le sue modificazioni limitatamente alla tipologia contrattuale considerata; nella Sezione 3 si presenta la tipologia di dato utilizzato e si descrivono alcuni dei risultati dell'analisi; la Sezione 5 riassume le considerazioni conclusive.

## 2. Contratto di apprendistato: normativa di riferimento.

Il mercato del lavoro italiano ha subito nel corso degli ultimi vent'anni importanti mutamenti a seguito dell'emanazione di riforme economico-finanziarie e giuslavoristiche tutte finalizzate sia a sostenere l'occupazione della forza lavoro giovane e il ricollocamento nel mondo del lavoro dei soggetti rimasti privi di una occupazione, sia a promuovere competitività e produttività dell'impresa. Per quanto già sul finire del XX secolo la politica legislativa in materia di lavoro avesse cominciato a prevedere e definire forme di partecipazione flessibile e discontinua in risposta alle esigenze provenienti dal tessuto imprenditoriale del Paese (Legge n. 196/1997), è

[4]Tale puntualizzazione risulta necessaria ai fini di futuri approfondimenti del fenomeno.

principalmente a partire dagli anni 2000 che il governo italiano revisiona gradatamente e profondamente le disposizioni normative in materia di occupazione definendo nuove tipologie contrattuali e revisionando le esistenti perché le imprese possano utilizzare il lavoro con modalità più flessibili in base alla necessità imprenditoriali. Pur rischiando di fornire un'interpretazione eccessivamente semplicistica delle motivazioni alla base delle scelte di politica del diritto ai fini dell'ema–nazione di tali interventi e revisioni dell'ordinamento giuridico in materia di occupazione, è importante contestualizzare dal punto di vista economico i mutamenti messi in atto. Questi risultano condizionati dalle continue modificazioni dell'appa–rato produttivo mondiale, e quindi anche italiano, in risposta alla natura mutata dei mercati oramai globalizzati e sempre più da variabili oggettive instabili che pongono le imprese nelle condizioni di dover affrontare situazioni in continua evoluzione in un contesto geografico oramai allargato e più competitivo per riuscire a mantenere o occupare quote di mercato preziose. Conseguentemente, il contesto sociale che ne deriva vede il lavoratore come il soggetto più vulnerabile, per certi versi, rispetto ai mutamenti strutturali in atto che necessariamente, come conseguenza inevitabile, coinvolgono anche i rapporti tra datore di lavoro e prestatore d'opera al punto da richiedere interventi di revisione delle norme per ridefinire il quadro delle tutele giuslavoristiche interessato.

Il contratto di apprendistato, oggetto del presente lavoro, viene da sempre concepito nella normativa italiana come una modalità di avviamento al lavoro rivolta ai giovani e giovanissimi caratterizzata dall'obbligatorietà per il datore di lavoro di garantire formazione in azienda per il lavoratore assunto perché possa conseguire le capacità tecniche e professionali per la propria crescita lavorativa futura. Risale al 1955 l'impianto normativo che per la prima volta disciplina nell'ordinamento giuridico italiano l'apprendistato come forma "alternativa" di rapporto di lavoro rivolto ai giovanissimi in età 16 – 20 anni della durata massima di cinque anni (calcolabili anche in forma cumulata se previsti periodi di interruzione temporanei e comunque non superiori all'anno) durante i quali le ore destinate all'insegnamento complementare all'insegnamento scolastico sono considerate, a tutti gli effetti, ore lavorative e computate nell'orario di lavoro (Legge n. 25/1955). Quindi, a tutti gli effetti, il contratto di apprendistato nasce come una più attuale forma di contratto a termine e/o flessibile precorrendo i tempi rispetto all'impianto giuridico del mercato del lavoro che si sarebbe realizzato molti anni dopo e che definisce l'odierna strutturazione. È solo nel 2003[5], circa cinquant'anni dopo, che con il D. Lgs. n.

[5] Nel 1997, rispetto a quanto definito dalla Legge n. 25/1955, la Legge n. 196 (il cosiddetto Pacchetto Treu) interviene marginalmente sul contratto di apprendistato e limitatamente alla ridefinizione del

276/2003 (la cosiddetta Legge Biagi) il governo italiano procede con una prima sostanziale e completa riformulazione della disciplina e del rapporto di apprendistato, nelle intenzioni del legislatore diretta a migliorare la formazione iniziale del lavoratore con l'obiettivo di consolidarne la posizione lavorativa/professionale attraverso l'aumento delle competenze. In realtà, per quanto le intenzioni del legislatore fossero le migliori, a tutti gli effetti le disposizioni inserite nel quadro giuridico italiano con il D. Lgs. n. 276/2003, facenti riferimento a diverse forme contrattuali di lavoro atipico, cosiddette parasubordinate e non standard, paiono rispondere a finalità che vanno più nella direzione di tutela dell'interesse del datore di lavoro, al quale si autorizza per legge una gestione più flessibile e conveniente dal punto di vista economico della forza lavoro, piuttosto che del lavoratore (Pinto, 2004)[6]. Ad ogni modo, riguardo al rapporto di lavoro di apprendistato, il D. Lgs. n. 276/2003, pur mantenendo sempre l'impianto normativo che caratterizza tipicamente un contratto di lavoro a termine, introduce una diversificazione del rapporto rispetto alla regolamentazione vigente al momento di entrata in vigore della riforma (ossia, la Legge n. 196/1997) in ragione degli obiettivi formativi in base ai quali viene stipulato il rapporto, che ne determinano anche la durata. Sostanzialmente, il D. Lgs. 276/2003 prevede tre tipologie di contratto di apprendistato, sempre rivolte alle fasce più giovani della popolazione, come di seguito brevemente descritte:

1. apprendistato per l'espletamento del diritto-dovere di istruzione e formazione, rivolto ai giovanissimi di età compresa nella classe 15 – 18 anni per consentire loro di espletare il proprio diritto/dovere di istruzione e formazione come previsto in materia di riordino dei sistemi di istruzione e formazione professionale (Legge n. 53/2003), della durata non superiore a 3 anni determinata in considerazione della qualifica da conseguire, dal titolo di studio, dai crediti professionali e formativi acquisiti;
2. apprendistato professionalizzante, rivolto ai giovani di età compresa nella classe 18 – 29 anni per il conseguimento di una qualifica professionale attraverso un'attività di formazione sul campo "e l'acquisizione di competenze di base, trasversali e tecnico-professionali" (art. 49, comma 1), della

limite superiore di età di utilizzo del rapporto, che difatti passa a 24 anni, ovvero a 26 anni nelle aree di cui agli obiettivi n. 1 e 2 del regolamento (CEE) n. 2081/1993 e successive modificazioni, e della durata del rapporto che, comunque, viene ridefinita su un intervallo temporale 18 mesi – 4 anni.

[6] Per completezza di informazione è bene ricordare che il D. Lgs. n. 276/2003 introduce nell'ordinamento giuridico italiano varie forme contrattuali di lavoro atipico, i cosiddetti contratti di lavoro parasubordinato e non standard, oltre quelli già esistenti, introdotti nelle precedenti legislature, che lo stesso decreto riforma (apprendistato, collaborazione coordinata e continuativa, somministrazione, appalto e distacco, lavoro intermittente, lavoro ripartito, lavoro occasionale, lavoro a progetto, contratto di inserimento).

durata, stabilita dai contratti collettivi stipulati dalle associazione dei datori di lavoro e lavoratori più rappresentative, non inferiore a 2 anni e comunque non superiore a 6 anni;

3. apprendistato per l'acquisizione di un diploma o per percorsi di alta formazione, rivolto sempre ai giovani in età 18 – 29 anni, in accordo con le associazioni dei datori di lavoro e lavoratori, le Università e gli altri enti di formazione[7].

Ciò che accomuna le tre tipologie di rapporto di apprendistato descritte dianzi è, come previsto dal D. Lgs. 276/2003, la natura "tandem" delle stesse, nel senso che l'apprendistato è una forma di rapporto di lavoro la cui regolamentazione è rimessa a due soggetti istituzionali: lo Stato da un lato, che ne definisce l'impianto giuridico di applicazione per il tramite della norma, e le Regioni e /o Province autonome di Trento e Bolzano, a cui la stessa norma assegna la regolamentazione dei profili formativi dell'apprendistato d'intesa con le associazione dei datori di lavoro e lavoratori più rappresentative sul piano regionale. Chiaramente, una tale impostazione della norma di applicazione del rapporto di apprendistato, che ne disciplina interamente l'utilizzo, definisce già in fase di impianto una sorta di disomogeneità e disparità territoriale a livello nazionale, che rimarrà confermata anche nel successivo, cronologicamente parlando, importante intervento giuslavoristico che il legislatore nel 2011 approverà in materia di apprendistato. Il Testo Unico dell'Apprendistato (D. Lgs. n. 167/2011) inquadra meglio il rapporto di apprendistato all'interno dell'impianto giuridico che regolamenta il mercato del lavoro, pur confermando la definizione delle tre tipologie di contratto e la loro natura "tandem" di regolamentazione di comune accordo con gli enti periferici dello Stato, ma apporta anche modifiche importanti rispetto alla norma in vigore al momento della sua approvazione. La più importante di queste è indubbiamente la ridefinizione del contratto di apprendistato come un "contratto di lavoro a tempo indeterminato finalizzato alla formazione e alla occupazione dei giovani" (art. 1, comma 1), tranne nei casi in cui si tratti di attività stagionali per le quali si prevedono margini di svolgimento del contratto di apprendistato anche a tempo determinato. Le altre modifiche apportate dal D. Lgs. n. 167/2011 al D. Lgs. 276/2003 riguardano principalmente le modalità di utilizzo delle tre tipologie di contratto di apprendistato, che per l'occasione vengono anche rinominate nel modo seguente:

[7] Il limite di età si abbassa a 17 anni se il giovane dovesse aver già conseguito la qualifica professionale sia nel caso dell'apprendistato professionalizzante, sia nel caso dell'apprendistato per l'acquisizione di un diploma o per percorsi di alta formazione.

1. apprendistato per la qualifica e per il diploma professionale, rivolto ai giovani di età compresa nella classe 15 – 25 anni, la cui durata può non essere superiore ai 3 o 4 anni a seconda della durata del corso di diploma frequentato;
2. apprendistato professionalizzante o contratto di mestiere, rivolto ai giovani di età compresa nella classe 18 – 29 anni, la cui durata non può essere superiore ai 3 anni;
3. apprendistato di alta formazione, rivolto sempre ai giovani in età 18 – 29 anni, che vede una estensione del campo di applicazione del rapporto anche ai dottorati di ricerca[8].

Un'importantissima integrazione alle modalità di applicazione del rapporto di apprendistato che il D. Lgs. n. 167/2011 introduce nell'ottica di poter incidere anche sul ricollocamento di quei lavoratori che hanno perso una precedente occupazione riguarda i limiti di età di applicazione dianzi riportati che vengono superati e non costituiscono vincolo di utilizzo del rapporto di apprendistato laddove si tratti di lavoratori beneficiari di indennità di mobilità, per i quali è possibile l'assunzione con un rapporto di apprendistato al fine di garantirne la qualificazione o riqualificazione professionale (art. 7, comma 4).

Per quanto nel 2014 venga approvata la nuova e attuale, seppur con le ultime modifiche apportate dal D. L. n. 87/2018, riforma del lavoro JA, il D. Lgs. n. 167/2011 continua a rappresentare un punto di riferimento normativo importante nel panorama giuridico per quel che riguarda il rapporto di lavoro di apprendistato. Il JA si concretizza sostanzialmente di otto decreti legislativi ed un decreto correttivo emanati durante gli anni 2015 e 2016 al fine di poter consentire al Governo di adoperarsi nelle numerose ed ampie deleghe assegnategli proprio nel 2014 dalla Legge Delega n. 183/2014, e viene approvato ed emanato proprio a causa del perdurare dell'incertezza economica e della grave crisi occupazionale del Paese. Il JA interviene sulla disciplina del mercato del lavoro principalmente modificando e/o integrando il quadro normativo in vigore riguardante le tipologie di contratto di lavoro a termine e atipico, facendo buon uso di quanto disciplinato nelle precedenti legislature, e apportando novità di rilevanza assoluta nel panorama normativo italiano riguardanti il contratto di lavoro subordinato a tempo indeterminato. All'interno degli interventi più propriamente rivolti al rapporto di apprendistato, il JA non apporta sostanziali modifiche se non allargare anche a coloro che percepiscono un sussidio di disoccupazione la possibilità di potersi reinserire nel mercato

[8] Idem nota 7.

del lavoro anche attraverso un contratto di apprendistato senza vincoli di età (D. Lgs. n. 81/2015).

All'interno del disegno del legislatore, al fine di rendere "appetibile" per i datori di lavoro il contratto di apprendistato e di favorire l'ingresso nel mondo del lavoro dei giovani, sia il D. Lgs. n. 167/2011, sia il JA prevedono agevolazioni sia in termini di contributi da versare, sia in termini di retribuzione da erogare. Ciò si base, principalmente, sugli stessi impianti normativi suddetti che prevedono che l'apprendista possa essere inquadrato fino a due livelli inferiori rispetto alla categoria spettante ai lavoratori con stesse mansioni o funzioni. Inoltre, la retribuzione dell'apprendista può essere stabilita in misura percentuale e graduale all'anzianità di servizio. Infine, a partire dal 2017, il JA, relativamente alle assunzioni in "apprendistato", viene affiancato e rafforzato da eccezionali provvedimenti di natura incentivante che consentono a tutte le imprese non agricole di poter usufruire di un esonero totale del versamento dei contributi previdenziali complessivi a carico dei datori di lavoro secondo due modalità distinte:

1. per un importo massimo annuo di € 3.250, della durata triennale per ogni nuovo contratto di apprendistato di I livello o di III livello[9] stipulato nel corso dell'anno finanziario 2017 (provvedimento inserito nella Legge di Bilancio 2017 L. 232/2016 - art. 1, commi 308–310);
2. per un importo massimo annuo di € 8.060, della durata annuale per ogni nuovo contratto di apprendistato professionalizzante stipulato nel corso dell'anno finanziario 2017 per l'assunzione di giovani in età 16 – 29 anni iscritti al programma Garanzia Giovani[10] (provvedimento inserito nel Decreto Direttoriale n. 394 del 2 dicembre 2016).

## 3. Analisi dei dati sui contratti di apprendistato.

La RCFL rappresenta una delle fonti statistiche di informazione relativamente alle modalità e al grado di partecipazione al mercato del lavoro degli individui costituenti la popolazione. In particolare, a partire dal 2004 la rilevazione ISTAT è stata

[9] Nel JA, per apprendistato di I livello si intende l'apprendistato per la qualifica ed il diploma professionale, il diploma di istruzione secondaria superiore e il certificato di specializzazione tecnica superiore, mentre per apprendistato di III livello si intende l'apprendistato di alta formazione e ricerca.

[10] La Garanzia Giovani è il Piano Europeo per la lotta alla disoccupazione giovanile, che prevede finanziamenti per i Paesi membri con tassi di disoccupazione superiori al 25% utilizzabili per politiche attive di orientamento, istruzione e formazione e inserimento al lavoro, a sostegno dei giovani che non sono impegnati in un'attività lavorativa, né inseriti in un percorso scolastico o formativo (i cosiddetti NEET).

sottoposta ad una significativa ristrutturazione in modo tale da rispondere a quanto stabilito dal Regolamento n. 577/98 del 9 marzo 1998 dell'Unione Europea relativamente all'organizzazione delle indagini sulle Forze di Lavoro.

Uno degli aspetti più interessanti intervenuti nella nuova formulazione della RCFL ha riguardato il carattere continuativo assunto dall'indagine attraverso la ripartizione delle settimane di rilevazione su tutto l'arco dell'anno, permettendo, in tal modo, di tener opportunamente conto dell'evoluzione dinamica del mercato del lavoro. In tal senso, nella RCFL si pone una maggiore attenzione sulla rilevazione di maggiori e più significative indicazioni circa l'evoluzione dinamica della situazione lavorativa degli intervistati, sia in termini di situazione occupazionale, sia di mobilità territoriale, attraverso la richiesta, in apposite sezioni del questionario, di una serie di informazioni riguardanti sia la condizione lavorativa retrospettiva nei periodi precedenti la settimana di rilevazione, sia delle eventuali variazioni di residenza intervenute nello stesso periodo.

Relativamente all'oggetto del presente studio, occorre innanzitutto rilevare come all'interno del questionario ISTAT sulla RCFL si distingua tra la stipula di un contratto di apprendistato e la posizione di apprendista. Infatti, della sezione C del questionario viene espressamente chiesto agli intervistati sia se essi abbiano stipulato un contratto di apprendistato, sia se essi ricoprano la posizione lavorativa di apprendista. Inoltre, nell'ultima strutturazione del questionario datata I trim. 2018, alcune modifiche apportate alla sequenza di domande ed al numero di modalità di risposta previste, permettono di cogliere in maniera più precisa alcuni aspetti del rapporto di apprendistato.

In particolare, sino al IV trim. 2017 alla domanda C9 veniva chiesto all'intervi–stato quale figura professionale rivestisse, ed in particolare se egli fosse un apprendista. In seguito, alla domanda C23A veniva chiesto all'intervistato in quale tipologia rientrasse il suo contratto, una delle quali era quella di apprendistato. A partire dalla I trim. 2018, invece, alla domanda C9 è stata aggiunta una domanda di controllo, la C9D, in cui si chiede espressamente all'intervistato se ha un contratto di apprendistato, mentre la domanda C23A è stata sostituita dalla C21A in cui le opzioni di risposta alla richiesta su quale tipologia di contratto abbia stipulato l'intervistato sono più ampie che nei questionari precedenti.

È importante, comunque, osservare come nel questionario ISTAT il contratto di apprendistato venga fatto rientrare nella tipologia dei contratti a tempo determinato come implicitamente si rileva dalla domanda C22 nei questionari sino al IV trim. 2017 e C22α e C22β nei questionari a partire dal I trim. 2018, in cui si formula la richiesta di indicare la ragione per la quale il proprio lavoro è a termine e si include

nelle modalità di risposta lo svolgimento di un apprendistato. Sebbene da un punto di vista giuridico il rapporto di apprendistato risulti rientrare nella tipologia dei contratti a tempo indeterminato sin dal 2011 (cfr. par. 2), la motivazione alla base della scelta dell'ISTAT di rilevarlo come rientrante nella tipologia di contratto a termine è, molto probabilmente, legata alla modalità di *autopercezione* da parte dell'intervistato della propria posizione lavorativa, la quale presenta a tutti gli effetti una certa ambiguità in relazione alla ufficiale durata del contratto, che effettivamente si può concludere alla sua effettiva scadenza nei limiti previsti dalla legge.

Ai fini dell'analisi del dato, quindi, nel prosieguo del lavoro si è analizzata contemporaneamente sia la tipologia di contratto stipulato dagli intervistati, isolando quello di apprendistato, sia la condizione effettiva di posizione di apprendista, sia la motivazione alla base della stipula del rapporto di lavoro da parte degli intervistati con contratto di apprendistato. Al fine, inoltre, di cercare di analizzare separatamente l'apprendistato di I livello e l'apprendistato professionalizzante, si sono suddivisi gli intervistati secondo le classi di età. In considerazione, però, del fatto che il contratto di apprendistato di I livello può essere stipulato dagli individui in età 15 – 25 anni e quello professionalizzante dagli individui in età 18 – 29 anni, e che, inoltre, la normativa vigente prevede la possibilità di stipulare contratti di apprendistato anche per individui oltre i 29 anni, si sono suddivisi gli individui in 4 classi di età, per tener conto della sovrapposizione di alcune classi, nel seguente modo: 15 – 17 anni, 18 – 25 anni, 26 – 29 anni e 30 – 35 anni. Per quanto il quadro normativo di riferimento non preveda limiti di età per la stipula dei rapporti di apprendistato nei casi specifici indicati (cfr. par. 2), l'ISTAT circoscrive la rilevazione ai lavoratori in età 15 – 35 anni.

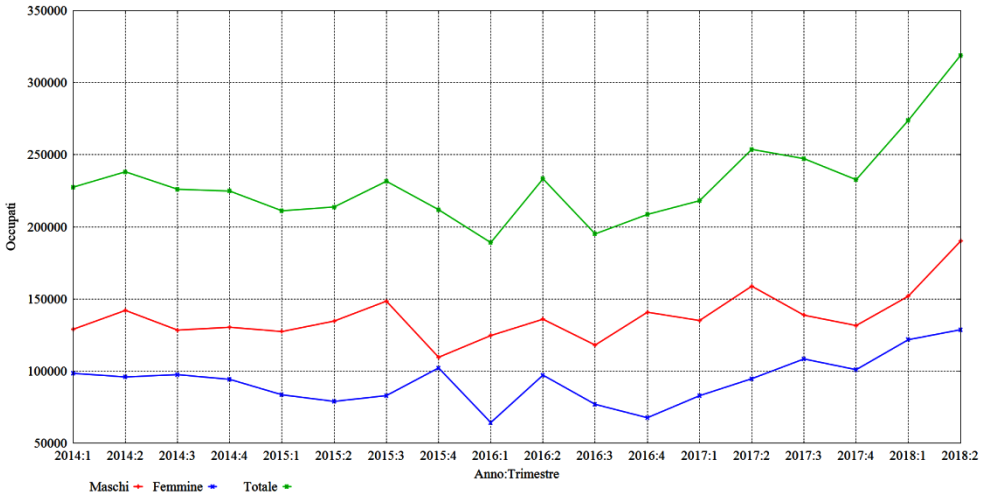
### *3.1 Analisi del dato.*

Dall'analisi dei dati rivenienti dalla RCFL si può osservare come nel periodo analizzato il numero di contratti di apprendistato abbia registrato un aumento di circa 91 mila unità (+40,1%), passando da circa 228 mila unità a circa 319 mila unità. Mediamente, quasi il 60% dei titolari di contratto di apprendistato sono di sesso maschile (Tab. 1, in Appendice). La maggior parte dei contratti di apprendistato si registra nel Nord Italia, con regioni, quali la Lombardia, il Veneto, l'Emilia Romagna e il Piemonte, che investono maggiormente su questa modalità di ingresso per i giovani nel mondo del lavoro, mentre per il Centro e, in particolare, il Sud dell'Italia il numero di contratti di apprendistato risulta essere poco significativo.

Per quanto il numero di contratti si mostri in aumento, il trend del fenomeno nel tempo presenta delle significative variazioni, mostrando un sostanziale calo fino al

I trim. 2016, per poi crescere a partire dal III trim. 2016 (Figura 1). Tra questi due periodi, si ravvisa un incremento anomalo nei contratti di apprendistato nel II trim 2016 (+23,4%). Nel secondo periodo, altrettanto anomalo su trend tendenzialmente crescente è un significativo calo registrato nel II-III trim. 2017 (-8,4%). In particolare, è da rilevare una consistente crescita dei contratti di apprendistato negli ultimi due trimestri del periodo considerato (+34,1%). L'analisi di trend differenziata secondo il sesso mostra sostanzialmente andamenti simili a quelli del trend generale, ad eccezione dei trimestri a cavallo della fine degli anni 2016-2017 e nel III trim. 2017, dove si registrano significative variazioni in direzione opposta.

**Figura 1** *Contratti di apprendistato secondo il sesso. Periodo I trim. 2017–II trim. 2018.*




Se si guarda all'inquadramento degli intervistati (Tab. 3, in Appendice), si può vedere come, mediamente, nel periodo analizzato meno del 50% abbia un effettivo inquadramento come apprendista (46,5% in media), mentre, sempre in media, il 24,3% è inquadrato come impiegato e il 29,2% è, invece, inquadrato come operaio.

Una significativa differenza si riscontra tra i due sessi, in quanto le donne con contratto di apprendistato sono nella maggior parte dei casi inquadrate come impiegate (60,4%), mentre per gli uomini l'inquadramento prevalente è quello di operaio o apprendista (71,7% e 62,7%, rispettivamente). Un dato interessante emerge a partire dal I trim. 2018 e riguarda l'inquadramento dei lavoratori con contratto di apprendistato come quadri d'azienda. Per quanto tale dato sia abbastanza limitato, è risultato fortemente incrementato nel II trim. 2018. È importante, in ogni caso, evidenziare come per questa tipologia di inquadramento la quasi totalità degli individui sia di sesso maschile.

La stipula dei contratti di apprendistato da parte delle aziende è legata al soddisfacimento di diverse necessità in ordine sia all'organizzazione, sia alle indicazioni di legge sull'avviamento dei giovani al lavoro e alla loro qualificazione.

Le motivazioni principali per cui viene stipulato un contratto di apprendistato (Tab. 2, in Appendice) sono principalmente legate alla formazione (59%, in media nel periodo) e all'effettuazione di un periodo di prova finalizzato alla futura assunzione del lavoratore (20,4%, in media nel periodo). È interessante notare come la motivazione legata al periodo di formazione sia notevolmente incrementata nei primi due trimestri del 2018, arrivando a costituire oltre il 75% dei contratti di apprendistato. Di contro, la motivazione legata al periodo di prova ha mostrato un decremento a partire dal II trim. 2017. In tal senso, si deve però notare come nei primi due trimestri del 2018 la quota di intervistati che ha indicato come periodo di prova la motivazione per la stipula di un contratto di apprendistato sia fortemente diminuita, arrivando quasi ad azzerarsi. Una spiegazione di tale risultato è da ricercare nell'introduzione, a partire dal I trim. 2018, di un'ulteriore risposta alla domanda C22 del questionario, ovvero che il contratto di apprendistato "È l'unico che è stato proposto (compreso eventuale lavoro senza contratto e chi dice non c'è una ragione specifica)". Tale circostanza ha determinato un sostanziale spostamento delle risposte degli intervistati verso quest'ultima modalità a scapito della motivazione legata al periodo di prova, come può facilmente rilevarsi sommando le percentuali delle due modalità di risposta. Ciò fa emergere come l'urgenza lavorativa, principalmente nella fascia giovanile d'età, porti ad accettare un tipo di contratto che si ritiene comunque non soddisfacente, come il periodo di prova in apprendistato, pur di avere un'occupazione che si spera possa stabilizzarsi.

Sostanzialmente in linea con il dato sulla motivazione della stipula di un contratto di apprendistato risulta il tipo di inquadramento aziendale (Tab. 3, in Appendice). Infatti, il 45,4% degli individui, in media, che dichiarano di essere in periodo di formazione è effettivamente inquadrato come apprendista, mentre la rimanente parte è quasi equamente suddivisa tra operai e impiegati (28,3% e 26,2%, rispettivamente). È da osservare che i pochi lavoratori con qualifica di quadro che hanno stipulato un contratto di apprendistato hanno indicato come motivazione quella di essere in periodo di formazione o in periodo di prova.

Situazione sostanzialmente analoga si ritrova anche tra coloro che hanno dichiarato di aver accettato un contratto di apprendistato perché in periodo di prova, con il 48,9% che dichiara di essere inquadrato come apprendista, il 31,9% come operaio e il 19,2% come impiegato.

In totale, il 79,5% degli individui ha indicato le due motivazioni prima analizzate come alla base del contratto di apprendistato. Delle rimanenti motivazioni, soltanto quella relativa ad un lavoro occasionale discontinuo presenta una percentuale di una certa rilevanza, pari al 9,5% del totale dei contratti di apprendistato.

Come era da aspettarsi considerato l'impianto normativo che regolamenta il rapporto di apprendistato, la tipologia di contratto di apprendistato e la qualifica assunta dal lavoratore in azienda sono legati anche al livello di educazione scolastica posseduta dal lavoratore. Al riguardo, dall'analisi di Tab. 4 (in Appendice) si può rilevare come nel periodo analizzato il 65,4% degli individui possegga un diploma di scuola media superiore sia con durata biennale o triennale (11,6%), sia con durata quadriennale o quinquennale (53,8%), principalmente del ramo tecnico commerciale, tecnico industriale, geometra. Soltanto meno dell'1% in media, degli individui possiede un titolo di licenza elementare o nessun titolo, mentre il 17,7% possiede un diploma di licenza media inferiore e il 16,6% risulta laureato.

Se si considera il dato in relazione alla posizione nella professione, si può vedere come i diplomati costituiscano, in media nel periodo, la maggioranza all'interno di tutte le qualifiche: 57,5% degli impiegati, 69,8% degli operai, 66,9% degli apprendisti, ad eccezione dei quadri (5,9%). I laureati, invece, costituiscono una quota rilevante degli impiegati (38,6%), una stretta minoranza nella qualifica di operaio (3,8%) e la quasi totalità dei quadri (97,1%). I possessori di diploma di scuola media inferiore, invece, presentano quote significative tra gli operai (26% in media) e gli apprendisti (19,7%), mentre risultano minoritari tra gli impiegati (4,2%).

Oltre al titolo di studio, anche l'età risulta fondamentale per la tipologia di contratto di apprendistato stipulabile. Dall'analisi di Tab. 5 (in Appendice), si può notare come nel periodo la classe 18 – 25 anni sia quella che presenta la maggiore percentuale di individui, con circa il 64,7% in media di lavoratori con contratto di apprendistato, mentre la classe di età 26 – 29 anni rappresenta circa il 26,1%, la classe 30 – 35 anni circa il 7% e, infine, la classe 15 – 17 anni appena l'1% circa. Pertanto, dai risultati emerge come in realtà, con altissime probabilità, tra le varie forme di apprendistato quello professionalizzante sia quella più utilizzata dalle imprese, piuttosto che l'apprendistato di I livello finalizzato al conseguito del titolo di diploma professionale.

La composizione percentuale generale si ripropone sostanzialmente simile se si analizza la ripartizione per classe di età dei titolari di contratto di apprendistato secondo la motivazione della stipula.

Se si analizza, invece, il tipo di inquadramento si può rilevare come, in media nel periodo, la classe 18 – 25 anni costituisca il 71,1% degli operai e il 67,2% degli

apprendisti, mentre la percentuale si abbassa al 52,4% degli impiegati. La classe 26 – 29 anni, invece, presenta la percentuale maggiore tra gli impiegati con, in media nel periodo, il 38,8%, mentre per quanto riguarda la categoria degli operai e degli apprendisti la percentuale scende al 21% e al 24,9%, rispettivamente.

## 4. Conclusioni

Il fenomeno dell'occupazione continua a rappresentare nella discussione scientifica un argomento di grande importanza e ampio dibattito. Numerosi continuano a essere gli interventi di natura giuridica ed economico-finanziaria finalizzati a sostenere il mercato del lavoro e le imprese, numerose continuano ad essere ancora le difficoltà di analisi che si riscontrano non solo per il quadro giuridico che lo governa, ma anche per una disponibilità a tutt'oggi ancora limitata di dati che lo rappresentino. Nel presente lavoro, ferme restando le tante difficoltà suddette, si è proceduto, attraverso l'utilizzo del dato ISTAT riveniente dalla Rilevazione Continua delle Forze di Lavoro, ad analizzare le potenzialità del rapporto di apprendistato quale modalità di ingresso nel mondo del lavoro dei giovani in età 15 – 29 anni. Per quanto sempre approssimativo il dato raccolto dall'ISTAT riguardo alle numerose tipologie di contratti esistenti nel panorama giuslavoristico italiano ed in parte non sempre preciso in relazione al quadro normativo, quanto emerso dai dati evidenzia un utilizzo significativo del contratto di apprendistato all'interno della sua categoria di contratto "atipico", per quanto genericamente sottoutilizzato nel quadro complessivo delle assunzioni. Quanto emerge in maniera sostanziale è un dato territoriale di estrema importanza che vede un uso del contratto di apprendistato soprattutto nelle regioni del Nord Italia, ed un uso praticamente nullo nelle regioni meridionali. Tale dato riflette in realtà la natura "tandem" del contratto di apprendistato così come previsto dalla normativa di riferimento, che vede l'organizzazione e la responsabilità della formazione a carico delle Regioni e Province autonome. Sono principalmente i giovani di sesso maschile in età 18 – 25 anni ad essere oggetto di apprendistato, per inquadramenti che vanno dal ruolo di operaio a quello di apprendista, mentre le donne trovano collocazione come apprendiste soprattutto nel ruolo di impiegate. Tra le motivazioni che sono alla base della stipula di un contratto di apprendistato si ritrovano formazione ed espletamento di un periodo di prova, per quanto con l'ultimo questionario usato per la rilevazione, datato 2018, tra le motivazioni emerga che molti intervistati abbiano accettato tale inquadramento perché il contratto di apprendistato ha rappresentato

l'unica forma contrattuale proposta. Pur avendo un andamento tendenzialmente in crescita a partire dal II trim. 2016, l'analisi ha evidenziato come a partire dal I trim. 2018 i contratti di apprendistato registrino significativi incrementi che lasciano intuire come, terminati gli effetti degli esoneri contributivi correlati alle assunzioni a tempo indeterminato a sostegno del JA, i datori di lavoro in via cautelativa preferiscano assumere i giovani piuttosto ricorrendo al contratto di apprendistato come fosse un periodo di prova "normato" piuttosto che assumersi costi di assunzioni più impegnative. Stante la sempre più necessaria esigenza di poter disporre di database sempre più completi ed esaustivi per l'analisi del fenomeno occupazione, le conclusioni raggiunte dal presente lavoro si ritengono soddisfacenti.

## Riferimenti bibliografici

Calzaroni M. (2008). *Le fonti amministrative nei processi e nei prodotti della statistica ufficiale*, in Atti della Nona Conferenza Nazionale di Statistica.

Camera dei Deputati (2016). *Relazione sull'attività dell'istituto nazionale di statistica, sulla raccolta, trattamento e diffusione dei dati statistici della pubblica amministrazione e sullo stato di attuazione del programma statistico nazionale (Anno 2015) e Rapporto della commissione per la garanzia della qualità dell'informazione statistica (Anno 2015)*, Atti Parlamentari, XVII Legislatura Doc. LXIX, n. 4, Camera dei Deputati, 27 maggio 2016

CEPR (2015). *No Country for Young People? Youth Labour Market Problems in Europe*, CEPR Press, London.

Commission of the European Communities (1998). *Council Regulation (EC) No 577/98 of 9 March 1998 on the organisation of a labour force sample survey in the Community*. Brussels.

*Decreto Legislativo 10 settembre 2003 n. 276, Attuazione delle deleghe in materia di occupazione e mercato del lavoro, di cui alla legge 14 febbraio 2003, n. 30*. Gazzetta Ufficiale n. 235 del 9 ottobre 2003 – Supplemento Ordinario n. 159. Roma.

*Decreto Legislativo 14 settembre 2011, n. 167, Testo unico dell'apprendistato, a norma dell'articolo 1, comma 30, della Legge 24 dicembre 2007, n. 247*. Gazzetta Ufficiale Serie Generale n. 236 del 10 ottobre 2011. Roma.

*Decreto Legislativo 15 giugno 2015 n. 81, Disciplina organica dei contratti di lavoro e revisione della normativa in tema di mansioni, a norma dell'articolo 1, comma 7, della legge 10 dicembre 2014, n. 183*. Gazzetta Ufficiale Serie Generale n. 144 del 24 giugno 2015 – Supplemento Ordinario n. 34. Roma.

*Decreto–Legge 12 luglio 2018 n. 87, Disposizioni urgenti per la dignità dei lavoratori e delle imprese*. Gazzetta Ufficiale Serie Generale n. 161 del 13 luglio 2018. Roma.

ISTAT (2006). *La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione*. Collana Metodi e norme n. 32 – 2006. Roma.

ISTAT (2014). *Rilevazione sulle Forze di Lavoro. Aspetti metodologici dell'indagine*. Roma.

*Legge 10 dicembre 2014 n. 183, Deleghe al Governo in materia di riforma degli ammortizzatori sociali, dei servizi per il lavoro e delle politiche attive, nonché' in materia di riordino della disciplina dei rapporti di lavoro e dell'attività ispettiva e di tutela e conciliazione delle esigenze di cura, di vita e di lavoro*. Gazzetta Ufficiale n. 290 del 15 dicembre 2014. Roma.

*Legge 24 giugno 1997 n. 196, Norme in materia di promozione dell'occupazione.* Gazzetta Ufficiale n. 154 del 4 luglio 1997 – Supplemento Ordinario n. 136. Roma.

*Legge 19 gennaio 1955 n. 25, Disciplina dell'apprendistato.* Gazzetta Ufficiale Serie Generale n. 36 del 14 febbraio 1955. Roma

Marini C., Nicolardi V. (*in pubblicazione*). Livelli e probabilità di transizione del mercato del lavoro: alcune evidenze degli effetti del JOBS ACT sull'occupazione in Italia, in E. Toma (a cura di), *Economia, Istituzioni, Etica e Territorio. Casi di studio ed esperienze a confronto*, Milano, Franco Angeli.

Marsden D., Ryan P. (2009). Institutional Aspects of Youth Employment and Training Policy in Britain, *British Journal of Industrial Relations*, 28: 3.

Nicolardi V., Marini C. (2017). Database del mercato del lavoro a confronto: possibile integrazione per una analisi dinamica dell'occupazione. In: *Metodi e Analisi Statistiche*, Università degli Studi di Bari "Aldo Moro", ISBN: 978-88-6629-013-1, pp. 127-149.

OECD (2017). *Economic Survey of Italy 2017 - OECD*, OECD Publishing, Paris.

Pacifico D., Browne J., Fernández R., Immervoll H., Neumann D., Thévenot C. (2018). *Faces of Joblessness in Italy: A People-Centred Perspective on Employment Barriers and Policies*, IZA Discussion Paper Series, Bonn.

Pinto V. (2004). La modernizzazione promessa. Osservazioni critiche sul Decreto Legislativo 10 settembre 2003, n. 276. In: *Lavoro e nuove regole, dal Libro Bianco al Decreto Legislativo n. 276/2003*, Roma, Ediesse, pp. 221 e seguenti, più appendici.

Quintini G., Manfredi T. (2009). *Going Separate Ways? School-to-Work Transitions in the United States and Europe*, OECD Social, Employment and Migration Working Papers, No. 90, OECD Publishing.

Runci M. C., Di Bella G., Galiè L. (2006). *Il sistema di integrazione dei dati amministrativi in Istat*, ISTAT Working papers n. 18

Sako, M. (1991). Institutional Aspects of Youth Employment and Training Policy: A Comment on Marsden and Ryan, *British Journal of Industrial Relations*, 29: 485-490.

## Appendice

**Tabella 1.** *Contratti di apprendistato secondo il sesso. Periodo I trim. 2014–II trim. 2018.*

| | Sesso | | |
|---|---|---|---|
| Trim. | Maschi | Femmine | Totale |
| *2014* | | | |
| I | 129090 | 98535 | 227625 |
| II | 142199 | 96047 | 238247 |
| III | 128504 | 97640 | 226144 |
| IV | 130501 | 94375 | 224875 |
| *2015* | | | |
| I | 127534 | 83678 | 211212 |
| II | 134771 | 79108 | 213879 |
| III | 148577 | 83114 | 231692 |
| IV | 109689 | 102292 | 211981 |
| *2016* | | | |
| I | 124735 | 64431 | 189166 |
| II | 136084 | 97296 | 233380 |
| III | 118139 | 77045 | 195185 |
| IV | 140911 | 67782 | 208693 |
| *2017* | | | |
| I | 135152 | 83059 | 218210 |
| II | 158970 | 94829 | 253799 |
| III | 138843 | 108517 | 247360 |
| IV | 131683 | 101079 | 232762 |
| *2018* | | | |
| I | 152038 | 121861 | 273899 |
| II | 190171 | 128800 | 318970 |

*Fonte: Archivi della Rilevazione Continua delle Forze di Lavoro. ISTAT.*

**Tabella 2.** *Contratti di apprendistato secondo la motivazione per la stipula del contratto. Periodo I trim. 2014–II trim. 2018.*

| Anno / Trim. | Periodo di formazione | Periodo di prova | Lavoro stagionale | Lavoro occasionale, discontinuo | Lavoro per la realizzazione di un progetto | Occupare un posto vacante | È l'unico che è stato proposto | Altro | Non sa | Totale |
|---|---|---|---|---|---|---|---|---|---|---|
| *2014* | | | | | | | | | | |
| I | 141817 | 45458 | 1670 | 20690 | 1363 | 1424 | - | 15204 | - | 227625 |
| II | 131117 | 54788 | 3710 | 27757 | 2007 | 1840 | - | 15476 | 1551 | 238247 |
| III | 110326 | 73629 | 2868 | 16163 | 2020 | 1193 | - | 19946 | - | 226144 |
| IV | 96935 | 63033 | 5302 | 25792 | 4513 | 2123 | - | 27177 | - | 224875 |
| *2015* | | | | | | | | | | |
| I | 115458 | 54829 | 2421 | 21124 | 964 | 3217 | - | 12584 | 616 | 211212 |
| II | 126907 | 46566 | 2782 | 19361 | 1920 | 770 | - | 15573 | - | 213879 |
| III | 125746 | 54258 | 4606 | 24240 | 1747 | 582 | - | 20513 | - | 231692 |
| IV | 121682 | 47981 | 1422 | 24809 | 4269 | 1746 | - | 10073 | - | 211981 |
| *2016* | | | | | | | | | | |
| I | 119657 | 37954 | 1450 | 13703 | 4311 | 2084 | - | 9749 | 258 | 189166 |
| II | 140254 | 52995 | 1538 | 23834 | 833 | 1259 | - | 12667 | - | 233380 |
| III | 111304 | 41909 | 5415 | 18440 | 1656 | 1806 | - | 14654 | - | 195185 |
| IV | 112390 | 45549 | 571 | 20564 | 4906 | 1062 | - | 23651 | - | 208693 |
| *2017* | | | | | | | | | | |
| I | 119691 | 54828 | 3019 | 26740 | 194 | 5071 | - | 8297 | 371 | 218210 |
| II | 153129 | 47081 | 3458 | 36734 | 984 | 1875 | - | 10077 | 461 | 253799 |
| III | 143792 | 42607 | 9540 | 31817 | 2719 | 1596 | - | 15288 | - | 247360 |
| IV | 152098 | 36577 | | 26781 | 2890 | 1156 | - | 13260 | - | 232762 |
| *2018* | | | | | | | | | | |
| I | 218620 | 11744 | 1050 | 2466 | 996 | 532 | 30071 | 3396 | 5025 | 273899 |
| II | 239620 | 11262 | 1239 | 3729 | 3060 | 3843 | 44089 | 704 | 11425 | 318970 |

*Fonte: Archivi della Rilevazione Continua delle Forze di Lavoro. ISTAT.*

**Tabella 3** *Contratti di apprendistato secondo il tipo di inquadramento e il sesso. Periodo I trim. 2014–II trim. 2018.*

| Anno / Trim. | Quadro | | | Impiegato | | | Operaio | | | Apprendista | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | M | F | Tot. | M | F | Tot. | Tot. | F | Tot. | M | F | Tot. |
| *2014* | | | | | | | | | | | | |
| I | - | - | - | 15572 | 35870 | 51442 | 50750 | 18999 | 69748 | 62769 | 43666 | 106434 |
| II | - | - | - | 17225 | 35404 | 52629 | 57714 | 19542 | 77256 | 67260 | 41102 | 108361 |
| III | - | - | - | 23387 | 32976 | 56363 | 49622 | 18073 | 67695 | 55495 | 46591 | 102086 |
| IV | - | - | - | 24979 | 41803 | 66783 | 46529 | 17295 | 63824 | 58993 | 35276 | 94269 |
| *2015* | | | | | | | | | | | | |
| I | - | - | - | 21555 | 34106 | 55661 | 39645 | 21360 | 61005 | 66334 | 28213 | 94547 |
| II | - | - | - | 18135 | 26802 | 44937 | 47142 | 18133 | 65275 | 69494 | 34174 | 103667 |
| III | - | - | - | 17676 | 28225 | 45900 | 47079 | 17194 | 64274 | 83823 | 37695 | 121518 |
| IV | - | - | - | 17810 | 42672 | 60481 | 33211 | 18321 | 51532 | 58669 | 41300 | 99969 |
| *2016* | | | | | | | | | | | | |
| I | - | - | - | 25383 | 20351 | 45734 | 34855 | 13404 | 48259 | 64496 | 30677 | 95173 |
| II | - | - | - | 21067 | 33355 | 54422 | 48940 | 22904 | 71844 | 66077 | 41037 | 107113 |
| III | - | - | - | 20986 | 28076 | 49062 | 42084 | 11269 | 53353 | 55070 | 37700 | 92770 |
| IV | - | - | - | 18412 | 26450 | 44862 | 52190 | 10510 | 62699 | 70310 | 30822 | 101132 |
| *2017* | | | | | | | | | | | | |
| I | - | - | - | 22699 | 24965 | 47664 | 45070 | 21730 | 66800 | 67383 | 36364 | 103747 |
| II | - | - | - | 36621 | 27147 | 63767 | 54046 | 21117 | 75163 | 68303 | 46566 | 114869 |
| III | - | - | - | 21928 | 31868 | 53796 | 52744 | 26713 | 79457 | 64171 | 49937 | 114107 |
| IV | - | - | - | 17254 | 35944 | 53198 | 44917 | 17699 | 62615 | 69513 | 47437 | 116949 |
| *2018* | | | | | | | | | | | | |
| I | 281 | - | 281 | 29380 | 52660 | 82040 | 46490 | 25864 | 72354 | 75887 | 43337 | 119225 |
| II | 1648 | 103 | 1751 | 28930 | 56685 | 85615 | 80322 | 25166 | 105488 | 79272 | 46846 | 126117 |

*Fonte: Archivi della Rilevazione Continua delle Forze di Lavoro. ISTAT.*

**Tabella 4** *Contratti di apprendistato secondo il titolo di studio posseduto. Periodo I trim. 2014–II trim. 2018.*

| Anno / Trim. | Nessun titolo | Licenza elementare | Licenza media | Diploma 2-3 | Diploma 4-5 | Laurea | Totale |
|---|---|---|---|---|---|---|---|
| *2014* | | | | | | | |
| I | 679 | - | 43895 | 28078 | 119026 | 35947 | 227625 |
| II | - | 576 | 46181 | 33432 | 126673 | 31385 | 238247 |
| III | - | 1069 | 43368 | 27948 | 113777 | 39983 | 226144 |
| IV | - | - | 37405 | 28290 | 122468 | 36712 | 224875 |
| *2015* | | | | | | | |
| I | 1072 | 1481 | 38279 | 23935 | 113601 | 32844 | 211212 |
| II | - | - | 43379 | 23418 | 117846 | 29236 | 213879 |
| III | - | - | 49312 | 23024 | 110596 | 48760 | 231692 |
| IV | - | 420 | 32586 | 25461 | 109971 | 43544 | 211981 |
| *2016* | | | | | | | |
| I | - | - | 39387 | 18643 | 106200 | 24936 | 189166 |
| II | - | - | 57819 | 29177 | 117276 | 29108 | 233380 |
| III | - | - | 33212 | 26251 | 111762 | 23959 | 195185 |
| IV | - | - | 30135 | 25270 | 115825 | 37463 | 208693 |
| *2017* | | | | | | | |
| I | 1646 | 1120 | 38112 | 20521 | 113259 | 43553 | 218210 |
| II | 2163 | - | 28140 | 31013 | 152688 | 39795 | 253799 |
| III | 449 | - | 37215 | 38709 | 126611 | 44375 | 247360 |
| IV | - | 228 | 35764 | 23599 | 135661 | 37511 | 232762 |
| *2018* | | | | | | | |
| I | 579 | 632 | 41832 | 18344 | 162147 | 50365 | 273899 |
| II | 1352 | - | 58447 | 33209 | 159436 | 66528 | 318970 |

*Fonte: Archivi della Rilevazione Continua delle Forze di Lavoro. ISTAT.*

**Tabella 5** *Contratti di apprendistato secondo la classe di età in anni compiuti. Periodo I trim. 2014–II trim. 2018.*

| Anno / Trim. | 15-17 | 18-25 | 26-29 | 30-35 | Totale |
|---|---|---|---|---|---|
| *2014* | | | | | |
| I | 534 | 145368 | 72041 | 9683 | 227625 |
| II | 1409 | 150654 | 65034 | 21149 | 238247 |
| III | 1131 | 149991 | 55805 | 19217 | 226144 |
| IV | 6811 | 136734 | 57535 | 23795 | 224875 |
| *2015* | | | | | |
| I | 3646 | 132281 | 56844 | 18442 | 211212 |
| II | 3325 | 140272 | 47869 | 22414 | 213879 |
| III | 1197 | 134405 | 73903 | 22187 | 231692 |
| IV | 1227 | 133883 | 65600 | 11272 | 211981 |
| *2016* | | | | | |
| I | 2663 | 125071 | 50513 | 10919 | 189166 |
| II | 2560 | 149731 | 66953 | 14136 | 233380 |
| III | 2167 | 125969 | 59496 | 7553 | 195185 |
| IV | 976 | 149890 | 49901 | 7926 | 208693 |
| *2017* | | | | | |
| I | 3407 | 143896 | 50488 | 20420 | 218210 |
| II | - | 167380 | 73118 | 13301 | 253799 |
| III | 3100 | 158386 | 72850 | 13023 | 247360 |
| IV | 3333 | 157437 | 54363 | 17629 | 232762 |
| *2018* | | | | | |
| I | 1768 | 178306 | 76104 | 17721 | 273899 |
| II | 4120 | 207021 | 86110 | 21719 | 318970 |

*Fonte: Archivi della Rilevazione Continua delle Forze di Lavoro. ISTAT.*

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018*
DIPARTIMENTO DI ECONOMIA E FINANZA
ISBN: 978-88-6629-048-3
pp. 83-108

# Il nuovo sistema statistico *"a misura di comune"*: una applicazione sul territorio lucano.

**Antonella Bianchino[1], Agata Maria Madia Carucci[1], Giovanni Vannella[2*].**

*[1]Istat – Sede Territoriale per la Basilicata, [2]Università degli Studi di Bari Aldo Moro*

**Riassunto:** Con il presente lavoro si illustrano le potenzialità del nuovo sistema statistico *"a misura di comune"* analizzando una sua utilizzazione in riferimento all'analisi della relazione tra consumo del suolo e relative attività antropiche ed economiche, con una applicazione al territorio lucano.

**Keywords:** Indicatori regionali e comunali; integrazioni tra fonti amministrative e fonti statistiche; Sistema informativo statistico *"a misura di comune"*, BES.

## 1. Introduzione

Le informazioni statistiche presentano una fruibilità tanto maggiore quanto più, a parità di altre condizioni, riescano a fornire, tempestivamente, quadri dettagliati a livello territoriale. Tuttavia, alla maggiore tempestività del dato statistico si può associare, sovente, una minore accuratezza dello stesso.

Per tali motivi, da tempo l'Istat, così come gli altri Istituti di statistica nazionali, di concerto con l'Eurostat, affianca alle tradizionali informazioni statistiche, censuarie e campionarie, informazioni di natura amministrativa, con il triplice vantaggio di ottenere una maggiore celerità, un minore aggravio di costi e un maggiore spettro informativo. Esperienze di tal genere sono rinvenibili, ad esempio, nella produzione del FRAME – SBS (Carucci, 2016), o nell'implementazione dei conti satelliti ambientali (Carucci, 2017). Una interessante, ulteriore e recentissima applicazione di tale offerta informativa statistica "integrata" è fornita dal nuovo si-

[*] Autore corrispondente: Giovanni Vannella, e-mail giovanni.vannella@uniba.it
Sebbene il lavoro sia frutto del lavoro congiunto dei tre autori, ad A. Bianchino sono attribuiti i paragrafi 3 e 4, ad A. M. M Carucci sono attribuiti il paragrafo 5, a G. Vannella i paragrafi 1,2 e 6

stema statistico *"a misura di comune"*[1] implementato dall'Istat attraverso un progetto che, partito nel 2016, ha portato alla pubblicazione dei primi risultati nel– l'agosto 2018.

Nel presente lavoro, dopo aver illustrato i principali aspetti metodologici e descrittivi di tale sistema informativo, si illustra una applicazione statistica riferita al territorio lucano.

## 2. Perché un sistema informativo statistico per i comuni

In linea con il percorso intrapreso da Eurostat, l'Istat sperimenta l'utilizzo di nuove fonti e l'applicazione di metodi innovativi nella produzione di dati, offrendo i risultati delle sperimentazioni alla fruizione e alla valutazione degli utenti. Il Progetto "Misure di benessere e programmazione a livello comunale" rappresenta un passo particolarmente importante per l'Istituto, nell'ambito dell'implementazione di indicatori socio-demografici, economici e ambientali a livello comunale.

D'altra parte, la recente normativa in materia di armonizzazione dei sistemi contabili e degli schemi di bilancio delle Regioni, degli Enti Locali e dei loro Organismi (Decreto legislativo 23 giugno 2011, n. 118 e successive modificazioni) ha aperto nuove opportunità di utilizzo dell'informazione statistica territoriale con interessanti risvolti applicativi. Essa ha introdotto il "documento unico di programmazione" (DUP), che sostituisce la "relazione previsionale e programmatica" (art. 74) e che deve contenere l'analisi delle condizioni esterne ed interne all'Ente.

Nell'ambito del progetto è stata effettuata, nella primavera 2017, una ricognizione rapida sui DUP di un sotto-insieme di 40 comuni, evidenziando una forte eterogeneità delle soluzioni adottate, tanto nel grado di articolazione delle analisi, quanto nella scelta delle fonti statistiche e nella strutturazione tematica. Tali evidenze empiriche hanno indirizzato la definizione e la struttura di un set di indicatori, che può pertanto fornire un supporto informativo pertinente nella redazione dei DUP, caratterizzato da requisiti di qualità del dato e di comparabilità territoriale.

Nella dinamicità di tale contesto, nasce il sistema informativo statistico *"a misura di comune"*, con la finalità di colmare la lacuna informativa disponibile con riferimento al dettaglio territoriale comunale.

[1] Il sistema informativo statistico "a misura di comune" è consultabile all'indirizzo web http://amisuradicomune.istat.it/aMisuraDiComune/

**Figura 1.** *Quadro sinottico degli indicatori del Rapporto UrBes 2015* (a)

| SALUTE | ISTRUZIONE | LAVORO E CONCILIAZIONE TEMPI DI VITA | BENESSERE ECONOMICO | RELAZIONI SOCIALI | POLITICA E ISTITUZIONI |
|---|---|---|---|---|---|
| | Partecipazione scuola infanzia (N) | | | | Partecipazione elettorale per genere |
| | Persone con almeno il diploma superiore (N) | | | | Rappresentanza femminile |
| Speranza di vita | Persone con titolo universitario (N) | | Reddito disponibile | | Donne negli organi decisionali comunali (N) |
| Mortalità infantile | Uscita precoce dalla istruzione e formazione. (N) | Occupazione | DISTRIBUZIONE DEI REDDITI IRPEF (N) | Volontari nelle UL non profit (N) | Età rappresentanza |
| Mortalità per incidenti | Giovani che non lavorano e non studiano (N) | Mancata partecipazione al lavoro | Qualità abitazione (N) | Istituzioni non profit (N) | Età media organi decisionali comunali (N) |
| Mortalità per tumore | Competenza alfabetica | Infortuni mortali | Individui in famiglie senza occupati (N) | Cooperative sociali (N) | RENDICONTAZIONE SOCIALE DELLE IST. PUBBLICHE (N) |
| Mortalità per malattie croniche | Competenze numerica | Occupazione delle donne con e senza figli (N) | SOFFERENZE BANCARIE DELLE FAMIGLIE (N) | LAVORATORI RETRIBUITI C. SOCIALI (N) | Lunghezza procedimenti civili (N) |
| | | | | | |
| Omicidi | | Biblioteche pubbliche (N) | DISPERSIONE ACQUA POTABILE (N) | Brevetti | Servizi infanzia |
| Furti in abitazione (N) | | Musei (N) | Qualità aria urbana | Specializzazione produttiva (N) | SCUOLE CON PERCORSI ACCESSIBILI (N) |
| Borseggi (N) | | UTENTI BIBLIOTECHE (N) | Verde urbano | Connessione Internet a banda larga (N) | Rifiuti in discarica (N) |
| Rapine (N) | | VISITATORI MUSEI (N) | Aree naturali protette (N) | | Raccolta differenz. rifiuti |
| | | Verde storico | ORTI URBANI (N) | | Tempo mobilità (N) |
| | | Tessuto urbano storico | TELERISCALDAMENTO (N) | | Trasporto pubblico locale (N) |
| | | | INQUINAMENTO ACUSTICO (N) | | PISTE CICLABILI (N) |
| | | | AUTO CON STANDARD <EURO-4 (N) | | AREE PEDONALI (N) |
| | | | | | INFOMOBILITÀ (N) |
| | | | | | INCIDENTALITA' STRADALE (N) |
| | | | | | PEDONI VITTIME DI INCIDENTI (N) |
| SICUREZZA | BENESSERE SOGGETTIVO | PAESAGGIO E PATRIMONIO CULTURALE | AMBIENTE | RICERCA E INNOVAZIONE | QUALITÀ DEI SERVIZI |

(a) In minuscolo, gli indicatori utilizzati o riconducibili a quelli della misura nazionale del Bes, in maiuscolo, quelli aggiuntivi introdotti per il Bes delle città. (N) = indicatori nuovi rispetto al rapporto UrBes 2013.

Fonte: Urbes – Il Benessere equo e sostenibile nelle città 2015

Esso costituisce il risultato del Progetto "Misure di benessere e programmazione a livello comunale" avviato nel 2016 e rappresenta l'evoluzione di due precedenti esperienze realizzate dall'Istat:

- "UrBes – il benessere equo e sostenibile nelle città", progetto realizzato tra il 2012 e il 2015 in partnership tra Istat e 29 comuni, che ha portato alla pubblicazione di due Rapporti nel 2013 e nel 2015, avviando la misurazione degli indicatori Bes a livello comunale e introducendo anche ulteriori indicatori di benessere urbano (Fig. 1). Con riferimento all'aspetto territoriale, il progetto UrBes vede il coinvolgimento di una rete delle città che comprende le 10 Città Metropolitane (Torino, Genova, Milano, Venezia, Bologna, Firenze, Roma, Napoli, Bari e Reggio di Calabria), le 4 Città Metropolitane già previste ma non ancora costituite (Palermo, Messina, Catania e Cagliari) ed altri 15 Comuni (Brescia, Bolzano, Verona, Trieste, Parma, Reggio Emilia, Cesena, Forlì, Livorno, Prato, Perugia, Terni, Pesaro, Potenza, Catanzaro);
- la sperimentazione di nuovi indicatori comunali nell'ambito di una Convenzione tra Istat e Regione Basilicata, realizzata anche avvalendosi, per la prima volta, del patrimonio informativo costituito dall'Archivio Integrato Microdati Economici e Demo-sociali (ARCH.I.M.E.DE). La predetta convenzione tra Istat e Regione Basilicata, risponde, d'altra parte, a specifiche esigenze dell'ente Regione, che punta ad avere una conoscenza analitica e continuamente aggiornata delle condizioni sociali, economiche, demografiche, ambientali dell'area sulla quale va ad incidere con politiche di sviluppo del territorio.

Il progetto in esame rappresenta pertanto la prima attuazione da parte dell'Istat delle attività del protocollo d'intesa con Anci e Upi, che, tra i sui obiettivi, prevede lo «sviluppo di basi di dati e di sistemi informativi integrati e tra loro armonizzati, di elevato dettaglio territoriale».

## 3. Caratteristiche del sistema informativo "a misura di comune": struttura tematica, funzionalità e fonti dei dati

La strutturazione tematica, ideata come strumento per i *policy maker* ma anche per i cittadini, quindi di facile ed immediata fruibilità, presenta una forte coerenza con la classificazione degli indicatori di benessere equo e sostenibile, i domini Bes (Istruzione, Lavoro, Benessere economico, Politica e Istituzioni, Ricerca e Innovazione) sono stati mutuati su base comunale.

Al tempo stesso, si è tenuto conto di alcuni ambiti tematici particolarmente significativi e riconoscibili ai fini dell'utilizzo degli indicatori nel Documento Unico di Programmazione degli Enti Locali (Popolazione e famiglie, Cultura, comunicazione e tempo libero, Territorio e ambiente, Economia insediata, Infrastrutture e mobilità). Ad ora, sono 71 gli Indicatori calcolati e pubblicati a livello comunale, ripartiti in 10 Aree tematiche:

- Popolazione e famiglie
- Istruzione
- Lavoro
- Benessere economico
- Politica e istituzioni
- Cultura, comunicazione e tempo libero
- Territorio e ambiente
- Economia insediata
- Ricerca e innovazione
- Infrastrutture e mobilità

Sono attualmente in fase di studio ulteriori indicatori su Ambiente, Sicurezza dei cittadini e Mortalità.

La costruzione di questa molteplicità di indicatori è stata possibile grazie alla definizione di un sistema multi-fonte, nel quale sono state valorizzate fonti di carattere sperimentale accanto ad altre più consolidate. Tra le fonti sperimentali, un posto di rilievo spetta alle basi dati realizzate all'interno del progetto ARCH.I.M.E.DE, sviluppato dall'Istat di concerto con il SISTAN, con l'obiettivo "*di ampliare l'offerta informativa del SISTAN mediante la produzione di collezioni di dati elementari, ottenuti dall'integrazione di archivi amministrativi, utili alla programmazione territoriale e settoriale e alla valutazione delle politiche pubbliche anche a livello regionale e locale, in un'ottica di riduzione dei costi e del fastidio statistico"*[2] (Garofalo, 2014)

Il Sistema Integrato dei Microdati ARCH.I.M.E.DE ha consentito di elaborare numerosi indicatori a livello comunale, sia relativi a importanti aspetti strutturali sia attinenti a fenomeni osservati anche nell'ambito della misurazione del Benessere equo e sostenibile. Tali indicatori sono stati selezionati e denominati attraverso uno specifico percorso di verifica.

Un contributo significativo riguarda anche l'utilizzo degli Open Data resi disponibili da altri enti del Sistan, come il Ministero dell'Interno, il Ministero

[2] https://www.istat.it/it/files/2014/11/IWP-n.-9-2014.pdf

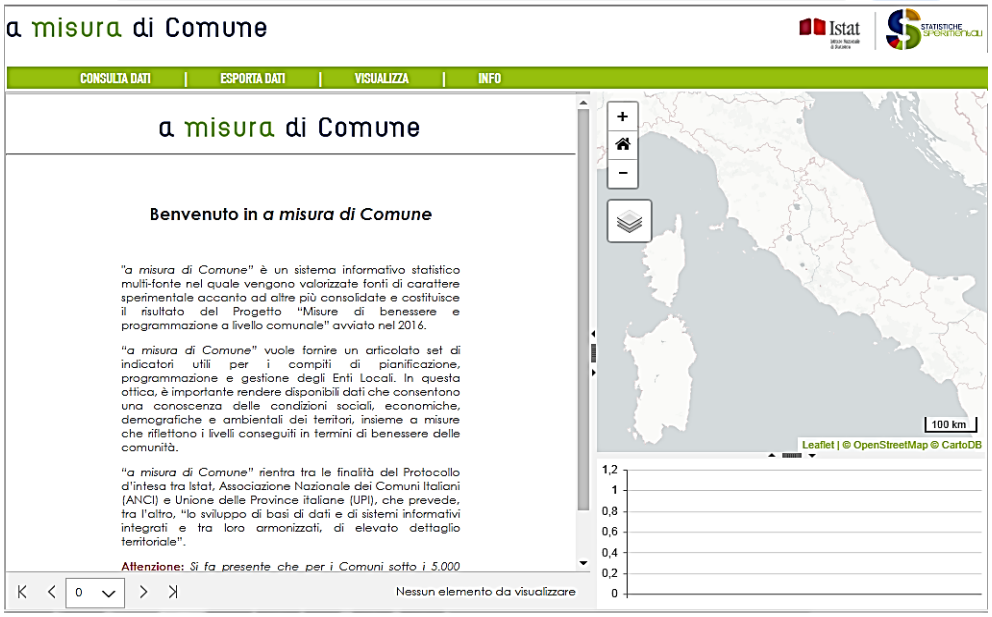
dell'Economia e delle Finanze, il Ministero dello Sviluppo Economico, l'Istituto superiore per la protezione e la ricerca ambientale (ISPRA), il Progetto sui fabbisogni standard di SOSE in collaborazione con IFEL.

Altre fonti rientrano invece nella produzione statistica corrente dell'Istat come, ad esempio, le statistiche demografiche, i registri statistici delle imprese attive, la rilevazione "Dati ambientali nelle città".

## 4 La navigazione nel sistema informativo

L'accesso al sistema informativo è libero per tutti gli utenti attraverso la sezione *"Statistiche sperimentali"* del sito istituzionale dell'Istat (Fig. 2).

**Figura 2.** *Schermata di accesso al sistema informativo statistico: "a misura di comune"*



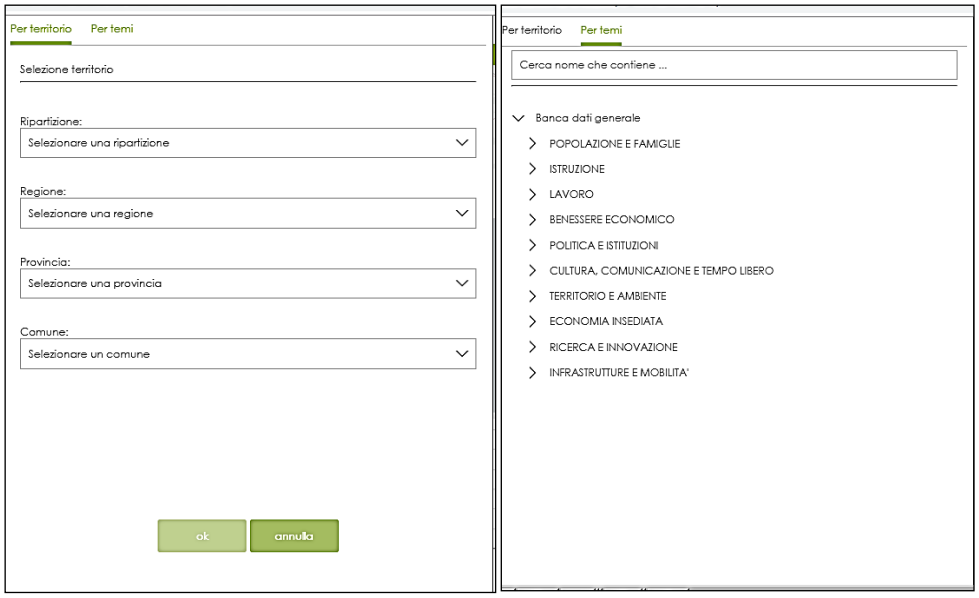

La consultazione degli indicatori è prevista attraverso due modalità (Fig. 3):

- per temi: selezionando un indicatore, è possibile scegliere la disaggregazione territoriale secondo diversi livelli di geografia amministrativa (Ripartizioni, Regioni, Province e Città Metropolitane, Comuni capoluogo, tutti i comuni), di classificazioni secondo caratteristiche territoriali (zone altimetriche, gradi di urbanizzazione, comuni litoranei e isolani ecc..) e di aree funzionali (sistemi locali del lavoro, distretti industriali, aree Eurostat).

La piattaforma di interrogazione consente di fare "partizioni utente" nella navigazione per tema, selezionando due o più aree (comuni, province ecc..) a fini di confronto.

- per territorio: selezionando un territorio o un'area sovra-comunale, si estrae una serie storica dei dati per tutti gli indicatori valorizzati.

**Figura 3.** *Schermata di consultazione dati nel sistema informativo statistico: "a misura di comune"*




Numerose sono le funzionalità del sistema informativo *"a misura di comune"*, di seguito si riporta una sintesi delle stesse:

- Navigazione interattiva del dato in modalità tabellare, grafica e cartografica;
- Dato n-dimensionale (tavole);
- Sistema alimentato dalle variabili di base al livello territoriale di riferimento;
- Indicatori calcolati in tempo reale a partire dalle variabili di base;
- Aggregazione in tempo reale per i livelli territoriali superiori;
- Sistema di controllo sull'indicatore computato (soglia, indicatore di qualità, etc.);
- Georeferenziazione statistica dinamica ed in tempo reale su geografie: amministrative, amministrative precedenti, funzionali, statistica europea, funzionali europee e caratteristiche del territorio;

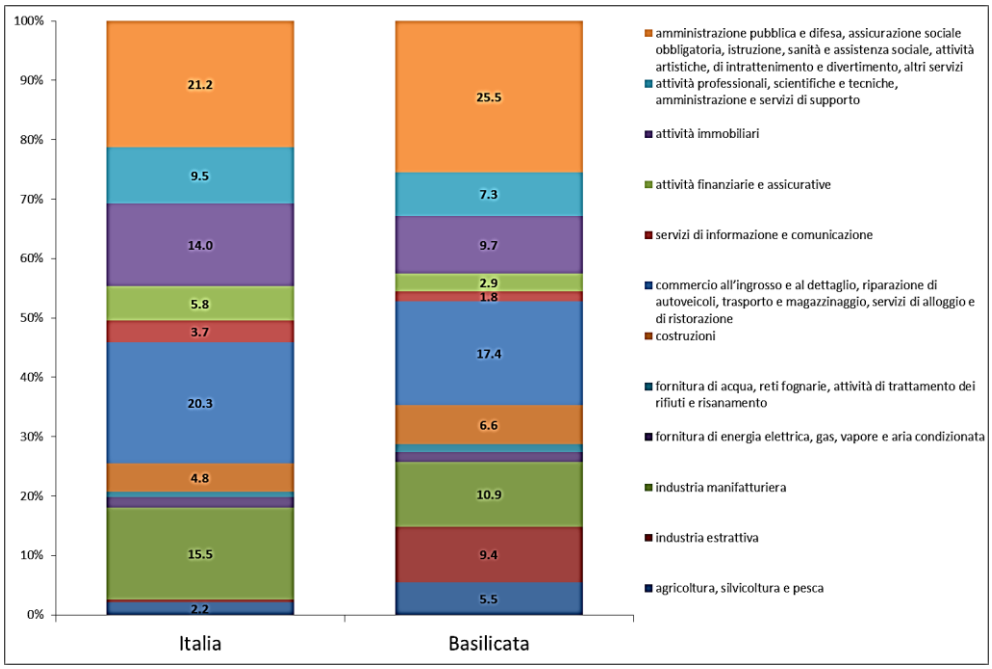
- Creazione di partizioni territoriali utente;
- Navigazione per tema e navigazione per territorio (con serie storica);
- Metadatazione statistica e territoriale;
- Funzionalità Gis;
- Esportazione in modalità CSV/Excel;
- Stampa grafica e cartografica in modalità PDF;
- Possibilità di scaricare gli shapefile di riferimento.

## 5. Un'analisi a partire da *"a misura di comune"*

Al fine di comprendere le potenzialità del nuovo sistema informativo statistico *"a misura di comune"*, si analizza, di seguito, il tessuto economico e produttivo della regione Basilicata, attraverso l'ausilio di alcuni indicatori introdotti nel nuovo sistema informativo e calcolati su base comunale.

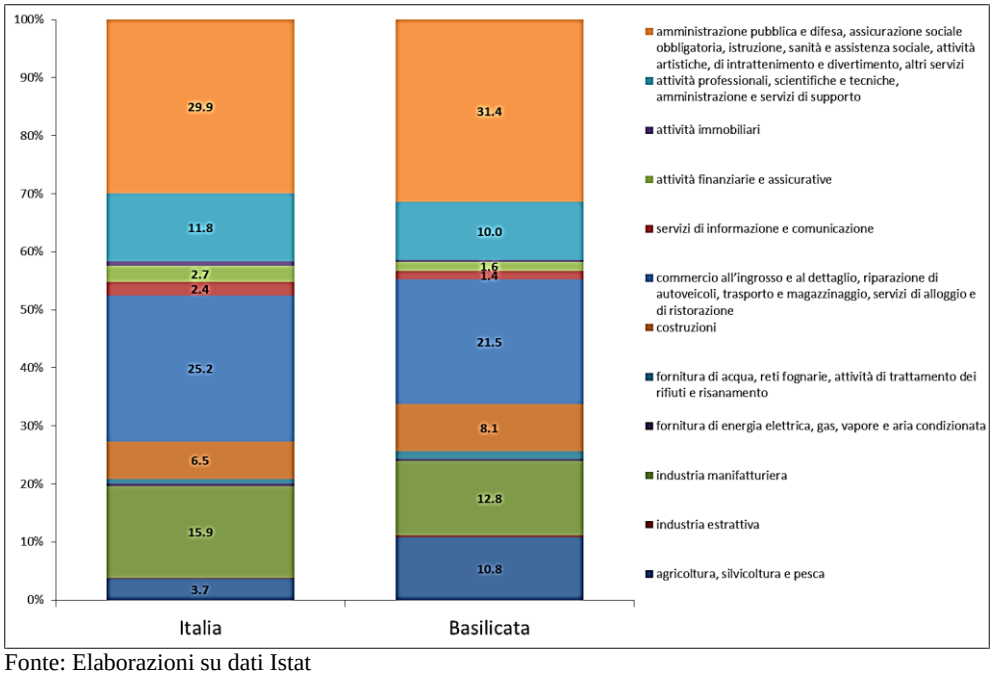
Come è noto, la Basilicata si caratterizza per la presenza di singolari attività economiche desumibili anche da un confronto tra la distribuzione di valore aggiunto e addetti per attività economica in Italia e Basilicata (Figg. 4 e 5).

**Figura 4.** *Composizione del valore aggiunto per attività economica: Italia e Basilicata. Anno 2015*




Fonte: Elaborazioni su dati Istat

**Figura 5.** *Composizione degli addetti per attività economica: Italia e Basilicata. Anno 2015*




Fonte: Elaborazioni su dati Istat

Nella regione, più importante rispetto all'intera economia è il contributo, in termini di valore aggiunto, dell'agricoltura, dell'attività estrattiva e della pubblica amministrazione. In termini occupazionali, la distribuzione delle attività economiche è particolarmente diversa: le attività estrattive, pur rappresentando quasi il 10% del valore aggiunto, rappresentano meno dell'1% degli occupati; di contro il 5,5% del valore aggiunto dell'agricoltura deriva da più del 10% di occupati nel settore.

In questo contesto, affiancare alla tradizionale analisi economica su base regionale, l'analisi su base comunale del sistema produttivo lucano e delle interazioni tra lo stesso ed il territorio e l'ambiente risulta particolarmente utile. Il primo indicatore, presente in *"a misura di comune"* ed analizzato in questa sede è "consumo di suolo". Esso spiega la relazione tra attività economiche ed antropiche persistenti sul territorio e il consumo dello stesso. L'indicatore consumo di suolo (C_suolo) è definito come:

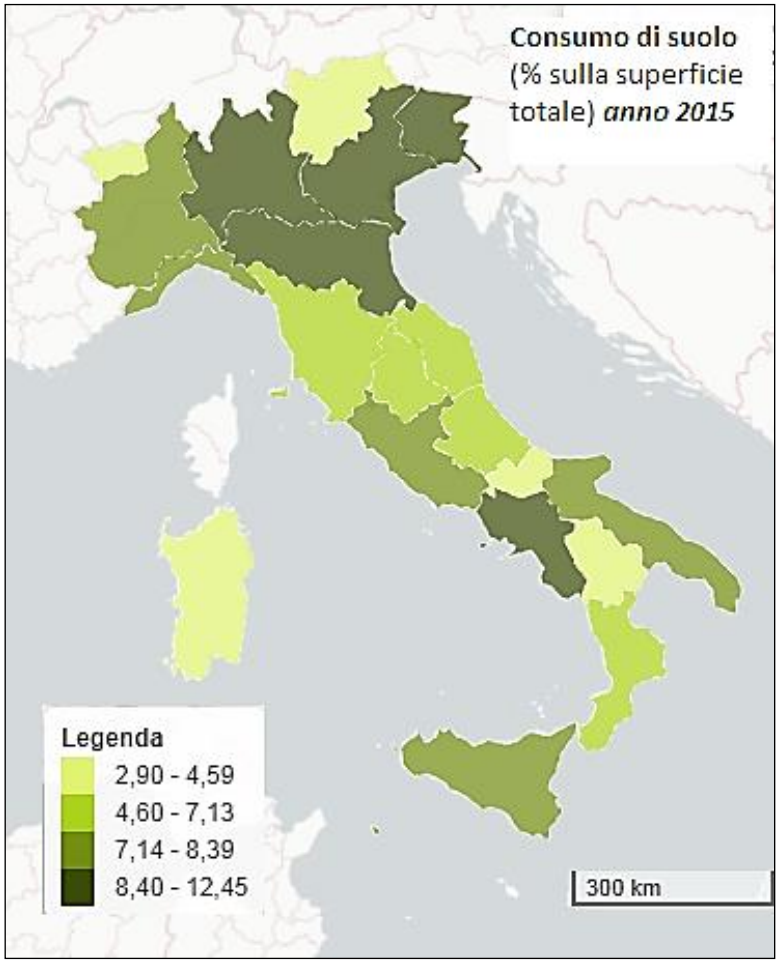
$$C_suolo = \frac{Suolo\ consumato}{Suolo\ consumato + Suolo\ non\ consumato + Suolo\ non\ classificato} \cdot 100\ .$$

Il consumo di suolo, calcolato su dati ISPRA, "è un fenomeno associato alla perdita di una risorsa ambientale fondamentale, dovuta all'occupazione di superficie originariamente agricola, naturale o semi-naturale. Il fenomeno si riferisce, quindi, a

un incremento della copertura artificiale di terreno, legato alle dinamiche insediative. Un processo prevalentemente dovuto alla costruzione di nuovi edifici, fabbricati e insediamenti, all'espansione delle città, alla densificazione o alla conversione di terreno entro un'area urbana, all'infrastrutturazione del territorio. Il consumo di suolo è, quindi, definito come una variazione da una copertura non artificiale (suolo non consumato) a una copertura artificiale del suolo (suolo consumato)"[3].

Dal cartogramma seguente (Fig. 6), risulta chiaramente come il dato regionale della Basilicata sia tra i più bassi della penisola, 3,38% per il 2015, secondo solo alla Valle d'Aosta 2,90%.

**Figura 6.** *Consumo di suolo su base regionale. Anno 2015*




Fonte: Sistema informativo *"a misura di comune"*

[3] Definizione riportata nella sezione metadati di http://amisuradicomune.istat.it/aMisuraDiComune/

In effetti, l'indicatore spiega solo in parte il tanto discusso *trade off* tra economia ed ambiente. Esso è associato in generale alle attività antropiche, correlate con la stessa struttura urbana e con numerosi fattori storici ed economici. Pertanto, questo tipo di aggregazione territoriale risulta essere poco significativa se non addirittura fuorviante.

Se da un lato indicatori specifici, in fase di pubblicazione (DMI regionale[4]) possano spiegare meglio il fenomeno, dall'altro una maggiore sensibilità territoriale può essere, senza dubbio, di aiuto almeno in questa fase esplorativa del lavoro in oggetto.

Calcolando l'indicatore su base provinciale prima, e comunale poi, emergono significative specificità. Prima di tutto, a livello provinciale si ha un'incidenza maggiore di consumo di suolo in provincia di Potenza rispetto alla provincia di Matera (Tab. 1). La variabilità dell'indicatore sul territorio emerge con maggiore intensità su base comunale (Fig. 7), presentando una maggiore concentrazione sulla costa ionica e nella parte settentrionale della provincia di Potenza.

In particolare, l'indicatore assume valori inferiori al 2% in aree morfologicamente complesse, interessate anche da gravi problemi di dissesto idrogeologico e carenze infrastrutturali (Craco o Aliano per fare qualche esempio) o strettamente montane e, per loro natura di difficile accesso (San Costantino Albanese, Terranova di Pollino, Cersosimo, San Chirico Raparo e Pietrapertosa) e valori superiori al 5%, con picchi del 10%, in aree più facilmente accessibili e in cui si sono, negli anni sviluppate le principali attività produttive. Questa forte variabilità, più alta in Basilicata rispetto ad altre regioni della penisola, ci conferma la necessità di esaminare il tessuto produttivo lucano su livelli territoriali più fini.

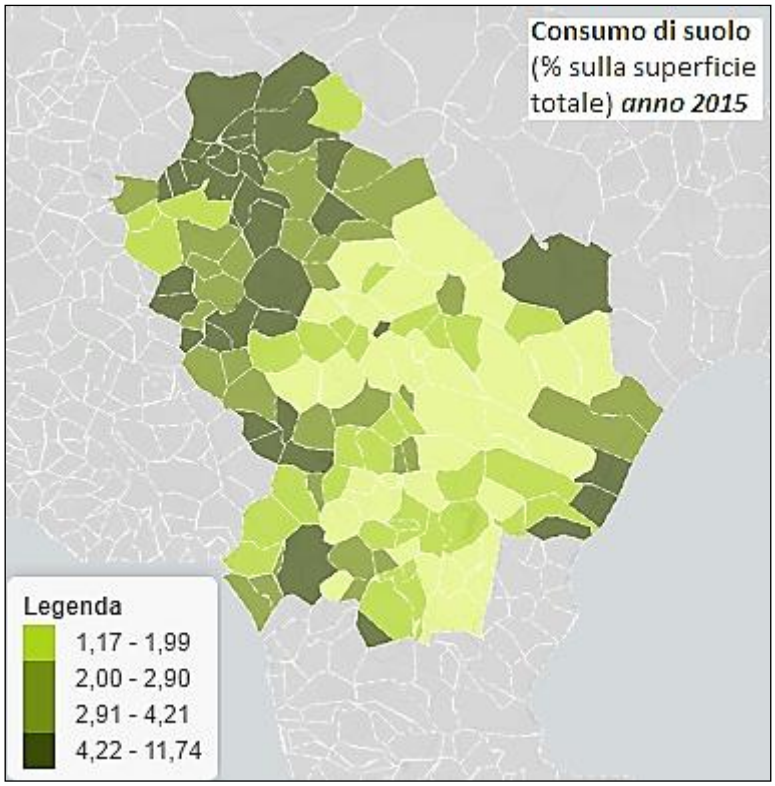
**Tabella 1.** *Consumo di suolo in Basilicata per provincia. Dati percentuali. Anno 2015*

| Provincia | Consumo di suolo (Percentuale sulla superficie totale) |
|---|---|
| Potenza | 3,65 |
| Matera | 2,87 |

Fonte: Sistema informativo *"a misura di comune"*

[4] Il Direct Material Input (DMI) misura le pressioni del sistema economico sull'ambiente naturale. È definito come quantità totale di risorse naturali estratte all'interno della regione e di risorse importate che entra nel sistema economico per essere successivamente trasformata e/o commercializzata, è dato dalla somma delle biomasse, dei combustibili fossili e degli altri minerali estratti dal suolo nazionale, nonché dei beni di ogni tipo importati dall'estero, inclusi nell'aggregato secondo il loro peso effettivo, indipendentemente dal grado di lavorazione.

**Figura 7.** *Consumo di suolo in Basilicata. Anno 2015*




Fonte: Sistema informativo *"a misura di comune"*

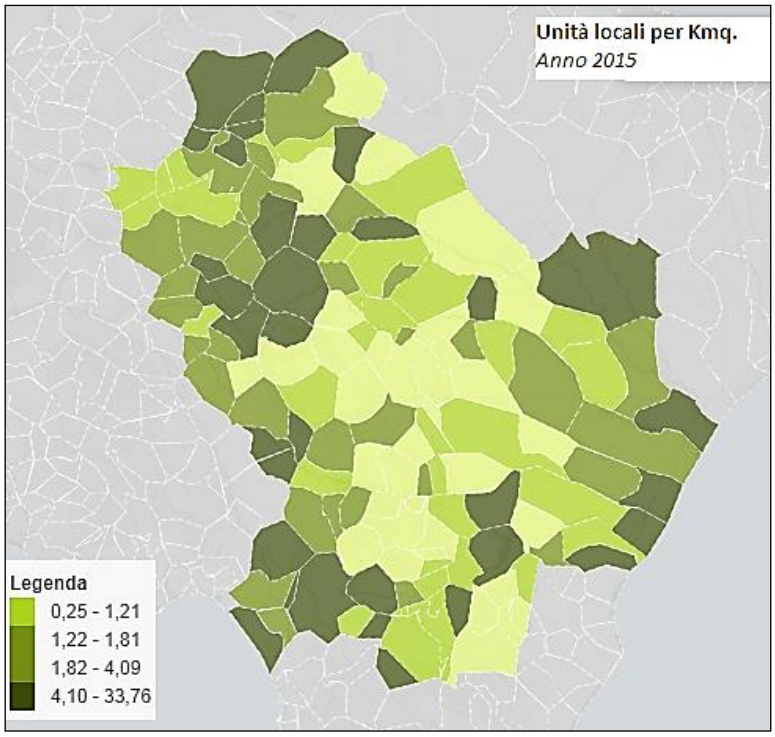
Appare doveroso, a questo punto, capire a livello comunale quali siano le attività produttive che persistono sul territorio, provando quindi a confrontare la distribuzione dell'indicatore Consumo di suolo con indicatori più specifici che descrivano il sistema produttivo locale e che sono classificati nel sistema informativo nell'area tematica *"economia insediata"*. Gli indicatori utilizzati sono, tutti di fonte ASIA[5] unità locali, i seguenti:

- Densità delle unità locali
- Addetti delle Unità locali per tipologia di attività economica (Valori percentuali)
- Unità locali per tipologia di attività economica (Valori percentuali)

Il primo indicatore è dato dal rapporto tra il numero di unità locali e la superficie comunale (Fig.8) e rappresenta un chiaro segnale della presenza sul territorio dei principali insediamenti produttivi.

[5] Archivio Statistico Imprese Attive

**Figura 8.** *Densità delle unità locali in Basilicata. Anno 2015*




Fonte: Sistema informativo *"a misura di comune"*

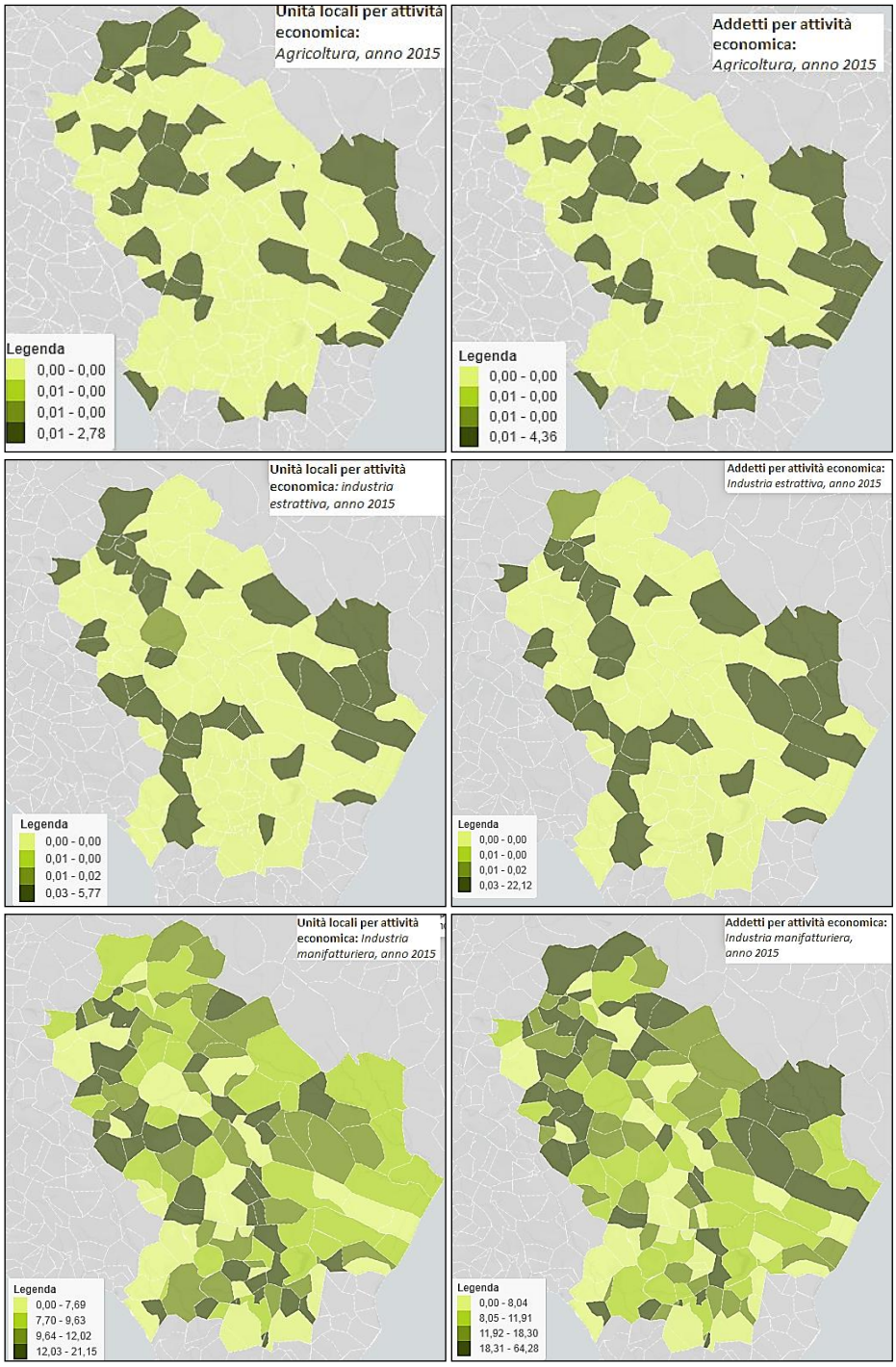
Il secondo indicatore è dato dal rapporto tra gli Addetti delle unità locali per tipologia di attività economica e il totale addetti (per cento); l'ultimo dal rapporto tra il Numero di unità locali per tipologia di attività economica e il totale unità locali (per cento).

La classificazione per tipologia riprende la classificazione utilizzata nell'analisi dei distretti industriali 2011 (Istat 2011[6]).

Per semplicità si riporta di seguito la distribuzione a livello comunale delle unità locali e degli addetti per le tre tipologie di attività economica che maggiormente sono legate al consumo di suolo: agricoltura, attività estrattiva ed industria manifatturiera. A livello comunale, vi è una distribuzione quasi sovrapponibile tra addetti ed attività economica sebbene la cosa non sia necessariamente vera a livello regionale per un effetto di composizione. È, d'altra parte, evidente la correlazione tra insediamenti produttivi e il consumo del suolo (Fig. 9).

[6] cfr. Istat, *I distretti industriali 2011*, p. 73. il volume è reperibile all'indirizzo: https://www.istat.it/it/archivio/150320

**Figura 9.** *Distribuzione delle unità locali e degli addetti per attività economica. Anno 2015*




Fonte: Sistema informativo *"a misura di comune"*

L'indicatore Consumo del suolo, come più volte sottolineato, non risente solo delle attività produttive in senso stretto, ma anche dello sviluppo di attività antropiche che portano il comune ad attrarre popolazione dalle regioni e dai comuni vicini o almeno a trattenere popolazione nel territorio e a favorire i flussi di pendolarismo in ingresso. Questi elementi possono essere esaminati con l'ausilio di due indicatori che misurano il potere attrattivo dei comuni lucani e degli spostamenti per studio e lavoro tra essi.

Gli indici analizzati, tratti dall'area tematica *"infrastrutture e mobilità"*, sono l'indice di attrazione e l'indice di autocontenimento.

L'indice di attrazione (I_attraz) è definito come:

$$I_attraz = \frac{\textit{Flussi in entrata nel comune per motivi di studio o lavoro}}{\begin{array}{c}\textit{(Flussi in entrata nel comune per motivi di studio o lavoro +}\\ \textit{+Flussi in uscita dal comune per motivi di studio o lavoro +}\\ \textit{+ Residenti che lavorano o studiano nel comune)}\end{array}} \cdot 100 .$$

Le informazioni sulla cui base viene sviluppato l'indicatore sono quelle contenute in diverse fonti amministrative, concernenti i luoghi di studio (Anagrafe degli studenti delle scuole primarie e secondarie – MIUR; Anagrafe delle iscrizioni universitarie – MIUR) e di lavoro (ASIA_Unità Locali; INAIL_PAT; Cedolini stipendiali; INPS_Autonomi Agricoltura) della popolazione residente (Liste Anagrafiche Comunali). I risultati, disponibili, forniscono un quadro della "mobilità" desunto da una geografia di fatto delineata dai segnali amministrativi di presenza nei luoghi di lavoro e studio dei residenti e come tali differiscono, seppure in misura non sostanziale, dai dati censuari descrittivi invece dei comportamenti rilevati tramite intervista[7]. L'indice di autocontenimento (I_autocont) è definito come:

$$I_autocont = \frac{\textit{Residenti che lavorano o studiano nel comune}}{\begin{array}{c}\textit{(Flussi in uscita dal comune per motivi di studio o lavoro +}\\ \textit{+ Residenti che lavorano o studiano nel comune)}\end{array}} \cdot 100 .$$

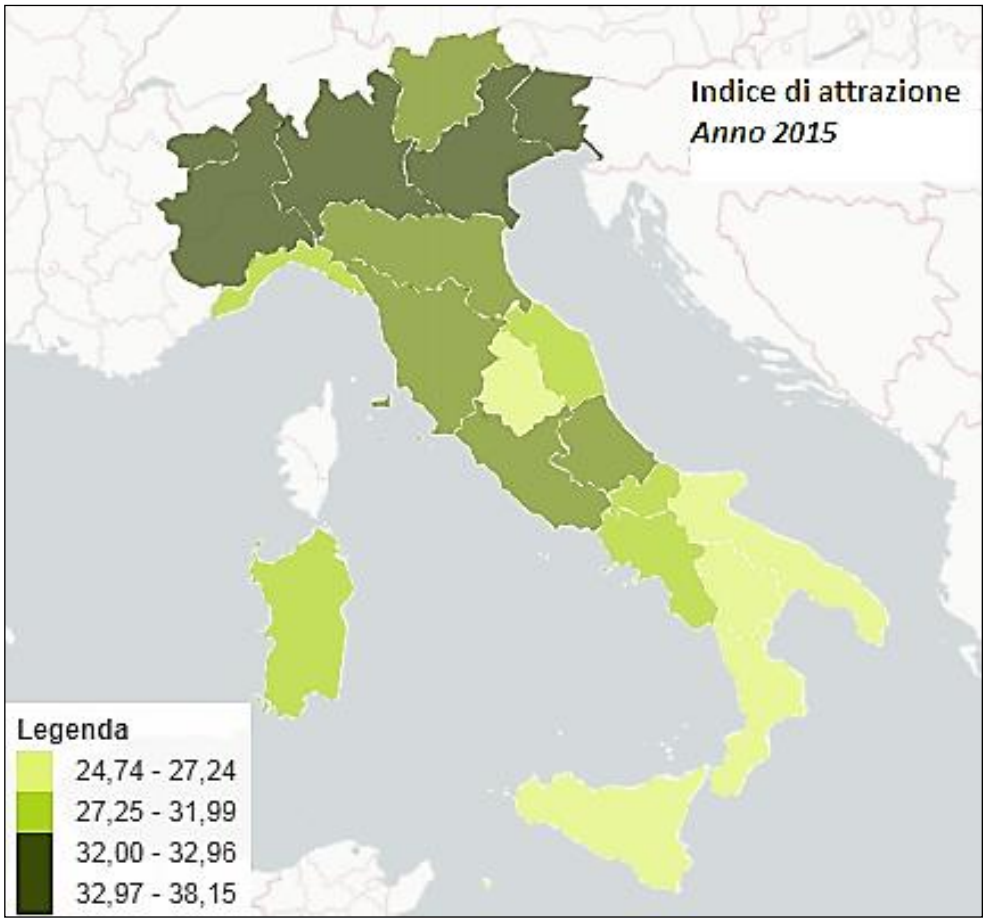
L'indice di autocontenimento è definito sulle stesse fonti di dati dell'indice di attrazione e fornisce un quadro della "mobilità" desunto da una geografia di fatto delineata dai segnali amministrativi di presenza nei luoghi di lavoro e studio dei residenti e come tali differiscono, seppure in misura non sostanziale, dai dati censuari descrittivi invece dei comportamenti rilevati tramite intervista[8].

[7] Per approfondimenti vedi https://www.istat.it/it/archivio/189907.

[8] Per approfondimenti vedi https://www.istat.it/it/archivio/189907.

In Fig. 10 e Fig. 11 si presentano ambo gli indicatori per l'intero Paese, a livello regionale.

**Figura 10.** *Indice di attrazione - Italia. Anno 2015*

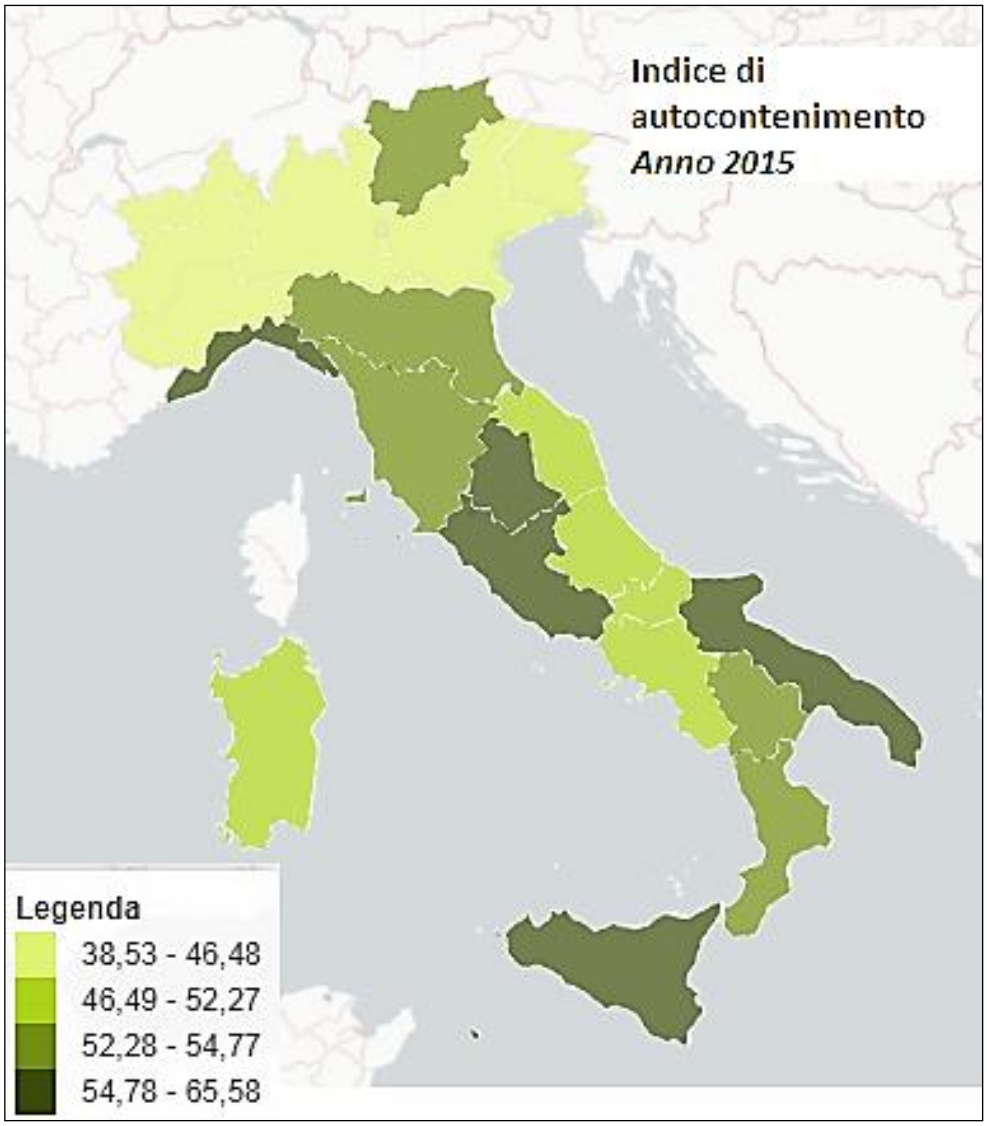



Fonte: Sistema informativo *"a misura di comune"*

Se con l'*indice di attrazione* classifichiamo la regione Basilicata tra i territori a più basso potere attrattivo, lo stesso non può dirsi per l'indice di *autocontenimento*.

Dal punto di vista del pendolarismo, molti individui si spostano dalle regioni circostanti, in Basilicata, principalmente per ragioni di lavoro. Anche in questo caso però, l'analisi su base regionale non dà una giusta fotografia delle caratteristiche specifiche della Basilicata.

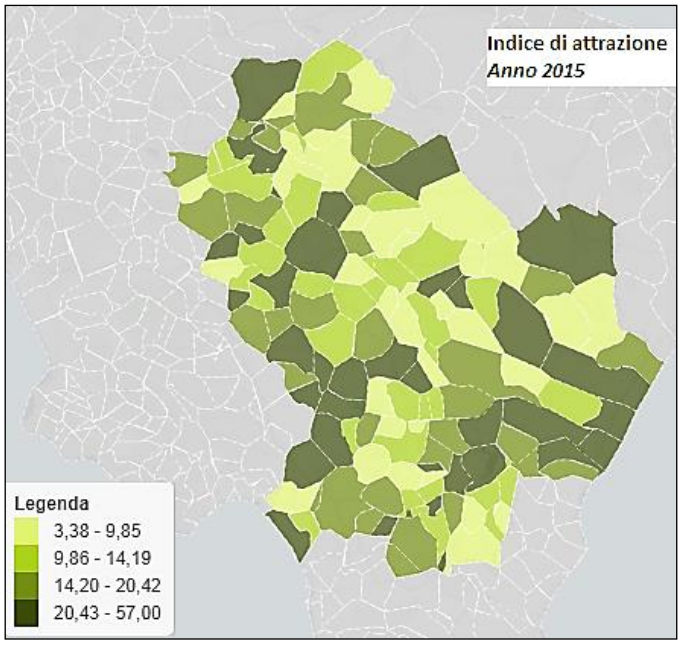
**Figura 11.** *Indice di autocontenimento - Italia. Anno 2015*




Fonte: Sistema informativo *"a misura di comune"*

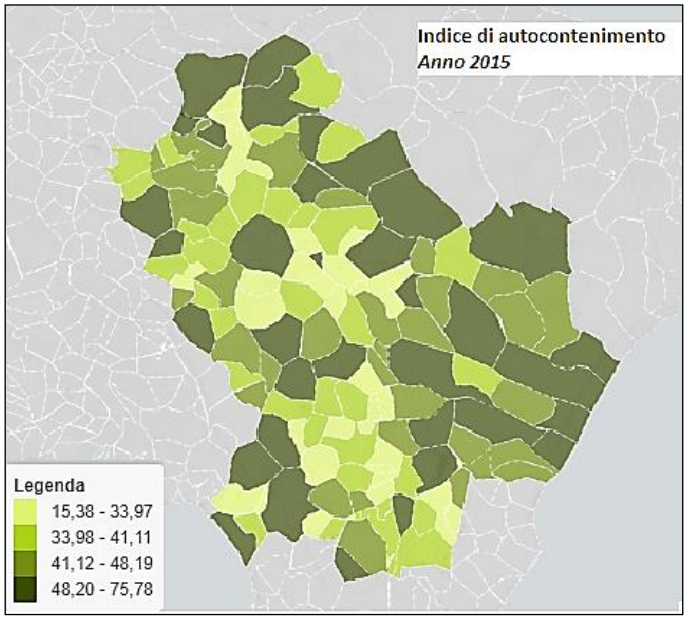
Più interessante, specialmente per la pianificazione delle politiche sul territorio e per l'analisi anche economica della regione, è approfondire i due fenomeni su base comunale. In Fig. 12 e in Fig. 13 si riportano i due indicatori in forma cartografica.

**Figura 12** *Indice di attrazione - Basilicata. Anno 2015*




Fonte: Sistema informativo *"a misura di comune"*

**Figura 13.** *Indice di autocontenimento - Basilicata. Anno 2015*




Fonte: Sistema informativo *"a misura di comune"*

Al fine di un'analisi puntuale dei comuni a più alto potere attrattivo e più interessati dal pendolarismo si riporta in Tab. 2 la graduatoria dei 20 comuni della Basilicata per i due indicatori.

**Tabella 2.** *Graduatoria dei comuni della Basilicata con indice di attrazione e autocontenimento più alti. Anno 2015*

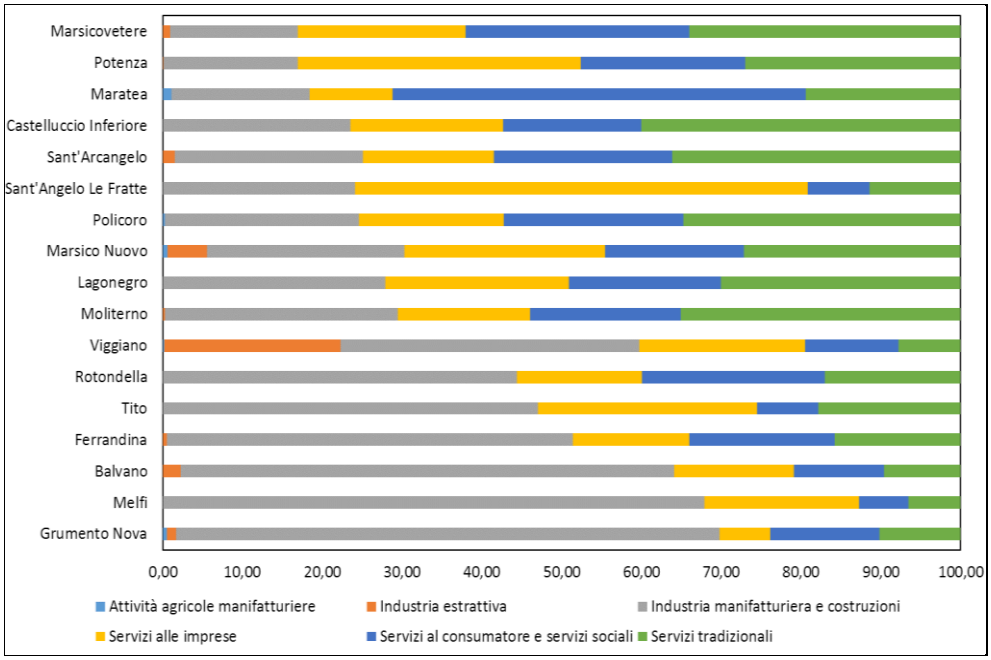
| Indice di attrazione | | Indice di autocontenimento | |
|---|---|---|---|
| Comune | Valore percentuale | Comune | Valore percentuale |
| Melfi | 57,00 | Matera | 75,78 |
| Guardia Perticara | 55,56 | Melfi | 74,96 |
| Viggiano | 53,07 | Potenza | 74,28 |
| Balvano | 46,68 | Sant'Arcangelo | 60,21 |
| Tito | 45,54 | Venosa | 59,83 |
| Potenza | 42,93 | Policoro | 59,73 |
| Grumento Nova | 40,42 | Pisticci | 58,8 |
| Lagonegro | 38,52 | Bernalda | 58,14 |
| Sant'Angelo Le Fratte | 34,64 | Ferrandina | 57,64 |
| Garaguso | 34,38 | Lavello | 57,63 |
| Moliterno | 31,80 | Genzano di Lucania | 57,27 |
| Castelluccio Inferiore | 29,96 | Stigliano | 56,1 |
| Policoro | 29,66 | Maratea | 55,42 |
| Rotondella | 28,62 | Moliterno | 54,81 |
| Ferrandina | 28,40 | Senise | 54,79 |
| Marsicovetere | 28,06 | Palazzo San Gervasio | 54,78 |
| Sant'Arcangelo | 27,58 | Francavilla in Sinni | 54,59 |
| Matera | 26,98 | Balvano | 54,08 |
| Maratea | 26,50 | Lauria | 54,07 |
| Marsico Nuovo | 24,99 | Viggiano | 53,64 |

Fonte: Sistema informativo *"a misura di comune"*

Analizzando le attività economiche prevalenti dei comuni a più grande potere attrattivo e maggiormente interessati dal pendolarismo (Figg. 14 e 15) si delineano le specificità economiche del territorio.

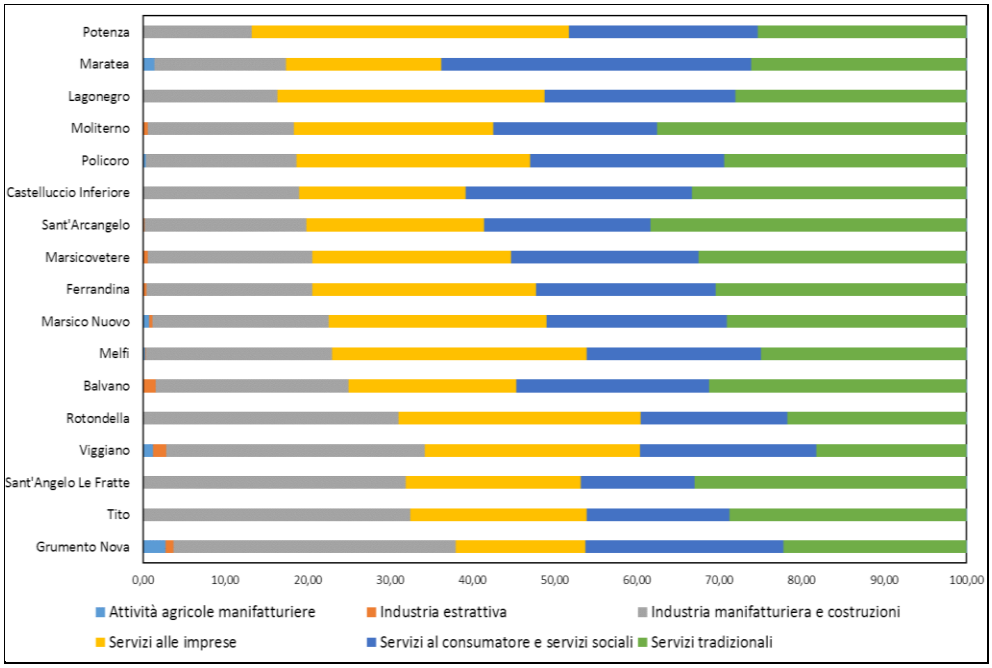
L'analisi non è sicuramente esaustiva, bisognerebbe da un lato considerare quanta parte dell'indicatore è spiegato dagli spostamenti per motivi di studio e dall'altra bisognerebbe depurare l'elenco riportato in Tab. 2 dai comuni la cui dimensione demografica è molto bassa. Inoltre, in territori come la Basilicata in cui l'incidenza di Valore Aggiunto derivante da attività legate alla Pubblica Amministrazione è significativamente alta, sarebbe opportuno analizzare anche la presenza degli uffici pubblici su base comunale; elemento che risulta avere un chiaro potere attrattivo sul territorio.

**Figura 14.** *Distribuzione degli addetti per attività economica. Anno 2015*




Fonte: Elaborazione su dati di fonte sistema informativo *"a misura di comune"*

**Figura 15.** *Distribuzione delle unità locali per attività economica. Anno 2015*




Fonte: Elaborazione su dati di fonte sistema informativo *"a misura di comune"*

Dall'analisi contestuale dei due indicatori è possibile individuare le ragioni "economiche" che portano i diversi comuni ad attrarre o trattenere la propria popolazione e i pendolari dai comuni vicini.

In Basilicata, si potrebbero individuare quindi almeno tre raggruppamenti di comuni caratterizzati da specifiche attività economiche e che attraggono, di conseguenza, popolazione:

- comuni a vocazione industriale/manifatturiera
- comuni a vocazione estrattiva
- comuni a vocazione terziaria.

Del primo raggruppamento fanno parte i comuni del corridoio Potentino-Melfese, del secondo i comuni della Val d'Agri e a seguire, i comuni del Lagonegrese insieme ai capoluoghi di provincia.

La domanda che ora ci si pone è: essere un polo attrattivo, porta nel comune anche un maggiore benessere economico?

Un indicatore di benessere economico (Fig. 16), calcolato su base comunale, su dati Archimede, è il reddito lordo pro capite (Red_pc) definito come:

$$Red_pc = \frac{Reddito\ complessivo\ lordo\ delle\ famiglie\ anagrafiche}{Numero\ componenti\ delle\ famiglie\ anagrafiche} \ .$$

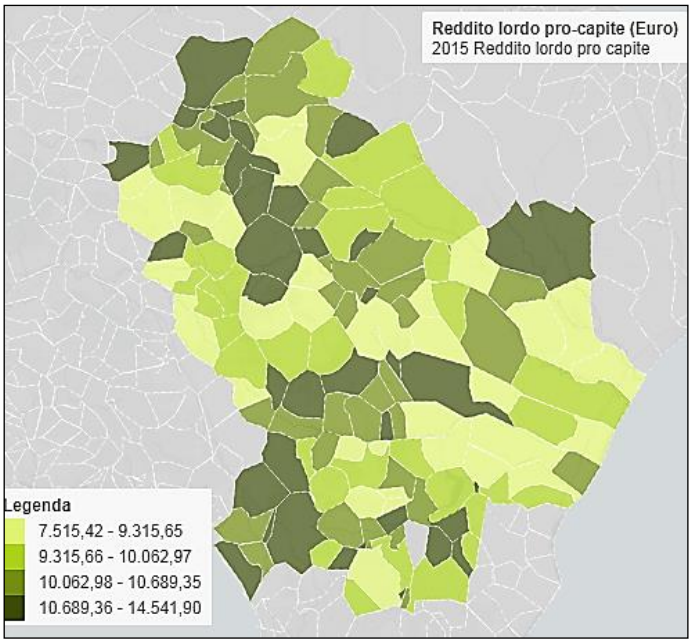
Il reddito complessivo al lordo della tassazione è inteso come somma di diverse componenti di natura economica percepite dai membri della famiglia anagrafica nel corso dell'anno.

Il reddito di una famiglia anagrafica è ottenuto sommando le voci di reddito, percepite da tutti i componenti della famiglia, desunte dalle fonti amministrative fiscali e previdenziali al lordo della tassazione.

L'informazione reddituale è infatti il risultato dell'integrazione di variabili presenti nei seguenti archivi: Banca Dati Reddituale del Ministero dell'Economia e Finanze, il Casellario dei Pensionati dell'Inps, l'archivio dei lavoratori domestici dell'Inps, l'Unico Persone Fisiche dell'Agenzia delle Entrate, l'archivio dei Trattamenti Monetari Non Pensionistici dell'Inps, l'archivio Uniemens dell'Inps e la Banca Dati dei Voucher dell'Inps.

Le principali voci incluse nel reddito lordo sono i redditi da lavoro dipendente e autonomo, da pensione, da capitale e da trasferimento pubblico e privato. Si segnala, inoltre, che per i comuni con meno di 5.000 abitanti l'indicatore potrebbe non essere sufficientemente robusto.

**Figura 16.** *Reddito lordo pro capite - Basilicata. Anno 2015*




Fonte: Sistema informativo *"a misura di comune"*

Nella graduatoria dei redditi lordi pro capite (Tab. 3) viene confermata la presenza di comuni quali Potenza, Matera, Melfi, Viggiano e Balvano, poli di attrazione lucani sebbene è sempre molto complicato fare delle analisi "economiche" sui soli redditi pro capite che, in comuni spesso molto piccoli, risentono di una forte variabilità. Al fine di tener conto anche di questo aspetto, nella sezione *"benessere economico"*, si riporta un indicatore di disuguaglianza a livello comunale "divari nel reddito al lordo delle imposte sul reddito" (I_div), calcolato come:

$$I_div = \frac{\textit{Reddito equivalente totale al lordo delle imposte posseduto dal 20\% degli iscritti in anagrafe con più alto reddito}}{\textit{Reddito equivalente totale al lordo delle imposte posseduto dal 20\% degli iscritti in anagrafe con più basso reddito}} .$$

Il reddito lordo equivalente è calcolato rapportando il reddito familiare lordo a un fattore di scala che tiene conto del numero e dell'età dei componenti della famiglia. La scala di equivalenza OCSE modificata per il calcolo è ottenuta assegnando un peso pari a 1 al primo componente adulto della famiglia anagrafica, 0,5 ad ogni altro adulto (di età maggiore o uguale a 14 anni) e 0,3 ad ogni componente di età inferiore a 14 anni.

**Tabella 3.** *Graduatoria dei comuni della Basilicata con reddito lordo pro capite più alto. Anno 2015*

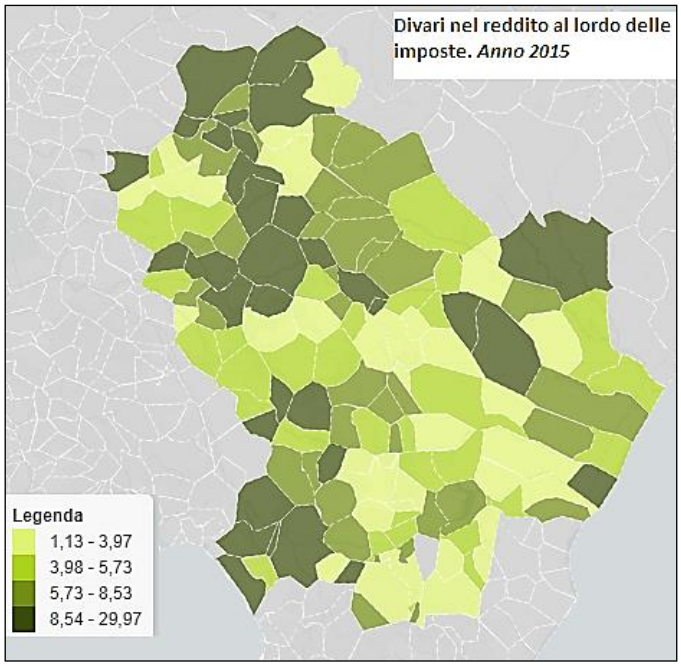
| Reddito lordo pro capite | |
|---|---|
| Comune | Valore percentuale |
| Potenza | 14.541,90 |
| Matera | 13.633,87 |
| Lagonegro | 13.477,68 |
| Pescopagano | 13.052,45 |
| San Paolo Albanese | 12.283,54 |
| Lauria | 12.011,08 |
| Rionero in Vulture | 11.870,90 |
| Pietragalla | 11.856,16 |
| Melfi | 11.745,71 |
| Corleto Perticara | 11.723,99 |
| Vaglio Basilicata | 11.651,53 |
| Castelluccio Inferiore | 11.496,77 |
| Maratea | 11.436,06 |
| Viggiano | 11.406,75 |
| Balvano | 11.341,10 |
| Marsicovetere | 11.331,10 |
| Avigliano | 11.283,82 |
| Pignola | 11.216,01 |
| Filiano | 11.180,66 |
| Ripacandida | 11.028,98 |

Fonte: Sistema informativo *"a misura di comune"*

Dalla Figura 17 risulta esservi una sostanziale sovrapposizione tra reddito pro capite e disuguaglianza e ciò a discapito di un concetto generale di benessere che dovrebbe portare ad un miglioramento delle condizioni generali delle famiglie piuttosto che ad un aumento della disuguaglianza tra famiglie.

D'altra parte parliamo di una regione con forti specificità ma con un tessuto economico e produttivo particolarmente debole. Ciononostante, a maggior ragione, risulta indispensabile un'analisi a livello comunale al fine di individuare le caratteristiche del territorio e di pianificare le azioni politiche ed amministrative sullo stesso.

**Figura 17.** *Divari nel reddito al lordo delle imposte - Basilicata. Anno 2015*




Fonte: Sistema informativo *"a misura di comune"*

## 6. Alcune osservazioni conclusive

Il presente lavoro permette di dedurre conclusioni di due ordini di livello: generale, sul nuovo sistema informativo e specifiche, relative al caso analizzato, con le seconde strettamente connesse alla struttura delle prime. Per quanto concerne le prime occorre osservare come il percorso di integrazione delle fonti statistiche amministrative con quelle "classiche" campionarie e censuarie stia permettendo un dettaglio informativo statistico prima impensabile, specialmente se inserito in progetti informativi atti a garantire continuità nel tempo, con le connesse conseguenze di poter avere informazioni nel tempo via via più robuste nonché confrontabili in ambito spazio-temporale.

Si aggiunga poi come l'esplicita finalità del progetto di valorizzare l'utilizzazione da parte dei vari policy maker, sia confluita in una oggettiva capacità di fornire dati di sintesi immediatamente fruibili.

Per quanto invece concerne le osservazioni specifiche sul caso analizzato, nel presente lavoro, è emerso che la potenzialità di tale strumento informativo possa essere foriera di condurre indagini che permettano di approfondire l'esistenza di un

*trade off* locale tra sfruttamento delle risorse e benessere socioeconomico o la possibilità di confutare o meno un utilizzo delle risorse compatibile con la salvaguardia ambientale.

D'altro canto, è emerso anche come i territori maggiormente interessati dallo sfruttamento delle risorse naturali risultino caratterizzati dalla presenza diffusa di piccoli e micro comuni, molti dei quali al di sotto di 5.000 abitanti.

Occorre tener presente come l'esigua dimensione dei comuni interessati potrebbe rendere non sufficientemente robusta l'informazione, specialmente a causa dell'influenza che potrebbero avere i dati anomali. L'elevato interesse verso la disaggregazione dei dati a livello comunale, unito alle parziali criticità emerse per la loro utilizzazione ai fini del presente lavoro, fanno ritenere quindi opportuno approfondire ulteriormente l'analisi allorquando, nell'immediato futuro, saranno disponibili i file dei microdati o ulteriori indicatori di analisi in tema di territorio e ambiente oppure, in alternativa allorquando il maggiore intervallo temporale di osservazione possa permettere di meglio validare le strutture dei fenomeni favorendo analisi in serie storica e su domini sempre più fini.

## Riferimenti bibliografici

Costanzo, L.; Ferrara, A., (2015). Indicatori di erosione dello spazio rurale basati su dati di censimento. *XXXVI Conferenza italiana di scienze regionali*

Bianchino, A.; Cariello, S; Fullone, F.; Mazziotta M. (2016). Indici compositi del benessere per i comuni della Basilicata. *Rivista Italiana di Economia Demografia e Statistica 53ma Riunione Scientifica.* Roma UNINT.

Carucci, A.M.M.; Vannella, G., (2016). Sull'integrazione tra fonti amministrative e statistiche per le imprese. *Metodi e analisi statistiche*, Università degli studi di Bari Aldo Moro

Carucci A.M.M.; Fullone F.; Vannella G. (2017). Dai conti economici ai conti satellite ambientali: Basilicata, un caso di studio. *Metodi e analisi statistiche*, Università degli studi di Bari Aldo Moro

Carucci, A.M.M.; Cavalli, L., (2017). Flussi interregionali nel reddito disponibile delle famiglie. *XXXVIII Conferenza italiana di scienze regionali*

Garofalo G. (2014). *Il progetto ARCHIMEDE obiettivi e risultati sperimentali*, Istat working paper, N.9 2014. https://www.istat.it/it/files/2014/11/IWP-n.-9-2014.pdf

Istat, (2011). *9° Censimento dell'industria e dei servizi e Censimento delle istituzioni non profit. I distretti industriali*. https://www.istat.it/it/archivio/150320

Istat, (2015). *Rapporto BES 2015: il benessere equo e sostenibile in Italia*. https://www.istat.it/it/files/2015/12/Rapporto_BES_2015.pdf

Istat, (2015). *Urbes - Il benessere equo e sostenibile nelle città*. https://www.istat.it/it/files//2015/04/UrBes_2015.pdf

## Sitografia

http://amisuradicomune.istat.it/aMisuraDiComune/

https://www.istat.it/it/archivio/189907

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018*
DIPARTIMENTO DI ECONOMIA E FINANZA
ISBN: 978-88-6629-048-3
pp. 109-115

# Le tecniche di autocorrelazione spaziale per l'analisi della povertà in Puglia

**Monica Carbonara***

*ISTAT, Ufficio territoriale per la Puglia*

**Riassunto:** Nel decennio corrente il rischio di povertà ed esclusione sociale è diventato centrale nel contesto delle politiche pubbliche sia a livello nazionale che europeo. Per analizzare la povertà in Puglia è stato calcolato un indicatore di povertà che misura a livello comunale la percentuale di famiglie anagrafiche con "reddito equivalente" inferiore all'importo dell'assegno sociale. I dati provengono dai microdati integrati dell'Istat "Condizioni socio-economiche delle famiglie" - Anno 2015. La dimensione territoriale del fenomeno è stata esaminata utilizzando tecniche di autocorrelazione spaziale, in particolare l'indice di Moran e il Local Indicator of Spatial Association (LISA) al fine di individuare eventuali cluster territoriali.

**Keywords:** Povertà; Reddito equivalente; Autocorrelazione spaziale; Local Indicator of Spatial Association.

## 1. Introduzione

Uno degli obiettivi specifici dell'Unione europea e degli Stati membri nell'ambito della politica sociale è la lotta contro la povertà e l'esclusione sociale. Tra le principali innovazioni introdotte nel 2010 dalla strategia Europa 2020 per una crescita intelligente, sostenibile e inclusiva, vi è l'obiettivo di ridurre del 25% il numero di europei che vivono al di sotto della soglia nazionale di povertà, facendo uscire dalla povertà più di 20 milioni di persone entro il 2020.

Nel 2015 in Italia la crescita moderata del reddito disponibile e del potere d'acquisto non ha modificato la disuguaglianza reddituale e non ha comportato una diminuzione dei livelli di povertà. La povertà assoluta, intesa come incapacità di ac-

* Autore corrispondente: mocarbon@istat.it .

quisire beni e servizi essenziali per evitare gravi forme di esclusione sociale, ha coinvolto 4 milioni e mezzo di persone a causa delle cattive condizioni delle famiglie più numerose, in particolare coppie con due figli e famiglie di stranieri. Queste condizioni di difficoltà risultano particolarmente diffuse nelle regioni del Mezzogiorno.

In Italia l'Indagine Istat sulle condizioni di vita delle famiglie (Eu-Silc)[1] è una delle principali fonti per la conoscenza delle condizioni economiche degli individui e delle famiglie. Dalla rilevazione derivano, infatti, gli indicatori ufficiali sui redditi e sul rischio di povertà, anche in funzione delle principali caratteristiche delle famiglie (numero di componenti, numero di percettori di reddito, tipologia familiare, presenza di minori o di anziani). Questi indicatori poiché si ricavano da indagini campionarie, consentono stime solo per domini territoriali fino al livello regionale. Con il progetto ARCH.I.M.E.DE (ARCHivio Integrato di Microdati Economici e Demografici) si è ampliata l'offerta formativa dell'Istat mediante la produzione di collezioni di dati elementari utili alla ricerca sociale ed economica, alla programmazione territoriale e settoriale, alla valutazione delle politiche pubbliche a livello nazionale, regionale e locale. Una peculiarità di tale progetto è la disaggregazione delle informazioni ad un livello territoriale comunale o provinciale, a differenza delle fonti tradizionali che forniscono stime per domini territoriali meno fini. La disponibilità di informazioni a livello comunale, sebbene non corrispondenti alle definizioni e classificazioni ufficiali, può rappresentare un valido strumento di supporto alla programmazione locale. All'interno del progetto ARCH.I.M.E.DE la base dati statistica "Condizioni socio-economiche delle famiglie" consente di analizzare congiuntamente vari aspetti delle condizioni socio-economiche delle famiglie residenti e di individuare eventuali fragilità.

Il presente lavoro si prefigge di mettere in luce, attraverso tecniche di autocorrelazione spaziale, l'eventuale presenza di associazione spaziale dei livelli di povertà in Puglia e di individuare possibili profili del disagio.

## 2. Materiali e metodi

I dati provengono dalla base dati "Condizioni socio-economiche delle famiglie - ArchIMEDe", Anno 2015[2]. Il principale indicatore in esame è dato dalle famiglie anagrafiche con reddito equivalente inferiore all'importo dell'assegno sociale in

[1] Precisamente, *Indagine sul reddito e le condizioni di vita (EU-SILC)*, il cui ultimo aggiornamento è riferito al biennio 2017-18 (https://www.istat.it/it/archivio/5663).
[2] I dati sono scaricabili dal Sistema informativo "A misura di Comune".

Puglia, indicatore proxy della povertà assoluta[3] calcolata dall'Indagine sulle spese delle famiglie condotta dall'Istat. Per confrontare le condizioni economiche delle famiglie con diversa numerosità e composizione, il reddito familiare è diviso per opportuni parametri (scala di equivalenza OCSE[4]), in modo da ottenere il cosiddetto reddito equivalente.

Per analizzare la dimensione territoriale del fenomeno sono state utilizzate tecniche di autocorrelazione spaziale, in particolare sono stati calcolati l'indice di Moran e il Local Indicator of Spatial Association (LISA). La caratteristica più interessante dell'autocorrelazione spaziale è la possibilità di analizzare allo stesso tempo la componente spaziale e di attributo dell'informazione (Goodchild, 1986). Di conseguenza, l'autocorrelazione spaziale è una tecnica molto efficace per esaminare la distribuzione spaziale di oggetti e valutare allo stesso tempo il grado di influenza e di relazione con gli elementi vicini. Questo concetto è riassunto nella cosiddetta "prima legge della geografia" (Waldo Tobler, 1970), in cui si afferma che "Tutti gli eventi sono legati tra loro, ma eventi vicini sono più collegati di quelli lontani".

Seguendo l'approccio di Goodchild (1986), Lee e Wong (2001) definiscono l'autocorrelazione spaziale come segue:

$$SAC = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} c_{ij} w_{ij}}{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}}$$

dove:

N è il numero degli oggetti;
i e j sono due oggetti diversi;
$c_{ij}$ è un grado di similarità degli attributi i e j;
$w_{ij}$ è un grado di similarità del luogo i e j.

Definendo $x_i$ come il valore dell'attributo dell'oggetto i, se $c_{ij}= (x_i-x)(x_j-x)$, l'indice $I$ di Moran (1948) è:

$$I = \frac{N \sum_i \sum_j w_{ij} \sum_i (x_i - \overline{x})}{\sum_i \sum_j w_{ij} \sum_i (x_i - \overline{x})^2}$$

L'indice di Moran è un indicatore globale di autocorrelazione spaziale e fornisce un'indicazione relativa della presenza di autocorrelazione. La localizzazione

[3] La povertà assoluta è la percentuale di persone appartenenti a famiglie con una spesa complessiva per consumi inferiore al valore soglia di povertà assoluta sul totale delle per-sone residenti.
[4] La scala di equivalenza OCSE modificata utilizzata per calcolare il reddito familiare equi-valente è pari alla somma di più coefficienti individuali (1 per il primo adulto, 0,5 per ogni altro adulto e 0,3 per ogni componente di età inferiore a 14 anni).

precisa di valori elevati di autocorrelazione è, invece, descritta dagli indici LISA o indicatori locali di autocorrelazione spaziale.

Uno degli indici LISA più utilizzati è quello proposto da Anselin (1988, 1995), calcolato secondo la seguente formula:

$$I_i = \frac{(x_i - \overline{x})}{S_x^2} \sum_{j=1}^{N} w_{ij}(x_j - \overline{x}) \, .$$

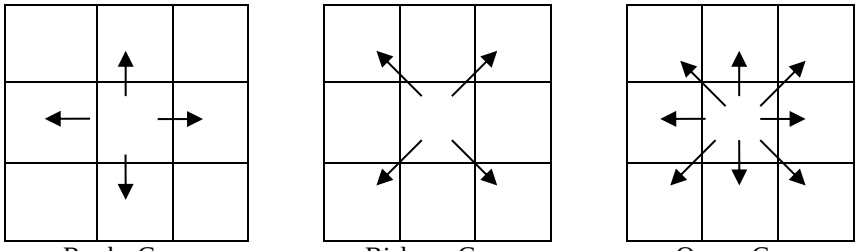
Esso è considerato come un indicatore locale di Moran. La somma di tutti gli indici locali, infatti, è proporzionale al valore di Moran:

$$\sum_i I_i = \gamma \cdot I \, .$$

L'indicatore LISA consente, per ogni luogo, di valutare la similarità tra ogni osservazione e gli elementi che la circondano. Si possono verificare cinque casi, in cui i diversi luoghi sono caratterizzati da:

- alti valori del fenomeno e alti livelli di similarità con il vicinato (Alto-Alto, colore rosso nella cartografia);
- bassi valori del fenomeno e bassi livelli di similarità con il vicinato (Basso-Basso, (colore blu nella cartografia);
- alti valori del fenomeno e bassi livelli di similarità con il vicinato (Alto-Basso), definiti come potenziali outliers (colore rosso chiaro nella cartografia);
- bassi valori del fenomeno e alti livelli di similarità con il vicinato (basso-alto), definiti come potenziali outliers (colore blu chiaro nella cartografia);
- completa assenza di autocorrelazione significativa (colore bianco nella cartografia).

**Figura 1**. *Matrice dei pesi spaziali e la metafora degli scacchi*



Rooks Case Bishops Case Queen Case

Gli indici LISA forniscono una misura efficace del grado di associazione spaziale relativa tra ogni unità territoriale e gli elementi che la circondano, consenten-

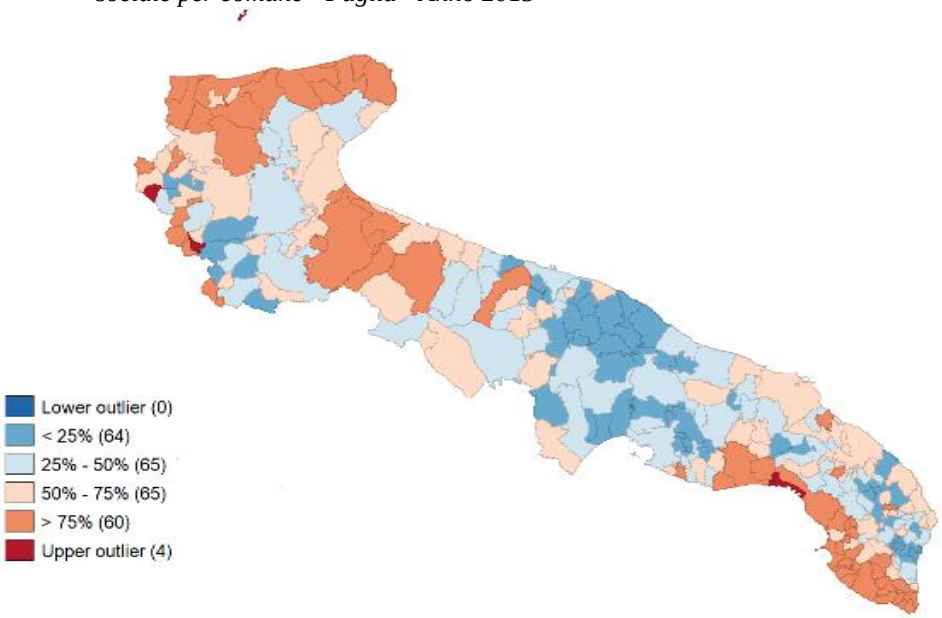
do di evidenziare raggruppamenti (o cluster) spaziali. Il sistema di contiguità è definito dalla matrice di contiguità, che presenta elementi $w_{ij}$ che assumono valore 1 se due generiche unità $i$ e $j$ sono confinanti e 0 se non lo sono. La contiguità può essere a bordo comune, a vertice comune e mista (vertice e bordo comune). Adottando la metafora del gioco degli scacchi (O' Sullivan e Unwin, 2002), la contiguità può essere considerata come il percorso consentito alla torre, all'alfiere e alla regina (Figura 1). In questo studio è stato adottato lo schema di contiguità mista (Queen case).

## 3. Risultati

Nel 2015 in Puglia la percentuale di famiglie anagrafiche con reddito equivalente inferiore all'importo dell'assegno sociale è del 18,3%, superiore a quella italiana (14,6%) e varia da un minimo di 17,0% nella provincia di Taranto a un massimo di 20,8% nella provincia di Barletta-Andria-Trani.

La distribuzione del fenomeno in Puglia (Box-map della Figura 2) mostra percentuali più alte nei comuni a nord del Gargano e dell'Alto Tavoliere delle Puglie, nei comuni a confine della provincia di Foggia e di Bari-Andria-Trani e in quelli della fascia sud ionica tarantina e salentina.

**Figura 2.** *Famiglie anagrafiche con reddito equivalente inferiore all'importo dell'assegno sociale per comune - Puglia - Anno 2015*

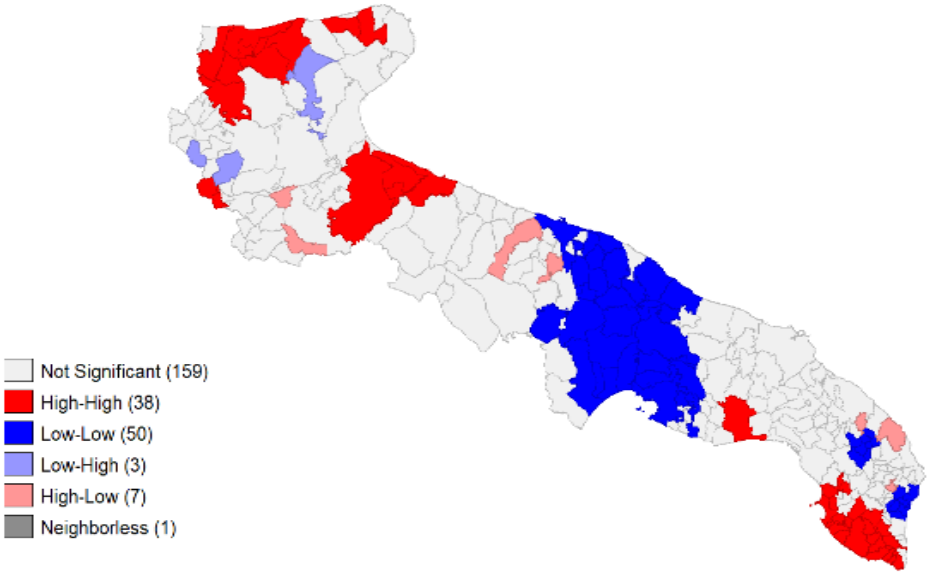


Valori assai meno intensi dell'indicatore si riscontrano, invece, nella parte sud orientale del subappennino dauno, nel sud barese, nei comuni a nord della provincia di Taranto e in isolati comuni nella parte orientale del basso Salento.

Il fenomeno in Puglia è caratterizzato da un'autocorrelazione spaziale positiva, messa in evidenza dal valore dell'indice di Moran (0,49). I cluster individuati dall'indice LISA consentono di identificare cinque aree con un'alta incidenza del fenomeno: la prima area comprende i comuni dell'Alto Tavoliere delle Puglie, la seconda i comuni foggiani di Ischitella e Vico del Gargano, la terza i comuni di Cerignola, Trinitapoli, San Ferdinando di Puglia, Margherita di Savoia e Barletta. A questi cluster si aggiungono Manduria che costituisce un cluster a se e i comuni ionici del tacco salentino.

Decisamente più contenuta l'incidenza nei comuni del sud barese e nella provincie di Taranto ad eccezione dei comuni di Ginosa e Laterza.

**Figura 3.** *Autocorrelazione spaziale - Famiglie anagrafiche con reddito equivalente inferiore all'importo dell'assegno sociale - Puglia - Anno 2015*




## 4. Conclusioni

L'analisi dei dati ha messo in luce la presenza di alcuni territori critici nei quali è più alta la percentuale di famiglie anagrafiche con reddito equivalente inferiore all'importo dell'assegno sociale. Fermo restando i limiti dettati dalla qualità dei da-

ti amministrativi e dalla non confrontabilità con le rilevazioni tradizionali dell'Istat, la disponibilità di informazioni ad un basso dettaglio territoriale può costituire un prezioso strumento di supporto alla programmazione delle politiche locali.

## Riferimenti bibliografici

COM/2010/2020 definitivo, *EUROPA 2020, Una strategia per una crescita intelligente, sostenibile e inclusiva.*

Istat (2017). *Rapporto BES 2017: Il benessere equo e sostenibile in Italia*.

Istat (2016). *Condizioni di vita, reddito e carico fiscale delle famiglie*. Report.

Istat (2018). *Indagine sul reddito e le condizioni di vita (EU-SILC)*. 2017-2018. https://www.istat.it/it/archivio/5663

Pratesi M. (2016). *Analysis of Poverty Data by Small Area Estimation*, New York: Wiley.

Quintano, C.; Castellano, R.; Punzo, G. (2007). Estimating Poverty in the Italian Provinces using Small Area Estimation Models, *Metodološki zvezki*, 4(1): 37-70.

Giuseppe Garofalo (2014). Il Progetto ARCHIMEDE obiettivi e risultati sperimentali, *Istat working papers*, n.9.

Anselin, L. (1995). Local Indicators of Spatial Association-LISA, *Geographical Analysis*, 27: 93-115.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*, Kluwer Academic, Boston - MA.

Geary, R. (1954), The Contiguity Ratio and Statistical Mapping, *The Incorporated Statistician*, 5, no. 3: 115-146.

Goodchild, M. F. (1986), *Spatial Autocorrelation*, Catmog 47, Geo Books, Norwich.

Lee, J.; Wong, D. W. S. (2001), *Statistical analysis with ArcView GIS*, John Wiley and Sons, New York.

Moran, P. (1948), The interpretation of statistical map, *Journal of the Royal Statistical Society*, Series B, 10: 243-251.

O'Sullivan, D.; Unwin, D. J. (2002), *Geographic Information Analysis*, John Wiley and Sons, New York.

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018* ISBN: 978-88-6629-048-3
DIPARTIMENTO DI ECONOMIA E FINANZA pp. 117-127

# Studio pilota su una coorte di pazienti con Lupus eritematoso sistemico

**Angela Maria D'Uggento*, Mara Turturo**

*Università degli Studi di Bari Aldo Moro*

**Riassunto:** Il lupus eritematoso sistemico è una malattia cronica autoimmune di natura infiammatoria caratterizzata da manifestazioni cutanee e fotosensibilità che può colpire gli organi interni, le articolazioni o il sistema nervoso. Allo stato attuale può essere assimilato ad uno stato patologico in quanto non prevede una guarigione ma alterna fasi acute a fasi di remissione. Il LES viene trattato con farmaci immunosoppressori e corticosteroidi e i pazienti vengono monitorati ad intervalli di tempo regolari per seguire l'andamento delle variabili cliniche e l'efficacia della terapia. Una terapia innovativa per la cura del lupus si basa su un anticorpo monoclonale, il *belimumab*, appartenente alla categoria farmaco-terapeutica degli immunosoppressori selettivi che riduce la severità della patologia e consente una miglior qualità della vita dei pazienti. Il presente lavoro illustra i primissimi risultati di un'analisi statistica condotta su una coorte di pazienti per definire le principali caratteristiche della patologia e valutare l'efficacia della terapia innovativa a cui sono stati sottoposti. Benché si tratti di una ricerca appena avviata, con un numero di casi ancora esiguo, i risultati ottenuti forniranno utili indicazioni cliniche e andranno ad integrare quelli di uno studio multicentrico a carattere nazionale.

**Keywords:** studio retrospettivo; lupus eritematoso sistemico; analisi statistica.

## 1. Introduzione

Il lupus eritematoso sistemico (d'ora in poi LES) è una malattia cronica autoim–mune, imputabile ad una autonoma reazione del sistema immunitario che produce

* Autore corrispondente: angelamaria.duggento@uniba.it

Il lavoro qui descritto è frutto di un progetto comune, tuttavia Angela Maria D'Uggento ha provveduto alla redazione dei paragrafi 1, 3 e 4 mentre Mara Turturo ha collaborato all'analisi dei dati ed ha redatto i paragrafi 2 e 3.1.

l'infiammazione dei tessuti e può colpire indifferentemente qualsiasi organo e apparato, con sintomi e situazioni cliniche variabili tra i pazienti. La malattia è variabile anche dal punto di vista delle fasi acute e di remissione, con riflessi sulla qualità della vita di chi ne viene colpito. Dagli studi clinici (Collins et al, 2016) emerge una maggior incidenza tra le donne, in rapporto di 9 a 1 rispetto agli uomini, con manifestazioni tra i 20 e i 45 anni indipendentemente dal genere, sebbene la malattia possa insorgere anche nell'infanzia o in età avanzata.

Il LES risulta caratterizzato da una predisposizione genetica, anche se non è classificata tra le malattie ereditarie: di conseguenza non è trasmissibile e, infatti, raramente colpisce più componenti di una stessa famiglia. Come si evince, si tratta di una patologia complessa, con eziopatogenesi sicuramente multifattoriale; probabili fattori di attivazione, oltre a quelli genetici, sono quelli ambientali, quali esposizione al sole, infezioni, stress acuti nonché fattori ormonali. Nei soggetti affetti da LES, il sistema immunitario, che normalmente produce anticorpi allo scopo di contrastare gli agenti infettivi, si attiva producendo una quantità di anticorpi in eccesso, detti autoanticorpi, che attaccano le cellule e i tessuti sani dell'organismo. Gli autoanticorpi sono responsabili dell'infiammazione e dei danni a carico di diversi organi.

Come anticipato, il LES è una malattia multiforme sia rispetto all'insorgenza che ai sintomi, che sono strettamente connessi all'organo bersaglio. Inizialmente, i sintomi più frequenti sono costituiti da disturbi aspecifici, quali stanchezza e malessere generale, tipici di un quadro clinico comune ad altre malattie, spesso associati a febbre, calo ponderale e astenia. Le manifestazioni più comuni sono quelle cutanee, prevalentemente causate da esposizione al sole e caratterizzate da eritemi, in particolare a livello del naso e delle guance (c.d. eritema "a farfalla") ma che possono interessare diverse parti del corpo, la fotosensibilità, le ulcerazioni delle mucose del naso e della bocca. Vi sono, inoltre, i dolori articolari, di differente intensità e associati a gonfiore delle articolazioni interessate ed il coinvolgimento degli organi interni, nei casi di malattia più severi, colpendo i polmoni, il cuore, i reni e il sistema nevoso.

Studi clinici hanno evidenziato che la malattia si avvia lentamente, con la comparsa progressiva di nuovi sintomi in tempi successivi (settimane, mesi o anche anni). Da quanto appena descritto si comprende quanto possa essere difficile effettuare la diagnosi di lupus eritematoso sistemico, anche perché non esiste, allo stato attuale, un test clinico in grado di individuare la presenza della malattia. Per formulare una diagnosi corretta e tempestiva, i medici, quindi, devono avvalersi delle informazioni che concorrono a definire il quadro clinico, compresi i criteri

elaborati dall'*American Rheumatism Association*, degli esiti degli esami di laboratorio ed utilizzare metodi statistici per l'analisi di tali dati.

Anche sul fronte delle cure la situazione non è incoraggiante. Allo stato attuale non esiste una cura definitiva per il LES, sia perché l'eziologia non è certa, sia per la molteplicità delle sue manifestazioni. Le terapie farmacologiche quindi, sono strettamente influenzate dai sintomi e dalla loro gravità e hanno come scopo la remissione della malattia, la riduzione della sua attività o, almeno, il contenimento dei danni agli organi interessati. I farmaci prevalentemente utilizzati sono i corticosteroidi, che hanno l'effetto di attenuare la reazione infiammatoria ma, nei casi più gravi, si ricorre anche agli immunosoppressori per ridurre la risposta immunitaria. Un utilizzo prolungato di tali farmaci comporta, tuttavia, l'insorgenza di effetti collaterali anche importanti.

Recentemente, l'attività della malattia e i danni agli organi sono tenuti più agevolmente sotto controllo con l'impiego di nuovi farmaci biologici che agiscono in maniera selettiva sulle cause generatrici del lupus (Prete et al., 2018). In particolare, i pazienti affetti dalle forme più gravi della malattia, con LES autoanticorpi-positivo con elevata attività della malattia quali, ad esempio anti-dsDNA positivi, vengono trattati con un anticorpo monoclonale, il *belimumab*, appartenente alla categoria farmacoterapeutica degli immunosoppressori selettivi. Il belimumab è il primo agente biologico approvato dalla Agenzia del farmaco americana per il trattamento del lupus. Tuttavia, poiché anche la somministrazione di tale molecola non è esente da reazioni di ipersensibilità, infezioni e reazioni da infusione severe o pericolose per la vita, esso deve essere somministrato in un ambiente in cui sia possibile intervenire immediatamente per il trattamento di tali reazioni. I pazienti rimangono sotto controllo clinico per un periodo prolungato di tempo, data la probabilità di una reazione ad inizio ritardato. Prima dell'infusione di belimumab può essere somministrato un pre-trattamento che include un antistaminico, anche se in questo caso la piena efficacia della sua azione non è garantita.

Gli studi clinici finalizzati a valutare l'efficacia di tale terapia innovativa hanno dimostrato apprezzabili miglioramenti nei pazienti ad essa sottoposti in confronto ai pazienti che ricevevano placebo. Con riferimento a risultati di laboratorio, rischio di eruzioni e infiammazioni, uso dei cortisonici e relativa qualità della vita dei pazienti affetti da LES, la maggior efficacia del belimumab è stata trattata in vari lavori scientifici (Collins et al., 2016; Iaccarino et al., 2017; Touma Z. et al, 2017). L'obiettivo del presente lavoro è quello di descrivere la malattia e, in particolare, i positivi effetti della terapia innovativa attraverso appropriate analisi statistiche dei dati rilevati su un campione di pazienti in cura presso l'A.U.O. Policlini-

co di Bari; dopo l'introduzione, nel paragrafo 2 sarà illustrata la metodologia statistica, nel paragrafo 3 si discuteranno i principali risultati per giungere alle sintetiche conclusioni illustrate nel paragrafo 4.

## 2. Materiali e metodi

Il campione analizzato comprende 38 pazienti affetti da LES residenti a Bari e provincia e in cura presso l'U.O Medicina Interna G. Baccelli, Dipartimento di Scienze biomediche e oncologia umana, dell'Azienda Universitaria Ospedaliera Policlinico di Bari. Tutti i pazienti sono stati trattati con belimumab, anche se la sua somministrazione avviene in momenti diversi per ciascun paziente.

Le manifestazioni che hanno comportato l'inclusione dei pazienti nel gruppo sottoposto a somministrazione di belimumab in aggiunta alla terapia già in atto si sono manifestate soprattutto a carico dell'apparato muscolo-scheletrico (63,1%), del distretto muco-cutaneo (31,6%), renale (26,3%) ed ematologico (13,1%).

Nel campione analizzato, il LES colpisce per il 78,9% le donne; l'età media dei pazienti si aggira intorno ai 44,3±3,9 anni, attestandosi su 44,5 anni per le femmine e 43,4 anni per i maschi.

Nella scheda di ciascun paziente partecipante allo studio sono disponibili le seguenti informazioni anagrafiche:

- data di nascita,
- genere,
- età, con calcolo aggiornato alla data dell'analisi (prima metà di giugno 2018),
- età a cui è stata diagnosticata la malattia,

e cliniche:

- profilo anticorpi (la presenza o assenza di un anticorpo è espressa assegnando valori di 1 (presenza) o 0 (assenza),
- pregresse terapie (terapie e farmaci pre- belimumab),
- pregressi impegni d'organo,
- comorbilità (patologie del paziente non associate al lupus),
- terapie concomitanti (terapie e farmaci usati contestualmente al belimumab),
- interessamento attivo (organi colpiti dalla malattia),
- interessamento prevalente, ossia impegni d'organo che hanno richiesto l'impiego di belimumab in aggiunta alla terapia già in atto.

Le suddette informazioni riguardano la situazione clinica dei pazienti prima del trattamento con belimumab; con T_0 si definisce il momento in cui è avvenuta la

prima somministrazione di belimumab e periodicamente, a partire da T_0, vengono misurate alcune variabili cliniche, il cui confronto permette di ottenere infor–mazioni sugli eventuali miglioramenti della condizione di malattia dei pazienti.

Per monitorare la condizione dei pazienti vengono misurate semestralmente le seguenti variabili cliniche:

- SLEDAI-2K, che valuta l'attività di malattia in nove organi (range 0-40),
- DAS28, che considera il numero di articolazioni dolenti (range 2-10),
- PGA, score soggettivo assegnato dal medico (range 0-3),
- SRI-4,
- VAS, SLICC-D.

Le ultime due variabili non contemplano un range predefinito; in particolare, la variabile SRI-4 comprime tre indici tra cui SLEDAI-2K e PGA mentre VAS è una misura soggettiva del dolore provato dal paziente, basata su una scala del dolore ad intervalli, di 10 punti, le cui estremità sono "Nessun dolore" e "Peggior dolore mai provato".

La variabile SLEDAI-2K valuta l'attività della malattia in nove organi, in particolare:

- coinvolgimento del sistema nervoso, vascolare, renale e muscolare a cui vengono associati al massimo quattro punti ciascuno;
- coinvolgimento cutaneo, immunologico e delle sierosa, a cui vengono associati al massimo tre punti ciascuno;
- coinvolgimento generale ed ematologico a cui viene associato al massimo un punto ciascuno.

Il valore SLEDAI-2K , come anticipato, può assumere valori compresi tra 0 e 40 secondo la seguente scala:

| *Score* | *Livello di attività* |
|---|---|
| SLEDAI-2K = 0 | Nessuna attività |
| $1 \leq$ SLEDAI-2K < 5 | Attività lieve |
| $5 \leq$ SLEDAI-2K <10 | Attività moderata |
| $10 \leq$ SLEDAI-2K < 20 | Attività elevata |
| SLEDAI-2K $\geq$ 20 | Attività molto elevata |

La variabile PGA è uno score soggettivo del paziente, compreso tra 0 e 3; esso è concorde con la variabile SLEDAI-2K.

La variabile DAS28 è calcolata considerando il numero di articolazioni dolenti su un totale di 28; il suo valore può essere compreso tra 2 e 10, secondo la seguente scala:

| *Score* | *Livello di attività* |
|---|---|
| DAS28 < 2,6 | Remissione |
| 2,6 ≤DAS28 < 3,2 | Attività lieve |
| 3,2 ≤ DAS28 <5,1 | Attività moderata |
| DAS28 ≥5,1 | Attività elevata |

Conformemente alla letteratura (Frieri et al, 2015), la variabile SRI-4 è influenzata da SLEDAI-2K e PGA e viene osservata dalla cinquantaduesima settimana di trattamento e fornisce informazioni sull'eventuale miglioramento del paziente. Se lo SLEDAI-2K si riduce di almeno quattro punti e il PGA registra un aumento inferiore a 0,3 allora l'indice SRI-4=1 e ciò indica un miglioramento della condizione di malattia del paziente.

Le analisi statistiche illustrate nel presente lavoro sono state realizzate con il software R. I test utilizzati sono stati quello di Shapiro-Wilk per verificare la normalità delle distribuzioni, il test di McNemar per campioni appaiati per comparare i valori medi della variabile PGA nei tempi $T_0$ e $T_6$, considerandola come variabile qualitativa per classificare il paziente a seconda dello "stato di salute migliorato" o "stato di salute peggiorato" , il test di Friedman per campioni dipendenti al fine di verificare l'ipotesi di uguaglianza delle medie delle variabili cliniche in tre tempi successivi, il test t di Student per il calcolo degli intervalli di confidenza delle stime dei valori medi dell'età.

Una delle criticità emerse nello svolgimento dell'analisi riguarda il periodo di inizio della somministrazione della terapia con belimumab. Infatti, dal momento che i pazienti hanno iniziato la terapia in tempi diversi tra loro ma comunque molto recenti, allo stato attuale si dispone ancora di poche informazioni cliniche relative ai tempi T_12, T_18,…,T_36 riguardo quei pazienti entrati nello studio nel 2017 o addirittura nel 2018. Ciò, ovviamente, non consente ancora di ottenere risultati "robusti" anche se sono state utilizzate le tecniche statistiche non parametriche; poiché lo studio proseguirà, tale criticità si risolverà con il trascorrere del tempo.

## 3. Risultati

Preliminarmente alle elaborazioni statistiche, le ipotesi di ricerca sono state formulate insieme all'equipe medica.

Una prima verifica della rappresentatività del nostro campione rispetto alla letteratura ha riguardato la distribuzione per genere per capire se la frequenza con cui

si sviluppa la malattia sia uguale negli uomini e nelle donne o, al contrario essa abbia maggiore probabilità di svilupparsi nelle donne.

Supponendo che il genere non influenzi la probabilità di insorgenza della malattia, nel campione la proporzione di pazienti di genere femminile, pari a 0,79, è risultata significativa al 95%. Dopo aver verificato l'ipotesi di normalità con il test di Shapiro-Wilk (p-value=0,2341) per poter utilizzare il test *T* di Student, dall'analisi esplorativa è emerso che l'età media dei pazienti attualmente in cura si attesta sui 44,3 ± 3,9 anni ($\alpha$=0,05). Prendendo in considerazione, invece, il dato relativo all'età a cui è stata diagnosticata la malattia per ciascun paziente, l'età media è risultata di 31,9 ± 4,1 anni.

L'età non si è rivelato un fattore che influisce sulla condizione di malattia dei pazienti e, infatti, come si evince dalle Figure 1 e 2, non emerge alcuna relazione tra l'età dei pazienti al tempo iniziale ($T_0$) e i valori delle variabili cliniche SLEDAI-2K e DAS28. Il tentativo di stima di un modello di regressione lineare mostra un valore di $R^2$ prossimo allo 0, fornendo ulteriore conferma. Passando ad un'analisi non parametrica, pur con un numero limitato di casi, il coefficiente $\rho$ di Spearman tra età e DAS28 è risultato pari a 0,303, evidenziando una bassa concordanza. Le variabili età e SLEDAI-2K sono risultate, invece, lievemente discordanti ($\rho$=–0,166).

**Figura 1.** *Relazione tra età dei pazienti in $T_0$ e valori SLEDAI-2K*

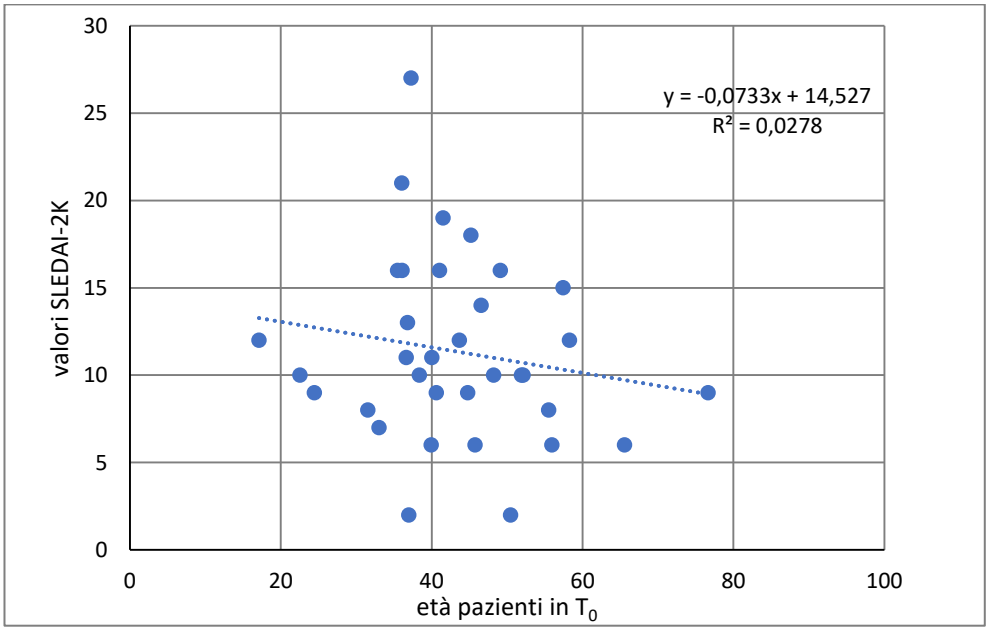
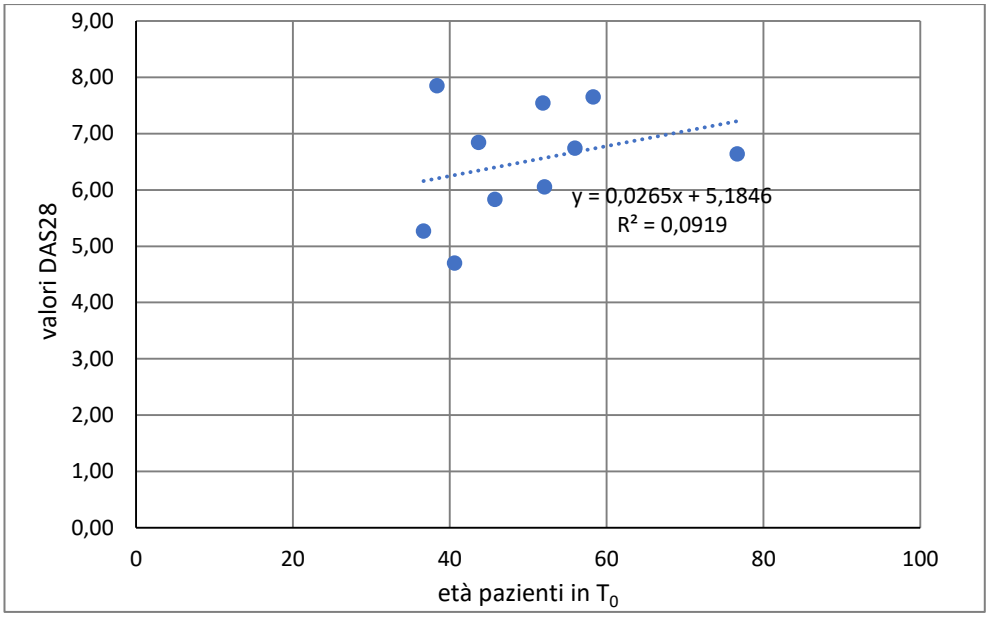


**Figura 2.** *Relazione tra età dei pazienti in $T_0$ e valori DAS28*




I dati sulle condizioni dei pazienti rilevati dall'equipe medica in tempi diversi, precisamente in $T_0$, $T_6$, $T_{12}$ hanno, altresì, consentito di verificare l'efficacia della terapia innovativa. La prima variabile considerata è PGA, lo score soggettivo attribuito dal medico sulla condizione di malattia del paziente; esso è compreso tra 0 e 3, valore quest'ultimo che indica uno stato di malattia molto grave.

Per valutare se tra i tre tempi ci sia stato un miglioramento delle condizioni dei pazienti, è stata verifica l'ipotesi di uguaglianza dei livelli medi della variabile PGA nei tre campioni dipendenti con il test di Friedman, non potendosi assumere la normalità della distribuzione; con probabilità pari a 0,95, il risultato del test consente di affermare che le condizioni di salute dei pazienti, in media, sono migliorate nel tempo (p-value=0,001805). Per individuare quali siano le medie dei campioni in $T_0$, $T_6$, $T_{12}$ a cui attribuire tale risultato, si è fatto ricorso al metodo dei confronti post-hoc, dal quale emerge che i valori nei tempi $T_0$ e $T_{12}$ sono diversi tra loro; in particolare, trascorsi 12 mesi, i livelli di PGA si attestano su valori più bassi, confermando quindi il miglioramento delle condizioni dei pazienti rispetto al momento di inizio della terapia.

### *3.1 Evoluzione della gravità della malattia nei primi sei mesi*

Per i clinici, la maggiore o minore rapidità di evoluzione di una malattia è una informazione molto importante ai fini della diagnosi e, soprattutto, della cura. Per

calcolare tale dato, si possono confrontare i valori assunti dalla variabile di interesse in due tempi di osservazione successivi tramite il test di McNemar. Il test verifica se il farmaco abbia indotto significativi cambiamenti rispetto alla gravità della malattia e quindi, nel campione, non si tiene conto dei pazienti la cui condizione è rimasta invariata in quanto, ai fini del test, non forniscono alcuna informazione sull'effetto del farmaco.

Se è vera l'ipotesi nulla, e quindi il farmaco non produce alcun effetto, un mutamento della condizione di malattia in quei pazienti che hanno sperimentato tale evento è di natura casuale. Di conseguenza, il numero (e la proporzione) di coloro che sono passati dalla condizione grave a quella non grave è equivalente al numero (e alla proporzione) di coloro che hanno cambiato nell'altra direzione, dal non grave a grave, con un margine di errore del 5%. La statistica test $\chi 2$ consente di affermare che, con una probabilità del 95%, la malattia ha registrato una riduzione in termini di gravità ascrivibile all'efficacia del farmaco.

**Tabella 1.** *Variazioni nelle gravità della malattia nei primi sei mesi*

| | PGA $T_6$ | | |
|---|---|---|---|
| PGA $T_0$ | Grave | Non grave | Totale |
| Grave | 16 | 10 | 26 |
| Non grave | 1 | 4 | 5 |
| Totale | 17 | 14 | 31 |

## 4. Conclusioni

Il lupus eritematoso sistemico (LES) è una malattia cronica autoimmune causata da un'attivazione incontrollata del sistema immunitario che comporta una infiammazione dei tessuti dell'organismo. I sintomi evidenti del LES sono la fotosensibilità e le manifestazioni cutanee, in particolare sulle guance e vicino al naso, con il tipico eritema a «farfalla». Il LES può colpire organi interni quali il cuore, i polmoni, il fegato, i reni e anche il sistema nervoso. Per il fatto di colpire le articolazioni, il LES è classificato anche tra le malattie reumatiche. Allo stato attuale non esiste una cura definitiva per il lupus, malattia che presenta alternativamente fasi acute a fasi di remissione. Il LES viene trattato con farmaci immunosoppressori e corticosteroidi, tuttavia le terapie standard non sono scevre da effetti collaterali importanti, da moderata capacità di contrastare l'attività della malattia e da ricorrenti eruzioni cutanee. Dalle analisi condotte sul campione di pazienti sono emersi alcuni interessanti risultati confermativi delle ipotesi di ricerca, per altro conformi alla letteratura. Il lupus colpisce soprattutto le donne, con

un'età media in linea con quella degli studi finora condotti, compresa tra i 40 e 48 anni; tuttavia, l'età è risultata non influenzata dalle tre variabili cliniche (SLEDAI-2K, PGA e DAS28). Analizzando l'andamento della variabile PGA, concorde con la variabile SLEDAI-2K, si è ottenuta la conferma del miglioramento del quadro clinico del paziente, con una riduzione della gravità della malattia, attestando così l'efficacia della nuova terapia con belimumab, almeno nel breve periodo.

Indubbiamente, il numero limitato di casi analizzati nel campione deve far ritenere i risultati illustrati nel presente lavoro del tutto preliminari, ancorché ottenuti adottando le tecniche statistiche per piccoli campioni. Tuttavia, lo studio è di tipo longitudinale, quindi tale limite sarà superato col trascorrere del tempo, allorquando le informazioni cliniche sui pazienti saranno più abbondanti e tali da poter condurre analisi statistiche più approfondite. Il contributo fornito dalle analisi statistiche al medico specialista è quello fornire un supporto nel formulare una diagnosi e nel predisporre una terapia quasi personalizzata per ciascun paziente, tenendo conto di una serie di parametri e cercando di salvaguardare la sua qualità della vita.

## Riconoscimenti

Gli Autori ringraziano il prof. Vito Racanelli e la dott.ssa Marcella Prete del Dipartimento di Scienze biomediche e oncologia umana, U.O. Medicina Interna G. Baccelli dell'Azienda U. O. Policlinico di Bari per la gentile concessione dei dati utilizzati, il supporto specialistico fornito per la conoscenza del fenomeno e la formulazione delle ipotesi di ricerca.

## Riferimenti bibliografici

Bland, M. (2000). *An introduction to medical statistics*. 3rd edition. Oxford University Press.

Delvecchio, F. (2015). *Statistica per lo studio dei fenomeni sociali*. CLEUP, Padova.

Triola M. M., Triola M. F. (2005). *Biostatistics for the Biological and Health Sciences*. Wiley.

Prete M., Leone P., Frassanito M. A., Desantis V., Marasco C., Cicco S., Dammacco F., Vacca A., Racanelli V. (2018. Belimumab restores Treg/Th17 balance in patients with refractoruìy systemic lupus erythematosus. *Lupus*, 27: 1926-1935.

Frieri M., Heuser W., Bliss J. (2015). Efficacy of novel monoclonal antibody belimumab in the treatment of lupus nephritis, *J Pharmacol Pharmacother*. 6: 71-76.

Collins C. E., Dall'Era M., Kan H. et al (2016). Response to belimumab among patients with systemic lupus erythematosus in clinical practice settings: 24-months results from the OBSErve study I the USA. *Lupus Sci Med*, 3: e000118.

Furie R., Petri M., Zamani O. et al (2011). A phase III, randomized, placebo-controlled study of belimumab, a monolocal antibody that inhibits B lymphocyte stimulator in patients with systemic lupus erythematosus, *Arthritis Rheum*, 63: 3918-3930.

Gatto M., Saccon F., Zen M. et al. (2016). Success and failure of biological treatment in systemic lupus erythematosus: A critical analysis. *J Autoimmun*, 74: 94-105.

Iaccarino L., Bettio S., Reggia R et al (2017). Effects of belimumab on flare rate and expected damage progression in patients with active systemic lupus erythematosus. *Arthritis Care Res (Hoboken)*, 69: 115-123.

Iaccarino L., Andreoli L, Bartoloni Bocci E. et al (2018). Clinical predictors of response and discontinuation of belimumab in patients with systemic lupus erythematosus in real life settings. Results of a large, multicentric, nationwide study. *Journal of Autoimmunity*, 86: 1-8.

Navarra S. V., Guzman R. M., Gallacher A. E. (2011). Efficacy and safety of belimumab in patients with active systemic lupus erythematosus*:* A randomized, placebo-controlled , phase 3 trial. *Lancet*, 377: 721-731.

Nikpour M., Urowitz M.B., Ibanez D., Gladman D. D. (2009). Frequency and determinants of flare and persistently active disease in systemic lupus erythematosus, *Arthritis Rheum*, 61: 1152-1158.

Touma Z., Sayani A., Pineau CA, et al (2017). Belimumab use, clinical outcomes and glucocorticoid reduction in patients with systemic lupus erythematosus receiving belimumab in clinical practice settings: Results from the OBSErve Canada Study. *Rheumatol Int*, 37: 865-873.

Urowitz M.B., Gladman D. D., Tom B. D., Ibanez D., Farewell V.T. (2008). Changing patterns in mortality and disease outcomes for patients with systemic lupus erythematosus. *J Rheumatol*, 35: 2152-2158.

## Sitografia

https://www.fondazioneveronesi.it/magazine/tools-della-salute/glossario-delle-malattie/lupus-eritematoso-sistemico

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018* ISBN: 978-88-6629-048-3
DIPARTIMENTO DI ECONOMIA E FINANZA pp. 129-145

# Indicatori per l'analisi della customer satisfaction in aeroporto

**Laura Antonucci[1], Corrado Crocetta[1], Francesco D. d'Ovidio[2*], Najada Firza[3], Ernesto Toma[2], Domenico Viola[2]**

*[1] Università degli Studi di Foggia (IT), [2] Università degli Studi di Bari Aldo Moro (IT), [3] Università Cattolica "Zoja e Këshillit të Mirë", Tirana (AL)*

**Riassunto:** L'obiettivo di questo studio è quello di indagare, partendo da un'indagine campionaria condotta negli aeroporti pugliesi di Bari e Brindisi, l'incidenza di ciascun servizio aeroportuale sulla soddisfazione del cliente e in che modo alcuni fattori osservabili e non osservabili possono influenzare tali percezioni. Innanzitutto, le relazioni tra le variabili osservate (valutazione dei servizi aeroportuali) e le dimensioni latenti sono state esplorate utilizzando l'Analisi Fattoriale Esplorativa. I risultati di questa sono stati poi utilizzati come punto di partenza per identificare le relazioni causali in un modello di Equazioni Strutturali, che può essere assunto come rappresentativo del comportamento del cliente nei confronti della qualità dei servizi aeroportuali. Infine, è stata eseguita un'analisi di invarianza di questo modello causale, discernendo per aeroporto e per stagione (alta o bassa).

**Keywords:** Servizi aeroportuali; Customer satisfaction; Analisi fattoriale esplorativa; Modelli di equazione strutturale; Analisi dell'invarianza.

## 1. Introduzione

Oggi gli aeroporti sono complessi centri multifunzionali che offrono una vasta gamma di servizi di viaggio, e stanno diventando sempre più fattori chiave per il progresso sociale ed economico. I passeggeri di tutto il mondo richiedono livelli sempre più elevati di servizio, e per questo motivo, le autorità di regolamentazione

---

[*] Corresponding Author: francescodomenico.dovidio@uniba.it (orcid: 0000-0003-1641-039X).
Questo studio è frutto di un progetto comune, ma L. Antonucci ha redatto la sezione 3, C. Crocetta la sezione 1, F. D. d'Ovidio le sezioni 5 e 6, N. Firza la § 4, E. Toma la § 7, D. Viola la § 2. Gli autori ringraziano la società "Aeroporti di Puglia S.p.A.", che ha finanziato e supportato l'indagine qui descritta, per l'autorizzazione all'utilizzo dei dati statistici raccolti.

prestano maggiore attenzione alla fornitura di servizi aeroportuali e alla qualità della fornitura di servizi.

Lo scopo di questo studio è di indagare sulla percezione dei passeggeri riguardo la qualità dei servizi aeroportuali, sulla base di un'indagine effettuata negli aeroporti pugliesi di Bari e Brindisi nel 2016, alla ricerca di un modello generale utile per la gestione dei servizi.

L'importanza soggettiva dei servizi aeroportuali è stata esplorata innanzitutto utilizzando semplici tecniche di classificazione statistica; successivamente, utilizzando l'Analisi Fattoriale Esplorativa sono state identificate alcune relazioni tra le variabili osservate (giudizi dei servizi) in grado di delineare l'esistenza di dimensioni latenti della qualità. I risultati della AFE sono stati quindi utilizzati per un'Analisi Fattoriale Confermativa allo scopo di trovare poi, con metodi totalmente "data driven", le relazioni causali in un Modello di Equazione Strutturale. Tale modello causale può essere assunto come rappresentativo del comportamento dei clienti nei confronti del servizi aeroportuali, utile per affrontare le sfide dell'orientamento strategico e quelle della gestione del marketing.

Ulteriori risultati di questa analisi sono l'identificazione di differenze e somiglianze tra le relazioni precedenti, distinte per aeroporto, e il discernimento dell'effetto della stagionalità sulle valutazioni fornite.

## 2. L'indagine sul campo

Nel 2016 è stata avviata un'indagine periodica sulla soddisfazione della clientela negli aeroporti pugliesi di Bari e Brindisi, entrambi gestiti da "Aeroporti di Puglia S.p.A.". Tale indagine, organizzata secondo le regole definite dall'Ente Nazionale per l'Aviazione Civile nel 2014, è stata effettuata in due tempi (stagione estiva ed invernale), al fine di valutare in modo estensivo la qualità dei servizi aeroportuali attraverso giudizi sui diversi servizi forniti nelle strutture e sulle caratteristiche di *Customer Satisfaction*.

Questo studio si basa su un questionario standardizzato, comprendente alcune informazioni personali e le restanti domande tutte richiedenti una valutazione in scala 1-10 punti (ove 1 indica il caso peggiore e 10 quello migliore), con l'obiettivo di valutare un'ampia gamma di servizi forniti dall'autorità aeroportuale o da altri soggetti autorizzati, seguendo appunto i criteri della circolare ENAC GEN-06 (2014).

Le dieci sezioni del questionario somministrato ai passeggeri in attesa di accesso nell'area di partenza degli aeroporti, cartaceo e composto da 2 facciate, sono:

**A.** *Dati personali* (Genere, Età, Titolo di studio, Motivi del viaggio);
**B.** *Servizi di sicurezza* (Servizio di controllo di sicurezza delle persone e dei bagagli a mano, Sicurezza personale e patrimoniale in aeroporto);
**C.** *Accuratezza e Puntualità* (Regolarità e puntualità dei servizi ricevuti in aeroporto);
**D.** *Pulizia e Igiene* (Pulizia e funzionalità dei servizi igienici, Pulizia dell'aeroporto);
**E.** *Comfort* (Disponibilità di carrelli portabagagli, Efficienza dei sistemi di trasferimento passeggeri, Efficienza degli impianti di climatizzazione, Comfort complessivo);
**F.** *Servizi Aggiuntivi* (Accessibilità e affidabilità di connessioni Wi-Fi, Disponibilità e reperibilità dei punti di ricarica per telefoni cellulari o computer portatili, Disponibilità di distributori di bevande e snack, Valutazione di ristoranti e altri negozi[1]);
**G.** *Servizi Informativi* (Efficacia dei punti informativi, Chiarezza ed efficacia della segnaletica interna, Professionalità del personale, Aggiornamento e facile lettura del sito web, Accessibilità e generale efficacia delle informazioni);
**H.** *Servizi Gate e Desk* (Servizi di biglietteria, Tempi di attesa ai banchi check-in, Tempi di attesa ai controlli di sicurezza);
**I.** *Rete di Trasporto* (Chiarezza ed efficacia della segnaletica esterna, Adeguatezza delle connessioni tra città e aeroporto);
**J.** *Soddisfazione complessiva* (Qualità attesa dei servizi aeroportuali, Qualità percepita dei servizi fruiti).

Sui questionari sono presenti ulteriori informazioni rilevate dagli intervistatori, ossia: data dell'intervista e aeroporto in cui essa è stata effettuata (Bari o Brindisi), compagnia aerea utilizzata dall'intervistato (Alitalia, Ryanair o altro vettore) e tipo di volo (nazionale o internazionale).

Nell'aeroporto di Bari sono stati intervistati in modo casuale sistematico (1 viaggiatore in partenza ogni 30 che attraversavano una "linea ideale" nell'area di accesso, successiva ai banchi del check in) 840 passeggeri in alta stagione, ossia nella settimana 27 giugno-3 luglio; ne sono stati invece intervistati 1.288 in periodo di bassa stagione (23-29 novembre), con la medesima cadenza; parallelamente e

[1] I giudizi riguardo "Valutazione di ristoranti ed altri negozi" sono in realtà espressi in 8 item distinti: *Assortimento, Qualità, Prezzi, Cortesia dello staff*, separatamente per negozi-edicole e per ristoranti-caffetterie.

con cadenza identica, 840 passeggeri sono stati intervistati a Brindisi nella settimana 4-9 luglio e 712 sono stati intervistati nella settimana 10-16 dicembre (Tabella 1). Sebbene la dimensione complessiva del campione sia molto cospicua (quasi 3.500 intervistati), la sua copertura della popolazione studiata risulta abbastanza ridotta (circa 0,5‰ in entrambi gli aeroporti); tuttavia, il campione, grazie anche alle modalità di rilevazione, garantisce stime affidabili dei parametri della popolazione dei passeggeri transitati negli aeroporti pugliesi.

**Tabella 1.** *Copertura della popolazione osservata tramite il campione.*

| | Passeggeri | | | | |
|---|---|---|---|---|---|
| Aeroporti | Popolazione (passeggeri nel 2016) | Alta stagione | Bassa stagione | Campione complessivo | Copertura (‰) |
| Bari | 4.312.496 | 840 | 1.288 | 2.128 | 0,49 |
| Brindisi | 2.323.298 | 640 | 712 | 1.352 | 0,58 |
| Aeroporti di Puglia | 6.635.794 | 1.480 | 2.000 | 3.480 | 0,52 |

## 3. Prime analisi statistiche dei servizi aeroportuali

Va qui premesso che, per motivi presumibilmente legati alla non disprezzabile lunghezza del questionario, solo 3.179 degli intervistati (83% del totale) hanno risposto alle domande finali, e specificatamente a quelle sulla soddisfazione generale, fondamentali per l'analisi statistica; pertanto, le analisi successive si concentrano proprio su questi passeggeri.

Poiché le domande con una grande quantità di valori mancanti producono risultati meno robusti e validi nelle analisi multivariate, è necessario procedere ad uno screening preventivo per poter utilizzare dati significativi. Pertanto, utilizzando tabelle incrociate multiple con eliminazione progressiva di domande critiche (quelle con valori mancanti, ma anche quelle scarsamente discriminanti o correlate alle altre, in questo con l'ausilio del coefficiente Alfa di Cronbach), è stato selezionato il massimo numero di item in grado di produrre un campione di numerosità superiore a 1.000 unità.

Il miglior risultato di questa procedura ha identificato 17 item: ciascuno di essi è stato valutato da quasi il 70% dei 3.179 intervistati che hanno completato il questionario, mentre ben 1.113 intervistati hanno risposto a tutte le domande.

Quindi, le analisi successive riguardano tutte le domande (o item) relative a *Servizio di Sicurezza*, *Precisione e Puntualità*, *Pulizia e Igiene*, *Comfort*, *Servizio Informazioni*, ma anche alcune domande relative ai *Servizi di Gate e Check-in*,

escluso il *Servizio Biglietteria*, a causa dell'uso ormai diffusissimo di prenotazione e acquisto via web, e una singola domanda sulle *Reti di Trasporto* (chiarezza ed efficacia della segnaletica esterna). Si veda in proposito la successiva Tabella 2. *Tutti* i servizi aggiuntivi risultano invece esclusi, perché molti passeggeri non li hanno utilizzati (in tutto o in parte) e quindi non hanno fornito alcun giudizio su di essi.

L'affidabilità dei 17 item selezionati è eccellente nel subcampione, come dimostra l'Alfa di Cronbach dell'insieme, pari a 0,905 (Cronbach, 1951). Come verifica di tale risultato, si è constatato che l'esclusione di altri item o il reinserimento di item precedentemente eliminati riduce, sia pur in misura limitata, l'indice Alfa.

**Tabella 2.** *Valutazione dei servizi aeroportuali da parte dei passeggeri intervistati nei due aeroporti.*

| Sez. | Indicatori dei servizi | Bari | Brindisi |
|---|---|---|---|
| B | Servizio di controllo sicurezza delle persone e dei bagagli a mano | 7,5 | 7,7 |
| | Sicurezza personale e patrimoniale in aeroporto | 7,7 | 7,7 |
| C | Regolarità e puntualità dei servizi ricevuti in aeroporto | 7,6 | 7,6 |
| D | Pulizia e funzionalità dei servizi igienici | 7,6 | 7,6 |
| | Pulizia dell'aeroporto | 7,9 | 7,8 |
| E | Disponibilità di carrelli portabagagli | 7,4 | 7,5 |
| | Efficienza dei sistemi di trasferimento dei passeggeri | 7,8 | 7,6 |
| | Efficienza degli impianti di climatizzazione | **8,0** | **7,9** |
| | Comfort complessivo | 7,9 | 7,7 |
| F | Accessibilità e affidabilità di connessioni Wi-Fi | *5,5* | 6,3 |
| | Disponibilità e reperibilità dei punti di ricarica per telefoni cellulari o computer portatili | 5,7 | *5,6* |
| | Disponibilità di distributori di bevande e snack | 7,0 | 7,0 |
| | Valutazione di ristoranti e altri negozi | 6,7 | 6,9 |
| G | Efficacia degli *info point* | 7,2 | 7,2 |
| | Chiarezza ed efficacia della segnaletica interna | 7,5 | 7,5 |
| | Professionalità del personale | 7,5 | 7,5 |
| | Aggiornamento e facile lettura del sito web | 7,2 | 7,2 |
| | Accessibilità e generale efficacia delle informazioni | 7,3 | 7,3 |
| H | Ticket service | 7,3 | 7,4 |
| | Tempi di attesa ai banchi check-in | 7,4 | 7,3 |
| | Tempi di attesa ai controlli di sicurezza | 7,6 | 7,5 |
| I | Chiarezza ed efficacia della segnaletica esterna | 7,7 | 7,7 |
| | Adeguatezza delle connessioni tra città e aeroporto | 7,4 | 7,5 |
| J | Qualità attesa dei servizi aeroportuali | 7,7 | 7,7 |
| | Percezione globale della qualità dei servizi aeroportuali fruiti | 7,8 | 7,8 |

Le valutazioni medie fornite dai passeggeri ai vari servizi fruiti (Tabella 2) risultano abbastanza simili tra i due aeroporti considerati, e anche abbastanza simili tra loro, con il massimo, assegnato all'efficienza degli impianti di climatizzazione, solo di qualche decimo di punto superiore agli altri item. Fanno eccezione i primi due item della sezione F dedicati alla tecnologia, sensibilmente inferiori agli altri.

## 4. Analisi Fattoriale Esplorativa

Al fine di identificare le dimensioni latenti della percezione della qualità del servizio in "Aeroporti di Puglia", è stata applicata una Analisi Fattoriale Esplorativa a tutti i 17 item precedentemente identificati. Si tenga conto che, pur essendo detti item espressi in punteggi di valutazione (quindi di per sé maggiormente prossimi a scale ordinali che non a scale cardinali), l'articolazione dei punteggi medesimi e la loro distribuzione consentono di considerarli *proxy* di variabili quantitative sottostanti.

Sono stati utilizzati, comunque, due differenti metodi di estrazione, quello delle Componenti Principali (Hotelling, 1933) e quello dei Fattorializzazione degli Assi Principali, noto come PAF, più robusto contro la violazione della normalità (Fabrigar et al., 1999). Entrambi hanno fornito risultati simili, ma qui vengono mostrati quasi sempre solo i risultati della seconda tecnica, perché la valutazione degli item fornita dagli intervistati è sensibilmente asimmetrica e ben lontana dalla distribuzione normale.

Per identificare un sistema fattoriale robusto e massimamente esplicativo degli indicatori utilizzati, è stata applicata una procedura di "backward elimination" degli item meno correlati agli altri, espressa nei seguenti passi:

1. in primo luogo, seguendo il criterio di Kaiser (1960) modificato come suggerito da Larsen e Warne (2010), ogni Componente Principale è stata selezionata solo se si verificava la condizione {(limite inferiore dell'intervallo di confidenza dell'autovalore) > 1,0};
2. dopo di ciò, è stato escluso l'item con la comunanza meno rilevante (cioè quello la cui quota di spiegata dalla soluzione fattoriale fosse inferiore alle altre e non maggiore di 0,5).

Dopo aver eliminato un item, la procedura andava reiterata dallo step 1, fino ad ottenere solo variabili con una buona comunanza. Tuttavia, in questo caso, essa si è interrotta dopo aver eliminato soltanto la valutazione dell'item "Chiarezza ed Efficacia della Segnaletica Esterna", sia con il metodo di estrazione PC che con il PAF.

Questo risultato ha consentito che le successive analisi potessero utilizzare le valutazioni fornite da un numero maggiore di soggetti, poiché alcuni erano stati precedentemente esclusi non avendo fornito risposta proprio a tale item.

Pertanto, il data base finale risulta essere composto da 1.164 unità. Sono state individuate tre componenti principali[2], in grado di spiegare oltre il 74% della variabilità totale espressa dai 16 item restanti; la statistica KMO di questa soluzione, pari a 0,891, è quasi ottimale, posto che Max(KMO) = 1,00 (Kaiser, 1972).

**Tabella 3.** *Analisi Fattoriale della qualità percepita di servizi, con rotazione obliqua Promax (k=3.5)*

| Componenti | Risultati iniziali ed estrapolati | | | *PC* | *PAF* |
|---|---|---|---|---|---|
| | Autovalori | Percentuale di Varianza | % Cumulativa di Varianza | *Autovalori Rotazione Promax* (*) | *Autovalori Rotazione Promax* (*) |
| ***1*** | **6,532** | **40,824** | **40,824** | ***5,621*** | ***5,371*** |
| ***2*** | **4,343** | **27,143** | **67,967** | ***5,400*** | ***5,150*** |
| ***3*** | **1,027** | **6,420** | **74,387** | ***3,684*** | ***3,420*** |
| *4* | 0,786 | 4,914 | 79,301 | | |
| *5* | 0,624 | 3,898 | 83,199 | | |
| *6* | 0,434 | 2,716 | 85,914 | | |
| *7* | 0,407 | 2,545 | 88,460 | | |
| *8* | 0,351 | 2,193 | 90,652 | | |
| *9* | 0,290 | 1,813 | 92,465 | | |
| *10* | 0,239 | 1,494 | 93,959 | | |
| *11* | 0,202 | 1,264 | 95,222 | | |
| *12* | 0,183 | 1,146 | 96,368 | | |
| *13* | 0,165 | 1,032 | 97,400 | | |
| *14* | 0,163 | 1,020 | 98,420 | | |
| *15* | 0,139 | 0,868 | 99,288 | | |
| *16* | 0,114 | 0,712 | 100,000 | | |

Le componenti identificate sono stati sottoposte ad un'analisi fattoriale obliqua mediante rotazione Promax con costante frazionale (Manly 1986): la Tabella 3 mostra gli autovalori della soluzione ruotata per entrambe le procedure di estrazione

[2] La robustezza di questa soluzione, così come quella del successivo modello causale, è stata verificata utilizzando una procedura di bootstrap (Efron, 1979), ottenendo risultati abbastanza buoni. Le mediane dei bootstrap degli autovalori sono molto vicine alla loro media (differenza dell'1% o meno), e anche gli intervalli percentili al 90% sono piuttosto stretti: in genere il 5° e il 95° percentile si discostano meno del ±15% dal valore mediano della rispettiva distribuzione e nella maggior parte dei casi sono prossimi a ±5%. I bootstrap sono distribuiti quasi normalmente, con un'asimmetria accettabile: minimo -0,29 per l'autovalore dell'estrazione della 2ª componente, massimo 0,81 per l'autovalore dell'estrazione del sesto componente (con errore standardizzato di circa 0,08). Si tenga conto che le distribuzioni non normali sono identificate da asimmetria maggiore di 1,0 in valore assoluto.

PC e PAF, mentre la Tabella 4 mostra i soli pesi fattoriali della matrice ruotata stimata dal PAF, che definisce le dimensioni latenti della qualità, sensibilmente correlate tra di loro.

**Tabella 4.** *Comunanze e pesi fattoriali (Pattern Matrix) degli item valutati.*

| Item | *Comu–nanze* | Fattori 1 | 2 | 3 |
|---|---|---|---|---|
| G2-Chiarezza ed efficacia della segnaletica interna | *0,809* | 0,904 | -0,064 | 0,049 |
| G3-Professionalità del personale | *0,809* | 0,903 | -0,022 | 0,004 |
| G5-Accessibilità/efficacia complessiva delle informaz. | *0,791* | 0,892 | -0,006 | -0,008 |
| G1-Efficacia degli *info point* | *0,729* | 0,855 | -0,068 | 0,074 |
| G4-Aggiornamento e facilità di lettura del sito web | *0,724* | 0,837 | 0,063 | -0,014 |
| H3-Tempo di attesa ai controlli di sicurezza | *0,644* | 0,807 | 0,013 | -0,060 |
| H2-Tempi di attesa ai banchi check-in | *0,646* | 0,788 | 0,101 | -0,068 |
| E4-Comfort complessivo | *0,671* | 0,023 | 0,837 | -0,040 |
| E1-Disponibilità di carrelli portabagagli | *0,623* | -0,049 | 0,825 | -0,047 |
| E3-Efficienza degli impianti di climatizzazione | *0,644* | 0,020 | 0,818 | -0,033 |
| E2-Efficienza dei sistemi di trasferimento dei passeggeri | *0,609* | 0,005 | 0,810 | -0,053 |
| D2-Pulizia dell'aeroporto | *0,642* | -0,018 | 0,766 | 0,063 |
| D1-Pulizia e funzionalità dei servizi igienici | *0,540* | 0,008 | 0,682 | 0,080 |
| C1-Regolarità e puntualità dei servizi ricevuti | *0,553* | 0,043 | 0,512 | 0,302 |
| B2-Sicurezza personale e protezione della proprietà | *0,883* | -0,006 | 0,116 | 0,867 |
| B1-Controllo sicurezza passeggeri e bagagli a mano | *0,736* | -0,019 | 0,058 | 0,824 |

A causa del metodo di rotazione scelto (motivatamente, per ottenere indicazioni utili alle successive analisi), infatti, il 2° e il 3° fattore sono fortemente correlati l'uno con l'altro (Tabella 5), in particolare a causa della loro relazione con "Regolarità e Puntualità dei servizi fruiti". Il primo fattore ha invece una leggera ma statisticamente significativa correlazione con il 2° fattore, ma anche con il terzo.

**Tabella 5.** *Matrice di correlazione dei fattori ruotati.*

| Fattori | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 1.000 | **0.213** | 0.146 |
| 2 | **0.213** | 1.000 | **0.600** |
| 3 | 0.146 | **0.600** | 1.000 |

Osservando ancora i pesi fattoriali mostrati in Tab. 4, si evince immediatamente che il 1° Fattore è molto rappresentativo della dimensione "Efficienza e Professionalità", mentre il 2° Fattore è chiaramente correlato agli "Aspetti Tangibili" e, ultimo ma non per importanza, il 3° Fattore è senza dubbio identificato con la "Affidabilità" dei servizi fruiti dai passeggeri in aeroporto. Non si ritiene necessario qui approfondire i risultati dell'analisi esplorativa, in quanto quasi replicati nel seguito.

## 5. Analisi Causale S.E.M.

L'Analisi Fattoriale Esplorativa ha identificato tre fattori latenti della Qualità dei Servizi, ma si ritiene in genere che l'interpretazione di questi fattori e le relazioni tra di loro sia più completa identificando un modello causale. Per raggiungere tale risultato, è stata applicata la procedura di Equazioni Strutturali con un modello LISREL-like (Jöreskog, 1973; Jöreskog and Sorbom, 1984)[3], identificato tramite il software AMOS con l'opzione *EmuLisrel.*

Al fine di identificare il miglior Sistema causale, è stata applicate un'Analisi Fattoriale Confermativa, partendo inizialmente dalla struttura fattoriale precedentemente descritta, la cui matrice di correlazione è illustrate in Tabella 5, per poi aggiungere una variabile latente (CS, ossia "Customer Satisfaction") misurata da entrambi gli item osservati "Qualità Attesa dei servizi aeroportuali" e "Qualità Percepita" (ossia la percezione globale della qualità dei servizi aeroportuali fruiti).

A causa delle notevoli violazioni di normalità delle variabili osservate (e di conseguenza, presumibilmente, anche delle variabili latenti, o LV), si è deciso di stimare pesi fattoriali e discrepanze tramite il metodo robusto *Asymptotically Distribution Free*[4] proposto da Browne (1984).

La Fattoriale Confermativa ha identificato un problema critico nella variabile osservata "Pulizia e Funzionalità dei servizi igienici", che è stata poi esclusa dalla successiva identificazione dei "rapporti causali" tra le variabili latenti al fine di ottenere un buon modello per l'analisi della *Customer Satisfaction* (utilizzando ancora un metodo "data driven" sotto il vincolo del miglior adattamento possibile).

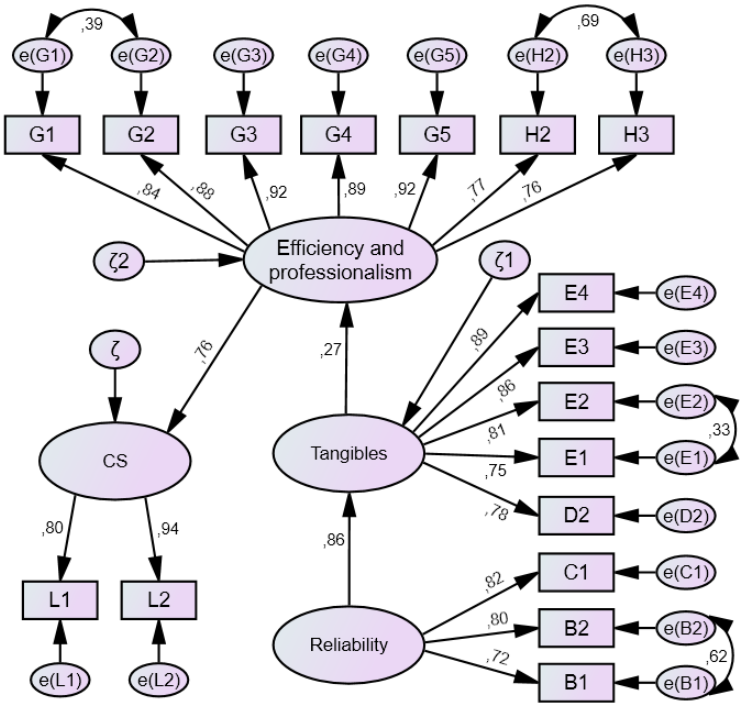
---

[3] Un Modello di Equazioni Strutturali (SEM) è descritto da due modelli matematici, che riguardano rispettivamente le variabili osservate e quelle latenti. Il primo modello (*modello di misurazione*) mappa le variabili osservate sulle proprie variabili latenti (LV): $x_h = \lambda_{ih}\xi_i + \delta_h$ se tale variabile è correlate ad un fattore causale $\xi_i$ ,$y_k = \lambda_{jk}\eta_j + \varepsilon_k$ se è collegato ad un LV causato ($\delta_h$ e $\varepsilon_k$ rappresentano I residui di ogni variabile osservata, le quali non sono correlate con nessuna LV). Il secondo modello (*structural model*) dimostra le relazioni reciproche tra le variabili latent (LV); in notazione matriciale: $\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}$ , dove **B** è la matrice (che ha gli zeri sulla diagonale principale) della regressione fattoriale pesata delle variabili endogene LV, $\boldsymbol{\Gamma}$ è la matrice dei pesi della regressione delle variabili esogene LV, e $\boldsymbol{\zeta}$ è il vettore dell'errore stocastico del modello di regressione in questione, che non è correlato con le LV. Tutte le LV e I residui hanno media zero, mentre I residui $\boldsymbol{\zeta}$ , $\boldsymbol{\varepsilon}$ e $\boldsymbol{\delta}$ sono incorrelati tra di loro. Riguardo la teoria SEM, vedere Jöreskog, 1973, 1977.

[4] Sebbene questo metodo fornisca statistiche di discrepanza più basse rispetto ad altri metodi, è molto più adeguato quando i dati non vengono misurati a livello di intervalli o non siano normalmente distribuiti (Muthèn & Kaplan, 1992). Tuttavia, esso può essere utilizzato con profitto solo quando la dimensione del campione è maggiore di $m\cdot(m+1)$, ove $m$ è il numero totale delle variabili esplicative presenti nel modello. Bentler e Chou (1987) suggeriscono, inoltre, un criterio minimo di 10 unità per ogni parametro libero (ossia, un modello con 50 d.f. deve essere misurato sulla base di almeno 500 unità statistiche).

Il software utilizzato fornisce indici di modifica, nonché indici di significatività e di adattamento, che aiutano a rimuovere o ad aggiungere relazioni al modello, in modo iterativo, fino a quando non viene perfezionato al meglio possibile. Il modello migliore, con coefficienti significativi per tutte le relazioni e la maggiore bontà dell'adattamento, è mostrato nel *path diagram* di Figura 1. Detto modello è stato poi convalidato utilizzando procedure di bootstrap (ottenendo distribuzioni quasi normali) al fine di escludere forti influenze di endogeneità e/o fattori di rumore.

**Figura 1.** *Struttura del modello SEM di Customer Satisfaction dei servizi aeroportuali*




***KEY:*** *variabili esplicative*

| **Affidabilità** | **Efficienza e Professionalità** |
|---|---|
| B1-Servizio di controllo sicurezza di passeggeri e bagagli a mano | G1-Efficacia degli *info point* |
| B2-Sicurezza personale e protezione delle proprietà | G2-Chiarezza ed efficacia della segnaletica interna |
| C1-Regolarità e puntualità dei servizi ricevuti | G3-Professionalità del personale |
| **Aspetti tangibili** | G4-Aggiornamento e facile lettura del sito web |
| D2-Pulizia dell'aeroporto | G5-Accessibilità/efficacia complessiva delle info |
| E1-Disponibilità di carrelli portabagagli | H2-Tempi di attesa ai banchi check-in |
| E2-Efficienza dei sistemi di trasferimento dei passeggeri | H3-Tempi di attesa ai controlli di sicurezza |
| E3-Efficienza degli impianti di climatizzazione | **Customer Satisfaction** |
| E4-Comfort complessivo | L1-Qualità attesa dei servizi aeroportuali |
| | L2-Percezione globale della qualità dei servizi aeroportuali fruiti |

Il modello identificato ha una struttura causale "a cascata" (o seriale), in cui ogni LV è causata dalla precedente e causa la LV successiva: logicamente, la LV *Affidabilità*, esogena, causa solo la seconda LV, *Aspetti Tangibili*, mentre la LV di interesse, *Customer Satisfaction*, è causata solo da *Efficienza e professionalità*. Tutti i coefficienti di regressione e covarianza, qui mostrati in valori standardizzati, sono fortemente significativi (p <0,001).

Gli indici di adattamento del modello, mostrati in Tabella 6, sono molto buoni[5], e dunque il modello si adatta decisamente bene ai dati osservati.

**Tabella 6.** *Indici di adattamento del modello di Customer Satisfaction identificato*.

| Indici di adattamento | Modello identificato | Modello saturato | Modello di Indipendenza |
|---|---|---|---|
| GFI-*Goodness of Fit Index* | 0.924 | 1,0 | 0.479 |
| AGFI-*Adjusted Goodness of Fit Index* | 0.896 | - | 0.414 |
| CFI-*Comparative Fit Index* | 0.920 | 1.0 | 0.000 |
| NFI-*Normed Fit Index* | 0,854 | 1,0 | 0.000 |
| TLI-*Tucker-Lewis Index* | 0.902 | - | 0.000 |
| RMR-*Root Mean Square Residual* | 0.121 | 0,0 | 1.164 |
| SRMR-*Standardized Root Mean Square Residual* | 0.102 | - | - |
| RMSEA-*Root Mean Square Error of Approximation* | 0.030 | - | 0.095 |

## 6. Analisi Causale Multigruppo

Partendo dal modello precedentemente descritto, sono state quindi eseguite due analisi multigruppo, al fine di rispondere a due ipotesi che hanno un certo interesse per "Aeroporti di Puglia S.p.A.":

- in primo luogo, un'analisi per aeroporto: Bari e Brindisi hanno ruoli diversi nella rete aerea della regione pugliese, dato che l'aeroporto di Bari accoglie settimanalmente molti voli internazionali, mentre quello di Brindisi ha una vocazione più nazionale e in ogni caso prevede un numero inferiore di voli; e questa differenza, nonostante le medie riportate nella precedente Tab. 2, potrebbe influire sulla qualità del servizio globale e sulla Customer Satisfaction;

[5] L'adattamento del modello è determinato usando indici di adattamento più corretti rispetto al test di adattamento della Discrepanza con Chi-square: cioè l'indice Adjusted Goodness-of-Fit (AGFI) di Jöreskog, l'indice di adattamento comparativo CFI (Bentler, 1990), l'indice TLI (Tucker and Lewis, 1973) e l'approssimazione della radice quadrata media d'errore (RMSEA), di Browne e Cudeck, 1993. Si noti che la massima bontà di adattamento (verificabile solo nel modello saturo sovrastimato) si verifica quando RMSEA=0.0 e gli altri indici sono pari a 1.0; TLI non ha limiti superiori, ma quando il suo valore è prossimo a 1,0, l'adattamento del modello è eccellente (Hu & Bentler, 1999).

- in seguito, un analisi stagionale: forse i passeggeri dei voli in bassa stagione (autunno), ove è presumibilmente molto meno rilevante la componente di turismo vacanziero, potevano avere esigenze e punti di vista diversi da quelli dei voli di alta stagione (estate), e quindi diverse dinamiche di soddisfazione.

**Tabella 7.** *Coefficienti del modello di misurazione e del modello strutturale, e correlazione dei fattori specifici, separatamente per gruppi; Test Z per l'equivalenza dei coefficienti all'interno di ciascun gruppo.*

| | Aeroporti | | | Stagione | | |
|---|---|---|---|---|---|---|
| | Bari (n=723) | Brindisi (n=441) | *Test Z ** | Alta (n=491) | Bassa (n=673) | *Test Z ** |
| Modello di misura | | | | | | |
| Reliability → B1 | 0,74 | 0,69 | *1,848* | 0,64 | 0,75 | ***-3,542*** |
| Reliability → B2 | 0,78 | 0,82 | *-1,654* | 0,70 | 0,79 | ***-3,253*** |
| Reliability → C1 | 0,78 | 0,92 | ***-8,594*** | 0,77 | 0,85 | ***-3,857*** |
| Tangibles → D2 | 0,73 | 0,83 | ***-4,299*** | 0,82 | 0,78 | *1,864* |
| Tangibles → E1 | 0,74 | 0,81 | ***-2,843*** | 0,80 | 0,74 | ***2,640*** |
| Tangibles → E2 | 0,80 | 0,89 | ***-4,808*** | 0,81 | 0,81 | *0,098* |
| Tangibles → E3 | 0,85 | 0,91 | ***-4,303*** | 0,83 | 0,87 | ***-2,512*** |
| Tangibles → E4 | 0,86 | 0,95 | ***-9,182*** | 0,87 | 0,91 | ***-3,346*** |
| Eff. & profess. → G1 | 0,85 | 0,88 | ***-2,324*** | 0,87 | 0,83 | ***2,565*** |
| Eff. & profess. → G2 | 0,88 | 0,91 | *-1,924* | 0,91 | 0,86 | ***4,168*** |
| Eff. & profess. → G3 | 0,94 | 0,91 | ***3,001*** | 0,96 | 0,90 | ***7,339*** |
| Eff. & profess. → G4 | 0,90 | 0,91 | *-0,989* | 0,90 | 0,89 | *0,809* |
| Eff. & profess. → G5 | 0,91 | 0,95 | ***-5,433*** | 0,92 | 0,92 | *0,544* |
| Eff. & profess. → H2 | 0,79 | 0,77 | *1,008* | 0,78 | 0,76 | *0,780* |
| Eff. & profess. → H3 | 0,78 | 0,75 | *1,085* | 0,80 | 0,74 | ***2,221*** |
| CS → L1 | 0,77 | 0,91 | ***-8,685*** | 0,84 | 0,78 | ***3,026*** |
| CS → L2 | 0,92 | 0,95 | ***-3,515*** | 0,92 | 0,96 | ***-6,104*** |
| Modello di struttura | | | | | | |
| Reliability → Tangibles | 0,88 | 0,87 | *0,284* | 0,91 | 0,85 | ***4,511*** |
| Tangibles → Eff. & profess. | 0,28 | 0,30 | *-0,378* | 0,31 | 0,24 | *1,258* |
| Eff. & profess. → CS | 0,74 | 0,79 | *-1,909* | 0,84 | 0,74 | ***4,609*** |
| Correlazioni dei fattori specifici | | | | | | |
| e(B1) ↔ e(B2) | 0,53 | 0,69 | ***-4,152*** | 0,77 | 0,61 | ***5,204*** |
| e(E1) ↔ e(E2) | 0,26 | 0,45 | ***-3,675*** | 0,61 | 0,23 | ***7,987*** |
| e(G1) ↔ e(G2) | 0,33 | 0,35 | *-0,393* | 0,33 | 0,40 | *-1,243* |
| e(H2) ↔ e(H3) | 0,69 | 0,70 | *-0,478* | 0,72 | 0,73 | *-0,213* |

* Valore critico della trasformazione Z: $z_{0,025} = 1,95996$

La Tabella 7 mostra un panorama sinottico di entrambe le analisi, vale a dire i coefficienti per ciascun aeroporto (colonne 2-3) e per ogni stagione (colonne 4-5). Sebbene molti coefficienti del modello di misurazione sembrano essere simili in entrambi i modelli aereoportuali ed in entrambi i modelli stagionali, in ciascuna

coppia di modelli solo poche misure sono statisticamente equivalenti, dato α=0.05, utilizzando l'appropriato test statistico[6].

Si noti, tuttavia, che la stima dell'ADF fornisce una buona consistenza interna (e, quindi il test di Wald sui singoli coefficienti è affidabile), ma la violazione di normalità delle variabili esplicative e delle variabili latenti rende meno potente il Test Z per la differenza dei coefficienti: per questo motivo l'importanza di tali differenze potrebbe essere discutibile.

Tuttavia è qui applicabile, avendo subcampioni di buone dimensioni, un metodo più robusto, pur se meno potente: il test del Chi quadrato con la statistica C-min delle minime discrepanze tra modelli causali teorici e i dati osservati. Poiché la struttura del modello non vincolato è risultata abbastanza stabile tra i gruppi, rispetto alle variabili date e alle relazioni identificate, siamo stati in grado di verificare le seguenti alternative correlate (cfr. Bollen, 1989; Kline, 1998):

- Invarianza della matrice dei coefficienti di misurazione (**Λ**):

$$H_\Lambda: \mathbf{\Lambda}_1 = \mathbf{\Lambda}_2\,;$$

- Invarianza della matrice dei coefficienti strutturali (**Φ**):

$$H_\Phi: \mathbf{\Phi}_1 = \mathbf{\Phi}_2\,;$$

- Invarianza della matrice di errore dei coefficienti ζ (ossia **Ψ**):

$$H_\Psi: \mathbf{\Psi}_1 = \mathbf{\Psi}_2\,;$$

- Invarianza della matrice dei residui (**Θ**):

$$H_\Theta: \mathbf{\Theta}_1 = \mathbf{\Theta}_2\,.$$

Tali ipotesi sono state verificate sia singolarmente che combinate, prefissando un livello di significatività α=0.05, per ogni coppia di modelli strutturali (aeroporto/stagione). I risultati sono mostrati in Tabella 8.

La prima sezione della tabella (ove si assume solo che il modello non vincolato sia corretto) evidenzia, come già osservato utilizzando il Test Z, che i coefficienti di misurazione di entrambi i gruppi sono *non invarianti*, e in questo scenario tutti i successivi test di invarianza falliscono. D'altro canto, assumendo che i coefficienti di misurazione siano corretti, tutte le restanti matrici della coppia di modelli "Ae-

[6]Dato che i pesi standardizzati nella regressione lineare sono pari ai coefficienti di correlazione di Pearson, per ogni *j*-ma coppia di coefficienti qui illustrati, l'ipotesi nulla $H_0$: $\lambda_{j1}=\lambda_{j2}$ vs. l'ipotesi alternativa $H_1$: $\lambda_{j1}\neq\lambda_{j2}$ è stata verificata utilizzando la trasformazione Z di *Fisher* per ogni coefficiente *r*, in modo da correggere la sua asimmetria e la sua eteroschedasticità (Fisher 1915, 1921): $z_i = ln[(1+r_i)/(1-r_i)]/2 = \text{arctanh}(r_i)$ se $|r_i|<1$, altrimenti $z_i = r_i$.

Pertanto, è qui utilizzato il test:

$$Z = (z_1 - z_2)/\sqrt{1/(n_1-3) + 1/(n_2-3)}\,,$$

che è normalmente distribuito per grandi campioni (quindi, fissato α, l'ipotesi nulla $H_0$ deve essere rifiutata se, in valore assoluto, $Z>z_{\alpha/2}$).

roporto" sono invarianti, ad eccezione dei residui di misurazione (identificabili come la matrice dei fattori specifici). Questo risultato si verifica in tutti i successivi passaggi. Le matrici relative alla coppia di modelli "Stagione", invece, falliscono i test di invarianza a quasi tutti i livelli.

**Tabella 8.** *Test di discrepanza per ogni coppia di modelli, assumendo che il modello gerarchicamente precedente sia corretto.*

| | Modello Aeroporti | | | Modello Stagioni | | |
|---|---|---|---|---|---|---|
| Model | g.d.l. | C-min | *P-value* | g.d.l. | C-min | *P-value* |
| **Assumendo che sia corretto il modello non vincolato (*Unconstrained*):** | | | | | | |
| Coefficienti di misurazione | 13 | 30.64 | *0.004* | 13 | 24.45 | *0.027* |
| Coefficienti strutturali | 16 | 33.55 | *0.006* | 16 | 44.06 | *0.000* |
| Covarianze strutturali | 17 | 35.26 | *0.006* | 17 | 49.34 | *0.000* |
| Residui strutturali | 20 | 39.10 | *0.006* | 20 | 56.31 | *0.000* |
| Residui di misurazione | 42 | 175.06 | *0.000* | 42 | 134.10 | *0.000* |
| **Assumendo che sia corretto il modello *Coefficienti di misurazione*:** | | | | | | |
| Coefficienti strutturali | 3 | 2.91 | *0.405* | 3 | 19.61 | *0.000* |
| Covarianze strutturali | 4 | 4.62 | *0.328* | 4 | 24.89 | *0.000* |
| Residui strutturali | 7 | 8.46 | *0.294* | 7 | 31.86 | *0.000* |
| Residui di misurazione | 29 | 144.42 | *0.000* | 29 | 109.65 | *0.000* |
| **Assumendo che sia corretto il modello *Coefficienti strutturali*:** | | | | | | |
| Covarianze strutturali | 1 | 1.71 | *0.191* | 1 | 5.28 | *0.022* |
| Residui strutturali | 4 | 5.54 | *0.236* | 4 | 12.25 | *0.016* |
| Residui di misurazione | 26 | 141.51 | *0.000* | 26 | 90.04 | *0.000* |
| **Assumendo che sia corretto il modello *Covarianze strutturali*:** | | | | | | |
| Residui strutturali | 3 | 3.83 | *0.280* | 3 | 6.97 | *0.073* |
| Residui di misurazione | 25 | 139.80 | *0.000* | 25 | 84.76 | *0.000* |
| **Assumendo che sia corretto il modello *Residui strutturali*:** | | | | | | |
| Residui di misurazione | 22 | 135.96 | *0.000* | 22 | 77.79 | *0.000* |

Una visione globale dei test di invarianza, quindi, fornisce i seguenti due importanti risultati:

- nonostante che in entrambi i gruppi di modelli causali i pesi di misurazione sembrano piuttosto vicini, così come lo erano i giudizi medi nel confronto precedente tra aeroporti, essi riflettono approcci diversi ai servizi aeroportuali, presumibilmente legati alla disponibilità o affidabilità di tali servizi (che in realtà non sono gli stessi negli aeroporti di Bari e Brindisi, e sicuramente sono diversi in alta e in bassa stagione, oltre ad aversi una differente congestione in aeroporto), oppure alle impressioni soggettive e spesso aleatorie, la cui modellazione è per definizione difficile o impossibile (e che possono essere considerati come fattori specifici, o residui di misura );

- a parte questi effetti diretti dell'interazione con il mondo concreto, la struttura latente della soddisfazione della qualità è pressoché stabile, ma l'importanza delle relazioni interne (abbastanza costante in entrambi gli aeroporti) risente grandemente dell'effetto stagionale.

## 7. Considerazioni conclusive

I cambiamenti di struttura e di appartenenza portano nuovi attori interessati e nuovi modelli di business nel settore aereoportuale, e la concorrenza tra gli aeroporti ha raggiunto nuove vette. Sono finiti i giorni in cui gli aeroporti erano solo i punti di partenza e di arrivo per viaggiatori disattenti.

In questo lavoro abbiamo cercato di comprendere come incrementare la soddisfazione dei passeggeri, e come i passeggeri valutano i servizi dell'Aeroporto di Puglia, quali aspetti rivestono un'importanza particolare e in che modo le percezioni e le priorità dei passeggeri cambiano nel tempo. Un'analisi esplorativa ha ridotto sensibilmente il numero di indicatori a solo 17 item utili e affidabili, ma la successiva analisi fattoriale ha evidenziato alcuni punti deboli in uno di questi indicatori, determinandone l'esclusione dal modello causale finale. Tale modello, ottenuto utilizzando la tecnica di stima robusta ADF, ha identificato un affidabile "percorso seriale" nella generazione della *Customer Satisfaction*, che è chiaramente influenzata dall'*Efficienza e Professionalità*, ma (indirettamente) anche dagli *Aspetti Tangibili* e dall'*Affidabilità* dei servizi.

Il modello identificato è stato infine sottoposto a due diverse analisi multigruppo, evidenziando la sua sostanziale invarianza rispetto ai due aeroporti considerati (e dunque alla localizzazione dei servizi), ma non rispetto alla stagione di viaggio: la valutazione dei servizi dunque risulta è considerevolmente diverso nei viaggi invernali (generalmente lavorativi) rispetto a quelli estivi, che fanno rilevare una quota consistente di viaggi per vacanza.

Questi risultati suggeriscono che le amministrazioni aeroportuali dovrebbero prestare comunque attenzione alla continuità temporale dei loro servizi interni, in particolare all'*Affidabilità* dei servizi (vedere le prime righe delle colonne di destra in Tabella 7), ma anche ad alcuni item della dimensione di *Efficienza e Professionalità*.

## Riferimenti bibliografici

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107: 238-246.

Bentler, P.M.; Chou, C.P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, 16: 78-117.

Bollen, K.A. (1989). *Structural Equations with Latent Variables*, John Wiley, New York.

Browne, M.W. (1984). Asymptotic distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37: 62-83.

Browne, M.W.; Cudeck,R. (1993). Alternative ways of assessing model fit. In: K.A. Bollen & J.S. Long (eds), *Testing structural equation models*, Sage, Newbury Park, CA: 136-162.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*. 16 (3): 297-334.

Efron B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, 7: 1-26.

Fabrigar, L. R.; Wegener, D. T.; MacCallum, R. C.; Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*. 4 (3): 272-299.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population. *Biometrika*. 10 (4): 507-521.

Fisher, R. A. (1921). On the "probable error" of a coefficient of correlation deduced from a small sample. *Metron*. 1: 3–32.

Hotelling, H.(1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24: 417-441, 498-520 (2nd part).

Hotelling, H.(1953). New light on the correlation coefficientand its transformation. *Journal of Royal Statistical Society*,Series B, 15: 193-232.

Hu, L.; Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1): 1-55.

Jöreskog, K.G. (1973). A general method for estimating a linear structural equation system. In: Goldberger A S., Duncan O. D. (eds) *Structural equation Models in the Social Sciences*, 85–112. Academic Press, New York.

Jöreskog, K.G. (1977). Structural equation models in the social sciences. In: P.R. Krishnaiah (ed.), *Application of Statistics*, North Holland, Amsterdam: 265-287.

Jöreskog, K. G.; Sorbom, D. (1984). *LISREL VI Analysis of Linear Structural Relation by Maximum Likelihood, Instrumental Variables, and Least Square Methods, User's guide*. Department of Statistics, University of Uppsala.

Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*. 20: 141-151.

Kaiser, H. F. (1974). An index of factor simplicity. *Psychometrika* 39: 31-36.
Kline, R. B. (1998). *Principles and Practice of Structural Equation Modeling*, Guilford Press, NY.
Larsen, R.; Warne, R. T. (2010). Estimating confidence intervals for eigenvalues in exploratory factor analysis. *Behavior Research Methods*, 42: 871-876.
Manly, B. F. J. (1986). *Multivariate Statistical Methods: a Primer*, London, Chapman & Hall: 77.
Muthén, B.; Kaplan, D. (1992). A comparison of some methodologies for the factor analysis of non-normal Likert variables: A note on the size of model. *British Journal of Mathematical and Statistical Psychology*, 45: 19-30.
Schwarz, G. E. (1978). Estimating the dimension of a model, *Annals of Statistics*, 6 (2): 461-464.
Tucker, L.R.; Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38: 1-10.

## Sitografia

ENAC-Ente Nazionale par l'Aviazione Civile (2014), *Circolare GEN-06*, in: https://www.enac.gov.it/la-normativa/normativa-enac/circolari/serie-gen/gen-06

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018* ISBN: 978-88-6629-048-3
DIPARTIMENTO DI ECONOMIA E FINANZA pp. 147-154

# Metodologie statistiche nei processi di sintesi dei nanomateriali

**Fabio Manca, Claudia Marin***

*Università degli studi di Bari "Aldo Moro"*

**Riassunto:** Nell'ambito delle nanoscienze gli scienziati hanno sempre dovuto affrontare la complessità dei materiali nanostrutturati, la polidispersività dei singoli campioni e la costante imperfezione dei nanomateriali. Le proprietà di questi ultimi possono dipendere dalle variazioni strutturali su scala atomica, il cui livello di precisione è, dal punto di vista sperimentale, inaccessibile. I nanomateriali posseggono qualità molto utili, ma purtoppo i loro processi di sintesi sono molto complicati ed estremamente sensibili a vari fattori. Fortunatamente esistono molteplici metodologie statistiche affidabili che sono applicabili ad essi, anche se possono presentare inconvenienti. Questo articolo propone un approccio statistico che potrebbe migliorare la capacità predittiva dei modelli fisici basata solo su supposizioni.

**Keywords:** nanoscienze; nanomateriali; processo di sintesi; modelli statistici.

## 1. Introduzione

L'obiettivo di questo articolo è di fornire una piccola disamina sugli sforzi fatti per affrontare il problema che le proprietà possedute dai nanomateriali non possano essere riprodotte su larga scala con tecniche economicamente vantaggiose.

Nella prima parte vengono mostrate le differenti categorie di approccio, quello fisico dei materiali e quello statistico, e i pro e i contro di ognuno. Nella seconda parte viene introdotto un modello sequenziale che presenta la potenzialità di delimitare i difetti dei nanomateriali. Sebbene possa sembrare un semplicissimo mo-

* Autore corrispondente: claudia.marin@uniba.it
Il lavoro qui descritto è frutto di un progetto comune, tuttavia F. Manca ha provveduto alla redazione dei paragrafi 1, 2 e 4, mentre C. Marin ha redatto il paragrafo 3 e 5.

dello di inferenza statistica, questo approccio segna la strada per la risoluzione di importanti problemi nel processo di sintesi dei nanomateriali e mostra come la statistica, in questo caso, si renda indispensabile per capire il meccanismo di crescita dei materiali nanostrutturati.

## 2. Big data e data science

Gli scienziati hanno generato per anni big data, che il più delle volte venivano sottovalutati, archiviati e ignorati. Il vero limite evidenziato da questi studiosi delle nanoscienze è rappresentato dalla incapacità di controllare in maniera indipendente un parametro strutturale o di processo; questa incapacità spesso nasconde le importanti caratteristiche fisico-chimico dei materiali nanostrutturati e rende difficile indirizzare il collegamento tra proprietà e performance.

Nelle nanoscienze la creazione di modelli è frutto di aree quali la fisica dello stato solido, i cui materiali sono continui e, di conseguenza, ben descritti da sottounità irriducibili, o anche la chimica, i cui campioni possono essere purificati e le cui molecole sono tutte uguali. Purtroppo, queste supposizioni non funzionano per i nanomateriali, in quanto le nanoparticelle non hanno una simmetria traducibile e, quindi, non possono essere descritte da sottounità irriducibili. Inoltre, i nanomateriali non possono essere purificati e la polidispersione (polidispersità) persiste a livello atomico indipendentemente dalla attenzione che si è avuta nel farli crescere, fabbricarli o processarli (Sun, Fernandez, Bernard, 2016).

Negli ultimi tempi i dati su questi materiali stanno aumentando sia in quantità che in complessità e si rende sempre più necessario trovare nuove strade per osservare, visualizzare e interrogare questa enorme quantità di informazioni.

La scienza dei dati (data science) è certamente un nuovo ambito di ricerca, che prevede una maggiore combinazione tra esperienza e conoscenza di più settori, e tra i suoi obiettivi ottempera l'individuazione di associazioni tra specifiche strutture e particolari proprietà.

### 2.1 Materials Genome Initiative

La MGI **(**Materials Genome Initiative**)** è una iniziativa "multi-agency" nata per creare una nuova era di politica, risorse e infrastrutture che possano supportare le istituzioni americane nella scoperta, manifattura e nell'uso di materiali avanzati. Con la collaborazione su larga scala di studiosi di statistica, di materiali e di esperti

di computer si intende superare la carenza di dati sistematici in materia in quanto i materiali avanzati sono fondamentali per la sicurezza economica e per l'essere umano.

Questa iniziativa è risultata come un esercizio su larga scala di data science: computer science, matematica e statistica e tutti i domini scientifici, quali chimica e fisica, collaborano in questa avventura (Sun, Fernandez, Barnard, 2016).

## 3. L'importanza della statistica nelle nanoscienze

Entriamo ora nel merito della statistica analizzando quali sono le tecniche maggiormente utilizzate in questo ambito di ricerca.

### *3.1 Tecniche di statistica descrittiva*

La statistica descrittiva agevola le nanoscienze già con il confronto tra medie, anche se la media semplice può essere applicata solo quando ogni struttura presenta lo stesso peso, la stessa ponderazione.

Nel 2015 Barnard e Wilson hanno studiato i punti quantici in silicio che sono fatti da un materiale luminescente versatile che può essere facilmente integrato con i dispositivi optoelettronici esistenti, ma con un migliore controllo dell'emissione ottica e della risoluzione spettrale. Si evince dal loro paper che, indipendentemente dalla distribuzione statistica presente nel campione, può essere sufficiente limitare semplicemente la polidispersione del campione nel modo giusto e che non è necessario ottenere una perfetta mono-dispersività, sebbene la risoluzione spettrale possa essere migliorata prendendo di mira forme specifiche. Infatti, in molti casi la struttura (e le proprietà) delle nanoparticelle presenta una distribuzione normale e si tratta di pesare in maniera conforme tutti i singoli dati e dividerli per la funzione di partizione canonica.

Altre distribuzioni statistiche possono essere applicate, comparate, facilmente analizzate e scelte all'occorrenza per rispettare alcune condizioni specifiche sperimentali o determinati processi industriali.

Anche il valore atteso e la varianza possono essere utilizzati per ottenere un Q-factor (fattore di qualità) che faciliti la classificazione delle nanoparticelle di diamante, utilizzate in ambito biomedico, e renda noto il disegno dei nanodiamanti, prevedendo come le diverse strategie sperimentali possano influenzare la performance. Nello studio di Lai e Barnard (2014) sono stati utilizzati alcuni semplici metodi statistici, in combinazione con simulazioni di strutture elettroniche, per pre-

vedere le proprietà di trasferimento di carica di diversi tipi di insiemi in cui sono state poste restrizioni sulla diversità delle strutture. Prevedendo fattori di qualità per una varietà di casi, si è evinta una chiara motivazione che ha spinto ad arricchire i campioni per aumentare la selettività e l'efficienza delle reazioni di trasferimento di carica, anche se i campioni non possano essere completamente purificati.

### *3.2 Metodi di campionamento e di riduzione*

Diverse tecniche di riduzione delle dimensioni vengono abitualmente applicate per trovare modelli in insiemi di dati ad alta complessità in diversi campi di ricerca, dalle scienze farmaceutiche a quelle terrestri. Ad esempio, l'analisi fattoriale è una tecnica statistica usata per riassumere e semplificare le relazioni di un insieme di variabili. Essa si propone di identificare le dimensioni principali di un fenomeno complesso analizzato mediante un elevato numero di variabili (q). Con essa si vuole appurare in quale misura ogni variabile realizza una ripetizione della descrizione effettuata dalle rimanenti (q-1) e se si riesce a raggiungere la stessa validità descrittiva con un numero minore o uguali di variabili (fattori). Mediante l'estrazione di pochi fattori latenti (fattori comuni a tutte le variabili osservate, o a gruppi di esse, che risultano tra loro correlate) è possibile ridurre la complessità del fenomeno pur conservando la maggior parte delle informazioni racchiuse nelle variabili originarie.

L'analisi dei fattori comuni si prefigge di ridurre la dimensione delle osservazioni ad un numero di variabili indirette più contenuto, in modo tale da non perdere le caratteristiche principali presenti nei dati osservati.

I metodi di estrazione più noti sono: il metodo delle componenti principali, il metodo degli assi principali, il metodo dei residui minimi, il metodo della massima verosimiglianza.

L'analisi delle componenti principali (ACP) viene spesso utilizzata per il riconoscimento di schemi comunemente usati per identificare i modelli intrinseci nei dati sui nanomateriali.

La *cluster analysis*, o analisi dei gruppi, comprende un gruppo numeroso di tecniche di analisi statistica multivariata. L'origine dell'analisi dei gruppi è legata all'esigenza di classificare nell'ambito delle scienze biologiche. Lo scopo di questa analisi è quello di ridurre le solitamente numerose unità statistiche iniziali in gruppi (cluster) esaustivi. Viene generalmente utilizzata per scopi pratici, organizzativi e di campionamento, ma anche per finalità esplorative volte ad individuare strutture di cluster sconosciute nei dati.

All'interno della classificazione dei metodi di analisi dei gruppi, possiamo dividere i metodi gerarchici (scissori e aggregativi) dai metodi non gerarchici (descrit-

tivi e probabilistici). Mentre i metodi gerarchici sono applicabili a matrici con poche unità statistiche (decine), i metodi non gerarchici, invece, sono impiegati su matrici di grandi dimensioni anche fino a 100.000 unità statistiche e, quindi, sono adottati spesso dagli studiosi di data science. Spesso sono utilizzati fissando a priori, non un certo numero di cluster, ma una successione di essi, in modo tale da ottenere una successione di partizioni delle unità statistiche osservate, come accade in un metodo gerarchico. Tra questi ricordiamo il metodo delle k medie con il quale, una volta fissato il numero dei cluster, vengono calcolati i centri iniziali in base ad un criterio statistico ragionato.

In alternativa, l'analisi archetipica (AA) è un metodo di fattorizzazione della matrice relativamente nuovo che risulta maggiormente flessibile dei cluster. Cutler e Breiman introdussero per la prima volta l'analisi archetipica nel 1994 per studiare l'inquinamento atmosferico e la forma dei crani degli uomini. Viene, anche, usata nel potenziamento del segnale e estrazione delle caratteristiche delle sequenze di immagini IR; nell'identificazione di genotipi umani estremi e rappresentativi all'interno della popolazione. Mørup and Hansen (2012) hanno dimostrato che il modello AA è rilevante per l'estrazione delle caratteristiche e la riduzione della dimensionalità per una grande varietà di problemi di apprendimento automatico derivati dalla visione artificiale, neuroimaging, chimica, text mining e filtraggio collaborativo che portano a rappresentazioni altamente interpretabili delle dinamiche nei dati. Più recentemente ancora, Palumbo e Ragozini (2017) hanno utilizzato l'analisi archetipale per individuare delle categorie in grado di riassumere l'intero dataset in gruppi omogenei. Gli archetipi vengono poi impiegati per identificare i prototipi e definire una categorizzazione rispetto a ciascun prototipo.

### *3.3 Data mining ed analytics*

Le metodologie di *data mining* sono sempre più diffuse in quanto sono fondamentali per estrapolare utili informazioni da grandi banche dati. Nel caso delle nanoparticelle le tecniche utilizzate sono basate sulla somiglianza strutturale che può essere rilevata da diversi tipi di descrittori strutturali a partire dalle semplici caratteristiche geometriche fino alle più sofisticate caratteristiche topologiche, elettroniche e quantum chimiche. Le tecniche e le strategie applicate alle operazioni di data mining sono per larga parte automatizzate, consistendo in specifici software e algoritmi adatti al singolo scopo.

Ad oggi, oltre alla cluster analysis e ai metodi di analisi fattoriale sopra descritti, si utilizzano tecniche di *machine learning*, quali gli alberi decisionali (DTree), l'al–goritmo *k-nearest neighbour* (kNN), la regressione lineare multipla (MLR) e anche

tecniche più sofisticate come le reti neurali (ANNs) e le macchine a vettore di supporto (SVM, *support vector machines*). Queste ultime due tecniche possono produrre accurati strumenti predittivi di QSPR (relazioni quantitative di struttura-proprietà che esprimono quantitativamente l'attività biologica di un farmaco, per esempio, in funzione di determinate caratteristiche chimico-fisiche o strutturali della materia), mentre tecniche di regressione lineare multipla e gli alberi decisionali possono rilevare semplici correlazioni e possono guidare il disegno razionale dei nanomateriali.

Da annoverare sono anche gli algoritmi genetici che rappresentano un modello differente di ottimizzazione stocastica ispirato da principi evolutivi. Per esempio, Gong e Yang (2004) hanno proposto un nuovo framework basato sull'algoritmo genetico a quattro quadranti che può essere applicato per risolvere molti dei problemi di visione dei computer. Le soluzioni dei problemi relativi alle immagini si ottengono riducendo la struttura quad-trees dell'algoritmo genetico tradizionale proponendo la segmentazione dell'immagine, la visione stereoscopica e la stima del movimento.

## 4. Un utile modello statistico

In questo nuovo ambito la previsione dei modelli statistici è spesso inattendibile, in quanto può capitare che infranga le leggi fisiche nella fase di estrapolazione dei dati, anche se questi modelli possono fornire ipotesi accurate e vicine ai dati oggetto di osservazione.

La metodologia qui proposta è basata su un modello sequenziale proposto da Joseph e Melkote (2009).

Consideriamo Y l'output degli esperimenti reali e $x = (x_1, .., x_p)'$ le variabili. A causa dell'esistenza del rumore che indicheremo con $\varepsilon$ e considerando che Y è una variabile casuale, otteniamo quindi

$$Y = \mu(x) + \varepsilon. \tag{1}$$

Dove $\mu(x)$ è la media di Y ad un dato punto $x$ che corrisponde ad una combinazione di parametri sperimentali in un processo reale. L'output dell'esperimento è definito da $f(x)$ e rispettivamente da $(x_1, y_1), \dots, (x_n, y_n)$. Assumiamo inoltre $\varepsilon \sim N(0, \sigma^2)$.

Definiamo quale misura di inadeguatezza del modello

$$MI = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{\lambda}^E(x_i)\right)^2, \tag{2}$$

dove $\hat{\lambda}^E(x_i) = f(x_i)$.

Dobbiamo confrontare MI con $\sigma^2$. Se MI è molto vicina a $\sigma^2$ ($MI \approx \sigma^2$), il modello funziona correttamente. Altrimenti, e quindi se MI non è vicina a $\sigma^2$, dobbiamo affidarci ai modelli di previsione e stimare $\sigma^2$dalle replicazioni.

Supponiamo di avere $r$ replicazioni per $d$ valori distinti di $x$, quindi $n = rd$.

Consideriamo $y_{ij}$ la $jma$ replica nell' $imo$ gruppo.

Definiamo anche $s^2 = \sum_{i=1}^{d}\left(y_{ij} - \bar{y}\right)^2/(r-1)$ la varianza dell'$i$.mo gruppo, e $s^2 = \sum_{i=1}^{d} {s_i}^2/d$ la varianza cumulata. Possiamo stimare $\sigma^2$ come $\bar{\hat{\sigma}}^2 = s^2$.

Allora modifichiamo la posizione-scala della equazione (3):

$$\mu_{(x)} - f(x) = \beta_0 + \beta_1\left(f(x) - \bar{f}\right) + \varepsilon \tag{3}$$

Dove $\bar{f} = \sum_{i=1}^{n} f(x_i)/n$, assumiamo che $\beta_0$ e $\beta_1$ sono indipendenti. Quindi la previsione è data da $\mu^C(x) = f(x) + \beta_0 + \beta_1\left(f(x) - \bar{f}\right)$. E calcoliamo la misura dell'inadeguatezza del modello con l'equazione (4):

$$MI = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{\mu}^C(x_i)\right)^2, \tag{4}$$

dove $\hat{\mu}^C(x_i)$ è la stima di $\mu^C(x)$. Utilizziamo questo modello per uso previsionale se MI è vicina a $\sigma^2$, mentre continuiamo con ulteriori modifiche se non lo è.

Successivamente utilizziamo un modello funzionale (5) per ovviare all'inadeguatezza della modifica della costante, che risulta essere:

$$\mu(x) - \mu^C(x) = \delta(x; a). \tag{5}$$

Definiamo a questo punto $\mu^F(x)$ il predittore. Generalmente determiniamo la forma di $\delta(x; a)$ attraverso l'analisi residuale di un modello costante. Le modifiche funzionali possono variare la forma del nostro modello, quindi bisogna essere attenti. Un metodo semplice consiste nell'usare un modello lineare dato da $\delta(x; a) = \sum_{i=0}^{m} a_i u_i(x)$, mentre Kennedy e O'Hagan (2010) usano Gauss per descrivere $f(x)$ e $\delta(x; a)$.

## 5. Conclusioni

Le nanoscienze e le nanotecnologie rappresentano una nuova direzione nella ricerca e la data science non fa altro che essere complementare alle altre scienze. I big data non devono essere considerati esclusivamente dei grandi database, ma devono essere valutati per identificare e comprendere le correlazioni che sarebbero, invece, difficili da individuare in un dataset limitato.

La statistica, da sempre, offre il suo supporto alle scienze, dalle semplici tecniche descrittive a quelle più sofisticate, ma presenta limiti, come i valori anomali e i valori overfitting (rischio di sovradattamento durante il processo di apprendimento induttivo), importanti per le nanoscienze.

L'utilizzo di un modello sequenziale potrebbe dare un contributo nel comprendere il controllo e l'ottimizzazione del processo di sintesi dei nanomateriali.

## Riferimenti bibliografici

Barnard A. S.; Wilson, H. F. (2015). Optical emission of statistical distributions of silicon quantum qots. *J. Phys. Chem. C.*, 119: 7969-7977.

Cutler, A.; Breiman, L. (1994) Archetypal analysis. *Technometrics*. Taylor & Francis online, Vol.36, Issue 4: 338-347.

Gong, M.; Yang, H. (2004). Quadtree-based genetic algorithm and its applications to computer vision. *Pattern recognition*. Elsevier, Vol 37, Issue 8: 1723-1733.

Joseph, V. R.; Melkote S. N. (2009). Statistical adjustments to engineering models. *Journal of Quality Technology*, 41: 362-375.

Lai, L.; Barnard A. S. (2014). Tuning the electron transfer properties of entire nanodiamond ensembles. *J. Phys. Chem. C.*, 118: 30209-30215.

Mørup, M.; Hansen, L. K. (2012) Archetypal analysis for machine learning and data mining. *Neurocomputing.* Elsevier, Vol. 80: 54-63.

Palumbo, F.; Ragozini, G. (2017) Statistical categorization through archetypal analysis. Analisi archetipale per la definizione di categorie. *Statistics and data science: new challenges, new generations*. Proceedings of the Conference of the Italian Statistical Society. ISBN (online) 978-88-6453-521-0.

Sun, B.; Fernandez, M.; Barnard, A. S. (2016). Statistics, damned statistics and nanoscience – using data science to meet the challenge of nanomaterial complexity. *Nanoscale Horizons*. Royal Society of Chemistry, Vol. 1: 89-95. DOI: 10.1039/c5nh00126a.

## Sitografia

https://www.mgi.gov/

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018*
DIPARTIMENTO DI ECONOMIA E FINANZA
ISBN: 978-88-6629-048-3
pp. 155-178

# L'analisi delle biografie come chiave di lettura dell'evoluzione delle convivenze di primo ordine. Il caso italiano.

**Thaís García-Pereiro, Roberta Pace***

*Università degli Studi di Bari Aldo Moro*

**Riassunto:** Questo articolo analizza le dinamiche delle prime convivenze in Italia attraverso la loro evoluzione nel tempo. La metodologia longitudinale applicata ai dati dell'Indagine Multiscopo sulle Famiglie e i soggetti sociali del 2003 consente di studiare le biografie di unione e gli eventi ad esse collegati, indagando sulle variabili che influenzano le diverse transizioni del processo di formazione e di dissoluzione delle prime convivenze. I risultati delle curve di incidenza cumulata mostrano come le prime convivenze compiano rapidamente la transizione al matrimonio. Inoltre, i modelli di regressione di rischi competitivi delineano profili ben differenziati tra le donne che trasformano la loro prima convivenza in matrimonio e coloro che scelgono di interromperla.

**Keywords:** analisi delle biografie; prime convivenze; matrimonio; separazione.

## 1. Introduzione

La letteratura recente ha dimostrato che tra le donne italiane prevale il matrimonio sulla convivenza come scelta per la formazione della prima unione, indipendentemente della loro età (García-Pereiro et al., 2014; García-Pereiro e Pace, 2016). Anche le convivenze sono aumentate dagli anni '90 (Rosina e Fraboni, 2004; Di Giulio e Rosina, 2007; Gabrielli e Hoem, 2010), ma non hanno ancora raggiunto i livelli di altri paesi, soprattutto nel Nord Europa. Tuttavia, permangono differenze importanti tra le generazioni, essendoci tra le più giovani un maggior numero di conviventi rispetto alle generazioni precedenti.

* Autore corrispondente: roberta.pace@uniba.it

Queste differenze intergenerazionali, insieme alla strategia di ritardare la formazione della coppia, potrebbero essere concepite come una conseguenza del ritardo di altri eventi del corso della vita che precedono l'unione e che registrano rilevanti ritardi. Il periodo di vita dedicato all'educazione è stato prolungato e, di conseguenza, si è allungato l'ingresso nel mercato del lavoro, il che comporta anche un prolungamento della permanenza nella casa dei genitori.

In questo lavoro confrontiamo le dinamiche delle prime convivenze in Italia, analizzando la loro evoluzione nel tempo. La prospettiva longitudinale ci permette di studiare in dettaglio le biografie di unione e gli eventi ad esse collegati seguendo sia un'approssimazione temporale di calendario sia un'approssimazione basata sull'osservazione delle coorti di nascita. L'elemento guida delle analisi empiriche sviluppate in questo lavoro è stato il corso della vita. Questo concetto implica l'idea dell'evoluzione temporale che rende le traiettorie di unione e dissoluzione processi dinamici che cambiano nel tempo. Ne consegue la scelta di privilegiare l'utilizzo di dati e di strumenti metodologici di analisi che tengano conto della natura propria di queste traiettorie.

La prospettiva del corso della vita riflette le interazioni esistenti tra l'individuo e la società e permette di conoscere i contesti sociali e temporali della pianificazione della biografia dell'individuo e allo stesso tempo mostra le molteplici forme di adattamento ai cambiamenti temporali e spaziali nelle diverse sfere della vita pubblica e privata (Heinz e Krüger, 2001). Non si può parlare di una teoria unica del corso della vita (Macmillan, 2005; Elder, 1994), tuttavia, gli sviluppi teorici inquadrati in questa prospettiva condividono concetti chiave e principi di base, essenziali per l'apprendimento dell'oggetto di studio: ovvero le biografie. Gli elementi temporali della prospettiva del corso della vita sono concepiti attraverso la distinzione tra le traiettorie, che riproducono le transizioni; le transizioni, che seguono una sequenza di cambi di stato al loro interno; e attraverso il calendario e la durata di tutti gli eventi che contraddistinguono la vita di un individuo (traiettorie, transizioni e stati).

Le traiettorie e le transizioni sono concetti comuni nella letteratura e costituiscono la base di tutti gli studi sul corso della vita, definendolo nella sua interezza, sia nel lungo che nel breve termine. Per studiare la natura delle prime convivenze è possibile individuare tre transizioni: dalla convivenza a) al matrimonio, b) alla separazione o morte del partner e c) alla prosecuzione della convivenza. Ognuno di questi passaggi può essere interpretato come un diverso tipo di unione in base al significato specifico che gli si attribuisce. Certamente, ogni risultato ha un'implicazione nel campo delle politiche pubbliche e della pianificazione familiare e può essere interpretato come un segno di cambiamento sociale.

## 2. Aspetti teorici ed ipotesi empiriche sull'evoluzione delle convivenze

È in costante crescita la letteratura che analizza le caratteristiche della trasformazione delle unioni iniziate attraverso la convivenza (Wu e Balakrishnant, 1995; Smock and Manning, 1997; Duvander, 1999; Brown, 2000; Litcher et al., 2006; Moors e Bernhardt, 2009). Al fine di considerare contemporaneamente i conviventi che trasformano la loro prima unione in matrimonio e coloro che hanno deciso di porre fine alla loro relazione, separandosi, sono necessarie due prospettive teoriche diverse - ma tra loro correlate - che corrispondono al verificarsi degli eventi in osservazione: una collegata alla formazione delle unioni, in particolare, del matrimonio; e l'altra relativa alla dissoluzione dell'unione.

Di seguito viene presentato il supporto teorico per entrambi i risultati (matrimonio e separazione) e le ipotesi empiriche di questo articolo, sulla base della selezione delle variabili incluse nell'analisi delle prime convivenze in Italia.

Le dinamiche tipiche della formazione della prima unione in Italia hanno mostrato una chiara tendenza alla diminuzione del numero di matrimoni e al ritardo del suo inizio (Kiernan, 2000; Rosina e Fraboni, 2004). Ricerche precedenti sull'argomento mostrano uno schema caratterizzato da una minore incidenza del matrimonio tra coorti di nascita (De Sandre et al., 1997; Barbagli et al., 2003; García-Pereiro et al., 2014), che danno alle giovani generazioni una posizione innovativa che consolida nuovi modelli di formazione della coppia. Secondo Thornton e Young-De Marco (2001), i giovani hanno sempre meno inibizioni rispetto al rimanere single o rompere un'unione e attribuiscono maggiori costrizioni alla vita matrimoniale.

- *Ipotesi 1 (coorte):* Sulla base della diminuzione del numero di matrimoni e l'aumento delle convivenze e delle separazioni (García-Pereiro et al., 2014), si prevede che l'incidenza cumulata del matrimonio post-convivenza diminuisca, mentre lo scioglimento della prima convivenza sarà più frequente nelle coorti più giovani rispetto a quelle più anziane.

  Seguendo le premesse teoriche della Seconda Transizione Demografica, le trasformazioni della famiglia sono strettamente collegate ai cambiamenti di ideali (Van de Kaa, 1987; Lesthaeghe, 1995). In questo senso, l'importanza acquisita dall'autonomia individuale nella sfera religiosa porta ad una crescente secolarizzazione della vita privata e ad un maggiore scetticismo nei confronti delle istituzioni.

  Data la forza delle credenze e delle tradizioni religiose e l'importanza della Chiesa e la sua influenza sulle questioni familiari in Italia, gli individui con un'affiliazione religiosa dichiarata potrebbero avere un atteggiamento maggior-

mente positivo verso il matrimonio. Seguendo questa linea, numerosi studi hanno dimostrato che le persone che si auto dichiarano credenti hanno una minor probabilità di vivere una relazione da conviventi o trovarsi nello status di separati (De Sandre e Dalla Zuanna, 1999; De Rose e Rosina, 1999; Dalla Zuanna et al., 2005; Castiglioni, 2004; Caltabiano et al., 2005; Di Giulio e Rosina, 2007).

- *Ipotesi 2 (credenze religiose):* per quanto riguarda la religiosità, intesa in contrapposizione alla secolarizzazione, in Italia - Paese caratterizzato da una forte tradizione cattolica - la secolarizzazione auto-dichiarata delle donne intervistate potrebbe essere legata ad una maggior probabilità di separazione e a una diminuzione della volontà di sposarsi, cioè, una transizione convivenza-matrimonio meno frequente.

  Nei paesi dell'Europa meridionale la formazione delle unioni è caratterizzata da: un aumento dell'età media, dal posticipo dell'uscita di casa dei genitori e dall'alta corrispondenza tra quest'ultimo e l'ingresso nel matrimonio (Billari et al., 2000; Baizán et al., 2003; Ongaro, 2003; Moreno Mínguez, 2003; De Rose et al., 2008). Liefbroer e De Jong Gierveld (1995) suggeriscono che le persone che vivono con i loro genitori hanno maggiori probabilità di essere coinvolte in relazioni matrimoniali e, di conseguenza, lasciano la casa dei genitori unicamente per sposarsi. D'altra parte, coloro che vivono in modo indipendente acquisiscono un maggior grado di autonomia personale che gli consente di valutare in modo diverso la relazione costo-beneficio dell'essere in unione.

- *Ipotesi 3 (indipendenza residenziale):* in Italia, dove una significativa percentuale di giovani adulti lascia la casa dei genitori per formare la prima unione (per lo più il matrimonio), le donne che hanno sperimentato una qualche forma di autonomia, vivendo da sole prima di iniziare le proprie convivenze, avranno una probabilità più bassa di sposarsi e saranno più propense a sciogliere le loro prime convivenze.

  La maggior parte degli studi precedenti relativi alla transizione al matrimonio tra i conviventi ha sottolineato i fattori economici che sostengono la decisione di sposarsi. Alcune teorie indicano l'importanza dell'accumulo di risorse economiche nel facilitare i processi di formazione della coppia (McLanahan e Casper, 1995; Goldstein e Kenney, 2001; Sweeney, 2002), mentre altri sostengono l'idea che l'accumulazione di capitale umano femminile tenda a distogliere le donne dall'entrare in unione (Becker, 1981). Altri studiosi hanno evidenziato che tra conviventi solo le caratteristiche economiche degli uomini sono importanti per compiere la transizione verso il matrimonio, mentre quelle delle donne

guadagnano rilevanza quando si parla di separazione (Oppenheimer, 1994; Smock e Manning, 1997; Brown, 2000). Questa tendenza inizia a mostrare segni di cambiamento, di fatto l'analisi di Sassler e Goldscheider (2004) conferma che il legame tra reddito degli uomini e matrimonio si stia indebolendo a causa della crescente presenza delle donne nel mercato del lavoro, con stipendi sempre più comparabili a quelli dei loro partner.

L'esperienza lavorativa è un indicatore delle competenze già acquisite, mentre l'attaccamento e le future opportunità nel mercato del lavoro, insieme al livello di istruzione, descrivono il potenziale reddito di un individuo. Il lavoro delle donne, inoltre, può essere considerato un indicatore di accumulazione di capitale umano sul mercato del lavoro che, allo stesso tempo, riflette sia l'effetto del reddito sulle transizioni, sia i costi di opportunità associati alla nascita di un figlio (Cigno e Ermisch, 1989; Kravdal, 1994). Per quanto riguarda lo scioglimento della coppia, Vignoli e Ferro (2009) confermano che il grado di indipendenza economica delle donne è un fattore che gioca un ruolo essenziale per poter affrontare una separazione.

- *Ipotesi 4 (autonomia economica femminile):* in contesti economici e residenziali sfavorevoli, l'acquisizione di una relativa stabilità economica potrebbe costituire un prerequisito necessario per la formazione del matrimonio. Le donne occupate guadagnano indipendenza e hanno l'opportunità di accumulare più risorse economiche rispetto alle donne che non hanno partecipato al mercato del lavoro. Pertanto, contrariamente alla teoria dell'indipendenza economica delle donne, è prevista una associazione positiva tra l'esperienza lavorativa delle donne ed entrambi gli eventi analizzati (matrimonio e separazione).

  La letteratura ha ampiamente dimostrato come le variabili legate alla fecondità delle coppie conviventi siano associate in modo significativo sia alla separazione sia al matrimonio (Smock e Manning, 1997; Wu e Balakrishnan, 1995; Lichter et al., 2006; Moors e Bernhardt, 2009). In questo senso, la concezione o la nascita di un figlio/a aumenta drasticamente il rischio di entrare in unione. Nell'Europa mediterranea, la maggior parte delle nascite avviene all'interno del matrimonio, un'istituzione che normalmente implica un elevato grado di impegno e una situazione associata ad una maggiore stabilità (Kiernan, 1999, 2001). In Italia, nonostante la crescente proporzione di nascite al di fuori del matrimonio, la prevalenza dei modelli istituzionali di nuzialità e di fecondità rimane fortemente legata alla coppia sposata, considerando la famiglia unita in matrimonio come l'ambiente più adatto per crescere e allevare i figli (Marí-Klose e Nos Colom, 1999; Barbagli et al., 2003; Baizan et al., 2003).

- *Ipotesi 5 (fecondità nella convivenza):* per quanto riguarda la nascita di un figlio/a, ci si aspetta un alto grado di interrelazione tra il concepimento e la transizione verso il matrimonio. Le donne che rimangono incinte o che hanno un figlio potrebbero trasformare la convivenza in matrimonio, pensando di offrire ai propri figli maggiore protezione sociale, emotiva ed economica, rispettando anche le pressioni normative e sociali. La stessa associazione è prevista per lo scioglimento della convivenza: la nascita di un figlio/a può dissuadere le coppie dal porre fine alla loro relazione. Infine, se la presenza dei figli era precedente alla formazione dell'unione, l'incentivo a sposarsi sarà probabilmente minore, riducendo la necessità di convertire la convivenza in matrimonio.

  Diversi autori hanno stabilito che l'età alla convivenza è correlata ai processi di transizione in varie forme (Brown, 2000; Smock e Manning, 1997; Wu e Balakrishnan, 1995), ed è per questo motivo che spesso si osservano differenze significative riguardo al rispetto delle norme sociali che portano a legittimare l'unione consensuale attraverso il matrimonio.

- *Ipotesi 6 (età all'inizio della convivenza):* un aumento dell'età alla prima convivenza sarà associato ad una contrazione della transizione al matrimonio. La pressione normativa e sociale per legittimare l'unione, infatti, varierà con l'età in cui la donna ha cominciato a convivere, essendo più alta in età più giovane. In linea con gli studi sulla dissoluzione del matrimonio, ci si aspetta che le donne che hanno iniziato le loro prime convivenze in giovane età abbiano una maggiore probabilità di scioglimento della coppia rispetto alle donne che hanno iniziato a convivere in età più avanzata (Wu e Balakrishnan, 1995; Brown, 2000).

  È stato dimostrato che il divorzio dei genitori influenza lo sviluppo del corso della vita dei propri figli. La letteratura utilizza tale assunto per spiegare le differenze relative ad attitudini e comportamenti circa la vita familiare delle persone che hanno vissuto l'esperienza del divorzio dei genitori rispetto a coloro che provengono da famiglie che non hanno vissuto i traumi della separazione. Le persone che hanno sperimentato il divorzio dei genitori sono più propense, infatti, a sciogliere le loro unioni (Teachman, 2002), perché non credono nel matrimonio come istituzione durevole e stabile, e tendono ad adottare più frequentemente comportamenti familiari meno tradizionali (Axinn e Thornton, 1996; Amato e De Boer, 2001; Ongaro e Mazzucco, 2009).

- *Ipotesi 7 (divorzio dei genitori):* La separazione dei genitori ha una discreta influenza sugli atteggiamenti e sui comportamenti relativi alla trasformazione della prima convivenza. Si ipotizza, infatti, che le donne con genitori divorziati ab-

biano sviluppato diversi livelli di tolleranza verso la vita coniugale: in primo luogo, credendo meno nella transizione al matrimonio e approvando la convivenza, e poi, essendo più propense a porre fine alla convivenza.

## 3. Dati e metodologia

I dati delle Indagini sulla fecondità e sulle famiglie (Fertility and Family Surveys FFS) consentono di studiare le coppie da una prospettiva longitudinale, indagando sulle variabili che influenzano le diverse transizioni all'interno del processo di formazione e di dissoluzione delle unioni. Questo tipo di fonte ci permette di approfondire i recenti cambiamenti demografici e, con le debite precauzioni, ci consente di contestualizzarli nel quadro della seconda transizione demografica e dell'analisi comparativa transazionale.

### *3.1 Indagini Multiscopo sulle Famiglie*

Il sistema di indagini multiscopo sulle famiglie comprende una serie di indagini svolte seguendo un disegno unico, il cui obiettivo è quello di indagare su vari aspetti della vita sociale. Le prime indagini sono state condotte negli anni Ottanta in risposta ad una crescente domanda di dati statistici nel campo sociale e per approfondire la conoscenza sulle trasformazioni in atto delle condizioni di vita della popolazione in Italia e usarli come base empirica per la formulazione di politiche pubbliche. Nel 1980 è stata effettuata la prima indagine sulle condizioni di salute della popolazione e l'assistenza ai servizi sanitari, due anni più tardi è stata condotta l'indagine sulle vacanze, i viaggi e gli sport e nel 1983 l'indagine sulla struttura e comportamenti familiari. Quest'ultima, insieme alla prima indagine sulla famiglia condotta nel periodo 1987-1991, faceva parte di un ciclo di interviste condotte ogni sei mesi con un sistema modulare per area tematica. Ogni ciclo comprendeva un campione di 23.000 famiglie su 123 comuni[†].

Il nuovo sistema di indagini multiscopo, avviato negli anni Novanta (1993), parte dall'esperienza maturata negli anni Ottanta e rappresenta un importante progresso informativo e metodologico. In primo luogo, la frammentazione dei diversi que-

[†] La sequenza dei sei cicli era la seguente: criminalità, incidenti domestici, disabilità e cultura; disabilità, attività sportiva, uso del tempo, storia delle unioni e della fecondità (donne) e cambio di residenza; disabilità, uso del tempo, viaggi, scuola e condizioni infantili; viaggi brevi, condizioni della popolazione anziana, disabilità, reti di relazioni familiari e interpersonali, assistenza ai servizi sanitari, malattie croniche; uso di servizi sanitari, malattie croniche, uso di droghe, medicina alternativa, peso e altezza, consumo di cibo; disabilità, consumo di cibo, eventi del ciclo di vita individuale e familiare, lavoro ed economia familiare.

stionari viene trasmessa al gruppo in una serie di unità tematiche ciascuna delle quali corrisponde ad un determinato sondaggio. Questo sistema è articolato in sette sondaggi: aspetti della vita quotidiana (annuale), viaggi e vacanze (trimestrali) e cinque sondaggi tematici che ruotano ogni anno, su: condizioni di salute e uso dei servizi sanitari; cittadinanza e tempo libero; sicurezza dei cittadini; uso del tempo; famiglia e soggetti sociali. A questi vengono aggiunti specifici moduli di approfondimento inseriti per coprire esigenze mirate di informazione. Un esempio è il modulo sugli aspetti del lavoro da una prospettiva di genere introdotto, in un momento successivo, nel sondaggio dedicato ai temi familiari e sociali. Questo sondaggio garantisce informazioni statistiche sulla formazione di nuove forme familiari come convivenze e famiglie ricostituite.

L'Indagine Multiscopo sulle famiglie e soggetti sociali è stata eseguita per la prima volta nel giugno del 1998 (Famiglia, soggetti sociali e condizioni dell'infanzia) ed era la principale fonte di statistiche sulla struttura familiare e sulle caratteristiche della famiglia in Italia. Le aree incluse nel sondaggio del 2003, permettono di ottenere informazioni su diversi argomenti legati alla vita quotidiana e alla storia degli individui[‡]. Alcune di queste informazioni sono state raccolte attraverso interviste dirette e se l'individuo non era disponibile al momento della rilevazione, le domande venivano fornite da un altro componente della famiglia. Il resto delle informazioni, invece, è stato raccolto attraverso questionari auto-compilati.

Fin dalla sua istituzione, l'indagine ha utilizzato una particolare strategia di campionamento. La popolazione di interesse (le unità statistiche) è composta dalle famiglie residenti in Italia e dagli individui che le compongono. Il periodo di riferimento considera gli ultimi dodici mesi prima dell'intervista, tuttavia, in alcune domande il riferimento temporale è il momento dell'intervista o della complicazione del questionario.

Così intesa, l'indagine permette di ottenere delle stime su diversi livelli territoriali: tutto il territorio nazionale, le cinque ripartizioni geografiche (Nord-Ovest, Nord-Est, Centro, Sud e Isole), le venti regioni (ad eccezione del Trentino-Alto Adige), e sei aree identificate in base alla tipologia sociodemografica dei comuni. Quest'ultimo livello è suddiviso nelle seguenti categorie: comuni nell'area metropolitana, vicino alle aree metropolitane ed i comuni al di fuori dell'area metropolitana. Quest'ultima è suddivisa a sua volta in quattro categorie: comuni fino a 2.000 abi-

[‡] Il corso della vita individuale e familiare, i rapporti familiari, le reti interpersonali, il supporto dato ai membri della famiglia, le reti di sostegno sociale, i genitori, il partner e il primo matrimonio, l'indipendenza residenziale di giovani adulti, la fecondità, la mobilità sociale, l'economia familiare e il lavoro domestico, i costumi e le tradizioni, la storia del lavoro e le caratteristiche dell'unità abitativa.

tanti, i comuni tra 2000-10000 abitanti, tra 10.001 e 50.0000, e i comuni con più di 50.000 abitanti.

Il disegno di campionamento è complesso e risponde a due diversi schemi. Il primo è applicato ai comuni auto rappresentativi che sono caratterizzati da una grande dimensione geografica. Tra questi, le unità di campionamento primarie sono rappresentate dalle famiglie che vengono estratte in modo casuale e sistematico dalle liste dei registri demografici dei comuni. Di ciascuna famiglia inclusa nella rilevazione - estratta senza re-immissione e con uguale probabilità - vengono indagate le caratteristiche di tutti i suoi componenti. Il secondo schema comprende il gruppo di comuni non rappresentativi, costituiti dai comuni non inclusi nel primo. All'interno di questo schema, una strategia stratificata viene adottata per le unità primarie in due fasi, stratificando i comuni. Le unità secondarie sono le famiglie, da cui vengono raccolte le informazioni fornite nel questionario. L'obiettivo di questa strategia di stratificazione è la riduzione dell'errore di campionamento, mediante la formazione di gruppi di unità statistiche caratterizzati da una grande omogeneità all'interno degli strati e una massima eterogeneità tra di essi.

Per definire la numerosità campionaria è stata utilizzata una strategia mista che risponde, sia a criteri organizzativi e di costo, sia a criteri statistici che riguardano il calcolo dell'errore di campionamento a diversi livelli di disaggregazione territoriale. Nello specifico: il campione nazionale non deve superare le 20.000 famiglie, per motivi legati al costo e all'organizzazione; e il numero di comuni interessati non dovrebbe essere maggiore di 900, al fine di controllare la rilevazione. La localizzazione del campione di famiglie e di comuni nel territorio nazionale è stata stabilita in modo tale da garantire la rappresentatività statistica ai diversi livelli geografici sopra descritti.

All'interno del Sistema di Indagine Multiscopo proposto dall'ISTAT, l'Indagine sulle Famiglie e i Soggetti Sociali del 2003, utilizza tre diversi questionari: uno generale che contiene le schede dei componenti e il questionario sulla famiglia; un questionario per individui di età superiore ai 18 anni; e uno per individui sotto i 18 anni di età. Il questionario generale include le caratteristiche sociodemografiche dei componenti della famiglia (età, sesso, livello di istruzione, professione, ecc.) così come il rapporto di parentela con il capofamiglia[§].

[§] Successivamente, le domande specifiche sono suddivise nelle seguenti sezioni: mobilità, limiti dell'autonomia familiare, fratelli e sorelle, figli e nipoti, genitori e nonni, altri parenti, amici e vicini di casa, volontariato, difficoltà economiche, corso di vita, convivenza, unioni prematrimoniali, unioni consensuali, economia familiare, matrimonio, alloggio, tra gli altri. Il questionario per coloro che hanno più di 18 anni contiene le sezioni sull'istruzione, il lavoro, il livello di istruzione e l'occupazione dei genitori, il lavoro domestico, l'indipendenza residenziale, l'organizzazione del tempo, la vita di

### *3.2 Analisi delle biografie e rischi competitivi*

Nella letteratura sono diversi gli studi sulle convivenze che hanno trattato il matrimonio come l'unica transizione possibile, applicando una metodologia focalizzata sulla transizione verso un singolo evento. Ma anche se il matrimonio può essere considerato il risultato più frequente in paesi in cui la diffusione della convivenza è ancora moderata, come l'Italia, condurre un'analisi che considera solo il matrimonio come via di evoluzione della prima unione è senza dubbio molto limitante. Il matrimonio non è l'unico risultato possibile della convivenza, infatti, le donne che iniziano un'unione consensuale possono anche decidere di porre fine all'unione o continuare la loro vita di coppia in convivenza.

Per questo motivo abbiamo scelto di analizzare i risultati delle prime convivenze in un quadro di rischi competitivi in cui più di un tipo di evento gioca un ruolo fondamentale nella transizione.

Quest'articolo affronta lo sviluppo della convivenza come prima unione in Italia. Il punto di partenza temporale è l'inizio della convivenza, considerando tre possibili eventi consequenziali: transizione al matrimonio, scioglimento della coppia e nessuna ulteriore trasformazione (continuazione della convivenza). Alle donne intervistate è stato chiesto sia l'inizio che la fine di ogni relazione che hanno vissuto fino al momento dell'intervista. Considerando solo le prime unioni che hanno avuto inizio come convivenza, sono state costruite una serie di variabili che indicano: l'anno in cui la convivenza è iniziata come il momento dell'evento origine, e l'anno in cui è finita la convivenza come l'anno in cui è stata: a) trasformata in matrimonio, b) interrotta o protratta (censurata); stabilendo per i casi censurati un limite di dieci anni dopo la formazione della convivenza. Inoltre, è stata costruita la variabile che contiene i rispettivi cambiamenti di stato.

In questa sede discutiamo la metodologia utilizzata sia per la transizione al matrimonio che per la separazione dopo la convivenza (trattati come eventi competitivi) poiché le considerazioni teoriche si concentrano sulla natura delle convivenze in base ai loro possibili risultati (Kiernan, 2002; Heuveline e Timberlake, 2004). Quindi la censura dei casi avviene dieci anni dopo la costituzione della convivenza, purché la coppia rimanga nello stato di convivenza. Questo metodo fornisce informazioni sulla tendenza della coppia a trasformare l'unione in matrimonio, in una situazione in cui potrebbero anche affrontare la dissoluzione dell'unione.

---

coppia, genitori e figli, e una parte speciale dedicata al single trai 18 e 34 anni che vivono ancora con almeno uno dei genitori. Il questionario dedicato ai minori comprende: scuola, relazioni familiari, giochi e televisione, tempo libero e amici, autonomia, lavoro occasionale e aiuto familiare.

Usando questo metodo una donna rimane nello stato di origine "convivente" fino al momento in cui avviene la transizione verso lo stato finale ovvero la "trasformazione della prima convivenza". Poiché vi sono due possibili obiettivi, vale a dire la transizione verso il matrimonio (matrimonio preceduto da convivenza) o la separazione, viene utilizzato un quadro metodologico di rischi competitivi. Nel momento in cui la prima convivenza si trasforma in matrimonio, la coppia non è più esposta al rischio di una separazione dopo la convivenza. Allo stesso modo, una volta interrotta la convivenza, la coppia non è più esposta al rischio di sposarsi dopo la prima convivenza. Quando né la separazione né il matrimonio avvengono, la biografia della convivenza viene censurata 10 anni dopo la formazione dell'unione.

Le transizioni analizzate sono:

1) la trasformazione della convivenza in matrimonio come evento di interesse e la separazione come evento competitivo;
2) lo scioglimento della prima convivenza come evento di interesse e il matrimonio come evento competitivo.

L'analisi dei rischi competitivi è stata condotta in due fasi: la prima descrive le funzioni d'incidenza cumulata dovute al matrimonio o alla separazione, mentre nella seconda vengono eseguiti i modelli di regressione semi-parametrica.

La prima parte dell'analisi è dedicata alle stime delle curve di incidenza cumulata per la transizione dalla prima convivenza al matrimonio e dalla convivenza alla dissoluzione per mostrare i cambiamenti sperimentati dalle prime convivenze in Italia. Il calcolo delle curve di incidenza cumulata in presenza di rischi competitivi, considera la funzione di occorrenza dei due eventi in competizione invece di considerare solo l'evento di interesse (Coviello e Boggess, 2004; Cleves et al., 2010).

In questo lavoro, l'attenzione si concentra sulla valutazione dell'impatto di alcune variabili sociodemografiche sulla trasformazione delle prime convivenze in matrimonio o sullo scioglimento della coppia, data l'esistenza di alternative. I rischi della sub-distribuzione all'interno di un quadro di eventi competitivi trattano questo problema in un modo molto più semplice, inoltre, questa metodologia è stata ampiamente utilizzata in ambito epidemiologico (Pintilie, 2007; Bakoyannis e Touloumi, 2011). In tali situazioni, i modelli di rischio competitivo con transizioni non ripetitive mutuamente esclusive costituiscono il metodo più appropriato per l'analisi delle biografie (Pintilie, 2007; Cleves et al, 2010).

Seguendo il metodo sviluppato da Fine and Gray (1999), le osservazioni degli eventi competitivi sono mantenute nel set di rischio con una ponderazione decrescente. In questo modo, il metodo fornisce un modello dei rischi di sub-distribuzione attraverso la stima della massima verosimiglianza che mostra le diffe-

renze attuali e reali tra gli eventi osservati in termini di rapporti di rischio di sub-distribuzione. Secondo gli autori, in presenza di eventi competitivi che impediscono il verificarsi dell'evento di interesse, un'analisi standard che utilizza la regressione di Cox calcola curve di incidenza che sono appropriate solo per un universo ipotetico in cui è possibile un solo tipo di evento. In presenza di rischi competitivi, è meglio concentrarsi sulla funzione di incidenza cumulativa (CIF) o sulla funzione di sub-distribuzione perché non è chiaro quale tipo di evento si verificherà fino a quando non si è verificato.

Nella seconda parte dell'analisi, l'obiettivo è misurare il tempo trascorso dalla prima convivenza al matrimonio o alla separazione in relazione a vari fattori sociodemografici.

Due modelli di rischi competitivi sono stati calcolati, il primo tratta il matrimonio come l'evento di interesse, mentre nel secondo l'attenzione è rivolta allo scioglimento della convivenza, in entrambi i casi l'evento opposto è considerato come l'evento competitivo. Usando il metodo di Fine e Gray, il rischio competitivo si riferisce alla possibilità che, al posto del matrimonio, venga osservato l'evento competitivo, in questo caso lo scioglimento dell'unione. Lo scioglimento della prima convivenza impedisce il verificarsi dell'evento di interesse (matrimonio preceduto dalla prima convivenza). Ciò non dovrebbe essere confuso con la censura, perché quando la prima convivenza non subisce trasformazioni, queste coppie si considerano ancora a rischio di trasformazione sia per il matrimonio che per la separazione. Considerando che la censura non consente l'osservazione del verificarsi dell'evento di interesse, un evento in competizione impedisce che si verifichi l'evento di interesse.

L’analisi dei rischi competitivi si basa su un modello di regressione semiparametrico che modella gli effetti di diverse variabili sulle funzioni di incidenza cumulata. A differenza della regressione di Cox, c'è un'ipotesi riguardo alla proporzionalità della sub-distribuzione (Cleves et al., 2010). Nei modelli stimati, l'ipotesi di proporzionalità viene mantenuta per tutte le variabili incluse nelle analisi.

Il modello raggruppa le variabili così descritte:

- Coorte (costante nel tempo): variabile con cinque categorie che raggruppa le donne in funzione alla loro data di nascita (nate prima del 1950, 1951-1960, 1961-1970, 1971-1980 e 1981-1990). L'ultima è una coorte incompleta e i modelli la escludono per evitare distorsioni nei risultati.
- Credenze religiose (costante nel tempo): una variabile fittizia in cui le categorie possono assumere solo due valori in base alla dichiarazione della situazione religiosa degli intervistati.

- Indipendenza residenziale (cambia nel tempo): misura se le donne hanno vissuto in modo indipendente (fuori dalla casa dei genitori) o meno almeno una volta prima di formare la prima unione attraverso la convivenza.
- Il più alto livello di istruzione raggiunto al momento dell'indagine (costante nel tempo): il record educativo è stato diviso in 3 categorie (1- istruzione primaria o meno, 2 - secondaria e 3 - studi universitari).
- Età alla prima convivenza (costante nel tempo): è una variabile categorica che indica se la prima convivenza è iniziata prima o dopo i 30 anni.
- Esperienza lavorativa (cambia nel tempo): è un indicatore dell'accumulo di capitale umano nel mercato del lavoro. È suddiviso nelle seguenti categorie: 0 - donne che non sono entrate nel mercato del lavoro e 1- per le donne occupate (con almeno una esperienza nel mercato di lavoro).
- La fecondità include la gravidanza e la nascita (cambia nel tempo): cattura sia il concepimento, sia la nascita di un figlio prima della formazione della convivenza e durante il suo sviluppo nel tempo. Il valore 0 viene assegnato alla donna senza figli; 1- alla donna che ha avuto un figlio/a prima di iniziare la convivenza (aveva già un figlio) e 2 - alla donna che ha avuto un figlio/a nato un anno prima, un anno dopo o l'anno in cui si è verificata la trasformazione dell'unione (gravidanza/parto).
- Divorzio dei genitori (costante nel tempo): misura se le donne provengono da famiglie separate o meno.

Dopo aver ripulito il dataset e aver escluso i casi con valori mancanti, il campione è pari a 1.410 prime convivenze a rischio di subire cambiamenti di stato nel tempo: la transizione al matrimonio o lo scioglimento della unione. Nello specifico, 777 prime convivenze sono state trasformate in matrimoni e 296 hanno subìto la separazione, totalizzando 1.073 eventi. Sono state stimate tre diverse specificazioni per ciascuno dei modelli costruiti in funzione all'evento di interesse (matrimonio o separazione): la prima include tutte le variabili descritte in precedenza, mentre la seconda e la terza escludono le variabili che nella prima non risultano statisticamente significative e si includono due indicatori per paragonare la loro bontà di adattamento (AIC e BIC).

## 4. Risultati

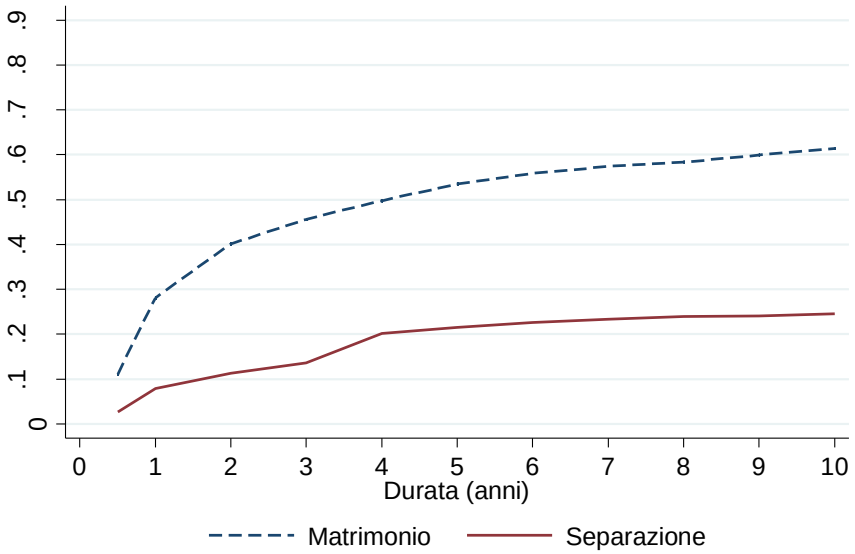
### *4.1 La trasformazione delle prime convivenze in Italia*

Nel quadro di analisi delle biografie in presenza di rischi competitivi, l'approccio più appropriato da seguire è la stima delle curve di incidenza cumulate, poiché con-

sentono di calcolare simultaneamente i rischi delle prime convivenze che vengono trasformate non solo nell'evento di interesse (matrimonio), ma anche nell'evento competitivo (dissoluzione della convivenza). Pertanto, questa curva è una funzione sia dell'incidenza e della durata del matrimonio sia della separazione e, di conseguenza, dipende dai cambiamenti prodotti in entrambi.

La Figura 1 mostra la curva di incidenza cumulata della transizione delle prime convivenze al matrimonio e alla separazione. Le rispettive curve mostrano che, 10 anni dopo la costituzione della convivenza, circa il 61% delle unioni è stata convertita in matrimonio, mentre il 24% delle prime convivenze hanno subìto la separazione e solo il 15% continua sotto la forma di prima convivenza.

**Figura 1.** *Incidenza cumulata della transizione al matrimonio e alla separazione (prime convivenze). Italia (n=1.410).*




*Fonte: elaborazioni proprie sui microdati dell'Indagine Multiscopo 2003.*

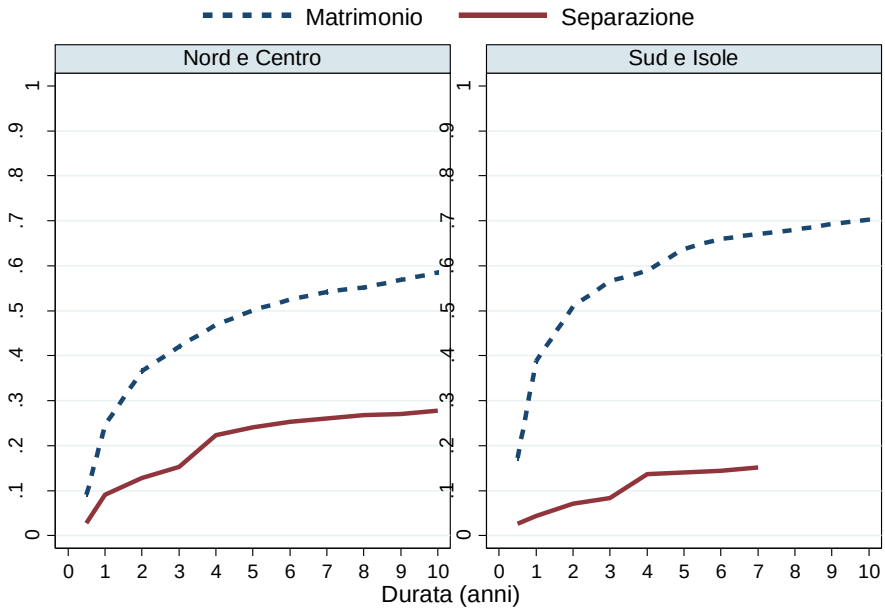
Risulta di gran importanza sottolineare che la transizione al matrimonio avviene molto più rapidamente della dissoluzione dell'unione. Infatti, già entro il secondo anno di vita il 40% delle convivenze hanno compiuto la transizione verso il matrimonio, cifra che raggiunge il 50% ancora prima di compiere il quarto anniversario.

All'interno del territorio italiano c'è una significativa eterogeneità che merita di essere evidenziata. La disaggregazione delle curve a livello di macro-ripartizioni (Figura 2) mostra differenze significative tra le zone del Centro-Nord e le regioni più tradizionali e meno sviluppate del Sud e delle Isole (De Rose e Vignoli, 2011; Castiglioni e Dalla Zuanna, 2009; Billari e Kohler, 2002).

Si osserva un chiaro schema che favorisce, da un lato, una maggiore incidenza della transizione al matrimonio nel Sud e Isole (70%) e, dall'altro, una maggiore incidenza di separazioni nel Centro-Nord (29%). Nonostante queste differenze, e come conseguenza di esse, la continuazione della convivenza è più omogenea tra le regioni. Infatti, la maggior eterogeneità è osservata negli eventi considerati come trasformazioni, non nella continuazione della prima convivenza.

È anche importante notare che la divergenza dell'incidenza delle convivenze seguite dal matrimonio è significativamente più bassa nel Nord e nel Centro. La riduzione della distanza tra le curve potrebbe indicare un carattere diverso della convivenza in queste aree che, sebbene favorisca ancora il matrimonio, conferisce alla dissoluzione dell'unione un ruolo importante.

**Figura 2.** *Incidenza cumulata della transizione al matrimonio e alla separazione (prime convivenze) per ripartizione geografica. Italia (n=1.410).*




*Fonte: elaborazioni proprie sui microdati dell'Indagine Multiscopo 2003.*

## *4.2 Sulle determinanti dell'evoluzione delle prime convivenze*

I risultati dei modelli di regressione di rischi competitivi sono riassunti nelle Tabelle 1 e 2. I parametri dell'effetto (Sub-Hazard Ratio SHR) indicano sia la forza dell'associazione tra le variabili (e le loro categorie) sia l'incidenza cumulata del matrimonio o della separazione rispetto al corrispondente evento competitivo. La categoria di riferimento delle variabili è indicata tra parentesi e, data la loro simili-

tudine, si commentano unicamente i coefficienti della prima specificazione di ciascuna tabella.

Riguardo alla trasformazione delle prime convivenze, l'incidenza cumulata della transizione al matrimonio diminuisce tra le coorti di nascita più giovani. Nonostante ciò, il valore del SHR legato al matrimonio risulta del 50% supriore tra le donne nate tra il 1961 e il 1970 in relazione alla categoria di riferimento (generazione <1951).

Contrariamente a quanto previsto, i risultati non indicano un'associazione negativa tra donne che hanno vissuto in modo indipendente (al di fuori della casa dei genitori) e l'incidenza cumulata della transizione al matrimonio.

**Tabella 1.** *Italia. Stima dei parametri dei modelli di regressione di rischi competitivi delle transizioni associate alle prime convivenze. Evento d'interesse: matrimonio. Evento competitivo: separazione.*

| | **Specificazione 1** | | **Specificazione 2** | | **Specificazione 3** | |
|---|---|---|---|---|---|---|
| Variabili indipendenti | SHR | Sign. | SHR | Sign. | SHR | Sign. |
| **Coorte di nascita** | | | | | | |
| (<1951) | | | | | | |
| 1951-1960 | 1,61 | ** | 1,62 | ** | 1,57 | ** |
| 1961-1970 | 1,50 | ** | 1,50 | ** | 1,46 | ** |
| 1971-1980 | 1,03 | | 1,03 | | 1,00 | |
| **Età alla convivenza 30+** | 0,56 | ** | 0,57 | ** | 0,55 | ** |
| **Non si dichiara religiosa** | 0,79 | ** | 0,79 | ** | 0,78 | ** |
| **Livello educativo** | | | | | | |
| (Primaria o minore) | | | | | | |
| Secondaria | 0,95 | | 0,95 | | - | - |
| Universitaria | 0,85 | | 0,85 | | - | - |
| **Indipendenza residenziale** | 1,03 | | - | - | - | - |
| **Occupata** | 0,78 | ** | 0,78 | ** | 0,76 | ** |
| **Fecondità** | | | | | | |
| (Senza figli) | | | | | | |
| Concepimento/nascita | 2,44 | ** | 2,43 | ** | 2,45 | ** |
| Figli anteriori alla convivenza | 0,58 | ** | 0,58 | ** | 0,59 | ** |
| **Famiglia intatta** | 1,52 | ** | 1,52 | ** | 1,50 | ** |
| n | 1.374 | | 1.374 | | 1.374 | |
| Evento d'interesse | 777 | | 777 | | 777 | |
| Log pseudolikelihood | -5167,7 | | -5167,8 | | -5168,4 | |
| AIC | 10359,4 | | 10357,5 | | 10354,8 | |
| BIC | 10422,1 | | 10415,0 | | 10401,9 | |

*** p≤0,001; *p≤0,05*

*Nota: Categorie di riferimento tra parentesi.*

*Fonte: elaborazioni proprie sui microdati dell'Indagine Multiscopo 2003.*

Le donne che non si sono auto-dichiarate come religiose hanno un'incidenza cumulata della transizione al matrimonio più bassa rispetto a quelle che si dichiarano religiose, confermando lo sfondo meno tradizionale delle donne che mantengono le convivenze nel tempo (come tipo di unione prediletta). L'impatto negativo della secolarizzazione sul SHR del matrimonio è pari al 21%.

Le donne che hanno formato le prime convivenze dopo il loro trentesimo compleanno mostrano minori rischi di transizione al matrimonio (44%) rispetto a coloro che hanno formato la convivenza prima del loro trentesimo compleanno.

Come ipotizzato, l'effetto della fecondità all'interno delle convivenze è significativo. Per esempio, in relazione alle donne che non hanno figli, il concepimento o la nascita di un figlio/a quasi triplica il rischio di trasformare la convivenza in matrimonio. Al contrario, se la presenza di bambini precede la formazione della convivenza, il rischio di sposarsi si riduce ad un 42%.

Le donne conviventi che sono riuscite a conseguire una certa esperienza nel mercato del lavoro mostrano una minore incidenza della transizione al matrimonio rispetto a coloro che non sono mai state occupate. L'SHR legato all'esperienza lavorativa è significativamente negativo (0,78).

Per quanto riguarda le caratteristiche della famiglia d'origine si osserva che le donne che non hanno vissuto il divorzio dei genitori sono più inclini a trasformare la loro convivenza in matrimonio, transizione che mostra un effetto positivo che raggiunge il 52%.

Per quanto riguarda lo scioglimento della prima coabitazione (Tabella 2), la sua incidenza risulta maggiore nell'ultima generazione osservata ma sempre minore rispetto alla generazione di riferimento (<1951).

Altre variabili che mostrano un effetto positivo sull'incidenza cumulata della separazione, sono l'orientamento verso la secolarizzazione e l'esperienza lavorativa accumulata nel mercato del lavoro.

Le donne che hanno iniziato la loro prima convivenza dopo i trent'anni hanno un maggior rischio di separazione rispetto a coloro che hanno iniziato in giovane età. Questo risultato indica che più giovane è l'età all'inizio dell'unione, minore sarà il rischio di dissoluzione.

In termini di fecondità, concepire e dare alla luce un figlio/a durante la prima convivenza diminuisce significativamente l'incidenza della separazione (77%). Anche la presenza di un figlio/a precedente alla convivenza esercita un effetto negativo sulla transizione alla separazione, ma la forza dell'associazione è inferiore rispetto alla categoria precedente (0,32).

**Tabella 2.** *Italia. Stima dei parametri dei modelli di regressione di rischi competitivi delle transizioni associate alle prime convivenze. Evento d'interesse: separazione. Evento competitivo: matrimonio.*

| | **Specificazione 1** | | **Specificazione 2** | | **Specificazione 3** | |
|---|---|---|---|---|---|---|
| Variabili indipendenti | SHR | Sign. | SHR | Sign. | SHR | Sign. |
| **Coorte di nascita** | | | | | | |
| (<1951) | | | | | | |
| 1951-1960 | 0,44 | ** | 0,44 | ** | 0,46 | ** |
| 1961-1970 | 0,49 | ** | 0,49 | ** | 0,51 | ** |
| 1971-1980 | 0,93 | | 0,93 | | 0,96 | |
| **Età alla convivenza 30+** | 1,65 | ** | 1,65 | ** | 1,69 | ** |
| **Non si dichiara religiosa** | 1,27 | * | 1,27 | * | 1,29 | * |
| **Livello educativo** | | | | | | |
| (Primaria o minore) | | | | | | |
| Secondaria | 1,06 | | 1,06 | | - | - |
| Universitaria | 1,29 | | 1,29 | | - | - |
| **Indipendenza residenziale** | 0,99 | | - | - | - | - |
| **Occupata** | 1,53 | * | 1,53 | * | 1,60 | * |
| **Fecondità** | | | | | | |
| (Senza figli) | | | | | | |
| Concepimento/nascita | 0,23 | ** | 0,23 | ** | 0,23 | ** |
| Figli anteriori alla convivenza | 0,32 | ** | 0,32 | ** | 0,31 | ** |
| **Famiglia intatta** | 0,49 | ** | 0,49 | ** | 0,49 | ** |
| n | 1.219 | | 1.219 | | 1.219 | |
| Evento d'interesse | 296 | | 296 | | 296 | |
| Log pseudolikelihood | -1985,6 | | -1985,6 | | -1986,5 | |
| AIC | 3995,2 | | 3993,2 | | 3991,1 | |
| BIC | 4057,9 | | 4050,7 | | 4038,1 | |

*** p≤0,001; *p≤0,05*
*Nota: Categorie di riferimento tra parentesi.*
*Fonte: elaborazioni proprie sui microdati dell'Indagine Multiscopo 2003.*

Tra le donne che hanno vissuto in una famiglia senza gli eventi del divorzio/separazione durante la loro infanzia, l'incidenza della separazione nelle prime convivenze si riduce del 51%.

## 5. Conclusioni

Oggi più che mai, la convivenza è un'alternativa sempre più comune al matrimonio come primo stadio della vita di coppia (Kiernan, 2001; Smock e Manning, 2004). Durante gli ultimi 50 anni la convivenza ha acquisito importanza (con diversi ritmi e livelli) in tutta Europa, rappresentando sempre più una valida alternativa all'unione di tipo matrimoniale.

Nell'ambito del corso di vita individuale, le dinamiche di formazione della coppia sono strettamente correlate con altre transizioni biografiche. Infatti, la formazione di una unione di tipo co-residenziale, sia la convivenza sia il matrimonio, è associata a diverse caratteristiche sociodemografiche proprie degli individui, delle coppie e dei loro genitori.

All'inizio del XXI° secolo, la centralità tipica data all'istituzione matrimoniale in Italia è diminuita significativamente. Allo stato attuale, la convivenza ha acquistato importanza nelle dinamiche di formazione della coppia e si è diffusa rapidamente, soprattutto tra i giovani adulti (Schröder, 2005; Di Giulio e Rosina, 2007; García-Pereiro et al., 2014; García-Pereiro e Pace, 2016).

L'analisi empirica qui presentata conferma la natura temporanea della prima convivenza in Italia, caratterizzata da una rapida trasformazione che porta al matrimonio. Effettivamente, dopo 4 anni insieme, il 50% delle convivenze ha sperimentato la transizione al matrimonio.

Le profonde differenze regionali sono in qualche modo simili alle trasformazioni demografiche storico-territoriali subite nel corso dei decenni, dove l'adozione di comportamenti innovativi si è prodotta prima nelle città del Centro-Nord.

Anche se il mainstream è la formalizzazione della convivenza attraverso la istituzione del matrimonio, queste difformità mostrano una non trascurabile incidenza cumulativa di convivenze che permangono tali. Questo risultato potrebbe essere interpretato come la nascita di un tipo di unione il cui significato risulta un po' più vicino a quello dell'istituzione del matrimonio perdendo, in qualche modo, il suo carattere transitorio.

Questa ricerca contribuisce anche alla comprensione delle dinamiche di transizione associate allo sviluppo della prima convivenza in Italia, individuando le differenze tra le donne conviventi che decidono di sposarsi e coloro che decidono di sciogliere le loro unioni. I risultati dei modelli di rischio competitivo hanno dimostrato come appartenere alle coorti più giovani, avere credenze secolari, aver costituito il matrimonio dopo i 30 anni, aver avuto esperienza nel mercato del lavoro ed aver avuto un figlio prima della formazione della convivenza, sono caratteristiche correlate negativamente con la transizione al matrimonio. Quest'ultima è positivamente influenzata dalla natura intatta della famiglia di origine e, soprattutto, dal concepimento o la nascita di un figlio. Inoltre, l'incidenza della separazione dopo una prima convivenza è più elevata tra le donne che avevano formato la loro unione dopo i 30 anni di età e direttamente coinvolte nel mercato del lavoro e senza figli.

I cambiamenti registrati nella prime convivenze con il passare degli anni servono a confermare la loro natura fragile e transitoria. Questa fase del processo di dif-

fusione della convivenza, in cui il matrimonio è il modello più praticato, descrive una transizione caratterizzata dalla formazione di una convivenza che sembra costituirsi come un test prematrimoniale o un preludio al matrimonio (Rindfuss e Vandenheuvel, 1990; Heuveline e Timberlake, 2005). La evoluzione delle prime convivenze è un processo di grande complessità che, tuttavia, sembra avere come obiettivo finale il matrimonio, specialmente dopo la nascita di un figlio.

Nell'analisi quantitativa qui proposta, le date, i tempi e la sequenza degli eventi, sono di grande rilevanza per determinare la complessità delle biografie di unione. Si ritiene in ogni caso che, sebbene le indagini cercano di registrare con precisione la data esatta del verificarsi degli eventi nella vita degli individui, sembra non rilevino ancora altre informazioni importanti sulla natura del processo (Antoine et al., 2006; Solsona, 2011). Nel caso specifico della formazione della coppia la più importante limitazione per le analisi è l'impossibilità di conoscere i processi per i quali si inizia (fidanzamento) o termina una relazione (crisi-scioglimento), la natura e la durata di questi processi, le decisioni e le motivazioni che hanno motivato i cambiamenti di stato (se fidanzati il passaggio a convivere o sposarsi, se conviventi la transizione al matrimonio o alla separazione), e come tali processi siano stati vissuti dagli stessi attori.

## Riferimenti bibliografici

Amato, P. R.; Deboer, D. (2001). The transmission of marital instability across generations: Relationship skills or commitment to marriage. *Journal of Marriage and Family, 63*: 1038–1051.

Antoine, P.; Adjamagbo, A.; Amétépé, F.; Béguy, D.; Binetou-Dial, F.; Ferrand, M.; Tichit, C. (2006). Unions et désunions: les histoires d'amour ne sont jamais simples! Groupe de réflexion su l'approche biographique, Etats flous et trajectoires complexes, INED. *Méthodes et savoirs*, 5: 61-86.

Axinn, W. G.; Thornton, A. (1996). The influence of parents' marital dissolution on children's attitudes toward family formation. *Demography*, 33: 66–81.

Baizán, P.; Aasve, A.; Billari, F. C. (2003). Cohabitation, marriage and first birth: The interrelationships of family formation events in Spain. *European Journal of Population*, 19: 147–169.

Barbagli, M.; Castiglioni, M.; Dalla Zuanna, G. (2003). *Fare famiglia in Italia. Un secolo di cambiamento*. Bologna: Il Mulino.

Billari, F. C.; Kohler, H. P. (2002). The impact of union formation dynamics on first births in West Germany and Italy: Are there signs of convergence? In E.

Klijzing e M. Corijn (Eds.), *Fertility and Partnership in Europe: Findings and Lessons from Comparative Research*, Volume 2. Geneva/New York: United Nations.

Billari, F.C.; Castiglioni, M.; Castro Martin, T.; Michielin, F.; Ongaro, F. (2000). *Household and union formation in a Mediterranean fashion: Italy and Spain*. Comunicazione presentata alla FFS Flagship conference, Bruselles, 29–31 Maggio.

Brown, S. (2000). Union Transitions among Cohabitors: The Significance of Relationship Assessments and Expectations. *Journal of Marriage and the Family*, 62(3): 833-846.

Caltabiano, M.; Dalla Zuanna, G.; Rosina, A. (2005). Interdependence between sexual debut and church attendance in Italy. *Demographic Research*, 14: 453–484.

Castiglioni, M. (2004). First sexual intercourse and contraceptive use in Italy. In G. Dalla Zuanna e C. Crisafulli (Eds.), *Sexual behaviour of Italian students* (pp. 19–39). Messina: Department of Statistics, University of Messina.

Castiglioni, M.; Dalla Zuanna, G. (2009). Marital and Reproductive Behavior in Italy After 1995: Bridging the Gap with Western Europe? *European Journal of Population*, 25: 1-26.

Cigno, A.; Ermisch, J. (1989). A microeconomic analysis of the timing of births. *European Economic Review*, 33, pp. 737-760.

Cleves, M.; Gutierrez, R.; Gould, W.; Marchenko, Y. (2010). *An introduction to survival analysis using Stata* (3rd ed.). Texas: Stata Press.

Coviello, V.; Boggess, M. (2004). Cumulative incidence estimation in the presence of competing risks. *Stata Journal*, *4*(2), 103–112.

Dalla Zuanna, G.; De Rose, A.; Racioppi, F. (2005). Low fertility and limited diffusion of modern contraception in Italy during the second half of the twentieth century. *Journal of Population Research*, *22*(1), 21–48.

De Rose, A.; Vignoli, D. (2011). Families all'italiana: 150 years of history. *Rivista Italiana di Economia, Demografia e Statistica*, Vol. LXV (2): 121-144.

De Rose, A.; Rosina, A. (1999). Scioglimento delle unioni. In P. De Sandre, A. Pinnelli e A. Santini (Eds.), *Nuzialità e fecondità in trasformazioni: Percorsi e fattori del cambiamento* (pp. 379–391). Bologna: Il Mulino.

De Sandre, I.; Dalla Zuanna, G. (1999). Correnti socioculturali e comportamento riproduttivo. In P. De Sandre, A. Pinnelli, & A. Santini (Eds.), *Nuzialità e fecondità in trasformazioni: Percorsi e fattori del cambiamento* (pp. 273–294). Bologna: Il Mulino.

De Sandre, P.; Ongaro, F.; Rettaroli, R.; Salvini, S. (1997). *Matrimonio e figli: Tra rinvio e rinuncia*. Bologna: Il Mulino.

Di Giulio, P.; Rosina, A. (2007). Intergenerational family ties and the diffusion of cohabitation in Italy. *Demographic Research*, 16: 441–468.

Di Giulio, P.; Rosina, A. (2007). Intergenerational family ties and the diffusion of cohabitation in Italy. *Demographic Research*, 16: 441-68.

Duvander, A. (1999). The transition from cohabitation to marriage: a longitudinal study of the propensity to marry in Sweden in the early 1990s. *Journal of Family Issues*, 20: 696–717.

Fine, J. P.; Gray, R. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, 94(446): 496–509.

Gabrielli, G.; Hoem, J. (2010). Italy's non-negligible cohabitational unions. *European Journal of Population*, 26: 33–46.

García-Pereiro, T.; Pace, R. (2016). Changes on first union dynamics: highlights of a comparative study between Italy and Spain. *Rivista Italiana di Economia, Demografia e Statistica*, 70(2): 129-140.

García-Pereiro, T.; Pace, R.; Didonna, M. G. (2014). Entering first union: the choice between cohabitation and marriage among women in Italy and Spain. *Journal of Population Research*, 31(1): 51-70.

Goldstein, J.R.; Kenney, C.T. (2001). Marriage Delayed or Marriage Forgone? New Cohort Forecasts of First Marriage for U.S. Women. *American Sociological Review*, 66(4): 506-519.

Heuveline, P.; Timberlake, J. M. (2005). The role of cohabitation in family formation: The United States in comparative perspective. *Journal of Marriage and the Family*, 67: 1214–1230.

Kiernan, C. (2002). The state of European unions: an analysis of partnership formation and dissolution. In Klijzing E. and M. Corijn (Eds.) *Fertility and Partnership in Europe: Findings and Lessons from Comparative Research* Vol. I (pp. 43-58). New York/Geneva: United Nations.

Kiernan, K. (1999). Cohabitation in Western Europe. *Population Trends, 96*, 25–32.

Kiernan, K. (2001). The rise of cohabitation and childbearing outside marriage in Western Europe. *International Journal of Law, Policy and Family, 15*(1), 1–21.

Kravdal, Ø. (1994). The Importance of Economic Activity, Economic Potential and Economic Resources for the Timing of First Birth in Norway. *Population Studies*, 48: 249-267.

Lesthaeghe, R. (1995). The Second Demographic Transition in Western European Countries: an interpretation. In K. Oppenheim Mason and A.-M. Jensen (Eds.) Gender and family change in industrialized countries (pp. 18-48). Oxford, Oxford University Press.

Lichter, D. L.; Qian, Z.; Mellot, L. M. (2006). Marriage or Dissolution? Union Transitions among Poor Cohabiting Women. *Demography*, 43: 223–40.

Liefbroer, A. C.; De Jong Gierveld, J. (1995). Standardization and individualization: The transition from youth to adulthood among cohorts born between 1903

and 1965. Population and family in the low countries 1994. V. d. B. J. C. and D. F. Dordrecht, Kluwer Academic Publishers: 57-80.

Macmillan, R. (2005). The structure of the life course: classic issues and current controversies. The structure of the life course: Standardized? Individualized? Differentiated? *Advances in life course research*, 9: 3-24.

Marí-Klose, M.; Colom, A. (1999). Itinerarios vitales: educación, trabajo y fecundidad de las mujeres. *Opiniones y Actitudes*, 27. Madrid: Centro de Investigaciones Sociológicas.

McLanahan, S. S.; Casper, L. M. (1995). Growing diversity and inequality in the American family. In R. Farley (ed.), *State of the Union: America in the 1990s*, vol.2. New York: Russell Sage Foundation.

Moors, G.; Bernhardt, E. (2009). Splitting Up or Getting Married? Competing Risk Analysis of Transitions Among Cohabiting Couples in Sweden. *Acta Sociologica*, Vol. 52, 3: 227–247.

Ongaro, F. (2003). Prima della scelta: La lunga transizione. Convegno 'La bassa fecondità tra costrizioni economiche e cambio di valori'. *Academia Nazionale dei Lincei*, 15–16 Maggio.

Ongaro, F.; Mazzuco, S. (2009). Parental separation and family formation in early adulthood: Evidence from Italy. *Advances in Life Course Research*, 14: 119–130.

Oppenheimer, V. K. (1994). Women's rising employment and the future of the family in industrial societies. *Population and development review*, 20:293-342.

Rindfuss, R.; VandenHeuvel, A. (1990). Cohabitation: precursor to marriage or an alternative to being single. *Population and Development Review*, 16(4): 703-726.

Rosina, A.; Fraboni, R. (2004). Is marriage losing its centrality in Italy? *Demographic Research*, 11: 149–172.

Sassler, S.; Goldscheider, F. (2004). Revisiting Jane Austen's theory of marriage timing: changes in union formation among American men in the late 20th century. *Journal of Family Issues*, 25:139-166.

Schröder, C. (2005). Cohabitation in Italy: Do parents matter? *Rostock: MPIDR working paper* WP-2005-030.

Smock, P. J.; Manning, W. D. (1997). Cohabiting partners' economic circum–stances and marriage. *Demography*, 34(3): 331-341.

Smock, P. J.; Manning, W. D. (2004). Living together unmarried in the U.S.: Demographic perspectives and implications for family policy. *Law & Policy*, 26(1): 87–117.

Solsona, M. (2011). Biografías de divorcio: Cambios de estado borrosos y trayectorias familiares complejas. *Notas de población*, 93: 77-106.

Sweeney, M. M. (2002). Two decades of family change: The shifting economic foundations of marriage. *American Sociological Review*, 67: 132–147.

Teachman, J. D. (2002). Childhood living arrangements and the intergenerational transmission of divorce. *Journal of Marriage and Family*, 64: 717–729.

Thornton, A.; Young Demarco, L. (2001). Four decades of attitudes toward family issues in the United States: The 1960s through the 1990s. *Journal of Marriage and the Family*, 63(4): 1009–1037.

Van de Kaa, D. J. (1987). Europe's Second Demographic Transition. *Population Bulletin*, 42(1): 3-59.

Wu, Z., Balakrishnant, R. (1995). Dissolution of premarital cohabitation in Canada. *Demography*, 3(2): 521-532.

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018* ISBN: 978-88-6629-048-3
DIPARTIMENTO DI ECONOMIA E FINANZA pp. 179-207

# Su qualche equazione alle differenze

**Mauro Bisceglia***

*Università degli Studi di Bari Aldo Moro*

**Riassunto:**Il presente lavoro è stato pensato con l'intento di esporre alcune semplici basi, al fine di consentire, attraverso l'utilizzo delle equazioni alle differenze di ordine *k* lineare a coefficienti costanti, studi su questioni di carattere sia ***finanziario*** che ***economico***. Naturalmente, per una idonea trattazione dell'argomento, si è reso necessario affrontare alcune questioniprelimiari di base; pertanto si è partiti con lo studio, in modo semplice, ma allo stesso tempo con il dovuto rigore che necessita, delle equazioni alle differenze (ordinarie) di ordine *k*, con i problemi che ne derivano, *dei valori iniziali* e di *Cauchy*, superati con l'ausilio dei *Teoremi di esistenza ed unicità*. Inoltre sono state poste le basi per applicare la *regola di risoluzione* delle equazioni alle differenze lineari. E solo dopo aver definito e studito qualche semplice operatore, si sono potute trattare le equazioni alle differenze del primo ordine lineare, per completare con quelle di ordine *k* lineare a coefficienti costanti.

**Keywords:**equazioni alle differenze; risoluzione delle equazioni alle differenze ordinarie; teoremi di esistenza ed unicità; equazioni alle differenze di ordine *k*.

## 1. Introduzione

Cominciamo con il considerare il seguente problema.

Siano: $X$ una parte di $R^{k+1}$ , $f$ una funzione reale definita in $N \times X$. Determinare tutte le successioni reali $(x_n)_{n \in N}$ tali che $\forall n \in N$ si abbia:

$$f(n, x_n, x_{n+1}, ..., x_{n+k}) = 0 . \qquad [1.1]$$

Il problema qui formulato si chiama *equazione alle differenze* (ordinarie) di ordine *k*.

*Autore corrispondente:maurogianfranco.bisceglia@uniba.it.

Si dice che la [1.1] è un'equazione alle differenze di ordine *k* di tipo normale, se essa può scriversi sotto la forma:

$$x_{n+k} = g(n, x_n, x_{n+1}, ..., x_{n+k-1}) \qquad [1.2]$$

$g$ essendo una funzione reale.

Si dice che la [1.2] è un'equazione alle differenze lineare di ordine *k* se esistono $k+1$ successioni reali, $(a_{1,n})_{n\in N}, (a_{2,n})_{n\in N}, ....., (a_{k,n})_{n\in N}$ ed $(s_n)_{n\in N}$, tali che la [1.2] possa scriversi sotto la forma:

$$x_{n+k} = a_{1,n}x_{n+k-1} + a_{2,n}x_{n+k-2} + ..... + a_{k,n}x_n + s_n \qquad [1.3]$$

Le successioni $(a_{1,n})_{n\in N}, (a_{2,n})_{n\in N}, ....., (a_{k,n})_{n\in N}$ si chiamano *coefficienti* del–l'equazione lineare [1.3] e la successione $(s_n)_{n\in N}$ si chiama *termine noto* dell'equa–zione stessa.

Se per ogni numero intero positivo *n* è $s_n = 0$, allora l'equazione [1.3] si dice omogenea.

*DEF.1.1.* Ogni successione reale, $(x_n)_{n\in N}$, che per ogni numero intero positivo *n*, soddisfa alla [1.1] si dice una soluzione particolare dell'equazione [1.1] o, più semplicemente, una soluzione dell'equazione [1.1].

Risolvere l'equazione [1.1] significa trovare tutte le sue soluzioni particolari.

*DEF.1.2.* Siano: *Y* una parte di $R^k$ e *w* una funzione reale definita in $N \times Y$.

Si dice che *w* è una soluzione generale dell'equazione [1.1], se per ogni elemento $(c_1, c_2, ....., c_k)$ di Y la successione $(w(n, c_1, c_2, ....., c_k))_{n\in N}$ è una soluzione particolare dell'equazione [1.1] e se, $(x_n)_{n\in N}$ essendo una soluzione particolare dell'equazione [1.1], esiste un solo elemento, $(c_1, c_2, ....., c_k)$, di Y tale che risulti : $(w(n, c_1, c_2, ....., c_k))_{n\in N} = (x_n)_{n\in N}$.

*OSSERVAZIONE 1.3.* Si noti che la funzione *w* è una soluzione generale dell'equazione [1.1] se

$$(c_1, c_2, ....., c_k) \in Y(w(n, c_1, c_2, ....., c_k))_{n\in N}$$

è una funzione biunivoca di *Y* sull'insieme delle soluzioni particolari dell'equazione [1.1].

*OSSERVAZIONE.* 1.4. Si noti che l'equazione [1.1] è risolta se si riesce a determinare una sua soluzione generale.

Relativamente all'equazione [1.2] si pongono diversi problemi (Albano 1998) i più comuni dei quali sono i seguenti due:

**Problema dei valori iniziali** *1.5*. Sia $(1, \bar{x}_1, \bar{x}_2, ....., \bar{x}_k)$ un elemento dell'insie–me di definizione della funzione *g*.

Determinare, se esistono, tutte le soluzione particolari, $(x_n)_{n \in N}$, dell'equazione [1.2] per cui risulti: $x_1 = \bar{x}_1, x_2 = \bar{x}_2, ..... x_k = \bar{x}_k$.

Il problema ora formulato si denota col simbolo:

$$\begin{cases} x_{n+k} = g(n, x_n, x_{n+1}, ..., x_{n+k-1}) \\ x_1 = \bar{x}_1, x_2 = \bar{x}_2, ..... x_k = \bar{x}_k \end{cases}$$

**Problema di Cauchy** *1.6*. Sia $(m, \bar{x}_1, \bar{x}_2, ....., \bar{x}_k)$ un elemento dell'insieme di definizione della funzione *g*.

Determinare se esistono tutte le soluzioni particolari $(x_n)_{n \in N}$, dell'equazione [1.2] per cui risulti: $x_m = \bar{x}_1, x_{m+1} = \bar{x}_2, ..... x_{m+k-1} = \bar{x}_k$.

Il problema ora formulato si denota col simbolo:

$$\begin{cases} x_{n+k} = g(n, x_n, x_{n+1}, ..., x_{n+k-1}) \\ x_m = \bar{x}_1, x_{m+1} = \bar{x}_2, ..... x_{m+k-1} = \bar{x}_k \end{cases}$$

*OSSERVAZIONE 1.7.* Si noti che il problema dei valori iniziali è il caso particolare del problema di Cauchy quando in questo si prende $m$ uguale ad uno.

## 2. Teoremi di esistenza e unicità

Cominciamo col dimostrare la seguente proposizione:

*2.1*. Se $g$ è una funzione reale definita in $N \times R^k$ ed $(\bar{x}_1, \bar{x}_2, ....., \bar{x}_k)$ è un elemento di $R^k$, allora il problema dei valori iniziali 1.9 ammette una ed una sola soluzione.

***Dim.*** Infatti, per *n* uguale ad uno, si ha: $x_{1+k} = g(1, \bar{x}_1, \bar{x}_2, ....., \bar{x}_k)$; supposto quindi d'aver calcolato $x_{n+k}$, risulta: $x_{n+1+k} = g(n+1, x_{n+1}, x_{n+2}, ..., x_{n+k})$, quindi, per induzione, resta definita una successione $(x_n)_{n \in N}$, che risulta essere evidentemente l'unica soluzione del problema dei valori iniziali considerato nell'enunciato della proposizione 2.1.

Conviene mettere qui in evidenza la seguente altra proposizione:

2.2. Siano: *g* una funzione reale definita in $N \times R^k$, $(\bar{x}_1, \bar{x}_2, ....., \bar{x}_k)$ un elemento di $R^k$ ed *m* un elemento di *N*.

Se esiste una funzione reale *h* definita in $N \times R^k$ tale che, per ogni numero intero positivo *n*, l'equazione $x_{n+k} = g(n, x_n, x_{n+1}, ..., x_{n+k-1})$ sia equivalente alla seguente altra: $x_n = h(n, x_{n+1}, x_{n+2}, ..., x_{n+k})$, allora il problema di Cauchy 1.10 ammette una ed una sola soluzione.

***Dim.*** Infatti successivamente si ha $x_{m-1} = h(m-1, \bar{x}_1, \bar{x}_2, ....., \bar{x}_k)$, $x_{m-2} = h(m-2, x_{m-1}, \bar{x}_1, ....., \bar{x}_{k-1})$ ed $x_1 = h(1, x_2, x_3, ..., x_{k+1})$ e di qui, in virtù della proposizione 2.1, consegue la 2.2.

## 3. Equazioni alle differenze lineari

Se $(a_{1,,n})_{n\in N}, (a_{2,n})_{n\in N}, ..., (a_{k,n})_{n\in N}$ ed $(s_n)_{n\in N}$ sono $k+1$ successioni reali, insieme all'equazione alle differenze di ordine *k* lineare (Fedele 2002).:

$$x_{n+k} = a_{1,n}x_{n+k-1} + a_{2,n}x_{n+k-2} + ... + a_{k,n}x_n + s_n \qquad [3.1]$$

considereremo l'altra:

$$x_{n+k} = a_{1,n}x_{n+k-1} + a_{2,n}x_{n+k-2} + ... + a_{k,n}x_n \qquad [3.2]$$

che chiameremo l'equazione omogenea associata all'equazione [3.1].

In questo paragrafo dimostreremo che l'equazione [3.1] ammette una soluzione generale (Albano 1998).

Cominciamo con l'osservare che, indicato con *S* lo spazio vettoriale reale delle successioni di numeri reali e posto:

$$L:(x_n)_{n\in N} \in S \to (x_{n+k} - a_{1,n}, x_{n+k-1,} - a_{2,n}x_{n+k-2} - ... - a_{k,n}x_n)_{n\in N},$$

*L* è un'applicazione lineare di *S* in *S* e, se denotiamo con 0 l'elemento nullo di *S*, l'equazione [3.2] può essere scritta sotto la forma: $L:(x_n)_{n\in N} = 0$; conseguentemente l'insieme delle soluzioni dell'equazione [3.2] coincide con il nucleo di *L*, quindi è un sottospazio vettoriale di *S* che nel seguito denoteremo con il simbolo $S_{\circ,k}$.

*3.1.* Siano: $(x_{1,n})_{n\in N}, (x_{2,n})_{n\in N}, ..., (x_{h,n})_{n\in N}$ *h* elementi di $S_{\circ,k}$.

Condizione necessaria e sufficiente affinché gli elementi $(x_{1,n})_{n\in N}, (x_{2,n})_{n\in N}, ..., (x_{h,n})_{n\in N}$ di $S_{\circ,k}$ siano linearmente indipendenti è che la matrice

$$\begin{Vmatrix} x_{1,1} & x_{2,1} & \dots & x_{h,1} \\ x_{1,2} & x_{2,2} & \dots & x_{h,2} \\ \dots & \dots & \dots & \dots \\ x_{1,k} & x_{2,k} & \dots & x_{h,k} \end{Vmatrix}$$

abbia caratteristica uguale ad *h*.

***Dim.*** Infatti se $c_1, c_2, ..., c_h$ sono *h* numeri reali, allora $\sum_{i=1}^{h} c_i \left(x_{i,n}\right)_{n\in N}$ è ancora un elemento di $S_{\circ,k}$ che, in virtù della proposizione 2.1, coincide con zero, se e solo se $\left(c_1, c_2, ..., c_h\right)$ è una soluzione del seguente sistema:

$$\begin{cases} x_{1,1}y_1 + x_{2,1}y_2 + ... + x_{h,1}y_h = 0 \\ x_{1,2}y_1 + x_{2,2}y_2 + ... + x_{h,2}y_h = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ x_{1,k}y_1 + x_{2,k}y_2 + ... + x_{h,k}y_h = 0 \end{cases}$$

e di qui, in virtù della proposizione sulla condizione necessaria e sufficiente dei sistemi determinati, consegue quanto asserito.

Siamo ora in grado di dimostrare la seguente proposizione:

*3.2.* $S_{\circ,k}$ è uno spazio vettoriale di dimensione *k*.

***Dim.*** La proposizione 3.1 assicura che la dimensione di $S_{\circ,k}$ non è maggiore di *k*, il nostro asserto sarà dimostrato se facciamo vedere che esistono *k* elementi di $S_{\circ,k}$ linearmente indipendenti.

Per questo, per ogni elemento $(i, j)$ di $\{1,2,..., k\}^2$, $\bar{x}_{i,j}$ sia un numero reale tale che il determinante della matrice $\left\|\bar{x}_{i,j}\right\|$ sia diverso da zero. In virtù della proposizione 2.1, per ogni elemento *i* di $\{1,2,..., k\}$ il problema dei valori iniziali:

$$\begin{cases} L\left(\left(x_n\right)_{n\in N}\right) = 0 \\ x_1 = \bar{x}_{i,1},\ x_2 = \bar{x}_{i,2}, ..., x_k = \bar{x}_{i,k} \end{cases}$$

ammette una sola soluzione; diciamo $\left(x_{i,n}\right)_{n\in N}$ tale soluzione.

È ora evidente che gli elementi $\left(x_{1,n}\right)_{n\in N}, \left(x_{2,n}\right)_{n\in N}, ..., \left(x_{k,n}\right)_{n\in N}$ di $S_{\circ,k}$ sono linearmente indipendenti e quindi costituiscono una base di $S_{\circ,k}$ .

Dalla proposizione 3.2 ora dimostrata consegue la seguente altra:

*3.3.* L'equazione [3.2] ammette una soluzione generale.

***Dim.*** Infatti se $(b_{1,n})_{n\in N}, (b_{2,n})_{n\in N}, ..., (b_{k,n})_{n\in N}$ è una base di $S_{\circ,k}$ esistente in virtù della proposizione 3.2, l'applicazione:

$$(c_1, c_2, ..., c_k) \in R^k \to c_1(b_{1,n})_{n\in N} + c_2(b_{2,n})_{n\in N} + ... + c_k(b_{k,n})_{n\in N} \in S_{\circ,k}$$

è, come volevasi, una funzione di $R^k$ biunivoca su $S_{\circ,k}$.

È ora evidente che:

*3.4.* Se $(x_{1,n})_{n\in N}$ ed $(x_{2,n})_{n\in N}$ sono due soluzioni particolari dell'equazione 3.1, allora la successione $(x_{1,n})_{n\in N} - (x_{2,n})_{n\in N}$ è una soluzione particolare dell'equazione [3.2].

Siamo ora in grado di dimostrare quanto sopra annunciato:

*3.5.* L'equazione [3.1] ammette una soluzione generale.

***Dim.*** Infatti se $(b_{1,n})_{n\in N}, (b_{2,n})_{n\in N}, ..., (b_{k,n})_{n\in N}$ è una base di $S_{\circ,k}$, esistente in virtù della proposizione 3.4, e $(v_n)_{n\in N}$ è una soluzione particolare dell'equazione [3.1], esistente della in virtù della proposizione 2.1, servendosi della proposizione 3.4 è immediato dimostrare che la funzione

$$(c_1, c_2, ..., c_k) \in R^k \to c_1(b_{1,n})_{n\in N} + c_2(b_{2,n})_{n\in N} + ... + c_k(b_{k,n})_{n\in N} + (v_n)_{n\in N}$$

è, come volevasi, un'applicazione di $R^k$ biunivoca sull'insieme delle soluzioni dell'equazione [3.1].

Possiamo quindi enunciare la seguente

**Regola di risoluzione dell'equazione** [3.1].

Si risolve prima l'equazione [3.2] omogenea associata all'equazione [3.1], quindi si calcola una soluzione particolare dell'equazione [3.1], la soluzione generale dell'equazione [3.1] si ottiene sommando alla soluzione generale dell'equ–zione [3.2] una soluzione particolare dell'equazione [3.1].

Osserviamo ora che un'altra conseguenza della linearità della funzione *L* è la seguente altra proposizione:

*3.6.* Siano: $(s'_n)_{n\in N}$ ed $(s''_n)_{n\in N}$ due elementi di *S*.

Se $(v'_n)_{n\in N}$ è una soluzione particolare dell'equazione $L((x_n)_{n\in N}) = (s'_n)_{n\in N}$, e $(v''_n)_{n\in N}$ è una soluzione particolare dell'equazione $L((x_n)_{n\in N}) = (s''_n)_{n\in N}$, allora, per ogni coppia (*a*,*b*) di numeri reali, la successione $a(v'_n)_{n\in N} + b(v''_n)_{n\in N}$ è una soluzione particolare dell'equazione:

$$L((x_n)_{n\in N}) = a(s'_n)_{n\in N} + b(s''_n)_{n\in N}.$$

***Dim.*** Infatti essendo $L((v'_n)_{n\in N}) = (s'_n)_{n\in N}$ ed $L((v''_n)_{n\in N}) = (s''_n)_{n\in N}$ l'asserto consegue quando si osservi che è: $aL((v'_n)_{n\in N}) + bL((v''_n)_{n\in N}) = L(a(v'_n)_{n\in N}) + b((v''_n)_{n\in N}) = a(s'_n)_{n\in N} + b(s''_n)_{n\in N}$.

Dalla proposizione 2.2 consegue evidentemente la seguente altra:

*3.7.* Se per ogni numero intero positivo $n$ è $a_{k,n} \neq 0$, allora, per ogni elemento, $(m, \bar{x}_1, \bar{x}_2, ..., \bar{x}_k)$, di $N \times R^k$ il problema di Cauchy:

$$\begin{cases} x_{n+k} = a_{1,n}\, x_{n+k-1} + a_{2,n}\, x_{n+k-2} + ... + a_{k,n}\, x_n + s_n \\ x_m = \bar{x}_1,\ x_{m+1} = \bar{x}_2, ...,\ x_{m+k-1} = \bar{x}_k \end{cases}$$

ammette una ed una sola soluzione.

Chiudiamo questo paragrafo dimostrando la seguente proposizione:

*3.8.* Se per ogni numero intero positivo $n\, a_{k,n} \neq 0$ , allora condizione necessaria e sufficiente affinché gli elementi $(x_{1,n})_{n\in N}, (x_{2,n})_{n\in N}, ..., (x_{k,n})_{n\in N}$ di $S_{\circ,k}$ siano linearmente indipendenti è che, per ogni numero intero positivo $n$ il determinante di Casorati:

$$C(n) = \begin{vmatrix} x_{1,n} & x_{2,n} & \cdots & x_{h,n} \\ x_{1,n+1} & x_{2,n+1} & \cdots & x_{h,n+1} \\ \cdots & \cdots & \cdots & \cdots \\ x_{1,n+k-1} & x_{2,n+k-1} & \cdots & x_{h,n+k-1} \end{vmatrix}$$

sia diverso da zero.

***Dim.*** La sufficienza della condizione consegue dalla proposizione 3.3.

Dimostriamo quindi la necessarietà della condizione. Per questo osserviamo che non è difficile rendersi conto che, per ogni numero intero positivo $n$, è: $C(n+1) = (-1)^{k-1} a_{k,n} C(n)$ e di qui, ricordando che in virtù della proposizione 3.1 è $C(1) \neq 0$, supposto che sia $C(n) \neq 0$ consegue che $C(n+1)$ è diverso da zero. Il nostro asserto resta quindi dimostrato per induzione.

## 4. Su qualche operatore

In questo paragrafo definiremo e studieremo alcuni operatori che ci saranno utili per lo studio delle equazioni alle differenze che effettueremo nei paragrafi successivi.

Cominciamo col porre la seguente definizione:

DEF.4.1 Sia $(x_n)_{n\in N}$ una successione di numeri reali.

Per ogni numero intero positivo *n* il numero reale $x_{n+1} - x_n$ si chiama la differenza prima della successione $(x_n)_{n\in N}$ in *n* e si denota col simbolo $\Delta x_n$

La successione $(x_{n+1} - x_n)_{n\in N}$ si chiama la differenza prima della successione $(x_n)_{n\in N}$ e si denota col simbolo $\Delta(x_n)_{n\in N}$.

D'immediata verifica sono le seguenti tre proposizioni:

4.2 Siano : $(x_n)_{n\in N}$ ed $(y_n)_{n\in N}$ due elementi di *S*.

Se *a* e *b* sono due numeri reali, allora risulta: $\Delta(ax_n + by_n) = a\Delta x_n + b\Delta y_n$.

4.3. Se $(x_n)_{n\in N}$ ed $(y_n)_{n\in N}$ sono due elementi di *S* è: $\Delta x_n y_n = x_{n+1}\Delta y_n + y_n \Delta x_n = y_{n+1}\Delta x_n + x_n \Delta y_n$.

4.4 Se $(x_n)_{n\in N}$ ed $(y_n)_{n\in N}$ sono due elementi di *S*. e per ogni numero intero positivo *n* è $y_n \neq 0$, allora è: $\Delta \dfrac{x_n}{y_n} = \dfrac{y_n \Delta x_n - x_n \Delta y_n}{y_n y_{n+1}}$.

Poniamo ora le seguenti due definizioni(Albano 1998):

DEF. 4.5. Sia $(x_n)_{n\in N}$ una successione di numeri reali.

Per ogni numero intero positivo *n* la differenza prima della differenza prima della successione $(x_n)_{n\in N}$ in *n* si chiama la differenza seconda della successione $(x_n)_{n\in N}$ in *n* e si denota col simbolo $\Delta^2 x_n$.

La successione $(\Delta^2 x_n)_{n\in N} = (x_{n+2} - 2x_{n+1} + x_n)_{n\in N}$ si chiama la differenze seconda della successione $(x_n)_{n\in N}$ e si denota col simbolo $(\Delta^2 x_n)_{n\in N}$.

DEF. 4.6. Siano $(x_n)_{n\in N}$ una successione di numeri reali e *k* un numero intero maggiore di uno.

Supposto di aver definito la differenza $(k-1)$-esima della successione $(x_n)_{n\in N}$ in *n* si chiama la differenza *k*-esima della successione $(x_n)_{n\in N}$ in *n* e si denota col simbolo $\Delta^k x_n$.

La successione $\Delta^k (x_n)_{n\in N}$ si chiama la differenza *k*-esima della successione $(x_n)_{n\in N}$ e si denota col simbolo $\Delta^k (x_n)_{n\in N}$.

Se $(x_n)_{n\in N}$ è una successione di numeri reali e per ogni numero intero positivo *n* poniamo:

$$\Delta^1 x_n = \Delta x_n \quad \text{e} \quad \Delta^0 x_n = \Delta x_n$$

è immediato dimostrare che:

4.7. Se $h$ e $k$ sono due numeri interi maggiori o uguali a zero risulta:

$$\Delta^h\left(\Delta^k x_n\right)=\Delta^k\left(\Delta^h x_n\right)=\Delta^{h+k}x_n .$$

Facciamo ora vedere che :

4.8. Se $k$ è un numero intero maggiore o uguale a zero ed $\left(x_n\right)_{n\in N}$ è un elemento di $S$ è:

$$\Delta^k x_n=\sum_{i=0}^{k}(-1)^i\binom{k}{i}x_{n+k-i} .$$

**Dim.** Dimostreremo l'uguaglianza 4.9 per induzione. Per questo cominciamo con l'osservare che la 4.9 è evidentemente vera per $k=0$, supponiamo quindi la 4.9 vera per $k-1$e facciamo vedere che essa è vera anche per $k$.

A tale scopo osserviamo che risulta:

$$\Delta^k x_n=\Delta\Delta^{k-1}x_n=\Delta\sum_{i=0}^{k-1}(-1)^i\binom{k-1}{i}x_{n+k-1+i}=$$

$$=\sum_{i=0}^{k-1}(-1)^i\binom{k-1}{i}\left(x_{n+k-i}-x_{n+k-1-i}\right)=\sum_{i=0}^{k-1}(-1)^i\binom{k-1}{i}x_{n+k-i}+\sum_{i=0}^{k-1}(-1)^{i+1}\binom{k-1}{i}x_{n+k-1-i}=$$

$$=x_{n+k}\sum_{i=1}^{k-1}(-1)^i\left(\binom{k-1}{i}+\binom{k-1}{i-1}\right)x_{n+k-i}+(-1)^k x_n$$

e di qui l'asserto .

Con un ragionamento analogo a quello fatto per dimostrare la proposizione 4.8 si dimostra la seguente altra(Binmore2004).:

4.10 (UGUAGLIANZA DI LEIBNITZ). Se $k$ è un numero intero maggiore o uguale a zero e $\left(x_n\right)_{n\in N}$ e $\left(y_n\right)_{n\in N}$ sono due elementi di $S$ allora è:

$$\Delta^k x_n y_n=\sum_{i=0}^{k}\binom{k}{i}\Delta^i x_n\Delta^{k-i}y_{n+1}$$

Poniamo ora la seguente altra definizione:

DEF.4.11. Siano: $k$ un numero intero maggiore o uguale a zero e $p$ il polinomio reale:

$$r\in R\rightarrow a_0r^k+a_1r^{k-1}+...+a_k .$$

Se $\left(x_n\right)_{n\in N}$ è una successione di numeri reali, per ogni numero intero positivo $n$ porremo:

$$p(\Delta)x_n=a_0\Delta^k x_n+a_1\Delta^{k-1}x_n+...+a_k\Delta^0 x_n ,$$

conseguentemente porremo: $p(\Delta)\left(x_n\right)_{n\in N}=\left(p(\Delta)x_n\right)_{n\in N}$ e quindi col simbolo $p(\Delta)$ denoteremo la funzione:

$$(x_n)_{n\in N} \in S \to p(\Delta)(x_n)_{n\in N}.$$

D'immediata verifica è la seguente proposizione:

4.12. Se $p_1$ e $p_2$ sono due polinomi reali e poniamo $p = p_1 p_2$,allora è: $p(\Delta) = p_1(\Delta) \circ p_2(\Delta)$-

Nel seguito porremo:

$$p(\Delta) = p_1(\Delta) \circ p_2(\Delta) = p_2(\Delta) \circ p_1(\Delta) = p_1(\Delta) p_2(\Delta) = p_2(\Delta) p_1(\Delta)\text{-}$$

Introduciamo ora il seguente altro operatore:

DEF 4.13. Siano: *k* un numero intero maggiore o uguale a zero ed $(x_n)_{n\in N}$ una successione di numeri reali.

Per ogni numero intero positivo *n* il numero reale $x_{n+k}$ si chiama l'anticipo kappesimo della successione $(x_n)_{n\in N}$ in *n* e si denota col simbolo $E^k(x_n)_{n\in N}$.

Nel seguito porremo: $E^1 = E$.

D'immediata verifica è la seguente proposizione:

4.14. Siano: *k* un numero intero maggiore o uguale a zero ed $(x_n)_{n\in N}$ ed $(y_n)_{n\in N}$ due successioni di numeri reali.

Se $(a,b)$è una coppia di numeri reali, allora risulta:

$$E^k(ax_n + by_n) = aE^k x_n + bE^k y_n.$$

Inoltre è: $E^k x_n y_n = E^k x_n E^k y_n$.

In fine, se per ogni numero intero positivo *n* è $y \neq 0$,allora risulta:

$$E^k \frac{x_n}{y_n} = \frac{E^k x_n}{E^k y_n}.$$

Poniamo ora la seguente altra definizione:

DEF 4.15. Siano: *k* un numero intero maggiore o uguale a zero e il *p* il polinomio reale:

$$r \in R \to a_0 r^k + a_1 r^{k-1} + ... + a_k.$$

Se c è una successione di numeri reali, per ogni numero intero positivo *n* porremo:

$$p(E)x_n = a_0 E^k x_n + a_1 E^{k-1} x_n + ... + a_k E^0 x_n,$$

conseguentemente porremo: $p(E)(x_n)_{n\in N} = (p(E)x_n)_{n\in N}$ e quindi col simbolo $p(E)$denoteremo la funzione:

$$(x_n)_{n\in N} \in S \to p(E)(x_n)_{n\in N}.$$

D'immediata verifica è la seguente proposizione (Albano 1998):

4.16. Se $p_1$ e $p_2$ sono due polinomi reali e poniamo $p = p_1 p_2$,allora è: $p(E) = p_1(E) \circ p_2(E)$.

Nel seguito porremo:

$$p(E) = p_1(E) \circ p_2(E) = p_2(E) \circ p_1(E) = p_1(E) p_2(E) = p_2(E) p_1(E).$$

È anche immediato convincersi che sussistono le seguenti due uguaglianze:

$$\Delta = E - E^0 \text{ ed } E = \Delta + \Delta^0 .$$

Da queste uguaglianze e dalle proposizioni 4.16 e 4.12 consegue che :

4.17. Se *k* è un numero intero maggiore o uguale a zero allora risulta:

$$\Delta^k = \sum_{i=0}^{k} (-1)^i \binom{k}{i} E^{k-i} \quad \text{e} \quad E^k = \sum_{i=0}^{k} \binom{k}{i} \Delta^{k-i}$$

OSSERVAZIONE 4.18. Si noti che la prima delle uguaglianze scritte nell'enunciato della proposizione 4.17 coincide con l'uguaglianza 4.9.

Poniamo in fine la seguente definizione:

DEF 4.19. Sia $(x_n)_{n \in N}$ una successione di numeri reali.

Si dice successione primitiva di $(x_n)_{n \in N}$ o, più semplicemente primitiva di $(x_n)_{n \in N}$ ogni successione di numeri reali $(y_n)_{n \in N}$, tale che: $\Delta y_n = x_n$ per ogni numero intero positivo *n*.

Di immediata verifica sono le seguenti due proposizioni:

4.20. Siano: $(x_n)_{n \in N}$ una successione di numeri reali e *c* un numero reale. Se $(y_n)_{n \in N}$ è una primitiva di $(x_n)_{n \in N}$, allora la successione $(y_n + c)_{n \in N}$ è ancora una primitiva della successione $(x_n)_{n \in N}$.

4.21. Siano: $(x_n)_{n \in N}$ una successione di numeri reali. Se $(y'_n)_{n \in N}$ ed $(y''_n)_{n \in N}$ sono due primitive della successione $(x_n)_{n \in N}$, allora la successione $(y'_n - y''_n)_{n \in N}$ è costante.

Facciamo ora vedere che:

4.22. Ogni successione di numeri reali è dotata di primitiva.

**Dim.** Infatti se $(x_n)_{n \in N}$ è una successione di numeri reali è immediato rendersi conto che la successione $n \in N \to \begin{cases} 0 & \text{se } n = 1 \\ \sum_{i=1}^{n-1} x_i & \text{se } n > 1 \end{cases}$ è una primitiva della successione $(x_n)_{n \in N}$ .

Dalle proposizioni 4.22, 4.21 e 4.20 consegue infine la seguente altra:

4.23. Se $(x_n)_{n\in N}$ è una successione di numeri reali, $(s_n)_{n\in N}$ è una primitiva della successione $(x_n)_{n\in N}$, *m* è un numero intero positivo ed *a* un numero reale, allora la successione $(s_n + a - s_m)_{n\in N}$ è l'unica primitiva della successione $(x_n)_{n\in N}$ che in *m* assume il valore *a*.

Osservazione 4.24. Si noti che se $(x_n)_{n\in N}$ è una successione di numeri reali l'insieme delle primitive di $(x_n)_{n\in N}$ è, in virtù delle proposizioni 4.20 e 4.22, non vuoto ed infinito.

Quanto sopra osservato ci dice che $\Delta$ è una funzione di *S* su *S* non biunivoca.

Osserviamo infine che l'insieme delle primitive di $(x_n)_{n\in N}$ è l'immagine reciproca tramite $\Delta$ del sottoinsieme di *S* costituito dal solo elemento $(x_n)_{n\in N}$.

Nel seguito, se $(y_n)_{n\in N}$ è una primitiva di $(x_n)_{n\in N}$, porremo: $\Delta^{-1}(\{(x_n)_{n\in N}\}) = \Delta^{-1}x_n = y_n$.

Evidentemente si ha che:

4.25. Se $(x'_n)_{n\in N}$ ed $(x''_n)_{n\in N}$ sono due successioni di numeri reali ed *a* e *b* sono due numeri reali, allora risulta: $\Delta^{-1}ax'_n + bx''_n = a\Delta^{-1}x'_n + b\Delta^{-1}x''_n$.

## 5. L'equazione alle differenze del primo ordine lineare

Siano $(a_n)_{n\in N}$ ed $(s_n)_{n\in N}$ due successioni di numeri reali e consideriamo l'equazione alle differenze (Ritelli 2005):

$$x_{n+1} = a_n x_n + s_n \qquad [5.1]$$

In ossequio a quanto stabilito nella regola di risoluzione di cui sopra, per risolvere l'equazione [5.1] dobbiamo risolvere prima l'equazione:

$$x_{n+1} = a_n x_n \qquad [5.2]$$

omogenea associata all'equazione [5.1].

A tale fine osserviamo che è immediato rendersi conto, servendosi dell'uguaglianza [5.2] e del principio di induzione, che la successione:

$$\bar{x} : n \in N \to \begin{cases} 1 & \text{se } n = 1 \\ \prod_{i=1}^{n-1} a_i & \text{se } n > 1 \end{cases} \qquad [5.3]$$

è la soluzione del seguente problema dei valori iniziali: $\begin{cases} x_{n+1} = a_n x_n \\ x_1 = 1 \end{cases}$ e conseguentemente la funzione

$$c \in R \to c(\bar{x}_n)_{n \in N} \qquad [5.4]$$

è una soluzione generale dell'equazione [5.2] (Albano 1998).

Sempre in ossequio alla regola di risoluzione di cui sopra l'equazione [5.1] sarà risolta se ci calcoliamo una su soluzione particolare.

A tale fine osserviamo che, servendosi dell'uguaglianza [5.1] e del principio di induzione, non è difficile rendersi conto che la successione:

$$x^{\circ} : n \in N \to \begin{cases} 0 & \text{se } n = 1 \\ s_1 & \text{se } n = 2 \\ s_{n-1} + \sum_{i=1}^{n-2} s_i \prod_{j=i+1}^{n-1} a_j & \text{se } n > 2 \end{cases} \qquad [5.5]$$

è la soluzione del seguente problema dei valori iniziali: $\begin{cases} x_{n+1} = a_n x_n + s_n \\ x_1 = 0 \end{cases}$ e quindi la funzione

$$c \in R \to c(\bar{x}_n)_{n \in N} + (x^{\circ}{}_n)_{n \in N} \qquad [5.6]$$

è una soluzione generale dell'equazione [5.1].

## 6. L'equazione alle differenze di ordine *k* lineare a coefficienti costanti

Per risolvere l'equazione alle differenze di ordine *k* lineare si potrebbe ripetere un ragionamento analogo a quello fatto per risolvere l'equazione alle differenze del primo ordine lineare.

Noi qui non faremo un tale ragionamento che per altro risulta essere formalmente complicato, ma ci limiteremo a dare un procedimento formalmente semplice per risolvere l'equazione alle differenze di ordine *k* lineare ed a coefficienti costanti, procedimento che inoltre risulterà essere idoneo per lo studio di alcuni problemi che vedremo in seguito.

Per comodità cominceremo con il considerare le equazioni alle differenze del secondo ordine, siano quindi *a* e *b* due numeri reali ed $(s_n)_{n \in N}$ una successione di numeri reali e consideriamo la seguente equazione:

$$x_{n+1} = a x_{n+1} + b x_n + s_n \qquad [6.1]$$

In ossequio alla regola di risoluzione di cui sopra,risolviamo prima l'equazione:

$$x_{n+2} = a x_{n+1} + b x_n \qquad [6.2]$$

omogenea associata all'equazione [6.1].

Per questo, se consideriamo la seguente equazione algebrica:

$$t^2 - at - b = 0 \qquad [6.3]$$

che chiameremo equazione caratteristica dell'equazione [6.2], possiamo dimostrare la seguente proposizione:

6.1. Se l'equazione [6.3] ha una sola radice uguale a zero o, ciò che è lo stesso, se è $a = b = 0$,allora le successioni $(\delta_{1,n})_{n\in N}$ e $(\delta_{2,n})_{n\in N}$ sono due soluzioni particolari dell'equazione [6.2] linearmente indipendenti (Albano 1998).

Se l'equazione [6.3] ha una radice uguale a zero e l'altra diversa da zero o,ciò che è lo stesso, se è $b = 0$ ed $a \neq 0$,allora le successioni $(\delta_{1,n})_{n\in N}$ e $(a^n)_{n\in N}$ sono due soluzioni particolari dell'equazione [6.2],linearmente indipendenti.

Se l'equazione [6.3], ha due radici reali distinte e diverse da zero o,ciò che è lo stesso,se è $a \neq 0$ e $a^2 + 4b > 0$,allora, dette $u$ e $v$ le due radici dell'equazione [6.3] le successioni $(u^n)_{n\in N}$ ed $(v^n)_{n\in N}$ sono due soluzioni particolari dell'equazione [6.2] linearmente indipendenti.

Se l'equazione [6.3] ha una sola radice reale e diversa da zero o,ciò che è lo stesso, se è $a \neq 0$ e $a^2 + 4b = 0$,allora, detta $w$ la radice dell'equazione [6.3], le successioni $(w^n)_{n\in N}$ ed $(nw^n)_{n\in N}$ sono due soluzioni particolari dell'equazione [6.2] linearmente indipendenti.

Se l'equazione [6.3] non ha radici reali o, ciò che è lo stesso, se è $a^2 + 4b < 0$, allora, indicato con $\alpha$ l'elemento di $[0,2\pi[$ tale che risulti: $\cos\alpha = \frac{a}{2\sqrt{-b}}$ e $sen\alpha = \frac{\sqrt{-a^2 - 4b}}{2\sqrt{-b}}$, le successioni:

$$\left(\left(\sqrt{-b}\right)^n \cos\alpha\right)_{n\in N} \quad \text{e} \quad \left(\left(\sqrt{-b}\right)^n sen\alpha\right)_{n\in N}$$

sono due soluzioni particolari dell'equazione [6.2] linearmente indipendenti.

**Dim.** I primi quattro punti dell'enunciato della proposizione 6.1 sono di immediata verifica.

Verifichiamo quindi l'ultimo punto dell'enunciato della proposizione 6.1.

Per questo cominciamo con il fare vedere che le successioni ivi considerate sono due soluzioni particolari dell'equazione [6.2], infatti sostituendo nella [6.2], per la prima successione, risulta:

$$\left(\sqrt{-b}\right)^{n+2} \cos(n+2)\alpha - a\left(\sqrt{-b}\right)^{n+1} \cos(n+1)\alpha - b\left(\sqrt{-b}\right)^n \cos n\alpha =$$
$$= \left(\sqrt{-b}\right)^{n+1}\left(\sqrt{-b}\cos\alpha\cos 2\alpha - \sqrt{-b}sen\alpha sen 2\alpha - a\cos\alpha\cos\alpha +\right.$$

$$+ asen\alpha sen\alpha + \sqrt{-b}\cos\alpha\Big) = \left(\sqrt{-b}\right)^{n+1}\Big(\cos\alpha\Big(\sqrt{-b}\left(\cos^2\alpha - sen^2\alpha\right)\Big) -$$
$$+ a\cos\alpha + \sqrt{-b}\Big) + sen\alpha\Big(asen\alpha - \sqrt{-b}\,2sen\alpha\cos\alpha\Big)\Big),$$

quindi è:

$$\left(\sqrt{-b}\right)^{n+2}\cos(n+2)\alpha - a\left(\sqrt{-b}\right)^{n+1}\cos(n+1)\alpha - b\left(\sqrt{-b}\right)^{n}\cos n\alpha =$$
$$= \left(\sqrt{-b}\right)^{n+1}\cos(n+1)\alpha\left(2\sqrt{-b}\cos\alpha - a\right).$$

In modo analogo si fa vedere che, per la seconda successione, è:

$$\left(\sqrt{-b}\right)^{n+2} sen(n+2)\alpha - a\left(\sqrt{-b}\right)^{n+1} sen(n+1)\alpha - b\left(\sqrt{-b}\right)^{n} sen\alpha =$$
$$= \left(\sqrt{-b}\right)^{n+1} sen(n+1)\alpha\left(2\sqrt{-b}\cos\alpha - a\right),$$

da queste due ultime uguaglianze, ricordiamo che è $\cos\alpha = \dfrac{a}{2\sqrt{-b}}$, consegue quanto volevasi (Albano 1998).

Risultando inoltre: $\begin{vmatrix} \sqrt{-b}\cos\alpha & \sqrt{-b}sen\alpha \\ -b\cos 2\alpha & -bsen2\alpha \end{vmatrix} = -b\sqrt{-b}sen\alpha$, le due successioni considerate sono, come volevasi, linearmente indipendenti.

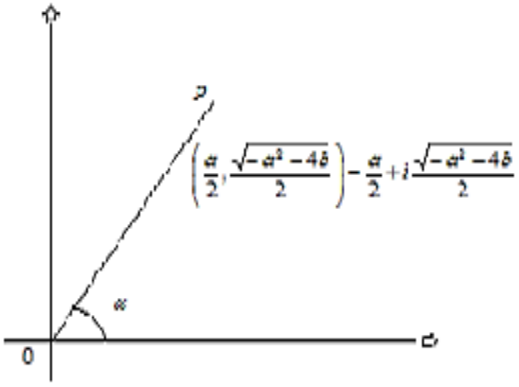
Osservazione 6.2. quando si ricordi la proposizione 3.1 e dopo quanto affermato nella proposizione 6.1, siamo ora in grado di scrivere una soluzione generale dell'equazione [6.2].

Osservazione 6.3. si noti che l'equazione [6.3] formalmente si ottiene imponendo che la successione $\left(t^n\right)_{n\in N}$ soddisfi l'equazione [6.2].

Osservazione 6.4. A completamento dell'osservazione 6.3 mettiamo qui in evidenza che, se si pone: $\sqrt{-I} = i$, le soluzioni dell'equazione [6.3] possono scriversi, nell'ipotesi che sia $a^2 + 4b < 0$ sotto la forma di numeri complessi nel seguente modo: $\dfrac{a}{2} + i\dfrac{\sqrt{-a^2-4b}}{2}$ e $\dfrac{a}{2} - i\dfrac{\sqrt{-a^2-4b}}{2}$, se quindi rappresentiamo sul piano il primo di tali numeri complessi, vedi la figura 1, risulta essere: $\overline{P0} = \sqrt{-b}$ ed $\alpha$ la misura in radianti dell'angolo che il segmento del piano avente per estremi i punti 0 e *P* forma con il semiasse delle ascisse positive, quindi è:

$$\frac{a}{2} + i\frac{\sqrt{-a^2-4b}}{2} = \sqrt{-b}\left(\cos\alpha + isen\alpha\right) \qquad [6.4].$$

**Figura 1.** *Rappresentazione dell'equazione [6.3] in numeri complessi*




In modo analogo si fa vedere che è:

$$\frac{a}{2} - i\frac{\sqrt{-a^2-4b}}{2} = \sqrt{-b}(\cos\alpha + isen\alpha) \qquad [6.5].$$

Facciamo ora vedere che per ogni numero intero positivo *n* è:

$$\left(\frac{a}{2} + i\frac{\sqrt{-a^2-4b}}{2}\right)^n = \left(\sqrt{-b}\right)^n(\cos\alpha + isen\alpha) \qquad [6.6].$$

Per questo osservando che l'uguaglianza [6.10] è vera per $n=1$, se quindi la supponiamo vera per un numero intero $h>1$, essendo:

$$(\cos\alpha + isen\alpha)^{h+1} = (\cosh\alpha + isenh\alpha)(\cos\alpha + isen\alpha) =$$
$$= (\cosh\alpha\cos\alpha - senh\alpha sen\alpha + i(senh\alpha\cos\alpha + \cosh\alpha sen\alpha)) =$$
$$= \cos(h+1)\alpha + isen(h+1)\alpha$$

essa è, come volevasi, vera, in virtù del principio di induzione, per ogni numero intero positivo *n* (Fedele2002).

In modo analogo si dimostra che per ogni numero intero positivo *n*, è:

$$\left(\frac{a}{2} - i\frac{\sqrt{-a^2-4b}}{2}\right)^n = \left(\sqrt{-b}\right)^n(\cos\alpha - isen\alpha) \qquad [6.7].$$

Le uguaglianze [6.6] e [6.7] ci permettono di asserire che, per ogni numero intero positivo *n*, risulta:

$$\left(\sqrt{-b}\right)^n\cos\alpha = \frac{1}{2}\left(\frac{a}{2} + i\frac{\sqrt{-a^2-4b}}{2}\right)^n + \frac{1}{2}\left(\frac{a}{2} - i\frac{\sqrt{-a^2-4b}}{2}\right)^n$$

$$\left(\sqrt{-b}\right)^n sen = \frac{1}{2i}\left(\frac{a}{2}+i\frac{\sqrt{-a^2-4b}}{2}\right)^n - \frac{1}{2i}\left(\frac{a}{2}-i\frac{\sqrt{-a^2-4b}}{2}\right)^n .$$

Per essere in grado di scrivere una soluzione generale dell'equazione [6.1] è sufficiente ora calcolarsi una sua soluzione particolare(Guerraggio 2004).

Per questo cominciamo con l'osservare che sussiste la seguente proposizione:

6.5. Siano: *m* un numero intero maggiore o uguale a zero, *q* un polinomio di grado *m*, *r* un numero reale diverso da zero ed $\alpha$ un elemento di $R-\bigcup_{h\in Z}\{h\pi\}$.

Se è $s_n = q(n)r^n$ per ogni numero intero positivo *n*, allora:

Se *r* non è radice dell'equazione [6.3], esiste un polinomio *p* di grado *m* tale che la successione $\left(p(n)r^n\right)_{n\in N}$ è una soluzione particolare dell'equazione [6.1].

Se *r* è radice semplice (risp. doppia) dell'equazione [6.3], esiste un polinomio *p* di grado *m* tale che la successione $\left(np(n)r^n\right)_{n\in N}$ (risp. $\left(n^2 p(n)r^n\right)_{n\in N}$) è una soluzione particolare dell'equazione [6.1].

Se è $s_n = q(n)r^n \cos\alpha$ oppure $s_n = q(n)r^n sen\alpha$ per ogni numero intero positivo *n*, allora:

Se, essendo *r* positivo (risp. negativo), il numero complesso $r(\cos\alpha + isen\alpha)$ (risp. $-r(\cos(\alpha+\pi)+isen(\alpha+\pi))$) non è radice dell'equazione [6.3], esistono due polinomi, $p_1$ e $p_2$, di grado *m* tali che la successione $\left(p_1(n)r^n \cos\alpha + p_2(n)r^n sen\alpha\right)_{n\in N}$ è una soluzione particolare dell'equazione [6.1].

Se, essendo *r* positivo (risp. negativo), il numero complesso $r(\cos\alpha + isen\alpha)$ (risp. $-r(\cos(\alpha+\pi)+isen(\alpha+\pi))$) è radice dell'equazione [6.3], esistono due polinomi, $p_1$ e $p_2$, di grado *m* tali che la successione $\left(np_1(n)r^n \cos\alpha + np_2(n)r^n sen\alpha\right)_{n\in N}$ è una soluzione particolare dell'equazione [6.1].

**Dim.** Siano: $a_0, a_1, ..., a_m$ $m+1$ numeri reali, *s* un numero intero maggiore o uguale a zero e *p* il polinomio $x \in R \to a_0 x^m + a_1 x^{m-1} + ... + a_m$ (Albano 1998).

Se si sostituisce nell'equazione [6.1] la successione $\left(n^s p(n)r^n\right)_{n\in N}$ si ottiene:

$$r^n \sum_{j=0}^{m} a_{m-j}\left(n^{j+s}\left(r^2 - ar - b\right) + r(j+s)n^{j+s-1}(2r-a) - \right.$$
$$\left. + r\sum_{i=0}^{j+s-2}\binom{j+s}{i} n^i \left(2^{j+s-i} r - a\right)\right) = q(n)r^n$$

e di qui in virtù del principio di identità dei polinomi, consegue il primo punto dell'enunciato della proposizione 6.5.

Siano ora $a'_0, a'_1, ..., a'_m$, $a''_0, a''_1, ..., a''_m$ $2(m+1)$ numeri reali, *s* un numero intero maggiore o uguale a zero e $p_1$ e $p_2$, rispettivamente, i polinomi $x \in R \to a'_0 x^m + a'_1 x^{m-1} + ... + a'_m$ e $x \in R \to a''_0 x^m + a''_1 x^{m-1} + ... + a''_m$.

Se si sostituisce nell'equazione [6.1] la successione $\left(n^s p_1(n) r^n \cos\alpha + n^s p_2(n) r^n sen\alpha\right)_{n\in N}$ in fine si ottiene:

$$\sum_{j=0}^{m}\Big(n^{j+s}\Big(a'_{m-j}\Big(r^{n+2}\cos(n+2)\alpha - ar^{n+1}\cos(n+1)\alpha - br^n\cos\alpha\Big)+$$
$$+a''_{m-j}\Big(r^{n+2}sen(n+2)\alpha - ar^{n+1}sen(n+1)\alpha - br^n sen\alpha\Big)\Big)+$$
$$+r^n\sum_{i=0}^{j+s-1}\binom{j+s}{i}n^i\Big(2^{j+s-i}r^2\Big(a'_{m-j}\cos 2\alpha + a''_{m-j}sen2\alpha\Big)-$$
$$+ar\Big(a'_{m-j}\cos\alpha + a''_{m-j}sen\alpha\Big)\Big)\cos\alpha + 2^{j+s-i}r^2\Big(a''_{m-j}\cos 2\alpha - a'_{m-j}sen2\alpha\Big)+$$
$$+ar\Big(a''_{m-j}\cos\alpha - a'_{m-j}sen\alpha\Big)\Big)sen\alpha\Big)\Big) = q(n)r^n\cos\alpha$$
$$(\text{oppure } = q(n)r^n sen\alpha$$

e di qui, sempre in virtù del principio di identità dei polinomi, consegue il l'altro punto dell'enunciato della proposizione 6.5.

La proposizione 6.5 da un metodo per il calcolo di una soluzione particolare dell'equazione [6.1] quando la successione $(s_n)_{n\in N}$ è di uno dei tipi particolari contemplati nel suo enunciato.

È quindi opportuno dare un metodo per il calcolo di una soluzione particolare dell'equazione [6.1] quando la successione $(s_n)_{n\in N}$ non soddisfa alle ipotesi contemplate nell'enunciato della proposizione 6.5, a questo scopo dimostriamo la seguente proposizione:

6.6. Siano: $(a_n)_{n\in N}$, $(b_n)_{n\in N}$ ed $(s_n)_{n\in N}$ tre successioni di numeri reali, $(u_{1,n})_{n\in N}$ e $(u_{2,n})_{n\in N}$ due soluzioni particolari linearmente indipendenti dell'equazione:

$$x_{n+2} = a_n x_{n+1} + b_n x_n \qquad [6.8]$$

Se per ogni numero intero positivo $n\, b_n \neq 0$, allora esistono due successioni di numeri reali, $(y_{1,n})_{n\in N}$ e $(y_{2,n})_{n\in N}$ tali che la successione $(y_{1,n}u_{1,n} + y_{2,n}u_{2,n})_{n\in N}$ sia una soluzione particolare dell'equazione:

$$x_{n+2} = a_n x_{n+1} + b_n x_n + s_n \qquad [6.9]$$

**Dim.** È immediato rendersi conto che l'equazione [6.15] può scriversi sotto la forma:

$$\Delta^2 x_n + (2 - a_n)\Delta x_n + (3 - a_n - b_n)x_n = 0 \qquad [6.10]$$

E conseguentemente la [6.16] sotto la forma:

$$\Delta^2 x_n + (2 - a_n)\Delta x_n + (3 - a_n - b_n)x_n = s_n \qquad [6.11]$$

Ciò osservato siano $(y_{1,n})_{n\in N}$ e $(y_{2,n})_{n\in N}$ due successioni di numeri reali tali che risulti:

$$u_{1,n+1}\Delta y_{1,n} + u_{2,n+1}\Delta y_{2,n} = 0 \qquad [6.12]$$

conseguentemente per ogni numero intero positivo *n*, è:

$$\Delta(y_{1,n}u_{1,n} + y_{2,n}u_{2,n}) = y_{1,n}\Delta u_{1,n} + y_{2,n}\Delta u_{2,n} \qquad [6.13]$$

e

$$\Delta^2(y_{1,n}u_{1,n} + y_{2,n}u_{2,n}) = y_{1,n}\Delta u_{1,n} + \Delta y_{1,n}\Delta u_{1,n+1} + y_{2,n}\Delta^2 u_{2,n} + \Delta y_{2,n}\Delta u_{2,n+1} \qquad [6.14]$$

imponendo quindi che la successione $(y_{1,n}u_{1,n} + y_{2,n}u_{2,n})_{n\in N}$ sia una soluzione particolare dell'equazione [6.9], tenendo conto delle [6.11], [6.10], [6.13] e [6.14], è immediato rendersi conto che per ogni numero intero positivo *n* deve essere:

$$u_{1,n+2}\Delta y_{1,n} + u_{2,n+2}\Delta y_{2,n} = s_n \qquad [6.15]$$

Le [6.12] e [6.15] costituiscono, per ogni numero intero positivo *n*, un sistema di due equazioni nelle due incognite $\Delta y_{1,n}$ e $\Delta y_{2,n}$, che in virtù della proposizione 3.8, è determinato e, se possiamo: $\begin{vmatrix} u_{1,n+1} & u_{2,n+1} \\ u_{1,n+2} & u_{2,n+2} \end{vmatrix} = C(n+1)$, risulta:

$$\Delta y_{1,n} = \frac{u_{2,n+1}s_n}{C(n+1)} \text{ e } \Delta y_{2,n} = \frac{u_{1,n+1}s_n}{C(n+1)} \qquad [6.16]$$

Passiamo ora allo studio dell'equazione alle differenze di ordine *k* lineari a coefficienti costanti. Per questo siano: $a_1, a_2, ..., a_k$ *k* numeri reali ed $(s_n)_{n\in N}$ una successione di numeri reali e consideriamo la seguente equazione:

$$x_{n+k} = a_1 x_{n+k-1} + a_2 x_{n+k-2} + ... + a_k x_n + s_n \qquad [6.17]$$

Anche in questo caso risolveremo prima l'equazione:

$$x_{n+k} = a_1 x_{n+k-1} + a_2 x_{n+k-2} + ... + a_k x_n \qquad [6.18]$$

Omogenea associata all'equazione [6.24] e quindi considereremo l'equazione:

$$t^k - a_1 t^{k-1} - a_2 t^{k-2} - ... - a_k = 0 \qquad [6.19]$$

che chiameremo equazione caratteristica dell'equazione [6.18].

Con ragionamenti analoghi a quelli fatti per dimostrare le proposizioni 6.1, 6.5 e 6.6 si dimostrano, rispettivamente, le seguenti altre tre proposizioni:

6.7. Se l'equazione [6.19] ha zero come radice di molteplicità *h*, *r* radici reali,distinte e diverse da zero, $b_1, b_2, ..., b_r$, rispettivamente di molteplicità $r_1, r_2, ..., r_r$ e $2s$ radici complesse, $c_1 + id_1, c_1 - id_1, c_2 + id_2, c_2 - id_2, ..., c_s + id_s, c_s - id_s$, rispettivamente di molteplicità $s_1, s_1, s_2, s_2 ..., s_s$ allora, se per ogni elemento *i* di $\{1,2,..., s\}$diciamo $\rho_i$ il numero reale positivo $\sqrt{c_i^{\,2} + d_i}^{\,2}$ e denotiamo con $\alpha_i$ l'elemento di $[0,2\pi[$ tale che risulti: $\rho_i \cos\alpha_i = c_i$ e $\rho_i sen\alpha_i = d_i$, le successioni:

$$\begin{array}{cccc}
\left(\delta_{1,n}\right)_{n\in N}, & \left(\delta_{2,n}\right)_{n\in N}, & ..., & \left(\delta_{h,n}\right)_{n\in N} \\
\left(b_1^{\,n}\right)_{n\in N}, & \left(nb_1^{\,n}\right)_{n\in N}, & ..., & \left(n^{\,r_1-1}b_1^{\,n}\right)_{n\in N} \\
\left(b_2^{\,n}\right)_{n\in N}, & \left(nb_2^{\,n}\right)_{n\in N}, & ..., & \left(n^{\,r_2-1}b_2^{\,n}\right)_{n\in N} \\
..., & ..., & ..., & ..., \\
\left(b_r^{\,n}\right)_{n\in N}, & \left(nb_r^{\,n}\right)_{n\in N}, & ..., & \left(n^{\,r_r-1}b_r^{\,n}\right)_{n\in N}
\end{array}$$

$$\begin{array}{cccc}
\left(\rho_1^{\,n}\cos\alpha_1\right)_{n\in N}, & \left(n\rho_1^{\,n}\cos\alpha_1\right)_{n\in N}, & ..., & \left(n^{\,r_1-1}\rho_1^{\,n}\cos\alpha_1\right)_{n\in N} \\
\left(\rho_1^{\,n}sen\alpha_1\right)_{n\in N}, & \left(n\rho_1^{\,n}sen\alpha_1\right)_{n\in N}, & ..., & \left(n^{\,r_1-1}\rho_1^{\,n}sen\alpha_1\right)_{n\in N} \\
\left(\rho_2^{\,n}\cos\alpha_2\right)_{n\in N}, & \left(n\rho_2^{\,n}\cos\alpha_2\right)_{n\in N}, & ..., & \left(n^{\,r_2-1}\rho_2^{\,n}\cos\alpha_2\right)_{n\in N} \\
\left(\rho_2^{\,n}sen\alpha_2\right)_{n\in N}, & \left(n\rho_2^{\,n}sen\alpha_2\right)_{n\in N}, & ..., & \left(n^{\,r_2-1}\rho_2^{\,n}sen\alpha_2\right)_{n\in N} \\
..., & ..., & ..., & ..., \\
\left(\rho_s^{\,n}\cos\alpha_s\right)_{n\in N}, & \left(n\rho_s^{\,n}\cos\alpha_s\right)_{n\in N}, & ..., & \left(n^{\,s_s-1}\rho_s^{\,n}\cos\alpha_s\right)_{n\in N} \\
\left(\rho_s^{\,n}sen\alpha_s\right)_{n\in N}, & \left(n\rho_s^{\,n}sen\alpha_s\right)_{n\in N}, & ..., & \left(n^{\,s_s-1}\rho_s^{\,n}sen\alpha_s\right)_{n\in N}
\end{array}$$

sono n soluzioni particolari linearmente indipendenti dell'equazione [6.18].

6.8. Siano: *m* un numero intero maggiore o uguale a zero, *q* un polinomio di grado *m*, *s* un numero intero positivo, *r* un numero reale diverso da zero ed $\alpha$ un elemento di $R - \bigcup_{h\in Z}\{h\pi\}$ (Fedele2002).

Se è $s_n = q(n)r^n$ per ogni numero intero positivo *n*, allora:

Se *r* non è radice dell'equazione [6.19], esiste un polinomio *p* di grado *m* tale che la successione $\left(p(n)r^n\right)_{n\in N}$ è una soluzione particolare dell'equazione [6.17].

Se *r* è radice di molteplicità *s* dell'equazione [6.19], esiste un polinomio *p* di grado *m* tale che la successione $\left(n^s p(n)r^n\right)_{n\in N}$ è una soluzione particolare dell'equazione [6.17].

Se è $s_n = q(n)r^n \cos\alpha$ oppure $s_n = q(n)r^n sen\alpha$ per ogni numero intero positivo *n*, allora:

Se, essendo *r* positivo (risp. negativo), il numero complesso $r(\cos\alpha + isen\alpha)$ (risp. $-r(\cos(\alpha+\pi)+isen(\alpha+\pi))$) non è radice dell'equazione [6.19], esistono due polinomi, $p_1$ e $p_2$, di grado *m* tali che la successione $\left(p_1(n)r^n \cos\alpha + p_2(n)r^n sen\alpha\right)_{n\in N}$ è una soluzione particolare dell'equazione [6.17].

Se, essendo *r* positivo (risp. negativo), il numero complesso $r(\cos\alpha + isen\alpha)$ (risp. $-r(\cos(\alpha+\pi)+isen(\alpha+\pi))$) è radice di molteplicità <u>s</u> dell'equazione [6.26], esistono due polinomi, $p_1$ e $p_2$, di grado *m* tali che la successione $\left(n^s p_1(n)r^n \cos\alpha + n^s p_1(n)r^n sen\alpha\right)_{n\in N}$ è una soluzione particolare dell'equazione [6.24].

6.9. Siano: $(a_{1,n})_{n\in N}, (a_{2,n})_{n\in N}, ..., (a_{k,n})_{n\in N}$ ed $(s_n)_{n\in N}$ $k+1$ successioni di numeri reali e $(u_{1,n})_{n\in N}, (u_{2,n})_{n\in N}, ..., (u_{k,n})_{n\in N}$ *k* soluzioni particolari linearmente indipendenti dell'equazione

$$x_{n+k} = a_{1,n}x_{n+k-1} + a_{2,n}x_{n+k-2} + ... + a_{k,n}x_n$$

Se per ogni numero intero positivo *n* è $a_{k,n} \neq 0$, allora esistono *k* successioni di numeri reali, $(y_{1,n})_{n\in N}, (y_{2,n})_{n\in N}, ..., (y_{k,n})_{n\in N}$ tali che, per ogni numero intero positivo *n* si abbia:

$$\begin{cases} u_{1,n+1}\Delta y_{1,n} + u_{2,n+1}\Delta y_{2,n} + ... + u_{k,n+1}\Delta y_{k,n} = 0 \\ u_{1,n+2}\Delta y_{1,n} + u_{2,n+2}\Delta y_{2,n} + ... + u_{k,n+2}\Delta y_{k,n} = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ u_{1,n+k}\Delta y_{1,n} + u_{2,n+k}\Delta y_{2,n} + ... + u_{k,n+k}\Delta y_{k,n} = s_n \end{cases}$$

e tali che la successione $\left(y_{1,n}u_{1,n} + y_{2,n}u_{2,n} + ... + y_{k,n}u_{k,n}\right)_{n\in N}$ sia una soluzione particolare dell'equazione:

$$x_{n+k} = a_{1,n}x_{n+k-1} + a_{2,n}x_{n+k-2} + ... + a_{k,n}x_n + s_n$$

OSSERVAZIONE 6.10.Si noti che il metodo indicato nell'enunciato della proposizione 6.9 per risolvere l'equazione [6.17] è applicabile anche quando non è $a_k \neq 0$. Infatti supposto che esista un elemento *r* di $\{1,2,...,k-1\}$ tale che sia $a_r \neq 0$ ed $a_{r+1} = a_{r+2} = ... = a_k = 0$ ,posto $x_{n+k-r} = y_n$ l'equazione [6.24] si scrive sotto la forma:

$$y_{n+r} = a_1 y_{n+r-1} + a_2 y_{n+r-2} + ... + a_r y_n + s_n \qquad [6.20]$$

che è un'equazione alle differenze di ordine $r$ lineare a coefficienti costanti alla quale è possibile applicare il metodo indicato nell'enunciato della proposizione 6.9. Detta quindi

$$\left(\bar{c}_1, \bar{c}_2, ..., \bar{c}_r\right) \in R^r \rightarrow \sum_{i=1}^{r} \bar{c}_i \left(\bar{u}_{i,n}\right)_{n \in N} + \left(\bar{v}_n\right)_{n \in N}$$

una soluzione generale dell'equazione [6.32] se per ogni elemento $i$ di $\{1,2,...,r\}$ poniamo:

$$u_i : n \in N \rightarrow \begin{cases} 0 \text{ se } n < k - r + 1 \\ \bar{u}_{i,n} \text{ se } n > k - r \end{cases}$$

e

$$v : n \in N \rightarrow \begin{cases} 0 \text{ se } n < k - r + 1 \\ \bar{v}_n \text{ se } n > k - r \end{cases}$$

, la funzione:

$$\left(c_1, c_2, ..., c_k\right) \in R^k \rightarrow \sum_{i=1}^{k-r} c_i \left(\delta_{i,n}\right)_{n \in N} + \sum_{i=k-r+1}^{k} c_i \left(u_{i-k+r,n}\right)_{n \in N} + \left(v_n\right)_{n \in N}$$

è una soluzione generale dell'equazione [6.17].

## 7. Qualche semplice applicazione a problemi economici

In questo paragrafo applicheremo quanto fin qui detto a qualche semplice problema Economico(De Giuli2008).

***Applicazione 1***

Sia $C$ un capitale disponibile all'inizio del tempo 1.

Si dice che $C$ è investito a regime di capitalizzazione semplice (risp. composta) con interesse $i$ se, detto $M_n$ il montante all'inizio dell'ennesimo periodo di investimento, è:

$$M_{n+1} = M_n + iC \quad \text{(risp. } M_{n+1} = (1+i)M_n\text{)} \qquad [7.1]$$

quindi il problema del calcolo del montante $M_n$ all'inizio dell'ennesimo periodo d'investimento è ricondotto al calcolo del valore in $n$ della soluzione del seguente problema dei valori iniziali:

$$\begin{cases} M_{n+1} = M_n + iC \\ M_1 = C \end{cases} \text{(risp. } \begin{cases} M_{n+1} = (1+i)M_n \\ M_1 = C \end{cases}\text{).}$$

Senza alcuna difficoltà è possibile stabilire che una soluzione generale dell'equazione [7.1] è la funzione:

$$c \in R \to c(c_N(n))_{n\in N} + (inC)_{n\in N} \text{ (risp. } c \in R \to c\left((1+i)^{n-1}\right)_{n\in N}\text{ )}$$

e che conseguentemente, per ogni numero intero positivo *n*, è: $M_n = C(1+(n-1)i)$ (risp. $M_n = C(1+i)^{n-1}$).

Si voglia ora pagare un debito gravato da un interesse *i*, contratto all'inizio del primo periodo e di importo *D* da estinguersi alla fine dell'h-esimo periodo in *h* rate costanti da pagarsi ognuna alla fine di un periodo a cominciare dal primo.

Se indichiamo con *S* l'importo di ogni singola rata il problema che si pone è il calcolo di *S*, per questo se, per ogni numero intero positivo *n*, indichiamo con $R_n$ il debito residuo all'inizio del periodo ennesimo, deve aversi:

$$R_{n+1} = (1+i)R_n - S \qquad [7.2]$$

quindi il problema del calcolo di $R_n$ è ricondotto al calcolo del valore in *n* della soluzione del seguente problema dei valori iniziali:

$$\begin{cases} R_{n+1} = (1+i)R_n - S \\ R_1 = D \end{cases}.$$

Anche in tal caso una soluzione generale dell'equazione [7.2] è la funzione:

$$c \in R \to c\left((1+i)^{n-1}\right)_{n\in N} + \left(S\frac{1-(1+i)^{n-1}}{i}\right)_{n\in N}$$

e che conseguentemente, per ogni numero intero positivo *n*, è: $R_n = (1+i)^{n-1} + S\dfrac{1-(1+i)^{n-1}}{i}$.

Per il calcolo dell'importo *S* della rata, ricordando che il debito deve essere estinto alla fine dell'h-esimo periodo, deve essere: $D(1+i)^h + S\dfrac{1-(1+i)^h}{i} = 0$ e quindi risulta:

$$S = \frac{iD(1+i)^h}{(1+i)^h - 1}.$$

Il problema è così completamente risolto.

***Applicazione II***

Quello che segue è un tentativo di costruire un semplice modello matematico di trattativa salariale tra lavoratori e datori di lavoro(Maddalena 2009).

Per questo supporremmo che inizialmente i lavoratori fanno una richiesta $\mathfrak{R}$ ed il datore di lavoro un'offerta *O* che naturalmente risulta essere minore di $\mathfrak{R}$ .

Nelle successive fasi della trattativa, nell'intento di giungere ad un accordo, i lavoratori fanno delle richieste via via minori, mentre il datore di lavoro fa delle offerte sempre maggiori.

Un semplice modo di formalizzare quanto su esposto è il seguente: Per ogni fase *n* della trattativa diciamo $L_n$ la richiesta dei lavoratori in tale fase è $D_n$ l'offerta del datore di lavoro nella stessa fase, siano quindi *a* e *b* due numeri reali positivi minori di uno e supponiamo quindi che, in ogni fase *n* della trattativa risulti:

$$\begin{cases} L_{n+1} = L_n - a(L_n - D_n) \\ D_{n+1} = D_n + b(L_n - D_n) \end{cases} \qquad [7.3]$$

Il problema del calcolo di $L_n$ e $D_n$ è ricondotto al seguente altro: Trovare due successioni di numeri reali, $(L_n)_{n\in N}$ e $(D_n)_{n\in N}$ soddisfacenti, per ogni numero intero positivo *n*, alle equazioni [7.3] e tale che risulti: $L_1 = \Re$ e $D_1 = O$ o, equivalentemente, trovare una soluzione particolare del sistema di due equazioni alle differenze lineari del primo ordine [7.3] soddisfacente alle condizioni iniziali: $L_1 = \Re$ e $D_1 = O$.

È quindi augurabile che il problema ora formulato ammette una ed una sola soluzione. Nel seguito faremo vedere che una tale circostanza si verifica.

A tale scopo cominciamo con l'osservare che dalla seconda delle equazioni [7.3] consegue che è: $D_{n+2} = (1-b)D_{n+1} + bL_{n+1}$, sostituendo quindi in questa espressione ad $L_{n+1}$ il suo valore che si ottiene dalla prima delle equazioni [7.3], si ha: $D_{n+2} = (1-b)D_{n+1} + abD_n + b(1-a)L_n$ ed osservando che per la seconda delle equazioni [7.2] è:

$$L_n = \frac{1}{b}D_{n+1} - \frac{1-b}{b}D_n \qquad [7.4]$$

In definitiva risulta:

$$D_{n+2} = (2-a-b)D_{n+1} - (1-a-b)D_n \qquad [7.5]$$

Il nostro problema, quando si osservi che è $D_1 = O$, in virtù delle [7.4], deve essere $D_2 = b\Re + (1-b)O$, è equivalente al seguente problema dei valori iniziali:

$$\begin{cases} D_{n+2} = (2-a-b)D_{n+1} - (1-a-b)D_n \\ D_1 = O \text{ e } D_2 = b\Re + (1-b)O \end{cases} \qquad [7.6]$$

e ciò dimostra come volevasi, che il nostro problema ha una sola soluzione.

Facilmente si osserva che una soluzione generale dell'equazione [7.5], se è $a + b \neq 1$, è la funzione:

$$(c_1, c_2) \in R^2 \to c_1(c_n(n))_{n\in N} + c_2\left((1-a-b)^n\right)_{n\in N}$$

Se invece è $a+b=1$, è la funzione:

$$(c_1, c_2) \in R^2 \to c_1(c_n(n))_{n\in N} + c_2\left(\delta_{1,n}\right)_{n\in N}$$

e quindi la soluzione del problema [7.5] se è $a+b\neq 1$, è la successione:

$$\left(\frac{aO+b\Re}{a+b}c_N(n) - \frac{b(\Re-O)}{a+b}(1-a-b)^{n-1}\right)_{n\in N}$$

Se invece è $a+b=1$, è la successione

$$\left((aO+b\Re)c_N(n) - (b(\Re-O))\delta_{1,n}\right)_{n\in N},$$

quindi, per ogni numero intero positivo *n*, [7.3], se $a+b\neq 1$, risulta:

$$L_n = \frac{aO+b\Re}{a+b} + \frac{a(\Re-O)}{a+b}(1-a-b)^{n-1} \quad \text{e} \quad D_n = \frac{aO+b\Re}{a+b} - \frac{b(\Re-O)}{a+b}(1-a-b)^{n-1}$$

se invece è $a+b=1$, risulta:

$$L_n = aO+b\Re+a(\Re-O)\delta_{1,n} \quad \text{e} \quad D_n = aO+b\Re-b(\Re-O)\delta_{1,n}$$

possiamo perciò asserire che se è $a+b\leq 1$, risulta:

$$\lim_n L_n = \lim_n D_n = \frac{aO+b\Re}{a+b}$$

se invece è $a+b>1$, risulta:

$$\lim_n L_n = +\infty \quad \text{e} \quad \lim_n D_n = -\infty.$$

***Applicazione III***

Il modello della ragnatela.

Si vuole descrivere l'andamento del mercato costituito da un solo bene *B*, e $D_n$, per ogni numero intero positivo *n*, denota il numero di unità del bene *B* richiesto dal mercato nel periodo ennesimo.

$O_n$, per ogni numero intero positivo *n*, denota il numero di unità del bene *B* offerto sul mercato nel periodo ennesimo.

$p_n$, per ogni numero intero positivo *n*, denota il prezzo di ogni unità del bene *B* nel periodo ennesimo.

In questo modello si fa l'ipotesi naturale che la quantità $D_n$ è, per ogni numero intero positivo *n*,funzione di *p*,porremo quindi:

$$D_n = \Phi(p_n) \qquad [7.7]$$

mentre $O_n$ è, per ogni numero intero *n* maggiore di uno, funzione di $p_{n-1}$, porremo quindi:

$$O_{n+1} = \Psi(p_n) \qquad [7.8]$$

L'uguaglianza [7.8] è dettata dall'ipotesi che i Produttori si aspettano che il prezzo unitario del bene *B* non vari nel periodo successivo, quindi programmano la produzione del bene *B* per il periodo successivo in base a tale prezzo.

Supponendo infine che, in ogni periodo, nessun produttore rimane con un prodotto invenduto e nessun consumatore rimane con una domanda insoddisfatta, dalle [7.7] e [7.8]consegue:

$$\Phi(p_{n+1}) = \Psi(p_n) \qquad [7.9]$$

La [7.9] è un'equazione alle differenze del primo ordine che noi discuteremo nelle ipotesi che le due funzioni $\Phi$ e $\Psi$ siano lineari.

Supporremo più precisamente $D_n = u + ap_n$ ed $O_{n+1} = w + bp_n$ dove *u*,*w*,*a* e *b* sono quattro costanti con *a* negativa e *b* positiva.

In queste ipotesi l'equazione [7.8] si scrive sotto la forma:

$$p_{n+1} = \frac{b}{a} p_n + \frac{w-u}{a} \qquad [7.10]$$

ed il problema del calcolo di $p_n$ è ricondotto al calcolo del valore in *n* della soluzione di un qualsiasi problema di Cauchy relativo all'equazione [7.10], per esempio se è noto il valore di $p_1^0$ del prezzo unitario del bene *B* nel primo periodo basta risolvere il seguente problema dei valori iniziali:

$$\begin{cases} p_{n+1} = \dfrac{b}{a} p_n + \dfrac{w-u}{a} \\ p_1 = p_1^0 \end{cases}$$

Non sarà complicato poter stabilire che una soluzione generale dell'equazione [7.10] è la funzione

$$c \in R \to c\left(\left(\frac{b}{a}\right)^{n-1}\right)_{n\in N} + \left(\frac{w-u}{a-b} c_N(n)\right)_{n\in N}$$

e che conseguentemente per ogni numero intero positivo *n*, è:

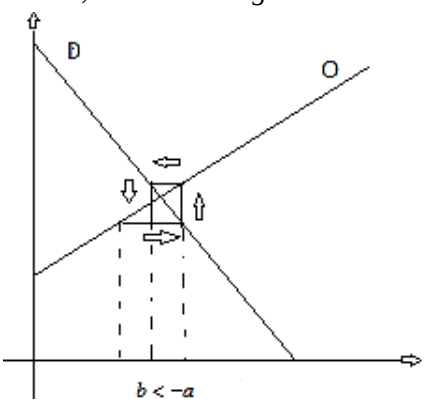
$$p_n = \left(p_1^0 - \frac{w-u}{a-b}\right)\left(\frac{b}{a}\right)^{n-1} + \frac{w-u}{a-b}$$

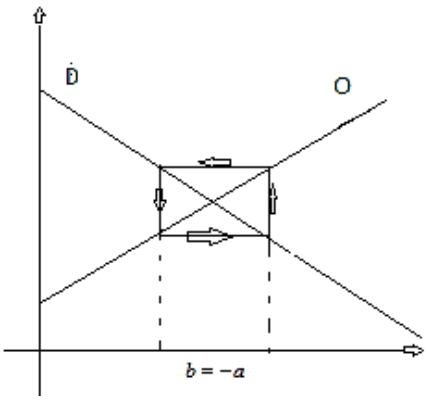
Analizziamo ora l'andamento di *p*.

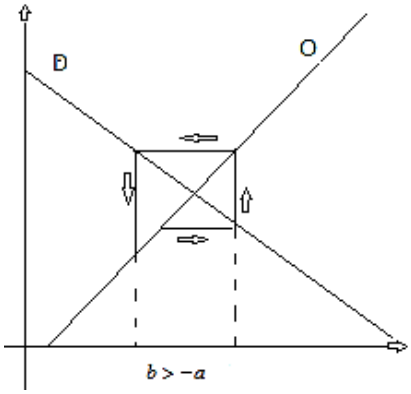
Se si suppone $p_1^0 = \frac{w-u}{a-b}$ allora $p_n$ è costante, in questa ipotesi quindi in ogni periodo il prezzo unitario del bene *B* è uguale a $p_1^0$, il modello è quindi stabile.

Supponiamo che $p_1^0 \neq \frac{w-u}{a-b}$, in questa essendo $\frac{b}{a} < 0$, il prezzo subirà delle oscillazioni variando da un valore maggiore di $\frac{w-u}{a-b}$ in un periodo ad un valore minore di $\frac{w-u}{a-b}$ nel periodo successivo. Più precisamente se $b < -a$ le oscillazioni sono smorzate e quindi col passare del tempo il valore del prezzo unitario si avvicinerà sempre di più a $\frac{w-u}{a-b}$, in questa ipotesi infatti la soluzione identicamente nulla è asintoticamente stabile per l'equazione $p_{n+1} = \frac{b}{a} p_n$ e quindi il $\lim_n p_n = \frac{w-u}{a-b}$, se è $b = -a$ le oscillazioni sono costanti e quindi con il passare del tempo il valore del prezzo unitario resterà a distanza costante da $\frac{w-u}{a-b}$ assumendo in ogni periodo dispari il valore $p_1^0$ ed in ogni periodo pari il valore $2\frac{w-u}{a-b} - p_1^0$, infine se $b > -a$ le oscillazioni sono esaltate e quindi col passare del tempo il valore del prezzo unitario si allontanerà sempre di più da $\frac{w-u}{a-b}$ risultando il $\lim_n p_{2n-1} = \left(p_1^0 - \frac{w-u}{a-b}\right)(+\infty)$ e $\lim_n p_{2n} = \left(p_1^0 - \frac{w-u}{a-b}\right)(-\infty)$.

Nelle tre figure che seguono è indicato l'andamento del prezzo unitario in ognuna delle tre ipotesi prese in esame e tale figure suggeriscono il nome dato a tale modello, detto della *ragnatela*.






Ritornando ora all'equazione [7.9] osserviamo che se si suppongono le funzioni $\Phi$ e $\Psi$ definite in *R* ed ivi derivabili due volte e se inoltre esiste un numero reale $\bar{p}$ tale che risulti $\Phi(\bar{p}) = \Psi(\bar{p})$ e $\Phi'(\bar{p}) \neq 0$, allora se, per ogni numero intero positivo *n*, si pone $p_n = \bar{p} + q_n$, esistono due numeri reali, $\xi$ e $\varsigma$, il primo appartenente a $]\min\{\bar{p}, \bar{p} + q_{n+1}\}, \max\{\bar{p}, \bar{p} + q_{n+1}\}[$ ed il secondo appartenente a $]\min\{\bar{p}, \bar{p} + q_n\}, \max\{\bar{p}, \bar{p} + q_n\}[$ tale che risulti:

$$\Phi(p_{n+1}) = \Phi(\bar{p}) + \Phi'(\bar{p})q_{n+1} + \Phi''(\xi)\frac{1}{2}q^2{}_{n+1}$$

e

$$\Psi(p_n) = \Psi(\bar{p}) + \Psi'(\bar{p})q_n + \Psi''(\xi)\frac{1}{2}q^2{}_n$$

quindi all'equazione [7.9] può sostituirsi in prima approssimazione l'equazione:

$$\Phi'(\bar{p})q_{n+1} = \Psi'(\bar{p})q_n,$$

una soluzione generale di quest'equazione è la funzione:

$$c \in R \to c\left(\left(\frac{\Psi'(\bar{p})}{\Phi'(\bar{p})}\right)^{n-1}\right)_{n \in N},$$

e se risulta: $|\Psi'(\bar{p})| < |\Phi'(\bar{p})|$, allora per ogni numero reale *c*, è $\lim_n \left(\bar{p} + c\left(\frac{\Psi'(\bar{p})}{\Phi'(\bar{p})}\right)^{n-1}\right) = \bar{p}$ e quindi la successione $\left(\bar{p} + c\left(\frac{\Psi'(\bar{p})}{\Phi'(\bar{p})}\right)^{n-1}\right)_{n \in N}$ è una prima approssimazione di una soluzione particolare dell'equazione [7.9].

## Riferimenti bibliografici

Albano, L. (1998). *Lezioni di Matematica Generale*, Cacucci Editore, Bari.

Binmore, K.; Davies, J. (2004). *Calcolo Differenziale di più variabili*, Casa Editrice Ambrosiana, Milano.

De Giuli, M. E.; Giorgi, G.; Maggi, M.; Magnani, U. (2008). *Matematica per l'economia e la finanza*. Zanichelli, Bologna.

Fedele, N. (2002). *Corso di Analisi Matematica*, Integrazione Equazioni Differenziali, Volume I1- Parte II, Liguori Editore, Napoli.

Guerraggio, A. (2004). *Matematica*, Mondadori, Milano.

Maddalena, L. (2009). *Matematica*, Giappichelli Editore, Torino.

Ritelli, D.; Bergamini, M.; Trifone, A. (2005). *Fondamenti di Matematica*, Zanichelli, Bologna.

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018*
DIPARTIMENTO DI ECONOMIA E FINANZA
ISBN: 978-88-6629-048-3
pp. 209-243

# L'apprezzo di Molfetta del 1417: analisi statistiche

**Vito Ricci***

*Università degli studi di Bari Aldo Moro*

**Riassunto:** Il contributo riporta i risultati di alcune analisi statistiche effettuate sui dati contenuti in un apprezzo di Molfetta dell'inizio del Quattrocento, fonte paradigmatica assai utile per la comprensione della realtà socioeconomica di questo centro di Terra di Bari. All'esame della distribuzione dei beni tra i contribuenti per quantità e valore estimale, segue l'analisi spaziale e temporale delle colture praticate. Se per un verso l'indagine conferma la tradizionale diffusione e l'importanza dell'olivicoltura, asse portante dell'economia molfettese, le finalità fiscali dell'apprezzo forniscono una buona *proxy* della ricchezza delle famiglie e pertanto possono essere impiegate per lo studio della disuguaglianza sociale. Nell'ultima parte, di natura comparativa, si esamina la compatibilità dei livelli di diseguaglianza molfettesi con quelli di altre località rurali dell'Italia Centro-Settentrionale.

**Keywords:** Paesaggio agrario; Distribuzione della ricchezza; Disuguaglianza; Modello log-normale; Power law; Fitting distribution.

## 1. Introduzione

L'apprezzo o estimo di Molfetta del 1417 (De Gennaro, 1963) costituisce uno dei pochissimi esempi di documento di natura fiscale nel Mezzogiorno medievale. Esso era lo strumento mediante il quale l'ammontare delle imposte dirette di un centro abitato veniva ripartito tra le unità fiscali (o "fuochi", in genere coincidenti con i nuclei familiari) in relazione al valore delle proprietà immobiliari e di specifici beni mobili (poste). Sono esclusi dalla rilevazione i nullatenenti, coloro i quali possedevano solo la propria abitazione, quella che in termini correnti possiamo defini-

* Autore corrispondente: vito.ricci@uniba.it

re come abitazione principale, e gli esentati. Per ciascuna posta è riportato il dato dell'apprezzo, ovvero il valore su cui calcolare l'imposta che il contribuente doveva pagare per quel bene. Sebbene non espressamente indicato, l'importo estimale corrispondeva al valore economico del bene (Alfani, 2015) e, pertanto, le stime dell'apprezzo costituiscono una buona *proxy* del valore di mercato del bene; da qui deriva che, sommando i valori relativi a ogni singolo contribuente, si ottiene una *proxy* della variabile ricchezza dell'unità fiscale. Dalla distribuzione degli apprezzi tra i singoli contribuenti si può ottenere con buona approssimazione quella della ricchezza; ovviamente si tratta di una distribuzione troncata, in quanto non vengono rilevati i nullatenenti, i proprietari della sola abitazione principale e gli esentati.

I dati contenuti nell'apprezzo costituiscono una buona base di dati per lo studio quali-quantitativo del paesaggio agrario di Molfetta e della distribuzione della ricchezza.

## 2. La base dei dati e le unità fiscali

L'apprezzo esaminato purtroppo risulta mutilo di alcune pagine iniziali: manca un piccolo numero di unità fiscali che più avanti si tenterà di stimare. Per ogni contribuente è riportato l'elenco delle poste corrispondenti alle proprietà; per ciascuna posta vi è breve descrizione del bene e il valore estimale espresso in once, tarì e grana, unità monetarie utilizzate nel Regno di Napoli nel Basso Medioevo. Un'oncia era pari a 30 tarì e un tarì a 20 grana. La maggior parte dei beni presenti nell'apprezzo sono terreni per i quali è indicato il toponimo, la coltura praticata (talvolta non è riportata) e i confini della proprietà. L'importo dell'apprezzo del bene era funzione diretta del valore del bene stesso; nel caso di appezzamenti di terra questo a sua volta dipendeva dal tipo di coltura, dall'ubicazione sul territorio, dall'estensione, dal numero di alberi presenti, dalla produttività, etc.: variabili che, purtroppo, ad eccezione dell'ubicazione e del tipo di coltura, non sono rilevate nell'apprezzo.

Partendo dalla trascrizione del documento medievale, si è costruito un database con le seguenti variabili:

- NomeContribuente
- Genere
- TipologiaContribuente
- Provenienza
- TipoColtura
- LocalitaBene

- FlagTerrenoChiuso
- TipoBene
- ImportoApprezzato.

Il numero di unità fiscali è risultato pari a 215 e quello delle poste a 925. Occorre precisare che per due contribuenti non è riportato alcun bene apprezzato e sono stati quindi esclusi dalle analisi miranti ad ottenere una stima della distribuzione della ricchezza; allo stesso modo alcune poste sono prive dell'importo estimale e non è stato quindi possibile utilizzarle nell'analisi di valore.

Esaminando i contribuenti secondo la tipologia e la provenienza (Tab. 1) si evidenzia come nel 71,6% si tratti di residenti a Molfetta, mentre il restante 28,4% risiede in centri vicini; quasi i due terzi (65,6%) appartengono alla categoria dei "cittadini", il 26,5% sono *exteri* (ovvero forestieri che dimorano in Molfetta) e il 7,9% chierici nobili (tassati per i loro patrimoni privati). La distinzione per tipologia e genere (Tab. 2) mette in evidenza la preponderante prevalenza di contribuenti maschi (78,1%), coincidenti quasi sempre con il capo-famiglia, mentre quelli di genere femminile rappresentano il 19,5% del totale; mancano all'appello 5 contribuenti dei quali non è stato possibile individuare il genere. Nella Tab. 3 si presenta la distribuzione delle unità fiscali secondo la provenienza e il genere; si noti la lieve maggiore incidenza del genere femminile tra i contribuenti *exteri* (21,3%) rispetto a quelli residenti a Molfetta (18,8%).

**Tabella 1**. *Distribuzione dei contribuenti secondo la tipologia e la provenienza*

| Tipologia di contribuente | Percentuali di riga | | | Percentuali di colonna | | |
|---|---|---|---|---|---|---|
| | Provenienza | | Totale | Provenienza | | Totale |
| | Molfetta | Altra | | Molfetta | Altra | |
| Chierico | 76,5 | 23,5 | 100,0 | 8,4 | 6,6 | 7,9 |
| Cittadino | 100,0 | 0,0 | 100,0 | 91,6 | 0,0 | 65,6 |
| Extero | 0,0 | 100,0 | 100,0 | 0,0 | 93,4 | 26,5 |
| Totale | 71,6 | 28,4 | 100,0 | 100,0 | 100,0 | 100,0 |

**Tabella 2**. *Distribuzione dei contribuenti secondo la tipologia e il genere*

| Tipologia di contribuente | Percentuali di riga | | | | Percentuali di colonna | | | |
|---|---|---|---|---|---|---|---|---|
| | Genere | | | | Genere | | | |
| | M | F | n.d. | Totale | M | F | n.d. | Totale |
| Chierico | 100,0 | 0,0 | 0,0 | 100,0 | 10,1 | 0,0 | 0,0 | 7,9 |
| Cittadino | 77,3 | 20,6 | 2,1 | 100,0 | 64,9 | 69,0 | 60,0 | 65,6 |
| Extero | 73,7 | 22,8 | 3,5 | 100,0 | 25,0 | 31,0 | 40,0 | 26,5 |
| Totale | 78,1 | 19,5 | 2,3 | 100,0 | 100,0 | 100,0 | 100,0 | 100,0 |

**Tabella 3**. *Distribuzione dei contribuenti secondo la provenienza e il genere*

| Provenienza | Percentuali di riga | | | | Percentuali di colonna | | | |
|---|---|---|---|---|---|---|---|---|
| | Genere | | | | Genere | | | |
| | M | F | n.d. | Totale | M | F | n.d. | Totale |
| Molfetta | 79,2 | 18,8 | 1,9 | 100,0 | 72,6 | 69,0 | 60,0 | 71,6 |
| Altra | 75,4 | 21,3 | 3,3 | 100,0 | 27,4 | 31,0 | 40,0 | 28,4 |
| Totale | 78,1 | 19,5 | 2,3 | 100,0 | 100,0 | 100,0 | 100,0 | 100,0 |

## 3. L'analisi delle poste

Una prima analisi attiene all'aspetto qualitativo e quantitativo delle poste indicate nell'apprezzo. In base al contenuto del documento fiscale è stato possibile classificare i beni come riportato nella Tab. 4. Appare lampante come la quasi totalità degli stessi (91,5%) fosse costituita da appezzamenti di terra, a conferma dell'economia prevalentemente agricola di Molfetta ai primi del XV secolo; seguivano, a notevole distanza gli impianti per la gestione delle acque (pozzi, cisterne, etc.) con il 3,1%, mentre gli immobili urbani (sovente seconde case affittate o a disposizione) costituivano il 2,5%.

**Tabella 4**. *Distribuzione del numero degli apprezzi in relazione alla tipologia dei beni e al genere del contribuente (percentuali di colonna, tra parentesi i valori assoluti)*

| Tipo di bene | Genere | | | Totale | % |
|---|---|---|---|---|---|
| | M | F | n.d. | | |
| Impianto di trasformazione | 1,6 (13) | 1,0 (1) | 4,2 (1) | 15 | *1,6* |
| Impianto idrico | 3,5 (28) | 1,0 (1) | | 29 | *3,1* |
| Mobile | 1,2 (10) | | | 10 | *1,1* |
| Terreno | 91,1 (731) | 92,9 (92) | 95,8 (23) | 846 | *91,5* |
| Immobile urbano | 2,4 (19) | 4,0(4) | | 23 | *2,5* |
| n. d. | 0,1 (1) | 1,0 (1) | | 2 | *0,2* |
| Totale | 802 | 99 | 24 | 925 | *100,0* |

Occorre ricordare come la casa di abitazione della famiglia fosse esclusa dall'imposizione fiscale. Il numero limitato di abitazioni apprezzate costituisce, di conseguenza, una testimonianza indiretta del possesso diffuso della casa tra i cittadini di Molfetta, indice di un certo grado di benessere. Bisogna anche tenere presente che nel Medioevo era comune la pratica enfiteutica e non è da escludere che diversi abitanti vivessero in case di Enti religiosi (chiese, monasteri, conventi, episcopi, confraternite, etc.) che le concedevano in enfiteusi dietro il pagamento di un canone in denaro. L'imposizione fiscale riguardava quasi esclusivamente i beni stabili: gli unici beni mobili colpiti erano i vasi per la conservazione dell'olio (*ve-*

*getibus oleariis*), a testimonianza del ruolo che l'olivicoltura e la produzione e il commercio dell'olio (con esportazioni anche fuori dei confini regnicoli) giocavano nell'economia molfettese. In un caso oggetto dell'apprezzo è la barca di un marinaio/pescatore. L'86,7% delle poste risulta far capo ad un contribuente di genere maschile, una prevalenza che emerge anche dalla relazione tra genere e tipologia del bene apprezzato.

Nella Tab. 5 si riporta la distribuzione degli apprezzi per tipo di bene e provenienza del contribuente. Le poste per l'86,5% sono riferibili ad abitanti di Molfetta e per il restante 13,5% a forestieri dimoranti in città. Lievemente maggiore l'incidenza dei terreni per quanto concerne questi ultimi (92,8% vs 91,3%). Nella Tab. 6 si incrociano i dati relativi alla tipologia dei beni apprezzati con le diverse tipologie di contribuente: i cittadini di Molfetta sono titolari dell'82,1% delle poste dell'apprezzo, seguono gli *exteri* con il 13,1% e i chierici con il 4,9%. L'incidenza dei terreni presenta lievi differenze tra le tre categorie, attestandosi sempre sopra il 90%.

**Tabella 5**. *Distribuzione del numero degli apprezzi in relazione alla tipologia del bene e alla provenienza del contribuente (percentuali di colonna, tra parentesi i valori assoluti)*

| Tipo di bene | Provenienza | | Totale | % |
|---|---|---|---|---|
| | Molfetta | Altra | | |
| Impianto di trasformazione | 1,8 (14) | 0,8 (1) | 15 | *1,6* |
| Impianto idrico | 3,5 (28) | 0,8 (1) | 29 | *3,1* |
| Mobile | 1,3 (10) | | 10 | *1,1* |
| Terreno | 91,3 (730) | 92,8 (116) | 846 | *91,5* |
| Immobile urbano | 2,1 (17) | 4,6 (6) | 23 | *2,5* |
| n.d. | 0,1 (1) | 0,8 (1) | 0 | *0,2* |
| Totale | 800 | 125 | 925 | *100,0* |

**Tabella 6**. *Distribuzione del numero degli apprezzi in relazione alla tipologia del bene e del contribuente (percentuali di colonna, tra parentesi i valori assoluti)*

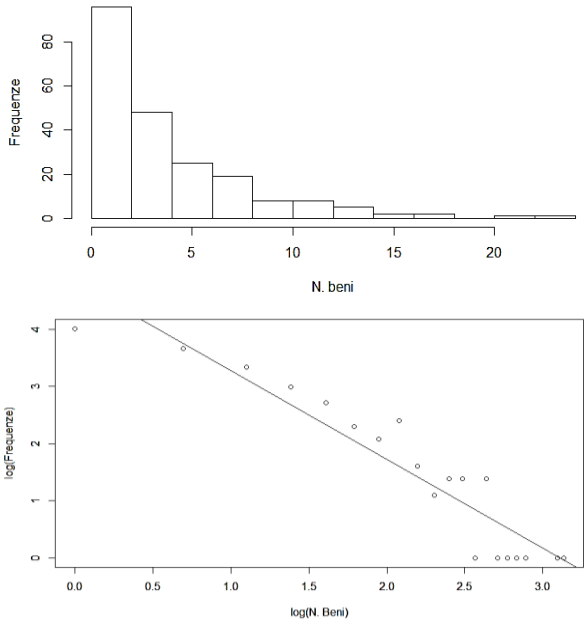
| Tipo di bene | Tipologia di contribuente | | | Totale | Totale |
|---|---|---|---|---|---|
| | Chierico | Cittadino | Extero | | |
| Impianto di trasformazione | 4, (2) | 1,6 (12) | 0,8 (1) | 15 | *1,6* |
| Impianto idrico | | 3,7 (28) | 0,8 (1) | 29 | *3,1* |
| Mobile | 2,2 (1) | 1,2 (9) | | 10 | *1,1* |
| Terreno | 93,3 (42) | 91,2 (692) | 92,6 (112) | 846 | *91,5* |
| Immobile urbano | | 2,2 (17) | 5,0 (6) | 23 | *2,5* |
| n.d. | | 0,1 (1) | 0,8 (1) | 2 | *0,2* |
| Totale | 45 | 759 | 121 | 925 | *100,0* |

Quanto alla distribuzione delle unità fiscali in base al numero dei beni apprezzati si vedano la Tab. 7 e la Fig. 1.

**Tabella 7**. *Distribuzione dei contribuenti secondo il numero di beni apprezzati*

| N. beni | N. contribuenti | % | Cumulativa | % |
|---|---|---|---|---|
| 1 | 55 | *25,8* | 55 | *25,8* |
| 2 | 39 | *18,3* | 94 | *44,1* |
| 3 | 28 | *13,1* | 122 | *57,3* |
| 4 | 20 | *9,4* | 142 | *66,7* |
| 5 | 15 | *7,0* | 157 | *73,7* |
| 6 | 10 | *4,7* | 167 | *78,4* |
| 7 | 8 | *3,8* | 175 | *82,2* |
| 8 | 11 | *5,2* | 186 | *87,3* |
| 9 | 5 | *2,3* | 191 | *89,7* |
| 10 | 3 | *1,4* | 194 | *91,1* |
| 11 | 4 | *1,9* | 198 | *93,0* |
| 12 | 4 | *1,9* | 202 | *94,8* |
| 13 | 1 | *0,5* | 203 | *95,3* |
| 14 | 4 | *1,9* | 207 | *97,2* |
| 15 | 1 | *0,5* | 208 | *97,7* |
| 16 | 1 | *0,5* | 209 | *98,1* |
| 17 | 1 | *0,5* | 210 | *98,6* |
| 18 | 1 | *0,5* | 211 | *99,1* |
| 22 | 1 | *0,5* | 212 | *99,5* |
| 23 | 1 | *0,5* | 213 | *100,0* |
| Totale | 213 | *100* | | |

**Figura 1**. *Distribuzione di frequenza del numero dei beni apprezzati in scala naturale e logaritmica*



Il valore modale si ha in corrispondenza di un unico bene: il 25,8% dei contribuenti possiede un solo bene apprezzato, seguono quelli con 2 e con 3; 3 beni è anche il valore mediano. I contribuenti fino a 4 beni apprezzati coprono i 2/3 del totale. Appare evidente una tendenza al frazionamento della proprietà.

Si noti come, nella prima parte della Fig. 1, la distribuzione di frequenza sembri seguire un andamento di tipo *power law*:

$$N_h = AB^{-h}$$

dove $N_h$ è la frequenza assoluta dei contribuenti con h beni; in scala doppio-logaritmica la relazione diventa lineare:

$$\log(N_h) = \log(A) - log(B)h$$

e i parametri A e B possono essere stimati con il metodo dei minimi quadrati (Piccolo, 1998) dando seguito a questi valori:

$$\log(N_h) = 4{,}83 - 1{,}55 * h$$

Il coefficiente di determinazione $R^2$ è pari a 0,88 e mostra un buon adattamento del modello della legge di potenza nel descrivere il fenomeno, come si vede anche graficamente dalla seconda parte della Fig. 1. L'inclinazione della retta logaritmica può essere interpretata come un indice sommario dell'uguaglianza nella distribuzione del numero dei beni; a maggiori valori corrisponde una distribuzione più egualitaria. Il suo aumento significa l'incremento dei valori più bassi e la diminuzione dei più elevati, invece la sua riduzione ha significato opposto. Possiamo calcolare il grado di concentrazione della distribuzione attraverso il calcolo dell'indice di Gini (Piccolo, 1998), misura utilizzata per i caratteri trasferibili. Ordinando il numero di beni in senso non decrescente, indichiamo con $p_i$ la quota dei primi i contribuenti e con $q_i$ la quota dei loro beni apprezzati. Si ha che l'indice di Gini assume la seguente formula:

$$G = \frac{\sum_{i=1}^{n-1}(p_i - q_i)}{\sum_{i=1}^{n-1} q_i}$$

con campo di variazione compreso tra 0 in caso di equi-distribuzione e 1 nel caso di massima concentrazione. L'indice G per la distribuzione del numero dei beni apprezzati da ciascun contribuente è pari 0,457.

Nella Tab. 8 sono riportati alcuni semplici indicatori statistici relativi al numero dei beni apprezzati secondo alcune variabili di classificazione.

**Tabella 8**. *Statistiche sul numero degli apprezzi per contribuente in relazione ad alcune variabili di classificazione*

| Variabile di classificazione | Media | Dev. Std. | Min | Max |
|---|---|---|---|---|
| *Genere* | | | | |
| M | 4,80 | 4,27 | 1 | 23 |
| F | 2,39 | 1,68 | 1 | 8 |
| n. d. | 4,80 | 4,07 | 1 | 12 |
| *Tipologia di contribuente* | | | | |
| Chierico | 2,65 | 1,71 | 1 | 6 |
| Cittadino | 5,41 | 4,35 | 1 | 23 |
| Extero | 2,14 | 2,17 | 1 | 12 |
| *Provenienza* | | | | |
| Molfetta | 5,22 | 4,23 | 1 | 23 |
| Altra | 2,07 | 2,11 | 1 | 12 |
| Totale | 4,33 | 4,02 | 1 | 23 |

Nel complesso mediamente ogni contribuente possiede 4,33 beni apprezzati, con un minimo di 1 e un massimo di 23. In base alla tipologia del contribuente, quelli con maggior numero medio di beni sono i cittadini di Molfetta, mentre gli *exteri* e i chierici ne possedevano meno; un discorso analogo può effettuarsi con riferimento alla provenienza. La circostanza che la media dei forestieri sia molto più bassa rispetto a quella dei cittadini di Molfetta è ovviamente intuibile: costoro possedevano in questo centro solo pochi beni, avendo verosimilmente concentrato le proprietà nel comune di residenza. Per quanto concerne il dato relativo ai chierici questo può essere spiegato con le esenzioni fiscali di cui gli ecclesiastici godevano nel Medioevo: nell'apprezzo erano tassati solo i beni rientranti nel patrimonio privato, mancavano quelli goduti a titolo di beneficio.

Sicuramente la variabile di maggiore interesse è quella del valore estimale del bene $x_{ij}$, espressione indiretta del suo valore economico. In particolare: sia n il numero dei fuochi, k il numero complessivo delle poste, $k_i$ il numero dei beni apprezzati dall'unità i.ma e $x_{ij}$ il valore estimale del bene j.mo dell'unità i.ma; ne deriva che la grandezza:

$$X = \sum_{i=1}^{n} \sum_{j=1}^{k_i} x_{ij}$$

costituisce l'ammontare complessivo dei valori estimali di tutti i contribuenti;

$$k = \sum_{i=1}^{n} k_i$$

è il numero totale delle poste. La formula della media del valore delle poste è la seguente:

$$\bar{x} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{k_i} x_{ij}}{k}$$

Per comodità i valori sono stati tutti convertiti in tarì e sono state escluse quelle poste (10) che non sono stimate nell'apprezzo. Nella Tab. 9 si riporta la distribuzione dei beni secondo il valore apprezzato. La classe è chiusa a destra e aperta a sinistra.

**Tabella 9**. *Distribuzione dei beni secondo la classe di valore di apprezzo* (in tarì)

| Classi valore di apprezzo | N. | % | Cumulativa | % | Importo medio |
|---|---|---|---|---|---|
| Fino a 0,5 | 59 | *6,4* | 59 | *6,4* | 0,33 |
| 0,51-1 | 61 | *6,7* | 120 | *13,1* | 0,81 |
| 1,1-2 | 93 | *10,2* | 213 | *23,3* | 1,58 |
| 2,1-3 | 100 | *10,9* | 313 | *34,2* | 2,77 |
| 3,1-4 | 54 | *5,9* | 367 | *40,1* | 3,70 |
| 4,1-5 | 88 | *9,6* | 455 | *49,7* | 4,61 |
| 5,1-6 | 76 | *8,3* | 531 | *58,0* | 5,73 |
| 6,1-7 | 27 | *3,0* | 558 | *61,0* | 6,75 |
| 7,1-8 | 48 | *5,2* | 606 | *66,2* | 7,64 |
| 8,1-9 | 43 | *4,7* | 649 | *70,9* | 8,75 |
| 9,1-10 | 22 | *2,4* | 671 | *73,3* | 9,75 |
| 10,1-15 | 117 | *12,8* | 788 | *86,1* | 12,63 |
| 15,1-20 | 53 | *5,8* | 841 | *91,9* | 17,57 |
| 20,1-25 | 31 | *3,4* | 872 | *95,3* | 22,26 |
| 25,1-30 | 20 | *2,2* | 892 | *97,5* | 27,54 |
| 30,1-35 | 8 | *0,9* | 900 | *98,4* | 31,91 |
| 35,1-40 | 8 | *0,9* | 908 | *99,2* | 36,56 |
| 40,1-50 | 3 | *0,3* | 911 | *99,6* | 40,92 |
| 50,1-80 | 4 | *0,4* | 915 | *100,0* | 62,75 |
| Totale | 915 | *100,0* | | | 7,92 |

Si osserva che la classe modale è data dall'intervallo 0,51-1, avendo la maggiore densità di frequenza dato che le classi sono di ampiezza diversa, mentre la mediana cade nella classe 5,1-6 tarì e il valore medio è pari a 7,92. La distribuzione appare caratterizzata da una decisa asimmetria positiva, come spesso accade per alcune grandezze economiche (reddito, ricchezza, patrimonio, etc.) e si può ricondurre ad un modello log-normale con funzione di densità della frequenza:

$$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{\log(x) - \mu}{\sigma}\right)^2\right]$$

Questo particolare induce ad utilizzare la trasformazione logaritmica per ottenere una distribuzione di tipo normale:

$$z_{ij} = \log(x_{ij})$$

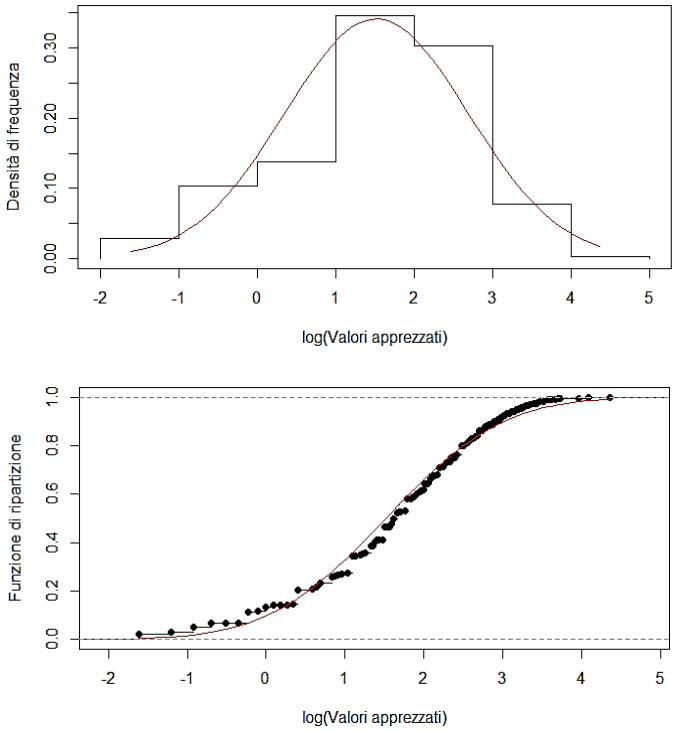
Essa infatti riduce notevolmente l'asimmetria, come si può vedere dalla Fig. 2 nella quale sono riportate anche la funzione di densità di frequenza e la funzione di ripartizione della distribuzione gaussiana. Gli indici di asimmetria e curtosi (Piccolo, 1998):

$$\gamma_1 = \frac{\sum_{i=1}^{n}(x_i - \mu)^3}{n\sigma^3}$$

$$\gamma_2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^4}{n\sigma^4} - 3$$

della distribuzione delle $z_{ij}$ sono è pari rispettivamente a -0,558 e -0,03, valori compatibili con il modello gaussiano, mentre $\gamma_1 = 2{,}51$ per la variabile X.

**Figura 2**. *Distribuzione di frequenza e funzione di ripartizione del logaritmo del valore d'apprezzo dei beni*



Se stimiamo i parametri della distribuzione log-normale con il metodo della massima verosimiglianza (Piccolo, 1998) otteniamo che $\hat{\mu} = 1{,}522$ e $\hat{\sigma} = 1{,}167$. Per misurare la bontà dell'adattamento del modello log-normale possiamo calcolare degli indicatori di grado di accostamento che permettono di verificare e valutare l'approssimazione tra le frequenze del carattere osservate e quelle calcolate con il modello teorico scelto. Nella fattispecie si è utilizzato questo indicatore di tipo relativo (Girone, 2009):

$$\delta = \frac{\sum_{i=1}^{n} |f_i - f_i^*|}{\sum_{i=1}^{n} f_i} * 100$$

dove $f_i$ sono le frequenze assolute osservate e $f_i^*$ quelle teoriche; quanto più l'indicatore è prossimo a 0, migliore risulta l'adattamento. Nel caso della distribuzione in questione si ha che $\delta = 23{,}4\%$, un risultato che mette in evidenza un discreto *fitting* del modello log-normale.

Nella Tab. 10 si riportano alcuni indicatori statistici di tipo descrittivo del valore estimale dei singoli beni secondo alcune variabili di segmentazione.

**Tabella 10**. *Statistiche sui beni apprezzati secondo alcune variabili di classificazione* (in tarì)

| Variabile di classificazione | N. | % | Valori Totale | % | Media | Dev. std. | Min | Max |
|---|---|---|---|---|---|---|---|---|
| *Genere* | | | | | | | | |
| M | 802 | *86,7* | 6.225,33 | *85,9* | 7,82 | 8,16 | 0,15 | 60,00 |
| F | 99 | *10,7* | 903,25 | *12,5* | 9,41 | 10,51 | 0,35 | 78,50 |
| n. d. | 24 | *2,6* | 117,50 | *1,6* | 5,11 | 3,27 | 0,75 | 13,00 |
| *Tipologia di contribuente* | | | | | | | | |
| Chierico | 45 | *4,9* | 353,63 | *4,9* | 7,86 | 7,15 | 0,50 | 30,00 |
| Cittadino | 759 | *82,1* | 5.746,95 | *79,3* | 7,64 | 7,84 | 0,15 | 52,50 |
| Extero | 121 | *13,1* | 1.145,50 | *15,8* | 9,71 | 11,40 | 0,50 | 78,50 |
| *Provenienza* | | | | | | | | |
| Molfetta | 800 | *86,5* | 6.083,08 | *83,9* | 7,67 | 7,82 | 0,15 | 52,50 |
| Altra | 125 | *13,5* | 1.163,00 | *16,1* | 9,53 | 11,26 | 0,50 | 78,50 |
| *Tipo di bene* | | | | | | | | |
| Impianto di trasformazione | 15 | *1,6* | 152,53 | *2,1* | 10,89 | 8,11 | 0,30 | 32,00 |
| Impianto idrico | 29 | *3,1* | 38,08 | *0,5* | 1,31 | 1,41 | 0,15 | 7,50 |
| Mobile | 10 | *1,1* | 93,38 | *1,3* | 9,34 | 7,17 | 0,95 | 21,00 |
| Terreno | 846 | *91,5* | 6.819,63 | *94,1* | 8,12 | 8,52 | 0,15 | 78,50 |
| Immobile urbano | 23 | *2,5* | 142,48 | *2,0* | 6,48 | 4,11 | 1,50 | 16,10 |
| n. d. | 2 | *0,2* | | *0,0* | | | | |
| Totale | 925 | *100,0* | 7.246,08 | *100,0* | 7,92 | 8,38 | 0,15 | 78,50 |

Si noti come il valore estimale medio delle unità fiscali di genere femminile sia superiore a quello delle unità di genere maschile; sebbene le donne vedano apprez-

zati un numero inferiore di beni rispetto agli uomini, questi beni sono di maggior valore. Per quanto attiene alla tipologia di contribuente non sembra esservi una sostanziale differenza tra cittadini e chierici, che si troverebbero quindi nelle stesse condizioni economiche. Un valore medio più elevato si riscontra invece per gli *exteri*: anche per costoro si può affermare che possedevano un numero inferiore di beni, ma di un certo pregio e valore. Allo stesso risultato conduce l'esame dei dati secondo la provenienza del contribuente. Per quanto riguarda la tipologia dei beni, quelli che presentano il valore estimale medio più elevato sono gli impianti di trasformazione che comprendono frantoi (*trappeti*), torchi (*mangani*) e mulini (*molendini*), molti dei quali ubicati nel *suburbio* della città. Molfetta, sebbene cittadina marittima, aveva un'economia basata sulla produzione dell'olio e quindi i *trappeti* costituivano una risorsa di grande valore e importanza economica: uno di questi era apprezzato per 1 oncia e 2 tarì. Seguivano i beni mobili, anche in questa circostanza legati al mondo dell'olio: si tratta di vasi e contenitori che erano valutati anche sino a una ventina di tarì; purtroppo mancano informazioni sul numero dei *vegetibus* e sulla loro capacità. Al terzo posto troviamo gli appezzamenti di terra, categoria numericamente più numerosa con il 94,1% dell'ammontare complessivo apprezzato. Gli immobili urbani, in genere case o altri immobili (forni, depositi) affittati avevano un valore estimale, e quindi una redditività, inferiore alla media. In tale categoria ricadeva anche la *stacione*, ovvero lo spazio pubblico occupato nella piazza dai venditori ambulanti, tassata per 3 tarì.

## 4. Il paesaggio agrario e le colture agricole

Come si è visto a conclusione del paragrafo precedente, il 91%, in termini di numerosità, e il 94%, in termini di valore, dei beni apprezzati è costituito da terreni, come è peraltro ragionevole aspettarsi in una società fortemente agricola. La descrizione degli appezzamenti è piuttosto scarna e finalizzata agli scopi fiscali del documento: oltre al valore estimale è indicata la coltura praticata, l'ubicazione sul territorio, eventuali proprietà confinanti; non è mai riportata l'estensione dell'ap–pezzamento (in alcune circostanze si precisa che si trattava della quota parte di un terreno, ma tali informazioni non sono di grande aiuto), assai raramente vi è l'indicazione del numero di alberi presenti. Tutte queste variabili, assieme alla produttività della terra e ad altri indicatori (vetustà degli alberi, eventuali danneggiamenti dovuti ad incendi, presenza o meno di recinzione, etc.), determinavano il valore economico del bene fondiario in base al quale era computato il valore estimale riportato nell'apprezzo.

Le proprietà fondiarie presenti nell'apprezzo del 1417 costituiscono una buona descrizione del paesaggio agrario di Molfetta agli inizi del XV secolo, sebbene occorra precisare che esse non coprono per intero il territorio di questo centro, in quanto dalla tassazione erano esclusi i terreni di proprietà degli Enti religiosi, esentati dalla contribuzione fiscale (talvolta appaiono menzionati nella descrizione dei confini delle proprietà apprezzate). Gli Enti religiosi erano tra i maggiori proprietari fondiari; tuttavia questo elemento, pur escludendo una parte non quantificabile di terreni, non inficia l'analisi del paesaggio agrario.

In primo luogo, nell'intento di ricostruire il paesaggio agrario di Molfetta ai primi del Quattrocento, si è proceduto a riclassificare le diverse colture riportate nell'apprezzo in modo da semplificarne la lettura. Nella Tab. 11 sono presentati alcuni dati di tipo esplorativo che evidenziano il peso della variabile colturale.

**Tabella 11**. *Statistiche su numero e valore dei terreni apprezzati in relazione alle colture praticate* (in tarì)

| Coltura | N. | *%* | Valori | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Totale | *%* | Media | Dev. Std. | Min | Max |
| Oliveto | 482 | *57,0* | 3.739,05 | *54,8* | 7,82 | 8,15 | 0,20 | 78,50 |
| Mandorlo e olivo | 89 | *10,5* | 1.055,85 | *15,5* | 12,00 | 9,58 | 0,23 | 60,00 |
| Mandorlo | 71 | *8,4* | 840,85 | *12,3* | 11,84 | 7,84 | 0,15 | 36,00 |
| Pastino | 36 | *4,3* | 168,00 | *2,5* | 4,67 | 3,80 | 0,68 | 19,25 |
| Orto con alberi | 26 | *3,1* | 414,08 | *6,1* | 15,93 | 11,82 | 3,50 | 40,50 |
| Vigneto | 24 | *2,8* | 104,65 | *1,5* | 4,36 | 2,41 | 0,50 | 9,00 |
| Coltura arborea promiscua | 21 | *2,5* | 251,13 | *3,7* | 11,96 | 10,26 | 1,50 | 36,00 |
| Pero | 10 | *1,2* | 24,35 | *0,4* | 2,44 | 4,25 | 0,15 | 15,00 |
| Macchia | 5 | *0,6* | 8,40 | *0,1* | 1,68 | 1,53 | 0,15 | 4,00 |
| Querceto | 4 | *0,5* | 2,80 | *0,0* | 0,70 | 0,50 | 0,20 | 1,50 |
| Fico | 3 | *0,4* | 14,40 | *0,2* | 4,80 | 3,00 | 1,50 | 8,75 |
| Frutteto | 3 | *0,4* | 36,50 | *0,5* | 12,17 | 14,04 | 1,50 | 32,00 |
| Non indicata | 72 | *8,5* | 159,58 | *2,3* | 2,25 | 4,03 | 0,15 | 30,00 |
| Totale | 846 | *100,0* | 6.819,63 | *100,0* | 8,12 | 8,52 | 0,15 | 78,50 |

L'olivo si conferma come la coltura più importante e diffusa sul territorio di Molfetta, occupando il 57% degli appezzamenti e il 54,8% in termini del loro valore estimale. Segue, a distanza, l'abbinamento mandorlo e olivo con il 10,5% degli appezzamenti e il solo mandorlo con l'8,4%. Le altre colture hanno un ruolo alquanto marginale rappresentando appena il 15,6% del totale: si tratta di vigneti[1], frutteti, querceti e di una parte residuale costituita da macchia mediterranea. Considerando il valore medio estimale, quello più elevato si ha per gli orti con alberi (15,93 tarì) che

[1] Tra i vigneti vi erano i pastini, terreni lavorati con viti di nuovo impianto, che risultavano avere un'incidenza maggiore rispetto ai vigneti di vecchio impianto 4,3% vs 2,8% elemento che evidenzia un'espansione della viticoltura i cui effetti si protrarranno al XVI secolo.

nell'apprezzo sono descritti come *covevoline*. Con questo termine si indicavano terreni ricchi di potassio che, "cucinando" le sementi, risultavano particolarmente adatti alla produzione di ortaggi e legumi di più facile cottura. Erano presenti tanto nelle contrade rurali quanto nelle vicinanze della città (*suburbio*), e molto spesso ospitavano anche alcune colture arboree: quelle più diffuse erano il mandorlo e il fico, più raramente l'olivo. Secondo in redditività è il frutteto (*jardeno*) con 12,17 tarì, sebbene si tratti di appena 3 attestazioni in prossimità del centro urbano.

Quasi appaiati sono i valori medi dei terreni coltivati a mandarlo e olivo, del solo mandorlo e delle altre colture arboree che si attestano intorno ai 12 tarì. Il mandorlo è una coltura che in Terra di Bari trova grande diffusione a partire del XV secolo, anche perché i suoi frutti trovavano ampio impiego nella cucina tardomedievale. L'oliveto possedeva una redditività inferiore, con una media di 7,82 tarì, ma con una significativa variabilità, testimoniata da un *range* di 78,3 tarì. I terreni meno redditizi in assoluto sono i pochi querceti, molto probabilmente utilizzati per ricavarne ghiande da destinare all'alimentazione animale. È chiaro che tali valori sono influenzati anche dall'estensione dei terreni, variabile della quale purtroppo non si può tenere conto.

Un'analisi alquanto interessante è quella relativa alla recinzione o meno della proprietà fondiaria. La *clausurea* (chiusura), terreno chiuso, circondato da muretti, costituiva la spia della realizzazione del possesso privato della terra e della sua valorizzazione colturale; è l'elemento che caratterizza il paesaggio agrario pugliese e raggiunge tra XIII e XIV secolo la massima diffusione. I muretti delimitano la proprietà e la proteggono dallo sconfinamento degli animali, dalle incursioni dei pastori e degli occasionali viandanti. In particolare, in questa sede si vuole esaminare la diffusione della chiusura tra le diverse colture praticate e verificare se vi era una differenza di valore tra il campo recitato e quello aperto.

Nella Tab. 12 è riportata la distribuzione dei terreni secondo la coltura e in base alla presenza o assenza di recinzione con muretti a secco (chiusura vs campo aperto), con il calcolo delle frequenze di riga e di colonna. Esaminando le percentuali di riga si riscontra come la maggiore incidenza di chiusure si abbia in corrispondenza delle terre coltivate a mandorlo (85,9%) e a mandorlo e olivo (83,1%), mentre piuttosto bassa è la percentuale di chiusure di oliveti (23,7%). Nel complesso l'appezzamento di terra circondato da muretti a secco costituiva il 35,2% del totale dei terreni apprezzati. Passando alle percentuali di colonna, il 38,3% del totale delle chiusure riguarda gli oliveti, seguono mandorlo e olivo (24,8%) e il solo mandorlo (20%). Nel complesso tali colture arboree interessano l'83,1% della distribuzione fondiaria.

**Tabella 12**. *Distribuzione degli appezzamenti in relazione alla tipologia colturale e alla distinzione tra terreni chiusi e aperti*

| Coltura | Valori assoluti | | | Percentuali di riga | | | Percentuali di colonna | | |
|---|---|---|---|---|---|---|---|---|---|
| | Tipologia di terreno | | Totale | Tipologia di terreno | | Totale | Tipologia di terreno | | Totale |
| | Chiuso | Aperto | | Chiuso | Aperto | | Chiuso | Aperto | |
| Oliveto | 114 | 368 | 482 | *23,7* | *76,3* | *100* | *38,3* | *67,2* | *57,0* |
| Mandorlo e olivo | 74 | 15 | 89 | *83,1* | *16,9* | *100* | *24,8* | *2,7* | *10,5* |
| Mandorlo | 61 | 10 | 71 | *85,9* | *14,1* | *100* | *20,5* | *1,8* | *8,4* |
| Pastino | | 36 | 36 | *0,0* | *100,0* | *100* | *0,0* | *6,6* | *4,3* |
| Orto con alberi | | 26 | 26 | *0,0* | *100,0* | *100* | *0,0* | *4,7* | *3,1* |
| Vigneto | 1 | 23 | 24 | *4,2* | *95,8* | *100* | *0,3* | *4,2* | *2,8* |
| Coltura arborea promiscua | 10 | 11 | 21 | *47,6* | *52,4* | *100* | *3,4* | *2,0* | *2,5* |
| Pero | 5 | 5 | 10 | *50,0* | *50,0* | *100* | *1,7* | *0,9* | *1,2* |
| Macchia | 1 | 4 | 5 | *20,0* | *80,0* | *100* | *0,3* | *0,7* | *0,6* |
| Querceto | | 4 | 4 | *0,0* | *100,0* | *100* | *0,0* | *0,7* | *0,5* |
| Fico | 2 | 1 | 3 | *66,7* | *33,3* | *100* | *0,7* | *0,2* | *0,4* |
| Frutteto | | 3 | 3 | *0,0* | *100,0* | *100* | *0,0* | *0,5* | *0,4* |
| Non indicata | 30 | 42 | 72 | *41,7* | *58,3* | *100* | *10,1* | *7,7* | *8,5* |
| Totale | 298 | 548 | 846 | *35,2* | *64,8* | *100* | *100* | *100* | *100* |

**Tabella 13**. *Valore medio estimale degli appezzamenti in relazione alla tipologia colturale e alla distinzione tra terreni chiusi e aperti* (in tarì)

| Coltura | Tipologia di terreno | | Totale |
|---|---|---|---|
| | Chiuso | Aperto | |
| Orto con alberi | | 15,93 | 15,93 |
| Frutteto | | 12,17 | 12,17 |
| Mandorlo e olivo | 12,29 | 10,58 | 12,00 |
| Coltura arborea promiscua | 11,35 | 12,51 | 11,96 |
| Mandorlo | 10,94 | 17,36 | 11,84 |
| Oliveto | 10,84 | 6,89 | 7,82 |
| Fico | 5,13 | 4,15 | 4,80 |
| Pastino | | 4,67 | 4,67 |
| Vigneto | 2,50 | 4,44 | 4,36 |
| Pero | 1,48 | 3,39 | 2,44 |
| Macchia | 0,50 | 1,98 | 1,68 |
| Querceto | | 0,70 | 0,70 |
| Non indicata | 3,69 | 1,19 | 2,25 |
| Totale | 10,25 | 6,96 | 8,12 |

Nella Tab. 13 si riportano i dati relativi al valore medio di apprezzo delle diverse colture tanto in campo aperto quanto nelle "chiusure". A livello complessivo si noti come i campi chiusi presentino un valore medio di 10,25 tarì contro i 6,96 del campo aperto il che permette di affermare come effettivamente la prima tipologia avesse un valore economico maggiore. Ma ciò non vale per tutte le colture: se

l'apprezzo medio delle chiusure è maggiore per gli appezzamenti a oliveto, a mandorlo e olivo e per i ficheti, non lo è per il mandorleto e le altre colture arboree promiscue. Occorre sempre precisare che tali valori vanno presi sempre come indicativi non potendo prendere in considerazione l'estensione degli appezzamenti, variabile importante nel determinare il valore di un terreno.

### *4.1 L'analisi temporale*

Per la città di Molfetta si dispone di una discreta serie di dati catastali relativi al XVI secolo, già oggetto di studio per la definizione del paesaggio agrario (Poli, 1990). È quindi possibile confrontare i dati ricavati dall'apprezzo del 1417 con quelli elaborati da Poli per alcuni anni del '500 e ricostruirne la dinamica temporale; in entrambe le analisi sono esclusi i patrimoni degli Enti religiosi in quanto esentati. Per poter comparare i dati, quelli del 1417 sono stati riclassificati secondo le categorie utilizzate nel lavoro di Poli. Anche questo Autore ha calcolato le incidenze percentuali delle diverse colture prendendo a riferimento la consistenza numerica degli appezzamenti di terra. I dati sono riportati in Tab. 14 e in Fig. 3.

**Tabella 14**. *Distribuzione percentuale del numero degli apprezzamenti per tipo di coltura. Anni 1417, 1509, 1523, 1542, 1561, 1578*

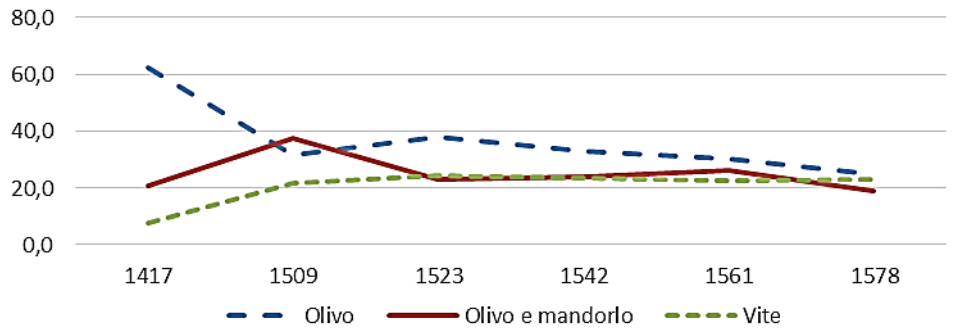
| Coltura | Anno | | | | | |
|---|---|---|---|---|---|---|
| | 1417 | 1509 | 1523 | 1542 | 1561 | 1578 |
| Olivo | 62,3 | 31,5 | 37,9 | 33,0 | 30,3 | 24,7 |
| Olivo e mandorlo | 20,7 | 37,3 | 22,8 | 24,0 | 26,3 | 19,0 |
| Vite | 7,8 | 21,6 | 24,4 | 23,5 | 22,6 | 23,1 |
| Orti e terre vacue | 3,4 | 3,4 | 6,3 | 5,9 | 3,5 | 5,1 |
| Seminativo, difese e macchie | 1,2 | 0,0 | 1,7 | 5,3 | 9,6 | 14,7 |
| Giardini | 0,4 | 0,0 | 2,3 | 1,2 | 1,4 | 0,9 |
| Altra | 4,4 | 6,3 | 4,7 | 7,1 | 6,5 | 12,5 |
| Totale | 100 | 100 | 100 | 100 | 100 | 100 |

Nota: i dati relativi agli anni 1509-1578 sono tratti da Poli, 1990.

Nel corso di poco meno di un secolo l'incidenza dei terreni coltivati ad olivo si sarebbe quindi dimezzata, passando dal 62,3% del 1417 al 31,5% del 1509, divenendo la seconda coltura per diffusione, preceduta dall'oliveto misto al mandorleto con il 37,3%. Il brusco calo della quota degli oliveti tra 1417 e 1509 potrebbe trovare una possibile spiegazione nella riduzione del numero degli appezzamenti di minore estensione (dall'apprezzo si nota una decisa polverizzazione della proprietà fondiaria) a favore della grande proprietà, processo documentato nel corso del Quattrocento per l'agro di Ostuni (Licinio, 1989) dove i piccoli coltivatori vendevano i loro terreni ad Enti religiosi. Non è neppure da escludere l'anticipazione al

XV secolo del fenomeno dell'affollamento contadino della terra descritto da Poli per il XVI secolo: si tenga presente che il numero di fuochi di Molfetta passò dai 223 del 1417 ai 455 del 1443-1447 (Mastrorocco, 2013).

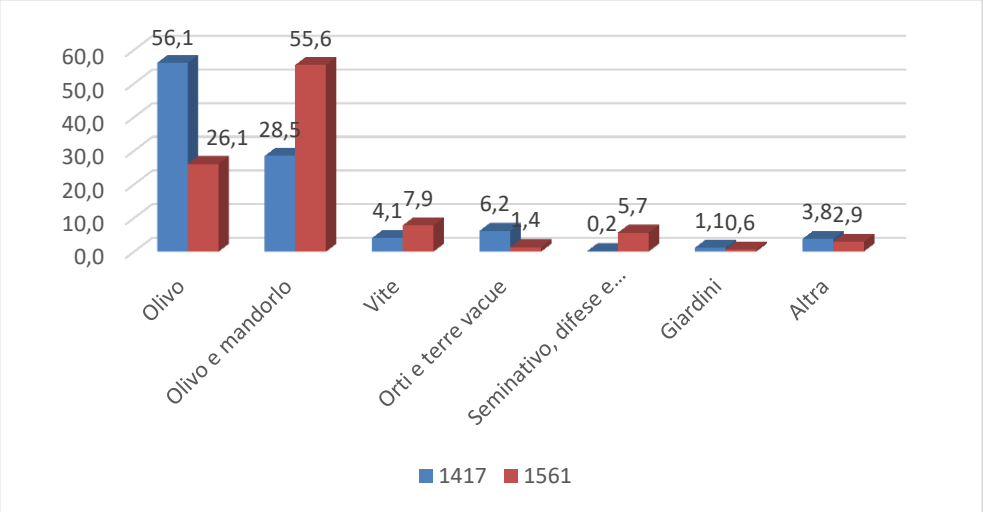
**Figura 3**. *Evoluzione temporale dell'incidenza della quota di terreni coltivati a olivo, mandorlo e olivo e vite dal 1417 al 1578*




In effetti nel corso del XV secolo si assiste ad una redditizia diffusione della coltura del mandorlo in tutta la Terra di Bari; l'associazione dell'olivo con il mandorlo era un binomio produttivo quasi obbligato dall'alternanza biennale della resa di queste due piante (Poli, 1990). Quanto alla viticoltura, la sua diffusione aumenta passando dal 7,8% al 21,6%. Nel corso del XVI secolo si assiste alla riduzione dello spazio destinato all'olivicoltura (Poli, 1990) che nel 1578 segna un 24,7%, restando tuttavia la coltura principale del territorio di Molfetta; anche la mandorlicoltura subisce un ridimensionamento, attestandosi a un livello simile a quello del 1417; stabile, intorno al 20%, nel tempo si mantiene la quota del vigneto (coltura praticata soprattutto dai ceti contadini più poveri, con una notevole parcellizzazione della proprietà e una produzione destinata al consumo locale), mentre nel 1578 si osserva l'incremento delle colture seminative, che raggiungono quasi il 15%, molto probabilmente a causa dell'aumento della domanda locale di prodotti cerealicoli. La riduzione progressiva dello spazio coltivato ad olivo, in termini di numerosità degli appezzamenti, era il risultato dell'aumentata pressione demografica sulla terra e delle scelte produttive imposte su scala locale (Poli, 1990). L'affol–lamento contadino della terra nel corso del tempo comportò la scomparsa degli appezzamenti di minore dimensione: difficoltà improvvise (un raccolto negativo, la morte improvvisa, i debiti, etc.) costringevano il piccolo proprietario ad alienare il proprio fazzoletto di terra, con la conseguente diffusione delle proprietà fondiarie più estese. Questo fenomeno ebbe un ruolo determinante nella riduzione del numero di appezzamenti censiti.

Per il catasto del 1561 (Poli, 1990) sono disponibili i dati relativi al reddito degli appezzamenti di terreno, il che consente una comparazione con i valori estimali del 1417, come riportato in Fig. 4. L'elemento più rilevante è che, in termini economici, nel 1417 l'olivo costituiva la coltura maggioritaria con il 56,1%, mentre nel 1561 era stata soppiantata dal connubio mandorlo/olivo con il 55,6%.

**Figura 4**. *Distribuzione percentuale del reddito o valore d'apprezzo dei terreni per tipo di coltura negli anni 1417 e 1561*




## *4.2 L'analisi spaziale*

Una delle variabili rilevate nell'apprezzo è la contrada rurale nella quale era ubicato l'appezzamento di terra. Questo particolare consente di effettuare un'analisi di tipo spaziale, sebbene occorra precisare che, nonostante nella toponomastica locale rimanga traccia di parte delle località menzionate nel documento del 1417, di altre si è persa completamente la memoria o quanto meno è difficile stabilirne l'esatta collocazione. Con l'analisi geografica si mira a ricostruire la distribuzione topografica delle diverse colture sul territorio molfettese allo scopo di verificare un'eventuale concentrazione spaziale dell'olivicoltura e delle altre tipologie colturali e di comprendere se il valore estimale variava significativamente da una località all'altra per la medesima coltura. Una prima sommaria analisi che si può proporre è il calcolo della percentuale di contrade (in totale sono menzionate 131 località) nelle quali era presente una data coltivazione. Da essa emerge che nel 75% delle contrade vi erano oliveti, nel 28,3% olivo e mandorlo, nel 25,2% mandorleti, nel 21,4% altre colture arboree, nel 13,7% vigneto e nell'8,4% le colture ortive. Già da

queste semplici percentuali si comprende l'ampia diffusione territoriale dell'olivi–coltura, mentre la vite e l'orto risultano concentrate in poche località.

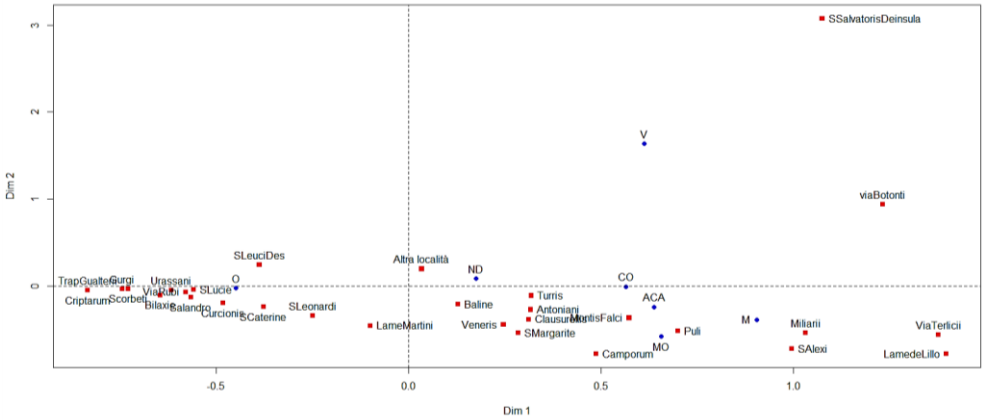
**Tabella 15**. *Distribuzione del numero dei terreni per tipo di coltura e località*

| Località | Colture | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Olivo | Mandorlo | Mandorlo e olivo | Altre colture arboree | Vite | Colture ortive | Non indicata | Totale |
| Antoniani | 7 | 2 | 2 | 4 | | | 5 | 20 |
| Baline | 6 | 1 | 2 | 1 | | | 5 | 15 |
| Bilaxie | 15 | | 1 | | | | 1 | 17 |
| Camporum | 5 | | 9 | | | | | 14 |
| Clausurellis | 4 | 2 | 2 | | | | 2 | 10 |
| Criptarum | 15 | | | | | | | 15 |
| Curcionis | 11 | 1 | 1 | | | | | 13 |
| Cutinarum | 12 | | | | | | | 12 |
| Gurgi | 12 | | | | | | 1 | 13 |
| Lame de Lillo | | 7 | 4 | 6 | | | | 17 |
| Lame Martini | 9 | | 5 | | | | | 14 |
| Lame Patroni | 23 | | | | | | | 23 |
| Miliarii | 1 | 2 | 9 | 1 | 1 | 1 | 2 | 17 |
| Montis Falci | 6 | 5 | 3 | 4 | | | 6 | 24 |
| Puli | 3 | 1 | 4 | 1 | | 3 | | 12 |
| Salandro | 10 | | 1 | | | | 1 | 12 |
| Sancte Caterine | 8 | 1 | 1 | | | | | 10 |
| Sancte Lucie | 16 | | | 1 | | | 3 | 20 |
| Sancte Margarite | 5 | 2 | 3 | | | | | 10 |
| Sancti Alexi | 2 | 5 | 4 | | | | 1 | 12 |
| Sancti Benedicti | 15 | | | | | | | 15 |
| Sancti Leonardi | 15 | | 5 | 1 | | | | 21 |
| Sancti Leuci deserti | 15 | 1 | | 1 | 2 | | | 19 |
| Sancti Salvatoris de insula | | | | | 12 | | 1 | 13 |
| Sancti Simeonis | | 6 | 1 | | | 2 | | 9 |
| Scorbeti | 10 | | | | | | 1 | 11 |
| Trappeti quondam dompni Gualterii | 10 | | | | | | | 10 |
| Turris | 5 | 1 | 3 | 1 | 1 | | 2 | 13 |
| Urassani | 17 | | | 1 | | | 2 | 20 |
| Veneris | 12 | 7 | 3 | | | | | 22 |
| Via Terlicii | | 5 | 1 | 2 | | 3 | | 11 |
| Via Botonti | | 3 | 1 | 1 | 4 | | 1 | 10 |
| Via Rubi | 21 | | | 2 | | | 2 | 25 |
| Altra località | 191 | 18 | 24 | 19 | 40 | 17 | 35 | 344 |
| Totale | 482 | 71 | 89 | 46 | 60 | 26 | 72 | 846 |

Nella Tab. 15 si riporta la distribuzione del numero degli appezzamenti per località e tipo di coltura; sono state prese in considerazione le contrade con almeno

10 osservazioni o un valore apprezzato complessivo di 100 tarì in modo da tenere in qualche considerazione anche l'estensione dei terreni, nell'ipotesi che da questa variabile dipendesse una parte consistente del valore d'apprezzo; le restanti contrade sono state aggregate nella modalità "Altra località".

**Figura 5**. *Grafico dell'analisi delle corrispondenze Ubicazione geografica vs Tipo di coltura*




Si osservi, nella Tab. 15, la presenza esclusiva di oliveti nelle contrade: *Criptarum, Cutinarum, Lame Patroni, Sancti Benedicti, Trappeti quondam dompni Gualterii*. Alcune di queste già nel nome rivelano un chiaro riferimento alla produzione olearia: *Criptarum* che indicava la presenza di grotte nelle quali verosimilmente avveniva la lavorazione dell'olio e *Trappeti quondam dompni Gualterii*, chiaro riferimento alla presenza di un frantoio. Le contrade *Lame Patroni* (4,8%) e *Via Rubi* (4,4%) sono quelle che presentano la maggiore incidenza percentuale sul totale degli oliveti. Nella Fig. 5 è riportato il *biplot* dell'analisi delle corrispondenze (Delvecchio, 1992) che sintetizza l'esame dei profili di riga e di colonna rinvenienti dalla Tab. 15. Si notino le tante contrade caratterizzate dalla olivicoltura (O), le poche segnate dalla presenza dal vigneto (V), e quelle a prevalente diffusione del mandorleto (M).

L'analisi geografica può raffinarsi con il ricorso al quoziente di localizzazione e agli indici di specializzazione e di concentrazione territoriale (Marbach, 1991):

$$L_i^j = \frac{k_{ij}/\sum_{j=1}^{c} k_{ij}}{\sum_{i=1}^{l} k_{ij}/\sum_{i=1}^{l}\sum_{j=1}^{c} k_{ij}}$$

$$S_i = \frac{1}{2}\sum_{j=1}^{c}\left|\frac{k_{ij}}{\sum_{j=1}^{c} k_{ij}} - \frac{\sum_{i=1}^{l} k_{ij}}{\sum_{i=1}^{l}\sum_{j=}^{c} k_{ij}}\right|$$

**Tabella 16**. *Quozienti di localizzazione delle colture nelle contrade di Molfetta*

| Località | Colture | | | | | | |
|---|---|---|---|---|---|---|---|
| | Olivo | Mandorlo | Mandorlo e olivo | Altre colture arboree | Vite | Colture ortive | Non indicata |
| Antoniani | 0,61 | 1,19 | 0,95 | **3,68** | - | - | 2,94 |
| Baline | 0,70 | 0,79 | 1,27 | 1,23 | - | - | **3,92** |
| Bilaxie | 1,55 | - | 0,56 | - | - | - | 0,69 |
| Camporum | 0,63 | - | **6,11** | - | - | - | - |
| Clausurellis | 0,70 | **2,38** | 1,90 | - | - | - | 2,35 |
| Criptarum | **1,76** | - | - | - | - | - | - |
| Curcionis | 1,49 | 0,92 | 0,73 | - | - | - | - |
| Cutinarum | **1,76** | - | - | - | - | - | - |
| Gurgi | 1,62 | - | - | - | - | - | 0,90 |
| Lame de Lillo | 0 | **4,91** | 2,24 | **6,49** | - | - | - |
| Lame Martini | 1,13 | - | **3,39** | - | - | - | - |
| Lame Patroni | **1,76** | - | - | - | - | - | - |
| Miliarii | 0,10 | 1,40 | **5,03** | 1,08 | 0,83 | 1,91 | 1,38 |
| Montis Falci | 0,44 | 2,48 | 1,19 | **3,07** | - | - | 2,94 |
| Puli | 0,44 | 0,99 | **3,17** | 1,53 | - | **8,13** | - |
| Salandro | 1,46 | - | 0,79 | - | - | - | 0,98 |
| Sancte Caterine | 1,40 | 1,19 | 0,95 | - | - | - | - |
| Sancte Lucie | 1,40 | - | - | 0,92 | - | - | 1,76 |
| Sancte Margarite | 0,88 | 2,38 | 2,85 | - | - | - | - |
| Sancti Alexi | 0,29 | **4,96** | **3,17** | - | - | - | 0,98 |
| Sancti Benedicti | **1,76** | - | - | - | - | - | - |
| Sancti Leonardi | 1,25 | - | 2,26 | 0,88 | - | - | - |
| Sancti Leuci deserti | 1,39 | 0,63 | - | 0,97 | 1,48 | - | - |
| Sancti Salvatoris de insula | - | - | - | - | **13,02** | - | 0,90 |
| Sancti Simeonis | - | **7,94** | 1,06 | - | - | **7,23** | - |
| Scorbeti | 1,60 | - | - | - | - | - | 1,07 |
| Trappeti quondam dompni Gualterii | **1,76** | - | - | - | - | - | - |
| Turris | 0,68 | 0,92 | 2,19 | 1,41 | 1,08 | - | 1,81 |
| Urassani | 1,49 | - | - | 0,92 | - | - | 1,18 |
| Veneris | 0,96 | **3,79** | 1,30 | - | - | - | - |
| Via Terlicii | - | **5,42** | 0,86 | **3,34** | - | **8,87** | - |
| Via Botonti | - | **3,57** | 0,95 | 1,84 | **5,64** | - | 1,18 |
| Via Rubi | 1,47 | - | - | 1,47 | - | - | 0,94 |
| Altra località | 0,97 | 0,62 | 0,66 | 1,02 | 1,64 | 1,61 | 1,20 |

$$C_j = \frac{1}{2}\sum_{i=1}^{l}\left|\frac{k_{ij}}{\sum_{i=1}^{l} k_{ij}} - \frac{\sum_{j=1}^{c} k_{ij}}{\sum_{i=1}^{l}\sum_{j=}^{c} k_{ij}}\right|$$

dove $L_i^j$ è il quoziente di localizzazione della coltura j.ma nella località i.ma, $S_i$ è l'indice di specializzazione della località i.ma, $C_j$ è l'indice di concentrazione territoriale della coltura j.ma, $k_{ij}$ è il numero di appezzamenti terra nella località i.ma con la coltura j.ma, l è il numero totale di località e c è il numero totale di colture.

Nella Tab. 16 sono riportati i quozienti di localizzazione delle singole colture nelle diverse località. Tale quoziente confronta, per ogni coltura, l'incidenza percentuale della località j.ma con quella di tutto il territorio di Molfetta; un valore superiore all'unità indica la specializzazione di quella coltura nella località considerata. Più elevato è il quoziente, maggiore è la distanza dal profilo medio di Molfetta e maggiore è la specializzazione. In grassetto sono evidenziati i valori più elevati.

Si noti che i quozienti di localizzazione dell'olivicoltura non sono mai superiori a 2, il che conferma come questa coltura non abbia una particolare specializzazione sul territorio di Molfetta, ma sia abbastanza diffusa, seppure con diverse gradazioni. I dati relativi alla mandorlicoltura indicano invece l'esistenza di diverse località specializzate: in alcune di esse il quoziente di localizzazione raggiunge e supera il valore 4 (*Lame de Lillo*, *Sancti Alexi*, *Via Terlicii*, *Sancti Simeonis*). Valori elevati anche per la coltura promiscua mandorlo e olivo, con due contrade che registrano un quoziente superiore a 5: *Camporum* e *Miliarii*. Tra le altre colture arboree si contraddistingue la contrada *Lame de Lillo* con quoziente uguale a 6,49, mentre a *Sancti Salvatoris de insula* il vigneto supera quota 13 e le colture ortive a *Sancti Simeonis*, *Puli* e *Via Terlicii* si attestano oltre quota 7.

**Tabella 17**. *Indici di specializzazione delle contrade di Molfetta*

| Località | Grado di spec. | Località | Grado di spec. | Località | Grado di spec. |
|---|---|---|---|---|---|
| Antoniani | 0,327 | Miliarii | **0,523** | Sancti Simeonis | **0,780** |
| Baline | 0,289 | Montis Falci | 0,421 | Scorbeti | 0,345 |
| Bilaxie | 0,313 | Puli | 0,476 | Trappeti quondam dompni Gualterii | 0,430 |
| Camporum | **0,538** | Salandro | 0,264 | Turris | 0,223 |
| Clausurellis | 0,326 | Sancte Caterine | 0,246 | Urassani | 0,295 |
| Criptarum | 0,430 | Sancte Lucie | 0,295 | Veneris | 0,265 |
| Curcionis | 0,276 | Sancte Margarite | 0,311 | Via Terlicii | **0,740** |
| Cutinarum | 0,430 | Sancti Alexi | **0,561** | Via Botonti | **0,606** |
| Gurgi | 0,353 | Sancti Benedicti | 0,430 | Via Rubi | 0,296 |
| Lame de Lillo | **0,757** | Sancti Leonardi | 0,277 | Altra località | 0,080 |
| Lame Martini | 0,325 | Sancti Leuci deserti | 0,254 | | |
| Lame Patroni | 0,430 | Sancti Salvatoris de insula | **0,852** | | |

Nella Tab. 17 sono riportati gli indici di specializzazione delle contrade rurali di Molfetta risultanti dall'apprezzo; in grassetto sono evidenziati i valori maggiori di 0,5. Tale indice può variare tra 0 nel caso di assenza di specializzazione, ovvero quando il profilo della località è esattamente identico a quello di tutto il territorio di Molfetta, e 1 nel caso di massima specializzazione. Solo poche contrade presentano un grado di specializzazione significativo: in particolare *Sancti Salvatoris de insula* (viticoltura), *Lame de Lillo* (mandorlicoltura e altre colture arboree), *Via Terlicii* (colture ortive) e *Sancti Simeonis* (mandorlicoltura e colture ortive). Quanto agli indici di concentrazione territoriale delle colture, i relativi valori sono riportati nella Tab. 18. L'indice varia tra 0, in assenza di concentrazione, e 1, che equivale alla massima concentrazione. L'olivicoltura registra il più basso valore di concentrazione, a ulteriore conferma della sua diffusione su tutto il territorio, mentre le colture ortive e la viticoltura appaiono concentrate in alcune contrade.

**Tabella 18**. *Indice di concentrazione territoriale delle colture*

| Coltura | Indice di concentrazione |
|---|---|
| Olivo | 0,166 |
| Mandorlo | 0,464 |
| Mandorlo e olivo | 0,390 |
| Altre colture arboree | 0,311 |
| Vite | **0,510** |
| Colture ortive | **0,534** |
| Non indicata | 0,298 |

Nella Tab. 19 si riportano i valori estimali medi per coltura e ubicazione geografica, da cui emerge una certa eterogeneità interna alle colture nelle diverse contrade rurali. Quelle con i valori complessivamente più alti sono *Sancti Simeonis*, *Sancti Benedicti*, *Sancte Margarite*, *Curcionis*, e *Sancti Alexi*.

**Tabella 19**. *Importo medio del valore apprezzato per tipo di coltura e ubicazione geografica* (in tarì)

| Località | Colture | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Olivo | Mandorlo | Mandorlo e olivo | Altre colture arboree | Vite | Colture ortive | Non indicata | Totale |
| Antoniani | 2,38 | 2,43 | 3,55 | 1,51 | | | 3,30 | 2,56 |
| Baline | 9,50 | 7,10 | 8,25 | 1,50 | | | 1,13 | 5,85 |
| Bilaxie | 5,72 | | 3,00 | | | | 1,50 | 5,28 |
| Camporum | 7,95 | | 5,69 | | | | | 6,50 |
| Clausurellis | 3,94 | 5,63 | 7,88 | | | | 1,25 | 4,53 |
| Criptarum | 6,72 | | | | | | | 6,72 |
| Curcionis | 12,25 | 21,00 | 15,75 | | | | | 13,19 |
| Cutinarum | 5,58 | | | | | | | 5,58 |
| Gurgi | 11,00 | | | | | | 1,50 | 10,27 |
| Lame de Lillo | | 9,11 | 13,13 | 9,17 | | | | 10,08 |
| Lame Martini | 4,66 | | 15,90 | | | | | 8,68 |
| Lame Patroni | 11,78 | | | | | | | 11,78 |
| Miliarii | 10,50 | 9,38 | 10,11 | 30,00 | 8,25 | 25,50 | 1,75 | 11,03 |
| Montis Falci | 4,61 | 10,20 | 9,00 | 1,73 | | | 1,05 | 5,12 |
| Puli | 4,00 | 7,15 | 11,81 | 32,00 | | 8,94 | | 10,44 |
| Salandro | 8,06 | | 9,00 | | | | 1,00 | 7,50 |
| Sancte Caterine | 8,59 | 15,75 | 15,75 | | | | | 10,03 |
| Sancte Lucie | 4,20 | | | 0,75 | | | 3,08 | 3,86 |
| Sancte Margarite | 9,15 | 27,38 | 14,75 | | | | | 14,48 |
| Sancti Alexi | 17,25 | 10,17 | 19,06 | | | | 3,00 | 13,71 |
| Sancti Benedicti | 14,52 | | | | | | | 14,52 |
| Sancti Leonardi | 7,43 | | 14,20 | 0,50 | | | | 8,71 |
| Sancti Leuci deserti | 13,05 | 5,00 | | 22,00 | 1,75 | | | 11,91 |
| Sancti Salvatoris de insula | | | | | 5,62 | | 1,25 | 5,28 |
| Scorbeti | 5,37 | | | | | | 10,50 | 5,83 |
| Trappeti quondam domp-ni Gualterii | 7,95 | | | | | | | 7,95 |
| Turris | 11,05 | 13,50 | 3,00 | 15,00 | 9,00 | | 0,93 | 7,97 |
| Urassani | 5,56 | | | 7,50 | | | 6,25 | 5,73 |
| Veneris | 5,20 | 15,82 | 19,38 | | | | | 10,09 |
| Via Terlicii | | 15,20 | 3,00 | 1,88 | | 16,00 | | 11,89 |
| Via Botonti | | 17,60 | 5,25 | 4,15 | 5,25 | | 0,15 | 8,34 |
| Via Rubi | 8,46 | | | 2,54 | | | 1,75 | 7,45 |
| Altra località | 7,37 | 11,07 | 15,08 | 7,76 | 4,09 | 16,51 | 2,23 | 7,79 |
| Totale | 7,82 | 11,84 | 12,00 | 7,34 | 4,54 | 15,93 | 2,25 | 8,12 |

## 5. Il valore estimale come *proxy* della ricchezza: l'analisi della disuguaglianza

Aggregando il valore estimale dei beni di ciascun contribuente si ottiene il valore estimale complessivo dell'unità fiscale. In termini di formule il valore estimale del contribuente i.mo è:

$$y_i = \sum_{j=1}^{k_i} x_{ij}$$

questa variabile sarà utilizzata come *proxy* della ricchezza del contribuente, in quanto essa può essere a ragione ritenuta proporzionale al patrimonio. La formula dell'estimo medio è la seguente:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

In letteratura i dati di apprezzi, estimi e catasti vengono utilizzati per misurare indirettamente la ricchezza delle famiglie e studiarne la distribuzione e la disuguaglianza (Alfani, 2015). Ovviamente si tratta di una forma di distribuzione troncata in quanto gli esenti (Enti religiosi *in primis*), i nullatenenti e i possessori della sola casa di abitazione erano esclusi dalla rilevazione e questo può comportare una sottostima o una sovrastima delle misure di disuguaglianza, come nel caso dell'indice di Gini. Inoltre, questo tipo di documento fiscale prende prevalentemente in considerazione la ricchezza fondiaria, trascurando quella mobiliare o relegandola a un ruolo marginale. Occorre osservare che in un'economia quasi esclusivamente agricola come quella medievale i beni immobili rappresentavano la principale fonte di ricchezza per la maggior parte degli abitanti, soprattutto in un contesto come quello dell'Italia meridionale (Alfani, 2015). Nonostante i suaccennati limiti, questa tipologia di dati probabilmente costituisce l'unica fonte per studiare la ricchezza e la sua distribuzione nelle società di *ancien régime*. L'apprezzo molfettese del 1417 rappresenta un caso molto raro nel Mezzogiorno medievale italiano, mentre nel Centro-Nord della Penisola questo tipo di documentazione fiscale era abbastanza diffuso anche nei secoli precedenti il XV.

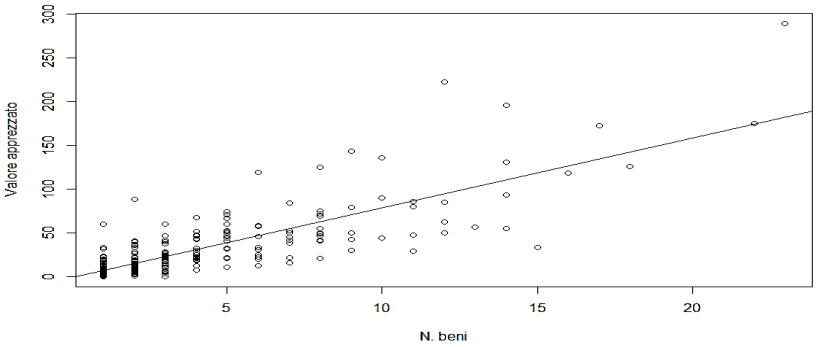
In primo luogo, si vuole esaminare la relazione esistente tra la variabile $y_i$ e il numero di beni apprezzati $k_i$. Dal grafico riportato nella Fig. 6 emerge l'esistenza di un legame di tipo lineare crescente, ma anche una certa variabilità di $y_i$ in corrispondenza del medesimo numero di beni posseduti. Si sono stimati con il metodo dei minimi quadrati i parametri della seguente retta di regressione:

$$y_i = a + bk_i$$

ottenendo: a=-0,34 e b=7,93 con un coefficiente di determinazione $R^2$=0,64 che mostra come il legame tra le due variabili sia sufficientemente forte. In sostanza possiamo affermare che ad ogni bene aggiuntivo il valore apprezzato del patrimonio del contribuente aumenta di circa 8 tarì. Il numero dei beni tuttavia spiega solo

il 64% della variabilità, mentre il restante 36% è imputabile a fattori di tipo qualitativo insiti nella natura del bene apprezzato: si tratta di quelle variabili che influenzano il valore economico del bene ma che non sono direttamente esplicitate dall'apprezzo (pregio, produttività, estensione, etc.).

**Figura 6**. *Valore dei beni apprezzati di ciascun contribuente vs numero di beni apprezzati*




Si è visto che la variabile $x_{ij}$ segue la distribuzione log-normale e si ha ragione di ritenere che, essendo:

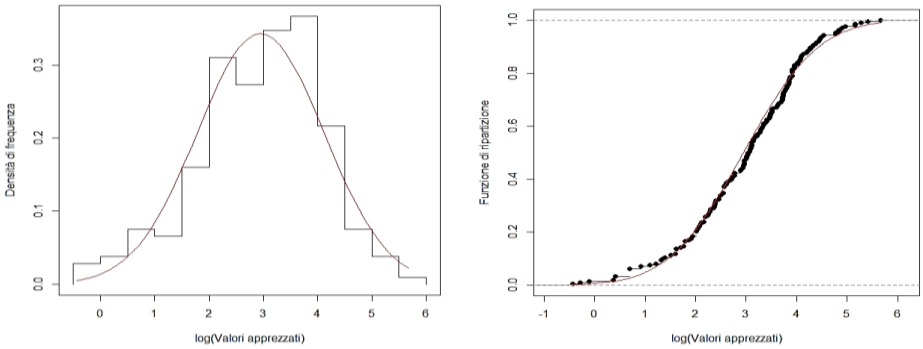
$$y_i = \sum_{j=1}^{k_i} x_{ij}$$

anche la variabile $y_i$ possa essere rappresentata analiticamente dalla stessa distribuzione teorica. Infatti, se consideriamo la trasformazione logaritmica:

$$w_i = \log(y_i)$$

questa presenta una distribuzione di tipo normale, come si vede dalla Fig. 7 e con $\gamma_1 = -0{,}4$ e $\gamma_2 = -0{,}05$.

**Figura 7**. *Distribuzione di frequenza e funzione di ripartizione del logaritmo del valore di apprezzo dei singoli contribuenti*



Le stime di massima verosimiglianza dei parametri del modello log-normale sono $\hat{\mu} = 2{,}949$ e $\hat{\sigma} = 1{,}161$; se si calcola l'indice di scostamento tra frequenze empiriche e quelle teoriche calcolate con la funzione log-normale si ottiene $\delta = 16\%$. Un indice di adattamento molto buono, particolare che conferma come la distribuzione dei valori estimali sia una *proxy* della ricchezza, variabile di tipo economico che segue la distribuzione log-normale.

Nella Tab. 20 sono presentati alcuni indicatori descrittivi del valore estimale dei contribuenti, segmentati secondo alcune variabili classificatorie. Nell'analisi per genere si nota come i contribuenti maschi registrino l'importo estimale medio più elevato, ma anche la variabilità e la disuguaglianza (misurata dall'indice di Gini G) maggiori.

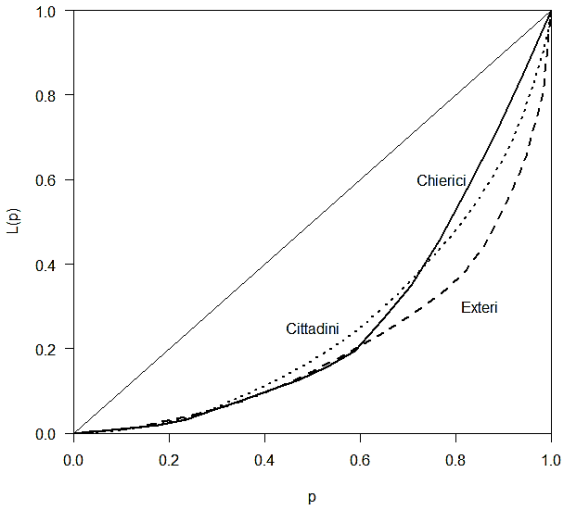
Scindendo l'analisi per tipologia di contribuente, si evince che i cittadini di Molfetta godevano di un valore estimale medio quasi doppio rispetto a chierici e forestieri, mentre gli *exteri* avevano il maggior grado di concentrazione. Si può affermare che i chierici non erano poi tanto più ricchi rispetto alla media degli altri cittadini, mentre per i forestieri la distribuzione presenta una maggiore disuguaglianza. Si noti come il livello di disuguaglianza nella distribuzione dei valori estimali, pari a 0,53, sia superiore a quello della distribuzione del numero dei beni (G=0,457), indice di una maggiore concentrazione.

**Tabella 20**. *Statistiche sul valore totale di apprezzo dei contribuenti secondo alcune variabili di classificazione* (in tarì)

| Variabile di classificazione | N. | *%* | Media | Dev.Std. | Min | Max | G |
|---|---|---|---|---|---|---|---|
| *Genere* | | | | | | | |
| M | 167 | *78,4* | 37,3 | 43,25 | 0,65 | 288,85 | 0,531 |
| F | 41 | *19,2* | 22,0 | 20,61 | 0,75 | 88,50 | 0,484 |
| n. d. | 5 | *2,3* | 23,5 | 15,06 | 6,00 | 50,25 | 0,342 |
| *Tipologia di contribuente* | | | | | | | |
| Chierico | 17 | *8,0* | 20,80 | 17,99 | 2,00 | 52,60 | 0,472 |
| Cittadino | 140 | *65,7* | 41,05 | 42,47 | 0,65 | 288,85 | 0,485 |
| Extero | 56 | *26,3* | 20,46 | 32,99 | 1,50 | 222,50 | 0,575 |
| *Provenienza* | | | | | | | |
| Molfetta | 153 | *71,8* | 39,76 | 41,17 | 0,65 | 288,85 | 0,483 |
| Altra | 60 | *28,2* | 19,38 | 32,13 | 1,50 | 222,50 | 0,583 |
| Totale | 213 | *100,0* | 34,02 | 39,90 | 0,65 | 288,85 | 0,530 |

Nella Fig. 8 sono riportati i grafici delle curve di Lorenz distinte in base alla tipologia del contribuente.

**Figura 8**. *Curve di Lorenz per tipologia di contribuente*




Come si è accennato nelle pagine iniziali, l'apprezzo del 1417 risulta mutilo di alcune pagine ed è possibile stimare in 8 il numero dei contribuenti (fuochi) mancanti: dal valore complessivo dell'intero apprezzo, riportato in coda al documento, si sottrae la somma dei singoli apprezzi effettivamente esistenti, poi tale differenza viene divisa per il valore medio estimale ottenendo un'attendibile stima del numero di contribuenti compresi nelle pagine perdute. In tal modo il numero di fuochi che verosimilmente erano oggetto di apprezzo nel 1417 ammontava a 213 (fuochi rilevati con $y_i > 0$), a cui si aggiungono 2 fuochi con $y_i=0$ e altri 8 stimati nel modo sopra indicato per un totale di 223 fuochi. Considerando una media di 4,75 (Mastrorocco, 2013) individui per fuoco si ha la stima della popolazione pari a 1.059 abitanti. Stime demografiche più accurate propendono per una popolazione inferiore al migliaio nel XV secolo (Malanima, 1998).

Negli ultimi anni si è sviluppato un notevole interesse per lo studio e la misurazione della disuguaglianza nella distribuzione del reddito e della ricchezza in età preindustriale. Ne è un esempio il progetto EINITE (*Economic Inequality across Italy and Europe, 1300-1800*) con il quale si è creata una base di dati relativa a diversi Paesi europei per un periodo che va dal tardo Medioevo sino agli anni precedenti la Rivoluzione industriale. Dall'esame di alcune recenti pubblicazioni (Alfani, Sardone, 2015) emerge che per il Mezzogiorno continentale italiano i primi dati disponibili risalgono solo al XVI secolo, mentre per altri contesti nazionali (ad esempio Toscana, Piemonte) si dispone di informazioni risalenti al XIV secolo. In un precedente lavoro (Ricci, 2017), per sopperire alla mancanza di dati fiscali, si è proposto di utilizzare come *proxy* della ricchezza le donazioni *pro anima* in denaro

di due obituari del Trecento relativi a Ruvo di Puglia e a Giovinazzo. Tralasciando Ruvo, centro di minori dimensioni, sembra pertinente un confronto tra Molfetta e Giovinazzo, città con molte caratteristiche simili (Ricci, 2017). L'indice di Gini stimato con le donazioni del necrologio di Giovinazzo è risultato pari a 0,64, mentre quello di Molfetta si è attestato a 0,53, con una contrazione della disuguaglianza dopo la Peste Nera che imperversò tra il 1347 e il 1352 e dopo le nefaste conseguenze delle guerre dinastiche interne alla casa d'Angiò che videro la Puglia come principale terreno di scontro. Entrambi i fattori ebbero un effetto perequatore nella distribuzione della ricchezza (Alfani, 2015; Ammannati, 2015). Si consideri, inoltre, che a Giovinazzo, prima della Peste Nera, il 64,6% della ricchezza era in mano al 10% dei contribuenti; nella Molfetta degli inizi del Quattrocento quella stessa percentuale è del 37,5%. La crisi demografica trecentesca causò un incremento dei salari, e quindi un aumento del potere di acquisto dei ceti popolari che rese più accessibile l'acquisizione di beni immobili, e una riduzione delle rendite fondiarie che comportò un parziale indebolimento dei ceti più elevati.

**Tabella 21.** *Molfetta (1417) e alcuni centri delle Sei Miglia (1411-1413): un'analisi comparativa della distribuzione percentuale dei valori apprezzati secondo il decile di appartenenza dei contribuenti*

| Decile di contribuenti | Molfetta | Sei miglia + comuni suburbani | Solo pievieri delle Sei miglia | Solo comuni suburbani | Arliano | Compito | Santo Stefano | Segromigno | Lunata |
|---|---|---|---|---|---|---|---|---|---|
| D1 | 0,70 | 0,32 | 0,30 | 0,86 | 0,18 | 0,18 | 0,35 | 0,56 | 0,60 |
| D2 | 1,75 | 0,87 | 0,88 | 1,39 | 0,65 | 0,53 | 0,81 | 1,34 | 1,09 |
| D3 | 2,64 | 1,76 | 1,84 | 2,46 | 1,86 | 1,41 | 1,21 | 2,37 | 1,34 |
| D4 | 3,67 | 2,83 | 2,98 | 3,30 | 3,05 | 2,71 | 2,67 | 3,75 | 2,96 |
| D5 | 5,68 | 4,37 | 4,62 | 4,53 | 4,62 | 4,53 | 5,82 | 5,35 | 3,86 |
| D6 | 7,05 | 6,77 | 6,95 | 6,07 | 7,09 | 7,35 | 10,03 | 8,45 | 6,98 |
| D7 | 10,42 | 9,61 | 9,67 | 6,79 | 10,56 | 10,15 | 11,84 | 10,76 | 13,34 |
| D8 | 13,26 | 13,33 | 13,46 | 12,03 | 14,52 | 12,91 | 12,99 | 13,85 | 11,66 |
| D9 | 17,36 | 19,78 | 19,70 | 21,00 | 19,66 | 22,51 | 19,94 | 19,87 | 24,83 |
| D10 | 37,48 | 40,35 | 39,61 | 41,55 | 37,82 | 37,71 | 34,35 | 33,71 | 33,36 |
| D | | 0,054 | 0,047 | 0,079 | 0,041 | 0,057 | 0,071 | 0,049 | 0,104 |
| G | 0,530 | 0,586 | 0,579 | 0,585 | 0,575 | 0,587 | 0,546 | 0,522 | 0,574 |
| N. fuochi | 223 | | | 235 | 122 | 105 | 96 | 182 | 104 |

Nota: i dati delle località toscane sono tratti da Ammannati, 2015.

È possibile effettuare alcune comparazioni tra la distribuzione della ricchezza (stimata attraverso la *proxy* del valore estimale) a Molfetta e in alcune località dell'Italia Centro-settentrionale studiate da altri Autori. I confronti hanno riguarda-

to non solo l'indice di Gini (G), ma anche la distribuzione delle percentuali secondo i decili dei contribuenti con l'utilizzo dell'indice di dissomiglianza:

$$D = \frac{1}{2}\sum_{i=1}^{s}|f_{1i} - f_{2i}|$$

dove $f_{1i}$ e $f_{2i}$ sono le frequenze relative delle due distribuzioni confrontate; D varia tra 0, nel caso in cui le due distribuzioni sono perfettamente somiglianti, e 1 nel caso di massima dissomiglianza.

Nella Tab. 21 si propone un confronto con alcune località delle Sei Miglia, area geografica nei dintorni di Lucca (si sviluppava secondo tre anelli concentrici: contrade, comuni suburbani, "pivieri" entro le sei miglia), a inizio XV secolo (Ammannati, 2015). Si può subito osservare come il coefficiente G di Molfetta non sia molto diverso da quello dei centri rurali delle Sei Miglia. Analogo risultato proviene dall'analisi della concentrazione della ricchezza nelle mani del decile di contribuenti maggiormente benestanti (D10). L'indice di dissomiglianza più basso si registra nel confronto con i soli "pivieri" delle Sei Miglia (D=0,047, anche in diversi "pivieri" collinari la coltura principale era rappresentata dall'olivo) e dei centri di Arliano (D=0,041) e Segromigno (D=0,049), sebbene occorra precisare che queste località censivano un minor numero di fuochi rispetto a Molfetta.

**Tabella 22.** *Molfetta e alcuni centri del Piemonte all'inizio del Quattrocento: un'analisi comparativa della distribuzione percentuale dei valori apprezzati secondo il decile di appartenenza dei contribuenti*

| Decile di contribuenti | Molfetta | Chieri | Moncalieri | Cherasco |
|---|---|---|---|---|
| D1 | 0,70 | 0,53 | 0,38 | 0,85 |
| D2 | 1,75 | 1,18 | 0,89 | 1,78 |
| D3 | 2,64 | 1,90 | 1,54 | 2,83 |
| D4 | 3,67 | 2,59 | 2,21 | 3,94 |
| D5 | 5,68 | 3,47 | 3,25 | 5,09 |
| D6 | 7,05 | 4,57 | 4,96 | 6,80 |
| D7 | 10,42 | 6,27 | 7,23 | 8,54 |
| D8 | 13,26 | 9,05 | 10,48 | 11,49 |
| D9 | 17,36 | 14,03 | 16,60 | 16,97 |
| D10 | 37,48 | 56,41 | 52,46 | 41,72 |
| | | | | |
| D | | 0,19 | 0,15 | 0,05 |
| G | 0,530 | 0,669 | 0,663 | 0,550 |
| N. abitanti | 1.059 | 6.700 | 3.830 | 3.570 |
| Anno | 1417 | 1437 | 1412 | 1395-1415 |

Nota: i dati delle località piemontesi sono tratti da Alfani, 2015.

Nella Tab. 22 il confronto è effettuato con alcuni centri del Piemonte, nello stesso frangente storico in cui fu redatto l'apprezzo di Molfetta. In questo caso sia l'indice di Gini che la distribuzione percentuale dei decili tra i contribuenti della cittadina pugliese sono distanti da quelli di Chieri e Moncalieri, realtà urbane di più ampie dimensioni; una maggiore similitudine si rileva con Cherasco, centro urbano di minori dimensioni.

**Tabella 23.** *Molfetta e alcuni centri della Toscana. Un'analisi comparativa dell'indice di Gini*

| Centro | Tipologia | Popolazione a inizio XV secolo | Anno | G |
|---|---|---|---|---|
| Antella | Rurale | 584 | 1394 | 0,523 |
| Borgo S. Lorenzo | Rurale | 1.068 | 1402 | 0,665 |
| Castel S. Giovanni | Rurale | 1.466 | 1402 | 0,592 |
| Castelfiorentino | Rurale | 1.353 | 1402 | 0,574 |
| Cerreto Guidi | Rurale | 946 | 1402 | 0,433 |
| Gambassi | Rurale | 618 | 1402 | 0,546 |
| Monterappoli | Rurale | 708 | 1402 | 0,567 |
| Poggibonsi | Rurale | 1.393 | 1402 | 0,528 |
| San Godenzo | Rurale | 698 | 1402 | 0,468 |
| S. Maria Impruneta | Rurale | 426 | 1402 | 0,449 |
| S. Martino alla Palma | Rurale | 589 | 1402 | 0,526 |
| Prato | Urbano | 3.533 | 1428 | 0,683 |
| San Gimignano | Urbano/Rurale | 3.138 | 1419 | 0,634 |
| Arezzo | Urbano | 4.152 | 1390 | 0,481 |
| Molfetta | Urbano/Rurale | 1.059 | 1417 | 0,530 |

Nota: i dati delle località toscane sono tratti da Alfani, Ammannati 2017.

Nella Tab. 23 si è comparato l'indice di Gini calcolato per Molfetta con quelli di alcuni centri della Toscana. Anche in questo caso la misura della disuguaglianza della cittadina pugliese appare sensibilmente diverso da quelli rilevati in grossi centri come Prato, Arezzo e San Gimignano, mentre è abbastanza prossimo ad alcuni centri rurali con dimensioni demografiche simili.

Una analoga tendenza emerge dalla Tab. 24, da cui si rileva una dissomiglianza nella distribuzione per decili di contribuenti; una certa somiglianza si verifica per il dato medio del Contado di Firenze, aggregato di località rurali limitrofe.

**Tabella 24.** *Molfetta e alcuni centri della Toscana agli inizi del Quattrocento. Un'analisi comparativa della distribuzione percentuale dei valori apprezzati secondo il decile di appartenenza dei contribuenti*

| Decile di contribuenti | Molfetta | Contado di Firenze | Arezzo | Prato | San Gimignano |
|---|---|---|---|---|---|
| D1 | 0,70 | 1,24 | 2,08 | 0,61 | 0,43 |
| D2 | 1,75 | 1,64 | 3,54 | 0,61 | 1,08 |
| D3 | 2,64 | 2,57 | 3,89 | 0,97 | 1,53 |
| D4 | 3,67 | 3,44 | 5,06 | 1,81 | 3,01 |
| D5 | 5,68 | 4,6 | 5,77 | 2,71 | 3,3 |
| D6 | 7,05 | 6,02 | 7,13 | 3,82 | 5,15 |
| D7 | 10,42 | 8,16 | 7,63 | 7,03 | 7,14 |
| D8 | 13,26 | 11,35 | 10,02 | 11,46 | 11,82 |
| D9 | 17,36 | 17,1 | 14,09 | 17,39 | 16,97 |
| D10 | 37,48 | 43,88 | 40,79 | 53,59 | 49,57 |
| | | | | | |
| D | | 0,069 | 0,093 | 0,161 | 0,121 |
| G | 0,530 | 0,570 | 0,481 | 0,683 | 0,634 |
| Abitanti | 1.059 | | 4.152 | 3.533 | 3.138 |
| Anno | 1417 | | 1390 | 1428 | 1419 |

Nota: i dati delle località toscane sono tratti da Alfani, Ammannati 2017.

**Tabella 25.** *Analisi comparativa dei rapporti inter-decile tra Molfetta e alcuni centri toscani e piemontesi*

| Centri | Rapporti inter-decile | | |
|---|---|---|---|
| | D10/D1 | D10/D5 | D5/D1 |
| Molfetta | 53,57 | 6,59 | 8,12 |
| | | | |
| Sei miglia + comuni suburbani | 126,09 | 9,23 | 13,66 |
| Solo pivieri delle Sei miglia | 132,03 | 8,57 | 15,40 |
| Solo comuni suburbani | 48,31 | 9,17 | 5,27 |
| Arliano | 210,11 | 8,19 | 25,67 |
| Compito | 209,50 | 8,32 | 25,17 |
| Santo Stefano | 98,14 | 5,90 | 16,63 |
| Segromigno | 60,20 | 6,30 | 9,55 |
| Lunata | 55,60 | 8,64 | 6,43 |
| | | | |
| Chieri | 106,43 | 16,26 | 6,55 |
| Moncalieri | 138,05 | 16,14 | 8,55 |
| Cherasco | 49,08 | 8,20 | 5,99 |
| | | | |
| Contado di Firenze | 35,39 | 9,54 | 3,71 |
| Arezzo | 19,61 | 8,06 | 2,77 |
| Prato | 87,85 | 19,77 | 4,44 |
| San Gimignano | 115,28 | 15,02 | 7,67 |

Oltre all'indice di Gini esistono in letteratura (OECD, 2016) altre misure di disuguaglianza della ricchezza e tra queste vanno sicuramente annoverati i rapporti inter-decilici (*interdecile ratios*). Seguendo la metodologia proposta da Alfani (Alfani, 2015) abbiamo:

$$\frac{D_x}{D_y} = \frac{(H_{Dx} * n)}{(H_{Dy} * n)} = \frac{H_{Dx}}{H_{Dy}}$$

dove $D_x$ e $D_y$ sono i diversi decili, $H_{Dx}$ e $H_{Dy}$ sono le quote di ricchezza posseduta dall'unità fiscale media in ciascun decile e n è il numero costante di unità fiscali presenti in ciascun decile, pari al 10% del totale delle unità. I rapporti più impiegati sono: D10/D1 (misura il rapporto esistente tra la fascia dei più ricchi e quella dei più poveri), D10/D5 (misura il rapporto tra la fascia dei più ricchi e quella del ceto mediano), D5/D1 (esprime il rapporto tra il ceto mediano e la parte di popolazione più povera). I valori di questi rapporti per Molfetta e gli altri centri precedentemente trattati sono presentati nella Tab. 25.

Per quanto riguarda il rapporto D10/D1 a Molfetta i più ricchi possedevano 53 volte la quota di ricchezza dei più poveri, un valore simile a quelli rilevati nei comuni suburbani delle Sei Miglia (48,31), a Lunata (55,6), a Segromigno (60,2) e a Cherasco (49,08), ma notevolmente inferiore a quelli registrati nei centri urbani di maggiore dimensione dove la sperequazione tra fascia alta e fascia bassa di ricchezza è maggiore. Nel confronto tra top 10% e ceto medio (D10/D5) il valore di Molfetta (6,59) si avvicina a quello di Segromigno (6,3) e Santo Stefano (5,9), mentre nel rapporto D5/D1 Molfetta (8,12) presenta una realtà somigliante a quella di Moncalieri (8,55) e San Gimignano (7,67).

Dalle diverse comparazioni effettuate emerge come Molfetta presenti un livello di disuguaglianza nella distribuzione della ricchezza che si avvicina maggiormente a quello dei centri rurali (per quanto concerne l'indice di Gini, la somiglianza nella distribuzione dei percentili e alcuni rapporti inter-decilici), mentre per alcuni aspetti, come ad esempio il rapporto D5/D1, mostra prossimità ad alcuni centri urbani. Una spiegazione potrebbe individuarsi nel carattere ibrido della cittadina pugliese: se da un lato è sede diocesana, cittadina portuale dedita ad attività marittime e commerciali e quindi con alcuni tratti di carattere decisamente urbano, dall'altro la sua economia è imperniata sull'olivicoltura praticata in un esteso entroterra, sovente in appezzamenti di piccole dimensioni e a conduzione familiare. Se nel corso del XIV secolo Molfetta aveva una popolazione di 7.000 abitanti, fino al XVI secolo essa si manterrà al di sotto del migliaio di residenti (Malanima, 1998), un aspetto che contribuiva a definirne la duplice natura urbana e rurale.

## 6. Conclusioni

L'apprezzo della città di Molfetta del 1417, oggetto d'indagine del presente contributo, costituisce una fonte particolarmente utile nella definizione della situazione socioeconomica di questo centro della Terra di Bari.

Partendo dalla trascrizione del documento fiscale, uno dei rari esempi del genere nel Regno di Napoli, si è costruito un database, punto di partenza delle successive analisi statistiche. Preliminarmente si è esaminata la distribuzione del numero dei beni apprezzati per ciascun contribuente e si è osservato come la *power law* descriva molto bene questo fenomeno, mentre il modello log-normale ha presentato un discreto adattamento nel definire la distribuzione del valore estimale dei beni. In seguito, si è passati all'analisi del paesaggio agrario e delle colture: le elaborazioni hanno a più riprese confermato come l'olivo fosse la pratica agricola più diffusa tra i terreni apprezzati; l'analisi temporale, riallacciandosi comparativamente ad alcuni dati catastali del XVI secolo, ha evidenziato nel corso del tempo un calo dell'incidenza percentuale degli appezzamenti coltivati ad olivo, dovuta in parte alla riduzione del numero complessivo di terreni per via di un processo di concentrazione della proprietà dettata dalle difficoltà economiche vissute dei piccoli coltivatori. L'analisi spaziale ha permesso di comprendere come l'olivicoltura nel XV secolo fosse diffusa uniformemente su tutto il territorio di Molfetta, mentre altre colture - mandorlo, vite e colture ortive - presentano aree di maggiore specializzazione.

L'ultimo aspetto esaminato è quello del valore estimale come *proxy* della ricchezza: il calcolo di misure di disuguaglianza della distribuzione ha consentito di effettuare delle comparazioni con analoghi dati di località dell'Italia centro-settentrionale all'inizio del Quattrocento. Il confronto ha messo in rilievo la somiglianza della distribuzione della ricchezza di Molfetta con quella di alcuni centri rurali di Toscana e Piemonte simili per ampiezza demografica, ma ha anche evidenziato un richiamo ad alcuni tratti di centri urbani di quelle stesse regioni, a testimonianza del carattere ibrido del centro pugliese. Un futuro sviluppo dell'indagine potrebbe consistere nella rilevazione dei dati presenti nei catasti di Molfetta del XVI secolo e nel successivo computo dei relativi indici di disuguaglianza al fine di studiarne l'evoluzione temporale nel corso della prima Età moderna.

## Riferimenti bibliografici

Alfani, G.; Sardone, S. (2015). *Long-term trends in economic inequality in southern Italy. The Kingdoms of Naples and Sicily, 16th-18th centuries: first results*. Economic History Association, annual meeting, Nashville, September 11-13, 2015 (consultato on line al seguente indirizzo: http://eh.net/eha/wp-content/uploads/2015/05/Alfani.pdf).

Alfani, G. (2015). Economic inequality in northwestern Italy: A long-term view (fourteenth to eighteenth centuries), *Journal of Economic History*, 75 (4), 1058-1096.

Alfani, G.; Ammannati, F. (2017), Long-term trends in economic inequality: the case of the Florentine state, c. 1300–1800, *Economic History Review*, 70 (4), 1-31.

Ammannati, F. (2015). La Peste Nera e la distribuzione della proprietà nella Lucchesia del tardo medioevo, *Popolazione e Storia*, 2, 1-45.

De Gennaro, G. (1963). *Il Liber appretii di Molfetta dei primi del Quattrocento*, Istituto di storia economica, Università di Bari, Bari.

Delvecchio, F. (1992). *Analisi statistica di dati multidimensionali*, Cacucci, Bari.

Girone, G. (2009). *Statistica*, Cacucci, Bari.

Licinio, R. (1989). Una coltura contrastata. Sviluppo e contraddizioni dell'olivicoltura medievale. In G. Ghionda (ed.), *L'ulivo e il trappeto. Riscatto di una civiltà nella lezione della storia*, Schena, Fasano 1989: 29-36

Malanima, P. (1998). Italian Cities 1300-1800. A quantitative approach, *Rivista di storia economica*, issue 2, 91-126.

Marbach, G. (1991). *Statistica economica*, UTET, Torino.

Mastrorocco, N. (2013), *Populus Apuliae. Atlante demografico del numero dei 'fuochi' (nuclei familiari) e della popolazione dei 258 comuni di Puglia dal XIII secolo ad oggi*, IPRES, Sammichele di Bari.

OECD (2016), Income inequality, *OECD Factbook 2015-2016: Economic, Environmental and Social Statistics*, OECD Publishing, Paris.

Piccolo, D. (1998). *Statistica*, Il Mulino, Bologna.

Poli, G. (1990), *Territorio e contadini nella Puglia moderna*, Congedo, Galatina.

Ricci, V. (2017), Una proposta per la stima della misura della disuguaglianza nella distribuzione della ricchezza nel Tardo Medioevo in Terra di Bari. In *Metodi e Analisi Statistich*e, Dipartimento di Economia e Finanza dell'Università di Bari, Bari: 263-282.

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018*
DIPARTIMENTO DI ECONOMIA E FINANZA
ISBN: 978-88-6629-048-3
pp. 245-259

# Analisi della customer satisfaction per i servizi erogati da Sanitaservice ASL FG s.r.l.

**Laura Antonucci*, Gianluca Bruno, Corrado Crocetta, Leonardo Di Gioia, Massimo Russo**

*Università degli studi di Foggia*

**Riassunto:** Il presente lavoro si propone di analizzare la customer satisfaction relativa ai servizi erogati dalla azienda Sanitaservice ASL FG s.r.l.. Mediante alcuni questionari, somministrati ai responsabili dei servizi dall'ASL FG si analizzano le performance dei diversi servizi erogati, costruendo un cruscotto di indicatori sintetici che permette di valutare la qualità dei diversi servizi erogati, secondi gli aspetti considerati per la valutazione.

**Keywords:** *Customer satisfacion*; indicatore sintetico; performance.

## 1. Introduzione

Questo rapporto sintetizza i risultati di una indagine statistica sulla *Customer satisfaction* dei servizi erogati da Sanitaservice ASL FG, nel 2017.
Tale indagine ha esaminato le diverse dimensioni valutative dei servizi erogati, individuando le principali aree di miglioramento degli stessi.

## 2. Materiali e metodi

Nel 2018, la società **Sanitaservice ASL FG s.r.l.** ha affidato al Dipartimento di Economia dell'Università di Foggia il compito di analizzare la soddisfazione del suo unico cliente: la ASL FG, rispetto ai principali servizi erogati per le seguenti aree:

a) ausiliariato, portierato;
b) pulizie e sanificazione ambientale;
c) manutenzione ordinaria immobili;

d) trasporto di farmaci e beni economali;
e) servizio di ausiliariato dell'UU.DD.TT.;
f) servizio infermieristico dell'UU.DD.TT.;
g) servizio di trasporto malati/feriti nella rete emergenza/urgenza (118).

Per ciascuna area sono state individuate le diverse dimensioni da valutare, definite in base al tipo di servizi erogati. Poiché l'azienda ha come unico cliente la ASL Fg, si è ritenuto opportuno rilevare le opinioni espresse dai responsabili dei diversi uffici Asl che sono, indubbiamente, dei testimoni privilegiati in grado di valutare tecnicamente i diversi aspetti analizzati. Lo strumento di rilevazione adottato è un questionario cartaceo con una struttura dinamica in grado di adattarsi alle esigenze dei diversi servizi considerati.

Tale studio rappresenta il primo stadio di uno studio più ampio per monitorare l'impatto che le strategie di efficientamento avviate, avranno sulla percezione della qualità dei servizi erogati. Gli obiettivi di tale analisi possono essere così sintetizzati:

- individuare i KPI per misurare il gradimento dei diversi servizi erogati;
- definire un livello di soddisfazione di riferimento rispetto al quale misurare gli scostamenti;
- analizzare le diverse dimensioni valutative;
- costruire un cruscotto di indicatori di performance in grado di monitorare nel tempo il livello di soddisfazione;
- individuare le criticità e le aree di miglioramento dei diversi servizi analizzati.

Preliminarmente è stata effettuata un'analisi della struttura organizzativa e delle procedure operative della Sanitaservice ASL FG s.r.l. per poter articolare al meglio lo studio della soddisfazione della clientela.

Il punto di partenza è stato il disciplinare unico dei servizi in house sottoscritto da ASL FG e Sanitaservice ASL FG riguardante:

1. servizio di trasporto delle persone malate e/o ferite soccorse nell'ambito del servizio 118 – rete emergenza/urgenza, prestazioni di trasferimento e trasporto secondario,;
2. servizio di pulizia e sanificazione ambientale dei locali destinati allo svolgimento delle attività istituzionali della ASL;
3. servizio di ausiliariato all'interno dei reparti e dei servizi diagnosi e cura;
4. servizio manutenzione ordinaria immobili;
5. servizio di gestione unità degenza territoriale (uu.dd.tt);
6. servizio trasporto di farmaci e beni economali.

Dovendo individuare dei testimoni privilegiati, in grado di misurare la soddisfazione rispetto ai servizi erogati, si è deciso di concentrare l'attenzione sui referenti di struttura che hanno la responsabilità di valutare i servizi ed eventualmente contestare le prestazioni erogate da Sanitaservice ASL FG s.r.l., per cui si è ritenuto di intervistare i direttori amministrativi e/o sanitari delle seguenti strutture ASL FG:

- aree servizi;
- dipartimenti;
- distretti;
- presidi ospedalieri.

I questionari sono stati suddivisi in 7 sezioni:

a. ausiliariato, portierato;
b. pulizie e sanificazione ambientale;
c. manutenzione ordinaria immobili;
d. trasporto di farmaci e beni economali;
e. servizio di ausiliariato dell'UU.DD.TT;
f. servizio infermieristico dell'UU.DD.TT;
g. servizio di trasporto malati/feriti nella rete emergenza/urgenza (118).

Per ogni sezione sono stati individuati i servizi principali oggetto di valutazione, da parte dei responsabili delle strutture. Per ciascun servizio sono stati individuati diversi aspetti da valutare a seconda della modalità operative previste:

- professionalità;
- comportamento;
- tempo di intervento;
- capacità relazionali;
- qualità dell'intervento;
- soddisfazione complessiva.

Per ciascun aspetto è stato attribuito un punteggio utilizzando una scala 1-10 dove il 10 rappresenta la massima soddisfazione.

La raccolta dei questionari è stata effettuata dal personale di Sanitaservice ASL FG, nel periodo 18 giugno 12 luglio 2018. Viste le dimensioni limitate del *frame* si è deciso di intervistare tutti i responsabili individuati e, infatti, i dati rilevati si riferiscono a tutta la popolazione oggetto di studio con l'eccezione di solo 3 rispondenti. Quando aveva senso, sono stati intervistati sia il direttore amministrativo che il direttore sanitario. In totale i rispondenti coinvolti nella rilevazione sono stati 28 con un tasso di copertura del 90,3%.

## 3. Risultati

### *3.1 Servizio di ausiliariato*

I rispondenti hanno assegnato un punteggio da 1 a 10 alla propria percezione di 4 aspetti oggetto di valutazione: professionalità, comportamento, capacità relazionali e soddisfazione complessiva. La numerosità dei rispondenti di questa sezione è pari a 14. La soddisfazione media complessiva è pari a 7,9 con una variabilità piuttosto bassa (scarto quadratico medio 0,2).

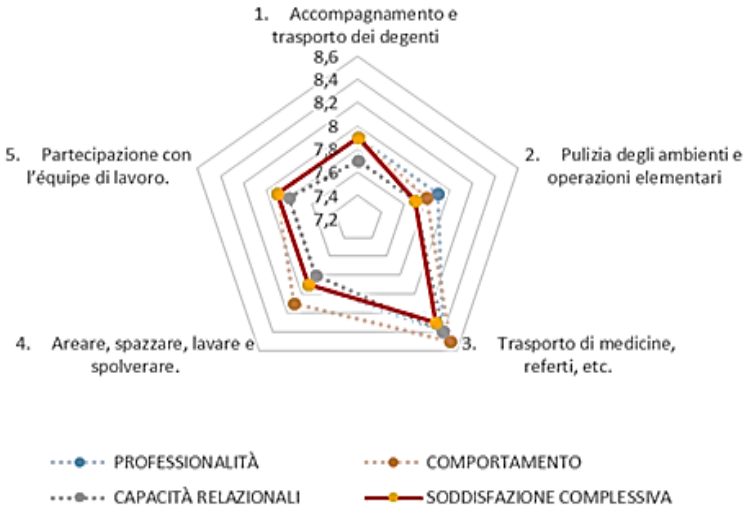
**Tabella 1** – *Punteggi in scala 1- 10 relativi alle attività di ausiliariato*

| ***ATTIVITÀ DA VALUTARE AUSILIARIATO*** | **ASPETTI DA VALUTARE** | | | |
|---|---|---|---|---|
| | PROFESSIONALITÀ | COMPORTAMENTO | CAPACITÀ RELAZIONALI | SODDISFAZIONE COMPLESSIVA |
| *Accompagnamento e trasporto dei degenti con mezzi adeguati* | **7,9** | **7,9** | **7,7** | **7,9** |
| *Pulizia degli ambienti e operazioni elementari e di supporto necessarie al funzionamento del reparto, quali lo spostamento dei ricoverati* | **7,9** | **7,8** | **7,7** | **7,7** |
| *Trasporto di medicine, referti, materiale biologico, sanitario ed economale, vitto, attrezzature, vestiario, biancheria, etc.* | **8,4** | **8,5** | **8,4** | **8,3** |
| *Areare, spazzare, lavare e spolverare tutti gli ambienti dell'unità operativa alla quale è addetto* | **7,9** | **8,1** | **7,8** | **7,9** |
| *Partecipazione con l'équipe di lavoro, limitatamente ai propri compiti* | **7,9** | **7,9** | **7,8** | **7,9** |
| *Controllo degli accessi delle persone.* | **7,7** | **7,8** | **7,8** | **7,7** |
| ***Media*** | **8,0** | **8,0** | **7,9** | **7,9** |

La massima soddisfazione media complessiva si registra per l'attività di trasporto di medicine, referti e materiale biologico, sanitario e economale (punteggio medio 8,3) mentre i valori medi più bassi sono stati assegnati all'attività di pulizia degli ambienti e alle operazioni elementari e di supporto necessarie al funzionamento

del reparto, ed all'attività di controllo degli accessi delle persone, in entrambi i casi il punteggio è di 7,7.

**Grafico 1** – *Punteggi in scala 1- 10 relativi alle attività di ausiliariato*




## *3.2 Servizio di pulizia e sanificazione ambientale*

I 15 intervistati responsabili del servizio di pulizia e sanificazione ambientale hanno analizzato delle attività diverse da quelle previste per il servizio di ausiliariato: la professionalità, la tempestività dell'intervento, la qualità dell'intervento e la soddisfazione complessiva. La soddisfazione media complessiva è pari a 7,7 punti con uno scarto quadratico medio di 0,3 punti.

**Grafico 2** – *Punteggi in scala 1- 10 relativi alle attività di pulizie*

**Tabella 2** – *Punteggi in scala 1- 10 relativi alle attività di pulizie*

| ***ATTIVITÀ DA VALUTARE PULIZIE*** | **ASPETTI DA VALUTARE** | | | |
|---|---|---|---|---|
| | PROFESSIONALITÀ | TEMP.INTERVENTO | QUAL.INTERVENTO | SOD. COMPLESSIVA |
| *Attività ordinarie di pulizia e sanificazione e disinfezione.* | **8,1** | **8,0** | **8,1** | **8,1** |
| *Decontaminazione e rimozione di eventuale materiale organico da tutte le superfici.* | **8,1** | **7,9** | **8,0** | **8,0** |
| *Raccolta di tutte le categorie di rifiuti presenti e successivo trasporto nel locale di stoccaggio provvisorio. Sostituzione del sacchetto, detersione e disinfezione dei contenitori pluriuso dei rifiuti.* | **8,1** | **7,9** | **8,1** | **8,0** |
| *Pulizia e disinfezione degli arredi e delle attrezzature mobili.* | **7,6** | **7,4** | **7,5** | **7,5** |
| *Spolveratura ad umido, detersione e successiva disinfezione, di tutte le superfici orizzontali e verticali (altezza 180 cm).* | **7,4** | **7,3** | **7,4** | **7,4** |
| *Detersione e disinfezione dei lavabi, accessori e arredi sanitari.* | **7,8** | **7,8** | **7,7** | **7,8** |
| *Pulizia a fondo e disinfezione degli arredi mobili (carrelli, tavoli con ruote, ecc.).* | **7,8** | **7,6** | **7,7** | **7,7** |
| *Disincrostazione sanitari, rubinetterie e zone limitrofe.* | **7,6** | **7,5** | **7,5** | **7,5** |
| *De-ragnatura.* | **7,4** | **7,4** | **7,4** | **7,4** |
| *Raccolta e trasporto dei rifiuti speciali fino al punto di deposito temporaneo.* | **8,3** | **8,3** | **8,3** | **8,3** |
| *Palestre: detersione materassini, letti per fisioterapia, parallele, scale, tapis roulant, cyclette, girelli, carrozzine, tappeti per deambulazione, ecc. (esclusi i giocattoli).* | **8,1** | **8,1** | **7,9** | **8,0** |
| *Media* | **7,9** | **7,8** | **7,8** | **7,7** |

La soddisfazione media complessiva più alta si registra in corrispondenza della raccolta e trasporto dei rifiuti speciali fino al punto di deposito temporaneo (voto medio 8,3) mentre quella più bassa corrisponde ai servizi di de-ragnatura e di spolveratura ad umido, detersione e successiva disinfezione, di tutte le superfici orizzontali e verticali. Entrambi questi aspetti hanno un punteggio medio di 7,4.

### *3.3 Servizio di manutenzione ordinaria immobili*

Con un punteggio in scala 1 -10 i 13 rispondenti responsabili del servizio di manutenzione ordinaria immobili hanno espresso la propria percezione rispetto ai quattro diversi aspetti analizzati, quali la professionalità, la tempestività dell'intervento, la qualità dell'intervento e la soddisfazione complessiva. La variabilità è bassa come indicato dallo scarto quadratico medio pari a 0,4.

**Tabella 3** – *Punteggi in scala 1- 10 relativi alle attività di manutenzione*

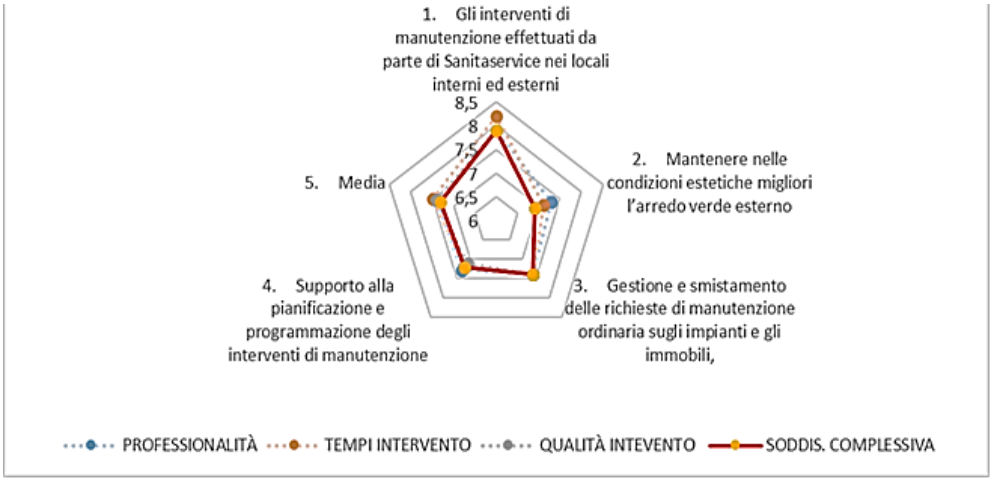
| ***ATTIVITÀ DA VALUTARE MANUTENZIONE*** | **ASPETTI DA VALUTARE** | | | |
|---|---|---|---|---|
| | PROFESSIONALITÀ | TEMP.INTERVENTO | QUAL.INTERVENTO | SOD. COMPLESSIVA |
| *Gli interventi di manutenzione nei locali interni ed esterni, (a titolo esemplificativo e non esaustivo: interventi su infissi, opere murarie, cancellate, impianti di scarico delle acquee, etc.).* | **7,9** | **8,2** | **7,9** | **7,9** |
| *Mantenere nelle condizioni estetiche migliori l'arredo verde esterno, intesa qualunque coltura arborea e floreale.* | **7,3** | **7,1** | **6,9** | **6,9** |
| *Gestione e smistamento delle richieste di manutenzione ordinaria sugli impianti e gli immobili, pervenuti sui protocolli informatici dei vari presidi ospedalieri e dai D.S.S. dell'ASL di Foggia.* | **7,4** | **7,4** | **7,4** | **7,4** |
| *Supporto alla pianificazione e programmazione degli interventi di manutenzione da eseguire presso le varie strutture sanitarie dell'ASL.* | **7,3** | **7,2** | **7,1** | **7,2** |
| *Media* | **7,4** | **7,5** | **7,4** | **7,3** |

I responsabili dei servizi manutenzione dalla ASL FG hanno assegnato il punteggio medio più alto, pari a 7,9, alla soddisfazione complessiva per gli interventi di manutenzione effettuati da parte di Sanitaservice nei locali interni ed esterni (ad esempio interventi su infissi, opere murarie, cancellate, impianti di scarico delle acquee, etc.).

La soddisfazione media complessiva registra il punteggio minimo in riferimento all'attività di supporto alla pianificazione e programmazione degli interventi di

manutenzione da eseguire presso le varie strutture sanitarie dell'ASL (voto medio 7,2).

**Grafico 3** – *Punteggi in scala 1- 10 relativi alle attività di manutenzione*




## *3.4 Servizio di trasporto di farmaci e beni economali*

Per misurare la soddisfazione dei responsabili ASL rispetto al servizio di trasporto di farmaci e beni economali è stata misurata la professionalità, la tempestività dell'intervento, la qualità dell'intervento e la soddisfazione complessiva rispetto ai servizi erogati dalla società in house. I 12 intervistati hanno espresso un punteggio medio per la soddisfazione pari a 8,1 con uno scarto quadratico medio 0,12.

**Grafico 4** – *Punteggi in scala 1- 10 per le attività di trasporto farmaci e beni economali*

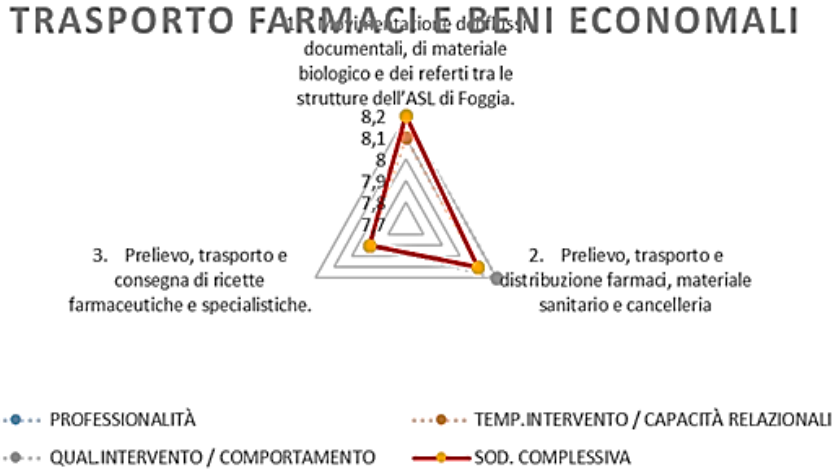


**Tabella 4** – *Punteggi in scala 1- 10 per le attività di trasporto farmaci e beni economali*

| ***ATTIVITÀ DA VALUTARE TRASPORTO BENI ECONOMALI*** | **ASPETTI DA VALUTARE** | | | |
|---|---|---|---|---|
| | PROFESSIONALITÀ | TEMP.INTERVENTO | QUAL.INTERVENTO | SOD. COMPLESSIVA |
| *Movimentazione dei flussi documentali, di materiale biologico e dei referti tra le strutture dell'ASL di Foggia.* | **8,2** | **8,1** | **8,2** | **8,2** |
| *Prelievo, trasporto e distribuzione farmaci, materiale sanitario e cancelleria* | **8,1** | **8,1** | **8,2** | **8,1** |
| *Prelievo, trasporto e consegna di ricette farmaceutiche e specialistiche.* | **7,9** | **7,9** | **7,9** | **7,9** |
| *Media* | **8,1** | **8,0** | **8,1** | **8,1** |

Il valore medio più alto, in termini di soddisfazione media complessiva, si rileva in corrispondenza dell'attività di movimentazione dei flussi documentali, di materiale biologico e dei referti tra le strutture dell'ASL di Foggia, con un punteggio di 8,2, mentre il valore più basso si ha in corrispondenza delle attività di prelievo, trasporto e consegna di ricette farmaceutiche e specialistiche con un punteggio di 7,9 punti.

## *3.5 Servizio di ausiliariato UU.DD.TT.*

Il servizio di ausiliariato UU.DD.TT. è stato misurato in base a quattro diversi aspetti, quali la professionalità, il comportamento, le capacità relazionali e la soddisfazione complessiva.

Le unità di degenza territoriale (UU.DD.TT.) operano nei comuni di Vieste, di Vico del Gargano e di Sannicandro Garganico e costituiscono l'alternativa all'assistenza domiciliare integrata laddove non ci sia una famiglia in grado di supportare il personale sanitario.

Le attività di ausiliariato analizzate nella tabella n. 1, corrispondono solo in parte a quelle considerate per valutare le attività di ausiliariato delle UU.DD.TT. poiché questo servizio viene erogato in modo diverso e prevede una diversa tipologia di adempimenti.

La numerosità media dei rispondenti di questa sezione è pari a 6 e la soddisfazione media complessiva è pari a 7,0 punti con una variabilità molto bassa (scarto quadratico medio 0,1).

**Grafico 5** – *Punteggi in scala 1- 10 relativi al servizio di ausiliariato UU.DD:TT.*

**Tabella 5** – *Punteggi in scala 1- 10 relativi al servizio di ausiliariato UU.DD:TT.*

| ***ATTIVITÀ DA VALUTARE AUSILIARIATO UU.DD.TT.*** | **ASPETTI DA VALUTARE** | | | |
|---|---|---|---|---|
| | PROFESSIONALITÀ | COMPORTAMENTO | CAPACITÀ RELAZIONALI | SODDISFAZIONE COMPLESSIVA |
| *Pulizia dell'ambiente dopo il pasto e aiuto nella distribuzione e somministrazione del pasto.* | **6,8** | **7,0** | **7,0** | **6,8** |
| *Collaborazione con l'inf. professionale per atti di accudimento semplici al malato.* | **7,0** | **7,0** | **6,8** | **7,0** |
| *Pulizia a manutenzione di utensili, apparecchi, presidi usati dal paziente e dal personale medico ed infermieristico per l'assistenza al malato.* | **7,0** | **7,0** | **7,0** | **7,0** |
| *Comunicazione all'infermiere professionale di quanto sopravviene durante il suo lavoro in quanto ritenuto incidente sull'assistito e sull'ambiente.* | **7,0** | **7,0** | **6,8** | **7,0** |
| *Rifacimento del letto non occupato.* | **7,2** | **7,2** | **7,0** | **7,2** |
| *Media* | **7,0** | **7,0** | **6,9** | **7,0** |

Il punteggio massimo rilevato per la soddisfazione media complessiva corrisponde al *Rifacimento del letto non occupato (comodino, letto e apparecchiature)* mentre il valore più basso corrisponde alla pulizia dell'ambiente dopo il pasto e all'aiuto nella distribuzione e somministrazione del pasto (6,8).

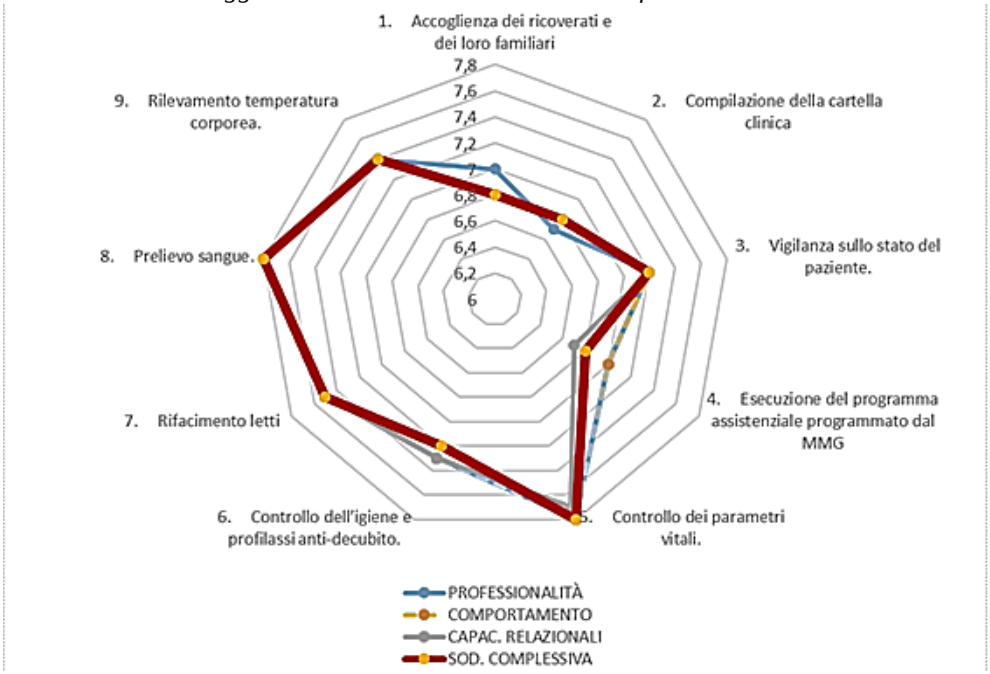
## *3.6 Servizio infermieristico dell'UU.DD.TT.*

Il servizio infermieristico dell'UU.DD.TT. viene misurato tendo conto della professionalità, del comportamento, delle capacità relazionali e della soddisfazione complessiva. La numerosità media dei rispondenti di questa sezione è pari a 6 e la soddisfazione media complessiva è pari a 7,4 con una variabilità molto bassa (scarto quadratico medio 0,4).

**Tabella 6** – *Punteggi in scala 1- 10 relativi al servizio infermieristico dell'UU.DD:TT.*

| ***ATTIVITÀ DA VALUTARE INFERMIEREUU.DD.TT.*** | **ASPETTI DA VALUTARE** | | | |
|---|---|---|---|---|
| | PROFESSIONALITÀ | COMPORTAMENTO | CAPAC. RELAZIONALI | SOD. COMPLESSIVA |
| *Accoglienza dei ricoverati e dei loro familiari informandoli sulle caratteristiche della struttura e dell'organizzazione assistenziale.* | **7,0** | **6,8** | **6,8** | **6,8** |
| *Compilazione della cartella clinica per quanto attiene alla parte infermieristica.* | **6,7** | **6,8** | **6,8** | **6,8** |
| *Vigilanza sullo stato del paziente.* | **7,2** | **7,2** | **7,2** | **7,2** |
| *Esecuzione del programma assistenziale programmato dal MMG* | **7,0** | **7,0** | **6,7** | **6,8** |
| *Controllo dei parametri vitali.* | **7,8** | **7,8** | **7,7** | **7,8** |
| *Controllo dell'igiene e profilassi anti-decubito.* | **7,3** | **7,3** | **7,3** | **7,2** |
| *Rifacimento letti e igiene dell'unità di vita del paziente.* | **7,5** | **7,5** | **7,5** | **7,5** |
| *Prelievo sangue.* | **7,8** | **7,8** | **7,8** | **7,8** |
| *Rilevamento temperatura corporea.* | **7,4** | **7,4** | **7,4** | **7,4** |

**Grafico 6** – *Punteggi in scala 1- 10 relativi al servizio infermieristico dell'UU.DD:TT.*




La soddisfazione media complessiva maggiore corrisponde all'attività di controllo dei parametri vitali (voto 7,8) e di prelievo sangue (voto 7,8), mentre i valori più bassi si riferiscono alle seguenti attività: compilazione della cartella clinica per quanto attiene alla parte infermieristica, esecuzione del programma assistenziale programmato dal MMG, collaborando attivamente nel raggiungimento degli obiettivi di salute prefissati e nell'accoglienza dei ricoverati e dei loro familiari informandoli sulle caratteristiche della struttura e dell'organizzazione assistenziale (6,8).

### *3.7 Servizio di trasporto malati/feriti nella rete emergenze/urgenza (118)*

Il servizio di trasporto malati/feriti nella rete emergenze/urgenza (118), viene valutato dai rispondenti avendo la possibilità di indicare, con un punteggio da 1 a 10 la propria percezione rispetto a quattro diversi aspetti, quali la professionalità, la tempestività dell'intervento, la qualità dell'intervento e la soddisfazione complessiva. La numerosità media dei rispondenti di questa sezione è pari a 3 e la soddisfazione media complessiva è pari a 8,2 punti con uno scarto quadratico medio di 0,3 punti.

Il valore massimo espresso dai rispondenti per quanto concerne la soddisfazione media complessiva è pari a 9,0 punti e corrisponde alle attività di coordinamento con gli altri servizi pubblici addetti alle emergenze. (9,0) ma anche le altre attività valutate hanno registrato punteggi medi piuttosto alti.

**Grafico 7** – *Punteggi in scala 1- 10 relativi al servizio di trasporto malati/feriti nella rete emergenze/urgenze 118.*

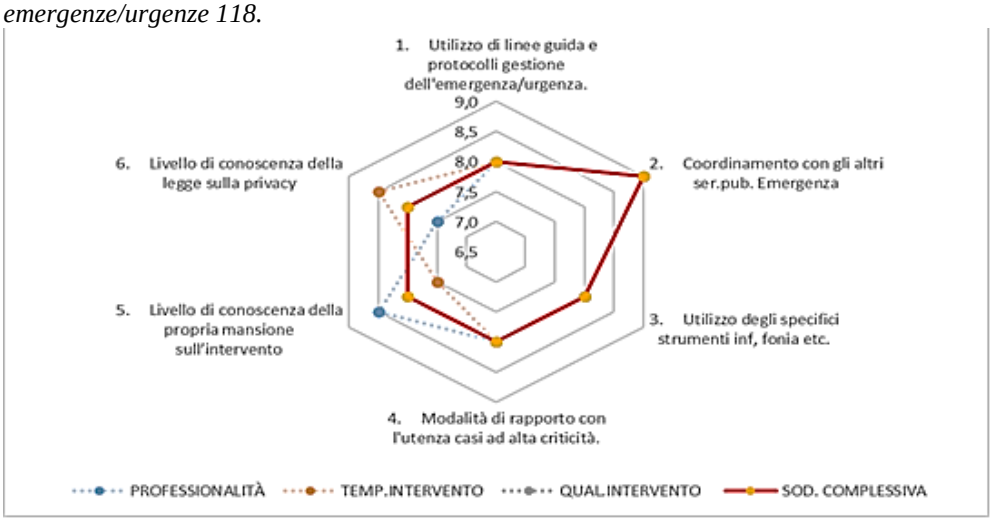



**Tabella 7** – *Punteggi in scala 1- 10 relativi al servizio di trasporto malati/feriti nella rete emergenze/urgenze 118.*

| ***ATTIVITÀ DA VALUTARE EMERGENZA URGENZA 118*** | **ASPETTI DA VALUTARE** | | | |
|---|---|---|---|---|
| | PROFESSIONALITÀ | TEMP.INTERVENTO | QUAL.INTERVENTO | SOD. COMPLESSIVA |
| *Utilizzo di linee guida e protocolli elaborati dalla Centrale operativa per la gestione dell'emergenza/urgenza.* | **8,0** | **8,0** | **8,0** | **8,0** |
| *Coordinamento con gli altri servizi pubblici addetti alle emergenze.* | **9,0** | **9,0** | **9,0** | **9,0** |
| *Utilizzo degli specifici strumenti informatici, di fonia, radiocomunicazione ed orografici.* | **8,0** | **8,0** | **8,0** | **8,0** |
| *Modalità di rapporto con l'utenza, soprattutto nei casi ad alta criticità.* | **8,0** | **8,0** | **8,0** | **8,0** |
| *Livello di conoscenza della propria mansione sull'intervento* | **8,5** | **7,5** | **8,0** | **8,0** |
| *Livello di conoscenza della legge sulla privacy in relazione all'intervento* | **7,5** | **8,5** | **8,0** | **8,0** |
| *Media* | **8,2** | **8,2** | **8,2** | **8,2** |

## 4. Conclusioni

Analizzando i valori medi dei servizi erogati da Sanitaservice ASL FG si nota come il livello di soddisfazione complessiva sia piuttosto elevato visto che oscilla fra un minimo di 7,0 ed un massimo di 8,2.

Il servizio che ha ricevuto la valutazione più bassa è quello di ausiliariato dell'UU.DD.TT. (7,0) seguito da quello di manutenzione ordinaria degli immobili (7,3) e infermieristico dell'UU.DD.TT. (7,4), mentre i servizi che hanno avuto le valutazioni più alte sono il trasporto di malati e feriti 118 (8.2) ed il trasporto dei beni economali (8,1).

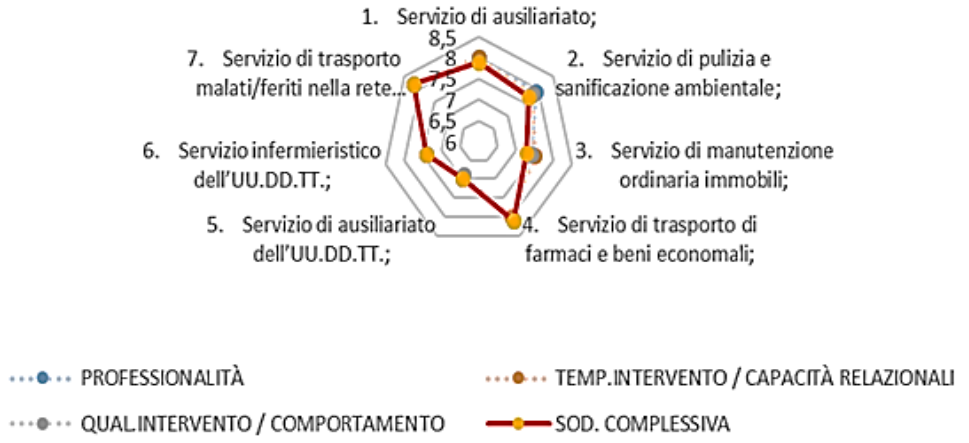
**Tabella 8** – *Punteggi medi in scala 1- 10 relativi ai diversi servizi erogati da Sanitaservice ASL FG s.r.l.*

| ***PUNTEGGI MEDI ATTRIBUITI ALLE DIVERSE ATTIVITÀ DA VALUTARE*** | **ASPETTI DA VALUTARE** | | | |
|---|---|---|---|---|
| | PROFESSIONALITÀ | TEMP.INTERVENTO / CAPACITÀ RELAZIONALI | QUAL.INTERVENTO / COMPORTAMENTO | SOD. COMPLESSIVA |
| *Servizio di ausiliariato* | **8,0** | **8,0** | **7,9** | **7,9** |
| *Servizio di pulizia e sanificazione ambientale* | **7,9** | **7,8** | **7,8** | **7,7** |
| *Servizio di manutenzione ordinaria immobili* | **7,4** | **7,5** | **7,4** | **7,3** |
| *Servizio di trasporto di farmaci e beni economali* | **8,1** | **8,0** | **8,1** | **8,1** |
| *Servizio di ausiliariato dell'UU.DD.TT.* | **7,0** | **7,0** | **6,9** | **7,0** |
| *Servizio infermieristico dell'UU.DD.TT;* | **7,4** | **7,4** | **7,4** | **7,4** |
| *Servizio di trasporto malati/feriti nella rete emergenze/urgenza (118)* | **8,2** | **8,2** | **8,2** | **8,2** |

Andando ad osservare i valori medi dei singoli aspetti da valutare, come professionalità, tempo di intervento, capacità relazionali, qualità dell'intervento e comportamento non si hanno molte informazioni, poiché tali valori si discostano dalla media complessiva al massimo di un decimale. Tale stabilità dei valori medi indica una forte coerenza interna delle valutazioni espresse dai vari referenti di struttura

ed un buon equilibrio fra le diverse dimensioni dei servizi analizzati. Pertanto per l'individuazione e l'analisi puntuale delle aree di miglioramento e per la valutazione dei punti di forza di debolezza è necessario riferirsi alle tabelle precedenti in cui è stata fatta una precisa valutazione dei vari servizi erogati.

**Grafico 8** – *Punteggi medi in scala 1- 10 relativi ai diversi servizi erogati da Sanitaservice ASL FG s.r.l.*




I valori riportati nelle tabelle precedenti rappresentano un sistema elementare di indicatori di performance (KPI) che può essere utilizzato come un primo cruscotto direzionale, in grado di individuare, attraverso un sistema di monitoraggio periodico, gli scostamenti rispetto agli standard di servizio definiti.

Attraverso tale cruscotto è possibile anche individuare le cause degli scostamenti in modo da poter intervenire tempestivamente. Tale sistema di monitoraggio deve essere integrato all'interno del sistema qualità dell'azienda e deve operare in una logica di miglioramento continuo.

# Piano nazionale Lauree Scientifiche

# 2017/2018

# Statistica

*Comitato scientifico:*
Proff. Nunziata RIBECCO (coordinatore), Ernesto Toma, Alessio Pollice, Francesco Domenico d'Ovidio, Angela Maria D'Uggento

# PNLS 2017/2018 - Questionario sul cyber-bullismo

## Atteggiamenti psicosociali sugli usi di mezzi tecnologici

Ti invitiamo a completare il questionario in ogni sua parte, esprimendo le tue personali opinioni. Leggi bene ciascuna domanda e dai la tua valutazione, barrando la casella che contrassegna la risposta prescelta. Il questionario è anonimo e destinato solo ad elaborazioni statistiche che verranno svolte in maniera aggregata ed in conformità alla vigente legge sulla privacy.

Comune di residenza_____

Scuola_____

**1. Genere:**
- .1 Maschio
- .2 Femmina

**2. Età (in anni compiuti): _____**

**3. Nazionalità:**
- .1 Italiana
- .2 Altra (specificare) _____

**4. Da chi è composta, oltre te, la tua famiglia?**
- .1 Madre
- .2 Padre
- .3 Fratelli/sorelle (specificare quanti): _____
- .4 Altri famigliari

**5. Titolo di studio dei tuoi genitori:**

| | Madre | Padre |
|---|---|---|
| .1 Licenza elementare | | |
| .2 Licenza media inferiore | | |
| .3 Licenza media superiore | | |
| .4 Laurea | | |

**6. Condizione lavorativa:**

| | Madre | Padre |
|---|---|---|
| .1 Occupato | | |
| .2 Disoccupato o in cerca di prima occupazione | | |
| .3 Casalinga | | |
| .4 Altra condizione | | |

**7. Quanto sei soddisfatto/a del tuo aspetto fisico?**
- .1 Per nulla
- .2 Poco
- .3 Abbastanza
- .4 Molto

**8. Quando hai qualche problema, con chi ti confidi solitamente?**
- .1 Nessuno
- .2 Amici
- .3 Famigliari
- .4 Insegnanti

**9. Hai amici "veri" di cui ti fidi?**
- .1 Sì
- .2 No

**10. Credi di avere buoni rapporti con i tuoi compagni di classe?**
- .1 Sì, ma con pochi compagni
- .2 Sì, con tutti o quasi tutti i compagni
- .3 No, con nessuno

**11. Credi di avere buoni rapporti con i tuoi genitori?**
- .1 Sì, con entrambi
- .2 Sì, solo con mia madre
- .3 Sì, solo con mio padre
- .4 No, con nessuno dei due

**12. Quali dei seguenti mezzi tecnologici usi solitamente?**
- .1 Cellulare di vecchia generazione
- .2 Smartphone
- .3 Tablet
- .4 Computer

**13. Hai accesso ad internet (tramite smartphone, tablet o pc):**

| | Sì | No |
|---|---|---|
| .1 In casa | | |
| .2 Fuori casa | | |

**14. Quando sei su internet, hai un adulto che controlla (anche se non sempre) su quali siti navighi?**
- .1 Sì
- .2 No

**15. Quali dei seguenti social network o mezzi di messaggistica istantanea usi ?**
- .1 Facebook
- .2 Twitter
- .3 Whatsapp/Viber/WeChat/Line ecc
- .4 Ask.fm
- .5 Instagram
- .6 Youtube
- .7 Altro (specificare): _____

**16. I tuoi genitori possiedono le password per accedere ai tuoi account dei social network?**
- .1 Sì
- .2 No

**17. Riesci a dire quanto tempo al giorno dedichi all'incirca ai social network o ai mezzi di messaggistica istantanea?**
- .1 Meno di 1 ora
- .2 Da 1 a 3 ore
- .3 Oltre 3 ore

**18. I tuoi genitori (almeno uno) sono tuoi amici sui social che usi?**
- .1 No
- .2 Sì, ma con restrizioni a ciò che possono vedere
- .3 Sì, senza restrizioni a ciò che possono vedere

***Il cyberbullismo indica atti di bullismo e di molestia effettuati tramite telefoni cellulari, e-mail, messaggistica istantanea, blog, social network, ecc.***

**19. Quali sono secondo te le motivazioni che inducono a mettere in atto episodi di cyberbullismo?**
- .1 Vendetta
- .2 Divertimento
- .3 Ferire i sentimenti di una persona antipatica
- .4 Perché non sono ritenuti episodi offensivi o gravi
- .5 Altro (specificare): _____

_____

**20. Sei mai stato vittima di cyberbullismo?**

.1 Sì, a scuola
.2 Sì, al di fuori della scuola
.3 Sì, sia a scuola che fuori
.4 No

**Se non sei mai stato vittima di cyberbullismo vai alla domanda n.25**

**21. Se sì, di che tipo di atti di cyberbullismo sei stato vittima?**

.1 Ho ricevuto foto/video spiacevoli sul cellulare
.2 Ho ricevuto insulti/minacce sul cellulare
.3 Ho ricevuto foto/video spiacevoli tramite mail o social network
.4 Ho ricevuto insulti/minacce tramite mail o social network
.5 Ho ricevuto scherzi telefonici
.6 Sono stato volutamente escluso da gruppi di amici/compagni di classe (su Fb, Whatsapp, ecc)
.7 Sono state pubblicate sui social delle mie foto/video senza il mio permesso
.8 Sono state pubblicate fake news
.9 Altro (specificare): _____

**22. Se sì, quante volte sei stato vittima di questi atti?**

.1 Solo una o due volte
.2 Una-due volte al mese
.3 Quasi ogni settimana
.4 Anche più volte a settimana

**23. Se sì, ci puoi dire se gli autori di questi atti nei tuoi confronti erano maschi o femmine?**

.1 Soprattutto una femmina
.2 Parecchie femmine
.3 Soprattutto un maschio
.4 Parecchi maschi
.5 Sia maschi che femmine
.6 Non so chi sia stato

**24. Cosa hai fatto quando sei stato/a vittima di atti di cyberbullismo?**

.1 Nulla, mi sono sentito impotente
.2 Ho ignorato cosa stava avvenendo, sperando che gli episodi non si ripetessero
.3 L'ho detto a un amico
.4 L'ho detto a un insegnante
.5 L'ho detto a un genitore/adulto che si prende cura di me
.6 Ho chiesto direttamente alla persona di smetterla
.7 Ho bloccato i messaggi/le telefonate/le e-mail
.8 Ho cambiato il mio numero di cellulare
.9 Ho cercato di fare a loro ciò che avevano fatto a me
.10 Altro (specificare): _____

**25. Hai assistito ad episodi in cui qualcuno è stato/a vittima di cyberbullismo?**

.1 Sì, a scuola
.2 Sì, al di fuori della scuola
.3 Sì, sia a scuola che fuori
.4 No

**26. Hai saputo di qualcuno che è stato/a vittima di cyberbullismo?**

.5 Sì, a scuola
.6 Sì, al di fuori della scuola
.7 Sì, sia a scuola che fuori
.8 No

**Se non hai mai visto o saputo di qualcuno vittima di cyberbullismo vai alla domanda n.29**

**27. Se sì, di che tipo di atti di cyberbullismo è stato/a vittima?**

.1 Ha ricevuto foto/video spiacevoli sul cellulare
.2 Ha ricevuto insulti/minacce sul cellulare
.3 Ha ricevuto foto/video spiacevoli tramite mail o social network
.4 Ha ricevuto insulti/minacce tramite mail o social network
.5 Ha ricevuto scherzi telefonici
.6 E' stato volutamente escluso da gruppi di amici/compagni di classe (su Fb, Whatsapp, ecc)
.7 Sono state pubblicate sui social delle sue foto/video senza il suo permesso
.8 Altro (specificare): _____

**28. Se hai visto o saputo di qualcuno che è stato/a vittima di cyberbullismo, qual è stata la tua reazione?**

.1 Ho consolato la vittima cercando di aiutarla
.2 Non ho dato peso o importanza all'episodio
.3 Ho cercato di intervenire affinché gli episodi non si ripetessero
.4 Ho riso dell'episodio, talvolta condividendolo con altri amici
.5 Altro (specificare): _____

**29. Hai mai preso parte ad episodi di cyberbullismo?**

.1 Sì, a scuola
.2 Sì, al di fuori della scuola
.3 Sì, sia a scuola che fuori
.4 No

**Se non hai mai preso parte ad episodi di cyberbullismo, il questionario finisce qui.**

**30. Negli episodi ai quali hai preso parte, eri:**

.1 Da solo/a
.2 In compagnia di coetanei
.3 In compagnia di amici più grandi

**31. Se sì, a che tipo di atti di cyberbullismo hai preso parte?**

.1 Ho inviato foto/video spiacevoli sul cellulare
.2 Ho inviato insulti/minacce sul cellulare
.3 Ho inviato foto/video spiacevoli tramite mail o soc network
.4 Ho inviato insulti/minacce tramite mail o soc network
.5 Ho fatto scherzi telefonici
.6 Ho volutamente escluso delle persone da gruppi amici/compagni di classe (su Fb, Whatsapp, ecc)
.7 Ho pubblicato sui social delle foto/video ser permesso
.8 Altro (specificare): _____

**32. Quali sono le motivazioni che ti hanno indotto a prendere parte ad episodi di cyberbullismo, ?**

.1 Vendetta
.2 Divertimento
.3 Ferire i sentimenti di una persona antipatica
.4 In compagnia di amici più grandi
.5 Altro (specificare): _____

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018* ISBN: 978-88-6629-048-3
DIPARTIMENTO DI ECONOMIA E FINANZA pp. 265-276

# Cyberbullismo@Euclide: i risultati di un'indagine tra gli studenti

**Floriana Rinaldi***

*Istituto Tecnico Tecnologico "EUCLIDE-CARACCIOLO" - Bari.*

**Riassunto:** In questo lavoro sono riportati i risultati di un'indagine statistica sul problema del cyberbullismo condotta nei primi mesi del 2018 tra gli studenti dell'Istituto Tecnico "Euclide" di Bari, nell'ambito delle attività proposte nel Laboratorio di Statistica del Piano Nazionale Lauree Scientifiche 17/18. I dati raccolti sono sintetizzati in tabelle e grafici e si pone l'attenzione su alcuni aspetti particolarmente interessanti emersi dall'indagine, confrontandoli con i risultati di analoghe indagini su scala nazionale e internazionale, fornendo, attraverso l'applicazio– ne dei metodi della statistica, una possibile interpretazione dei risultati ottenuti.

**Keywords:** Cyberbullismo; statistica; questionario; scuola; PNLS .

## 1. Introduzione

La diffusione tra i giovani delle tecnologie digitali, soprattutto degli smartphone, ha creato una generazione di adolescenti che sempre più sostituisce le relazioni personali "faccia a faccia" con quelle virtuali, attraverso l'uso di servizi vari di messaggistica e gruppi social sul web. Se questo cambiamento di stili di vita ha facilitato e allargato la possibilità di entrare in contatto con le persone, d'altra parte ha amplificato un fenomeno, come il bullismo, già notevolmente diffuso tra i giovani, facendone assumere la forma più subdola e vile del "cyberbullismo", cioè il bullismo in rete. La gravità di questo fenomeno spesso emerge solo quando le cronache riportano le tragiche conseguenze che talvolta provoca sulle vittime.

Già da una decina d'anni le istituzioni seguono con attenzione il preoccupante aumento dei casi di cyberbullismo tra la popolazione giovanile e numerose iniziati-

* Autore corrispondente: f_rinaldi@libero.it

ve sono state intraprese per tentare di evitarne la crescita. In particolare, il Miur ha recentemente pubblicato le "Linee di Orientamento per la prevenzione e il contrasto del Cyberbullismo" (Miur, 27/10/2017), che integra le precedenti linee guida sullo stesso tema pubblicate nell'Aprile 2015, alla luce dell'emanazione della legge dello stato n. 71/2017. Tale legge, infatti, affida al Miur un ruolo centrale attraverso il coordinamento delle diverse azioni di prevenzione e contrasto del bullismo e cyberbullismo nelle scuole, in sinergia con gli altri Enti ed Istituzioni previsti dalla normativa, come le forze di Pubblica Sicurezza e operatori della Giustizia minorile. Il Miur ha anche istituito un portale di riferimento per le azioni di prevenzione e contrasto del bullismo e del cyberbullismo al link www.generazioniconnesse.it.

La scuola, quindi, è chiamata in prima linea, in collaborazione con le famiglie, a promuovere l'educazione all'uso consapevole di Internet, anche mediante progetti attuati con la partecipazione di enti locali, servizi territoriali, organi di polizia ed altre associazioni o enti.

In questo contesto, la proposta di un'indagine statistica sul cyberbullismo tra la popolazione scolastica degli Istituti di appartenenza dei partecipanti al PLS Statistica 2017/2018, appare quanto mai coerente con le indicazioni ministeriali e costituisce uno strumento per conoscere e comprendere il fenomeno alla base, attraverso l'analisi dei dati raccolti e la ricerca di possibili connessioni.

Uno strumento con cui confrontarsi sono le indagini già condotte a livello nazionale e internazionale, come il questionario dal titolo "Il bullismo in Italia: comportamenti offensivi e violenti tra i giovanissimi" (ISTAT, 2015), in cui emerge come l'83% dei ragazzi tra gli 11 e i 17 anni utilizza Internet attraverso un telefono cellulare e il 57% naviga nel web senza un controllo degli adulti, una situazione che rende gli adolescenti più esposti sia a subire che a compiere atti di cyberbullismo. Le ragazze risultano più colpite dal bullismo virtuale (il 7,1% di quelle tra gli 11 e i 17 anni che si collegano a internet o dispongono di un telefono cellulare è stata vittima di atti di cyberbullismo, rispetto al 4,6% dei ragazzi). Anche a livello internazionale, un risultato analogo emerge dal report del Cyberbullyng Research Center (2016): dai dati raccolti si apprende che il 37% delle ragazze intervistate ammettono di essere vittima di cyberbullismo contro il 31% dei ragazzi. Dati provenienti da analoghe indagini condotte in Inghilterra confermano questi risultati, dove le ragazze figurano in prevalenza sia come vittime che, sorprendentemente, come bulli, soprattutto verso vittime dello stesso sesso. Inoltre, in un sondaggio condotto da Save the Children e Ipsos (2013), si legge che il 67% degli studenti intervistati indica che il motivo per cui un ragazzo/ragazza diviene oggetto di bullismo è da ricercare nelle sue caratteristiche fisiche.

L'importanza che l'aspetto fisico riveste per il genere femminile, soprattutto in fase adolescenziale, è un altro fattore che le rende potenzialmente vittime privilegiate di atti di cyberbullismo. Ancora da dati italiani si mette in risalto il rischio a cui sono maggiormente esposte le ragazze per l'uso compulsivo dello smartphone: dal report annuale dell'Osservatorio Nazionale Adolescenza (2017) emerge che il 15% degli adolescenti del campione totale, di cui il 73% sono ragazze, usa lo smartphone per più di 10 ore al giorno; il 18% dalle 7 alle 10 ore e la maggior parte di loro è di genere femminile (70%). Tra gli adolescenti intervistati che hanno ammesso di condividere nelle chat e nei loro profili tutto ciò che fanno durante la loro giornata, il 60% sono ragazze. Sul web molte persone cercano di mostrarsi diverse da quello che sono nella vita reale, cercando accettazione e apprezzamento, dietro cui, in realtà, spesso si nasconde insicurezza e vulnerabilità rispetto alle critiche.

L'indagine condotta nel nostro Istituto ha voluto mettere a confronto i nostri dati con quanto si trova in letteratura, per studiare eventuali analogie, fornire una interpretazione del fenomeno su base locale, valutarne l'entità e cercare le possibili cause, per mettere in atto le azioni preventive. La nostra analisi, in modo particolare, si è concentrata sulla definizione del profilo della vittima.

## 2. Materiali e metodi

Il gruppo di lavoro dell'Euclide è costituito da cinque studenti della classe IV, indirizzo Logistica, che hanno partecipato al progetto su base volontaria, accogliendo la proposta del docente referente, partecipante al PLS, e motivati dall'interesse verso un'attività pratica che potesse avviarli verso la conoscenza delle tecniche dell'analisi statistica quantitativa e qualitativa.

La fase preliminare si è svolta nella prima metà del mese di Gennaio con due incontri con gli alunni, durante i quali il docente ha illustrato gli obiettivi del progetto e presentato il questionario. Poiché il docente referente non appartiene al consiglio di classe, l'attività del gruppo di alunni è stata inserita nella programmazione di classe, come progetto extracurricolare di approfondimento. L'intera classe è stata costantemente aggiornata, attraverso gli studenti coinvolti, sulle varie fasi del progetto e sui risultati ottenuti.

La raccolta dei dati è avvenuta tramite la somministrazione del questionario, distribuito agli studenti dell'Euclide nella sola modalità on line. In questa fase, dal 22 Gennaio al 12 Febbraio, il gruppo di lavoro ha provveduto a diffondere il link al questionario attraverso sia i canali istituzionali (comunicazione del Dirigente Sco-

lastico) che informali (sms, email e chat). Nello stesso periodo sono stati svolti a scuola due incontri di formazione e preparazione alla fase operativa, in cui il docente ha fornito agli alunni gli elementi di base della Statistica descrittiva, sia attraverso la costruzione di semplici tabelle di dati e relativi grafici, sia con la consultazione del sito ISTAT e il report ISTAT 2015 sul Cyberbullismo. In questo modo gli studenti hanno familiarizzato con le tecniche di raccolta e presentazione di dati, per poterle poi applicare ai dati raccolti dalla nostra indagine.

Occorre precisare che l'obiettivo specifico del progetto non è quello di fornire risposte a un fenomeno complesso di natura sociopsicologica come quello del cyberbullismo, ma essenzialmente quello di far svolgere agli studenti un'esperienza formativa, avvicinandoli allo studio della Statistica e dei suoi metodi attraverso il contatto diretto con un problema concreto e con dati reali. In quest'ottica, i contenuti teorici sono stati forniti in modo essenziale e si è data prevalenza all'aspetto applicativo della disciplina, utilizzando una metodologia di tipo laboratoriale e servendosi degli strumenti di Excel con la componente aggiuntiva PHStat. Con questi strumenti si è potuto anche arrivare, dopo l'analisi esplorativa dei dati, a fornire una introduzione all'inferenza statistica attraverso l'utilizzo del test Chi-quadrato per stabilire il livello di significatività di una ipotesi, senza dover entrare nei dettagli della teoria, troppo complessa per il livello di partenza del gruppo. In particolare, si è utilizzata l'apposita funzione di PHStat per valutare, applicando la statistica Chi-quadrato, l'eventuale relazione tra due variabili qualitative, precedentemente evidenziata dalle tabelle di contingenza costruite con lo strumento Tabelle Pivot di Excel.

I risultati prodotti dalle elaborazioni sono stati commentati e confrontati anche negli incontri avvenuti in plenaria con tutto il gruppo (docenti e studenti) dei partecipanti al PLS.

## 3. Risultati

Il questionario, composto da 32 domande, è stato compilato da 319 studenti. I dati raccolti dalla prima sezione (18 domande) hanno contribuito a descrivere le principali caratteristiche socio-demografiche del campione.

Gli intervistati risultano prevalentemente di sesso maschile (87%) e residenti nella provincia di Bari (67%). Tali caratteristiche del campione rispecchiano quelle dell'intera popolazione scolastica, prevalentemente maschile e con un elevato pendolarismo.

Nella figura 1 osserviamo la distribuzione per età degli intervistati.

**Figura 1.** *Distribuzione dei soggetti intervistati per classi di età*

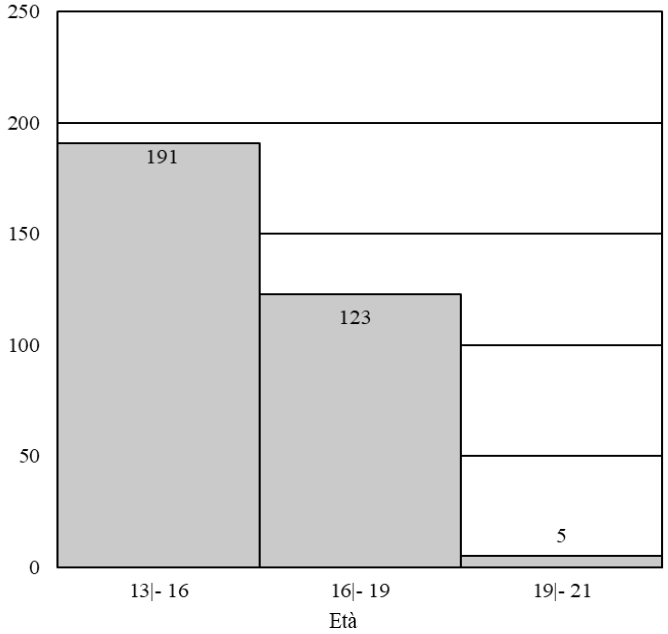
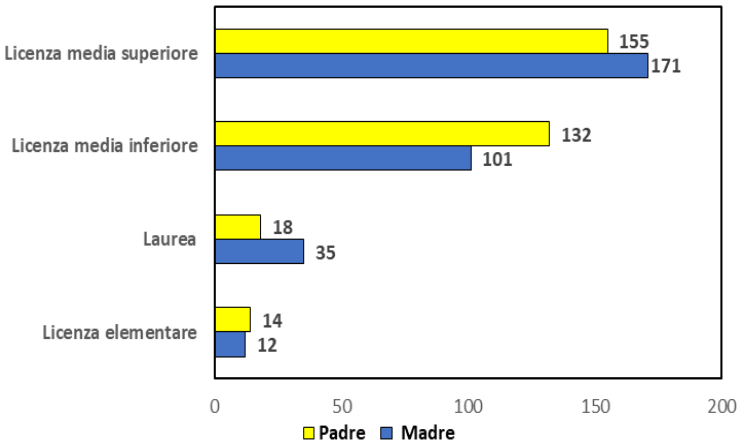


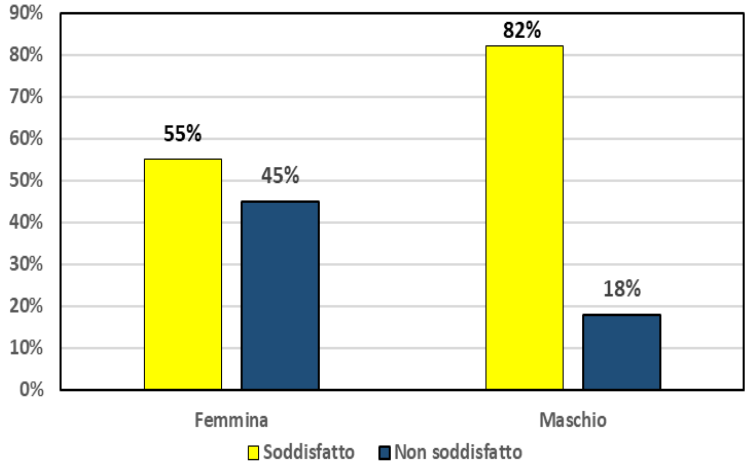

**Figura 2.** *Distribuzione dei soggetti intervistati secondo il titolo di studio dei genitori*




La figura 2 sintetizza le informazioni sul contesto familiare degli intervistati, che risulta mediamente istruito, con il 64,6% delle madri e il 54,8% dei padri con un titolo di studio di scuola superiore o laurea.

In riferimento ai rapporti con i genitori il 65% delle ragazze e l'84% dei ragazzi afferma di avere buoni rapporti con entrambi i genitori, anche se preferiscono confidarsi prevalentemente con gli amici. I rapporti con i compagni risultano buoni per il 69% dei ragazzi e il 50% delle ragazze.

**Figura 3** *Soddisfazione del proprio aspetto fisico.*




Nella Figura 3 sono riportati i risultati riguardanti la soddisfazione del proprio aspetto fisico.

Le informazioni relative al contesto familiare e al grado di soddisfazione del proprio aspetto fisico sono state oggetto di particolare attenzione, al fine di cercare eventuali relazioni e, quindi, delineare il profilo della vittima. In primo luogo, la distribuzione delle variabili qualitative *"Genere - Soddisfazione aspetto fisico"* ha suggerito la possibilità di analizzare l'esistenza di una connessione statistica, valutata attraverso il test Chi-quadrato, i cui risultati sono riportati nella Tabella 1.

**Tabella 1.** *Risultati Test Chi-quadrato Genere-Soddisfazione aspetto fisico*

| **Genere** | **Aspetto fisico** | | |
|---|---|---|---|
| | Soddisfatto | Non soddisfatto | Totale |
| ***Frequenze osservate*** | | | |
| Femmina | 22 | 18 | 40 |
| Maschio | 229 | 50 | 279 |
| Totale | 251 | 68 | 319 |
| ***Frequenze attese*** | | | |
| Femmina | 31,47 | 8,53 | 40 |
| Maschio | 219,53 | 59,47 | 279 |
| Totale | 251 | 68 | 319 |
| Test Chi-quadrato | 15,29 | | |
| Livello di significatività | 0,05 | | |
| Gradi di libertà | 1 | | |
| Valore critico Chi-quadro | 3,84 | | |
| *p*-Value | 0,00009 | | |
| ***Si rifiuta l'ipotesi nulla*** | | | |

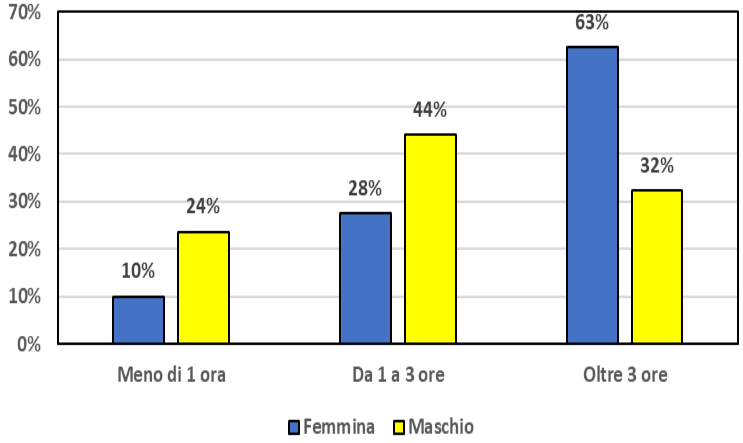
*(Elaborazione Excel con PHStat)*

Il test conferma l'ipotesi che esiste una connessione non casuale tra il genere e la soddisfazione verso il proprio aspetto fisico.

L'informazione ricevuta dalle risposte al questionario, in riferimento alle dotazioni tecnologiche degli intervistati, conferma il dato nazionale: nel nostro caso il 98% degli studenti possiede uno smartphone, il 95% utilizza la rete dentro e fuori casa e senza il controllo degli adulti (91%). Le applicazioni maggiormente utilizzate sono Whatsapp, You Tube e Instagram (scelti da oltre il 90% degli intervistati).

La Figura 4 rappresenta ancora una conferma di un risultato già emerso dalle indagini in scala nazionale e cioè che le ragazze sono, tra gli intervistati, i soggetti che trascorrono più tempo sui social o in chat.

**Figura 4.** *Tempo trascorso sui social o nelle chat*



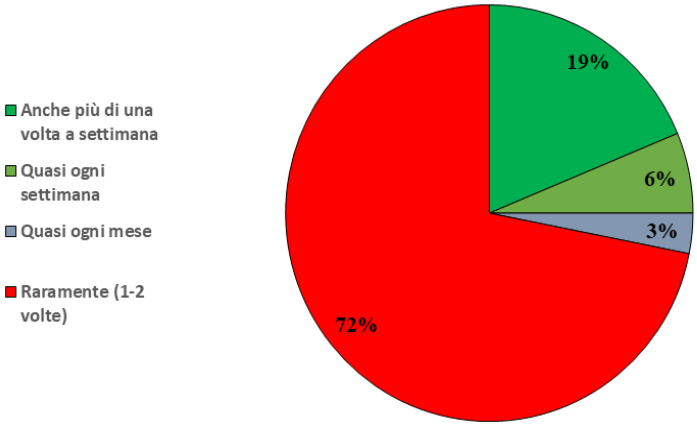

## *3.1 La vittima del cyberbullismo*

La domanda 20 del questionario discrimina gli intervistati rispetto all'essere vittima o no di cyberbullismo (Tabella 2). Notiamo come la percentuale delle vittime è pari al 10% degli intervistati.

**Tabella 2**. *Distribuzione delle vittime di cyberbullismo*

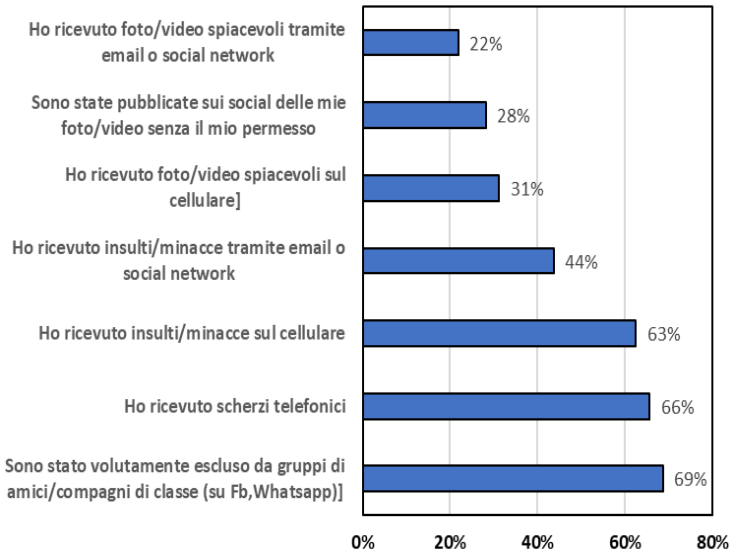
| Sei mai stato vittima di cyberbullismo? | Valori assoluti | Valori percentuali |
|---|---|---|
| No | 287 | 90,0 |
| Si a scuola | 9 | 2,8 |
| Si, al di fuori della scuola | 13 | 4,1 |
| Si, sia a scuola che fuori | 10 | 3,1 |
| Totale | 319 | 100,0 |

**Figura 5.** *Frequenza degli episodi di cyberbullimo*




Nella Figura 5 abbiamo i risultati sulla frequenza degli episodi di cyberbullismo, mentre nella Figura 6 sono sintetizzate le risposte sul tipo di azioni subite dalle vittime.

**Figura 6.** *Tipo di azioni subite dalle vittime*




Considerando le sole risposte “Si” alla domanda 20 si è potuto costruire un profilo della vittima. La Tabella 3 riporta, infatti, le principali caratteristiche della distribuzione dell’età delle vittime.

**Tabella 3**. *Principali valori caratteristici della distribuzione delle età delle vittime*

| | | | |
|---|---|---|---|
| Media | 15,5 | Varianza | 1,6129 |
| Mediana | 15,5 | Deviazione Standard | 1,2700 |
| Moda | 16 | Coeff. di Variazione | 8,19% |
| Valore minimo | 14 | Ind. di Asimmetria | 0,5543 |
| Valore massimo | 19 | Ind. di Curtosi | 0,1028 |
| Range | 5 | Errore Standard | 0,2245 |
| Totale casi | 32 | | |

*(Elaborazione Excel con PHStat)*

Alla domanda 20 il 25% delle ragazze intervistate ha dichiarato di essere stata vittima, contro l'8% dei ragazzi. Questo risultato ha suggerito una possibile relazione tra l'essere vittima e il genere e, come nel caso precedente, l'applicazione del test chi-quadrato ha confermato questo risultato, che si riporta in Tabella 4.

**Tabella 4**. *Risultati Test Chi-quadrato Genere-Vittima*

| **Genere** | **Vittima** | | |
|---|---|---|---|
| | No | Sì | Totale |
| ***Frequenze osservate*** | | | |
| Femmina | 30 | 10 | 40 |
| Maschio | 257 | 22 | 279 |
| Totale | 287 | 32 | 319 |
| ***Frequenze attese*** | | | |
| Femmina | 35,99 | 4,01 | 40 |
| Maschio | 251,01 | 27,99 | 279 |
| Totale | 287 | 32 | 319 |
| Test chi-quadro | 11,35 | | |
| Livello di significatività | 0,05 | | |
| Gradi di libertà | 1 | | |
| Valore critico | 3,84 | | |
| *p*-Value | 0,0008 | | |
| ***Si rifiuta l'ipotesi nulla*** | | | |

*(Elaborazione Excel con PHStat)*

Il risultato del test conferma l'ipotesi che esiste una connessione non casuale tra il genere e l'essere vittima. In particolare, osservando il grafico in Figura 7 si nota come la percentuale delle vittime che trascorrono più di 3 ore su chat e social sia ancora più alto (90%) per le ragazze rispetto al corrispondente valore del totale del campione (63%) .

Questo risultato è in linea con quanto è emerso dall'indagine dell'Osservatorio Nazionale Adolescenza, come riportato nell'introduzione, e mette in evidenza la possibilità di una relazione tra il genere delle vittime e il tempo trascorso sui social.

**Figura 7.** *Tempo trascorso su social e messaggistica (risposte delle vittime)*

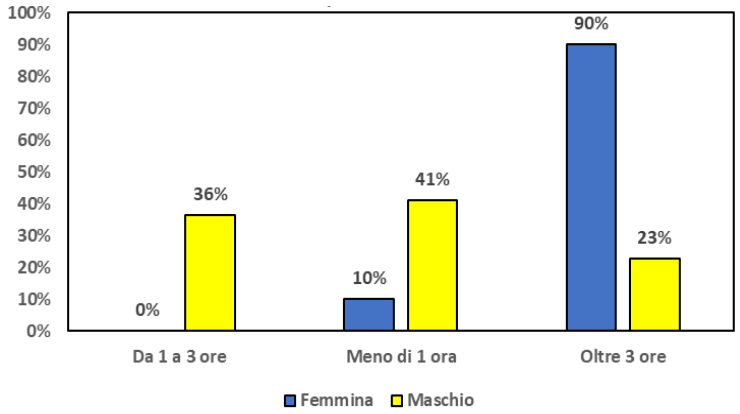



**Tabella 5.** *Risultati Test Chi-quadrato Genere delle vittime-Tempo sui social*

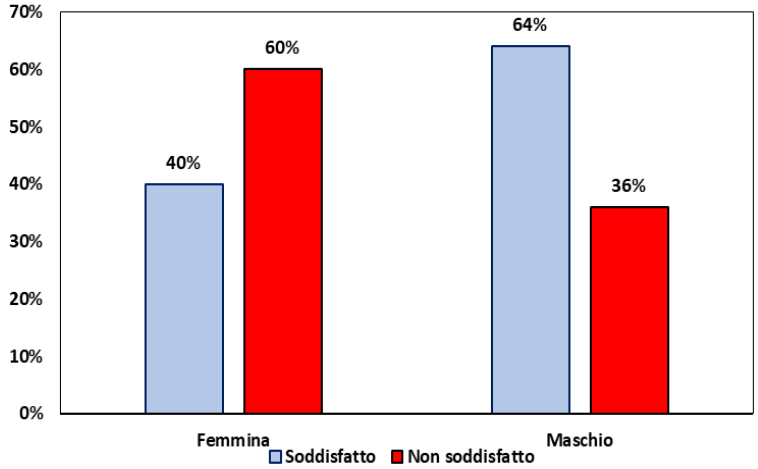
| | **Tempo sui social** | | |
|---|---|---|---|
| **Genere** | < di 3 ore | > di 3 ore | Totale |
| ***Frequenze osservate*** | | | |
| Femmina | 15 | 25 | 40 |
| Maschio | 189 | 90 | 279 |
| Totale | 204 | 115 | 319 |
| ***Frequenze attese*** | | | |
| Femmina | 25,58 | 14,42 | 40 |
| Maschio | 178,42 | 100,58 | 279 |
| Totale | 204 | 115 | 319 |
| Test chi-quadro | 13,88 | | |
| Livello di significatività | 0,05 | | |
| Gradi di libertà | 1 | | |
| Valore critico | 3,84 | | |
| p-value | 0,000195 | | |
| **Si rifiuta l'ipotesi nulla** | | | |

*(Elaborazione Excel con PHStat)*

Tale relazione è confermata dai risultati del test Chi-quadrato relativo ai caratteri *"Genere"* e *"Tempo sui social"*, quest'ultima variabile, per semplificare l'elaborazione, è stata resa dicotomica ed è stata classificata in due categorie: *"< di 3 ore"* e *"> di 3 ore"* (Tabella 5).

Con un procedimento analogo si sono incrociati i dati delle distribuzioni "*Genere delle vittime e Soddisfazione aspetto fisico*", evidenziando come il non essere soddisfatto del proprio aspetto fisico appare in percentuale maggiore nelle vittime femminili rispetto a quelle maschili (Figura 8).

**Figura 8.** *Genere delle vittime e soddisfazione del proprio aspetto fisico.*




Il test del Chi-quadrato in questo caso ha fornito un p-value pari 0,2155 che non consente di rifiutare l'ipotesi nulla, cioè si può ritenere che non esista connessione tra genere della vittima e soddisfazione del proprio aspetto fisico.

## 4. Conclusioni

L'indagine condotta tra gli studenti dell'Istituto Tecnico "Euclide" di Bari ha fornito alcune indicazioni sul fenomeno del cyberbullismo che trovano conferma in letteratura. Infatti, rispetto ai risultati presentati nel report ISTAT 2015, emerge anche nel nostro campione che il fenomeno è piuttosto diffuso tra i più giovani (nel nostro caso 14-15 anni), coinvolge circa il 10% del campione e soprattutto il genere femminile.

Un dato confortante è che il fenomeno è limitato a pochi episodi, mentre le azioni sulle vittime si concretizzano soprattutto attraverso l'esclusione per dispetto dai gruppi social, come si osserva anche nelle indagini nazionali ed internazionali.

Quello che è scaturito dalla ricerca di relazioni tra variabili è che l'essere maschio o femmina e l'essere soddisfatto o non soddisfatto del proprio aspetto fisico è statisticamente connesso a l'essere vittima o no e lo stesso vale per il tempo trascorso sui vari social. Il risultato conferma ciò che è stato rilevato da fonti ufficiali come l'Istat o da altre indagini anche internazionali ed è anche ciò che è suggerito dal senso comune: le ragazze adolescenti, solitamente molto critiche verso il proprio aspetto fisico, sono più sensibili e spesso, cercando conferme e apprezzamenti, si espongono postando immagini o messaggi di carattere personale. Questo comportamento può spiegare il motivo per cui risultano più predisposte a subire insulti

o dispetti attraverso i dispositivi sui quali, come emerge anche dall'indagine, trascorrono gran parte della loro giornata.

## Riconoscimenti

Si ringrazia tutto il team dei docenti universitari partecipanti al PLS Statistica 2017/2018 e, in modo particolare la coordinatrice, prof.ssa Ribecco, per l'impegno nell'organizzazione e il valido supporto nelle attività di formazione e ricerca. Un ringraziamento particolare alla dott.ssa Mancarella, per la guida nelle attività laboratoriali e il contributo nella revisione del presente lavoro. Si ringraziano gli studenti Abbinante, Albergo, Patano, Paparella, Sartorio della classe IV A Logistica, per la collaborazione prestata alla raccolta e all'elaborazione dei dati e tutti gli studenti dell'Euclide che hanno risposto al questionario.

## Riferimenti bibliografici

MIUR, 2015, *Linee di Orientamento per la prevenzione e il contrasto del bullismo e del cyberbullismo*, www.istruzione.it/allegati/2015/2015_04_13_16_39_29.pdf

*Legge 29 maggio 2017*, n.71, www.gazzettaufficiale.it/eli/id/2017/06/03/17G00085/sg

Nota Miur del 27/10/2017, *Linee di Orientamento per la prevenzione e il contrasto del cyberbullismo*

Istat 2015, *Il bullismo in Italia: comportamenti offensivi e violenti tra i giovanissimi*, www.istat.it/it/files/2015/12/Bullismo.pdf

*2016 Cyberbullying Data*, www.cyberbullying.org/2016-cyberbullying-data

*Bullismo in UK, Ditch The Label* - The Annual Bullying Survey 2017, www.ditchthelabel.org/research-papers/the-annual-bullying-survey-2017

*Safer Internet Day, Studio sugli effetti del Cyberbullismo in Italia*, 2014, www.savethechildren.it/cosa-facciamo/pubblicazioni/safer-internet-day-study--il-cyberbullismo

*Nella rete della rete, Report annuale Osservatorio Nazionale Adolescenza*, 2017, www.adolescienza.it/osservatorio/nella-rete-della-rete-report-annuale-osservatorio-nazionale-adolescenza-2017

## Sitografia

https://www.generazioniconnesse.it
https://www.Istat.it

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018* ISBN: 978-88-6629-048-3
DIPARTIMENTO DI ECONOMIA E FINANZA pp. 277-296

# Non solo cyberbulli a spasso in città

**Giuliana Gatti*, Cosimina Ligorio**

*Liceo G. Bianchi Dottula - Bari*

**Riassunto:** L'indagine ha avuto l'obiettivo di investigare sul cyber-bullismo che può essere considerato la nuova frontiera del bullismo tra gli adolescenti. In particolare, è stata orientata ad analizzare se gli studenti del Liceo G. Bianchi Dottula mettono in atto comportamenti responsabili piuttosto che essere autori, vittime o spettatori di atti di molestie, intimidazioni, esclusioni, con le modalità offerte dai nuovi media, nei confronti dei loro coetanei. Il problema è di grandi dimensioni anche se vittime e autori ne parlano ancora poco, solo l'11% degli studenti ha dichiarato di essere stata vittima e solo il 9% di essere stato autore. Le informazioni risultano più significative se a parlare è uno spettatore: il 52% degli intervistati dichiara infatti di aver assistito ad atti di cyberbullismo. Le motivazioni per atti di cyberbullismo sono varie, ma per il 36% degli autori possono addirittura verificarsi degli atti senza alcuna ragione. Gli atti più frequenti sono scherzi telefonici, minacce e insulti via telefono ed esclusione intenzionale delle vittime da gruppi di amici sui social. Risulta, infine, che circa il 3% degli intervistati è stato sia vittima che autore di atti di cyberbullismo. I dati raccolti sono stati confrontati qualitativamente con altre indagini sullo stesso tema presenti in letteratura.

**Keywords:** cyberbullismo, social, smartphone, adolescenti.

## 1. Introduzione

Si conosce il bullismo dei tempi più recenti, ma anche i cyber-bulli o le cyber-bulle insultano, minacciano e aggrediscono virtualmente a livello psicologico. Usano so-

* Giuliana Gatti: giuliana.gatti@liceobianchidottula.gov.it
Cosimina Ligorio: cosimina.ligorioi@liceobianchidottula.gov.it
Il lavoro qui descritto è frutto di un progetto comune, ma Giuliana Gatti ha provveduto alla redazione dei paragrafi 1 e 4, mentre Cosimina Ligorio ha redatto il paragrafo 2 e 3.

prattutto tecnologie digitali, ovvero l'invio di messaggi verbali, foto o video tramite cellulari, smartphones, pc, tablet (su social network, siti web, blog, ecc..), che rendono praticamente impossibile sottrarsi alle vessazioni. È una dinamica purtroppo comune tra bambini e adolescenti, che si lega strettamente ai bisogni della loro crescita espressi in modo problematico: ad esempio, la paura di essere esclusi o la ricerca dell'ammirazione degli altri. A differenza del bullismo tradizionale, qui la vittima può essere colpita 24 ore su 24 e ovunque si trovi. Nemmeno la casa è un rifugio sicuro. In più, il cyberbullo può avere un pubblico molto vasto, potenzialmente infinito, e continuare a rimanere anonimo, o come minimo non raggiungibile fisicamente. Questo spinge a colpire in modo ancora più aggressivo e violento, dicendo cose che dal vivo non si direbbero. Le conseguenze quindi possono essere gravi e persistenti come nel bullismo tradizionale, anche se non c'è contatto fisico.

Il cyberbullismo è avvertito dalla comunità degli adulti come un problema che colpisce i ragazzi minorenni in particolare gli adolescenti e i preadolescenti. Infatti, da alcuni fatti di cronaca particolarmente gravi, è nata l'esigenza di avere uno strumento normativo a tutela dei minorenni vittima del cyberbullismo. Il 29 maggio 2017 è stata varata la legge 71 in cui il cyberbullismo è definito: *"qualunque forma di pressione, aggressione, molestia, ricatto, ingiuria, denigrazione, diffamazione, furto d'identità, alterazione, acquisizione illecita, manipolazione, trattamento illecito di dati personali in danno di minorenni, realizzata per via telematica, nonché la diffusione di contenuti on line aventi ad oggetto anche uno o più componenti della famiglia del minore il cui scopo intenzionale e predominante sia quello di isolare un minore o un gruppo di minori ponendo in atto un serio abuso, un attacco dannoso, o la loro messa in ridicolo".*

L'obiettivo della legge è contrastare il fenomeno del cyberbullismo in tutte le sue manifestazioni, con azioni a carattere preventivo e con una strategia di attenzione, tutela ed educazione nei confronti dei minori coinvolti, sia vittime sia attori di atti illeciti. La legge pone l'attenzione sul ruolo della scuola come istituzione in prima linea per il contrasto di questi fenomeni. Il Ministero dell'Istruzione ha adottato infatti delle linee di orientamento per la prevenzione e il contrasto del cyberbullismo nelle scuole. Nelle linee guida è illustrato il progetto "Generazioni Connesse", con lo scopo di fornire alle istituzioni scolastiche una serie di strumenti didattici, di immediato utilizzo, tra cui: attività di formazione (online e in presenza) rivolte in maniera specifica alle comunità scolastiche e attività di informazione e sensibilizzazione realizzate in collaborazione con la Polizia di Stato per approfondire i temi della navigazione sicura in Rete.

Sullo studio di questo fenomeno, nell'ambito della nostra scuola, si basa l'indagine che è il risultato della partecipazione di alcuni studenti della III AE, della III BE e della IV AE del Liceo Economico Sociale "G. Bianchi Dottula", coordinati dalle docenti Giuliana Gatti e Cosimina Ligorio, al Piano Lauree Scientifiche Statistica 2017/2018.

Nel progetto PLS Statistica 2017/2018, a cui partecipa una rete di scuole di diverse province della regione Puglia, i docenti delle scuole secondarie superiori, in collaborazione con i tutor universitari, hanno individuato come tema di ricerca il *Cyberbullismo*, da analizzare attraverso un unico questionario comune a tutte le scuole sugli atteggiamenti psicosociali e sugli usi dei mezzi tecnologici. Sono state individuate due macro-sezioni per la formulazione del questionario:

I. informazioni sulle abitudini dell'uso dei mezzi tecnologici da parte degli studenti;
II. comportamenti e motivazioni del cyberbullismo.

La seconda sezione, a sua volta, contiene le dimensioni relative:

- all'essere stati autori, vittime o spettatori di atti di cyberbullismo;
- alla tipologia di atti a cui gli studenti abbiano eventualmente preso parte;
- alle reazioni agli atti subiti e/o osservati.

L'attività di progettazione del questionario ha coinvolto ciascuno studente che, attraverso l'attività di focus group svolta a scuola sotto il tutoraggio delle docenti, ha partecipato personalmente alla formulazione di proposte di quesiti da inserire nel questionario.

Il team dei docenti universitari, in rete con gli insegnanti delle scuole coinvolte nel PLS, ha successivamente elaborato il questionario nella sua forma definitiva: questo è stato quindi somministrato on line attraverso la piattaforma Google Doc a tutti gli studenti coinvolti nel progetto.

Dopo la fase di raccolta dei dati, insegnanti e studenti si sono dedicati all'analisi degli stessi, utilizzando in primo luogo metodi di analisi statistica descrittiva, per l'analisi univariata e bivariata dei dati.

Si sono quindi confrontati i dati raccolti con altre indagini presenti in letteratura sullo stesso tema; sebbene il confronto sia di carattere prevalentemente qualitativo poiché le indagini fanno riferimento a campioni diversi e utilizzano quesiti differenti si è cercato di estrapolare le informazioni comuni per ottenere l'andamento del fenomeno nel tempo e di far emergere le peculiarità del contesto analizzato.

Il gruppo di lavoro del liceo "G. Bianchi Dottula" ha svolto l'indagine allo scopo di conoscere le opinioni degli intervistati rispetto ai seguenti temi: il ruolo degli

studenti rispetto alla messa in atto di azioni di cyberbullismo; l'analisi della percezione di tali azioni; la gestione delle reazioni dei singoli studenti a tali azioni; il confronto tra i dati raccolti a scuola e quelli raccolti su scala nazionale.

Gli studenti intervistati gravitano, per residenza domestica o per frequenza scolastica, nella città di Bari. La nostra indagine, per quanto parziale, può essere considerata una prima analisi di modi e mezzi del cyberbullismo a Bari e nei comuni dell'area metropolitana.

## 2. Materiali e metodi

Nel liceo "G. Bianchi Dottula", la raccolta dei dati è avvenuta tramite somministrazione on line del questionario ad un gruppo di 425 studenti interessati alle tematiche proposte. Il campione è composto dagli studenti che per adesione volontaria hanno risposto alle domande e quindi, come tale, non può essere considerato un campione probabilistico sebbene il numero totale degli studenti iscritti sia 699. Pertanto il campione rappresenta circa il 61% dell'intera popolazione portandoci a ritenere che i dati raccolti siano sufficientemente significativi.

Le analisi compiute sono state, dapprima, a carattere descrittivo e univariato. In seguito, si è studiata l'associazione tra vari caratteri al fine di spiegare le potenziali relazioni di causa-effetto tra coppie di caratteri oggetto della rilevazione, attraverso tabelle di contingenza, mediante l'indice di connessione di Pearson (Chi-quadrato, $\chi^2$) e quindi, per una più facile lettura del risultato, di una sua versione normalizzata (indice V di Cramer $\varphi^2$).

Il software utilizzato per le analisi è Microsoft Excel, a disposizione della scuola che ne possiede le licenze ed è di facile utilizzo per gli studenti.

## 3. Risultati

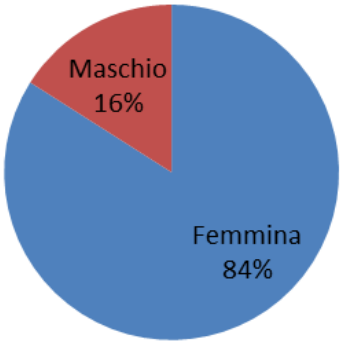
Dopo la caratterizzazione del campione attraverso l'analisi dell'età, del sesso e del comune di residenza, lo studio si è focalizzato su:

- dispositivi tecnologici più utilizzati dagli adolescenti, social network "usati" e tempo dedicato;
- motivazioni generiche del fenomeno cyberbullismo e confronto con le motivazioni dichiarate dagli attori;
- tipologie di azioni compiute e reazioni subite o osservate al cyberbullismo;
- suddivisione del campione tra attori, spettatori e vittime.

### *3.1 Caratterizzazione del campione*

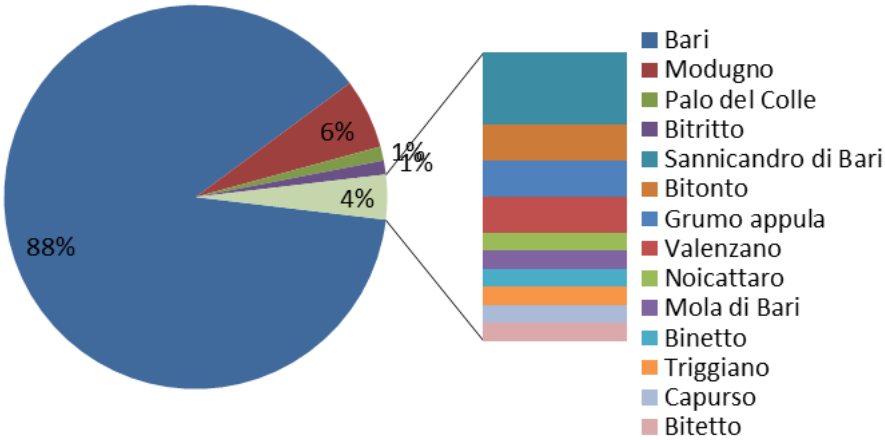
Nell'indagine sono stati intervistati 425 adolescenti, in prevalenza ragazze (84%), questo perché la scuola per la natura dei suoi indirizzi si connota come una scuola a prevalenza femminile (domanda 3 del questionario, Figura 1).

**Figura 1.** *Distribuzione degli intervistati per Genere.*



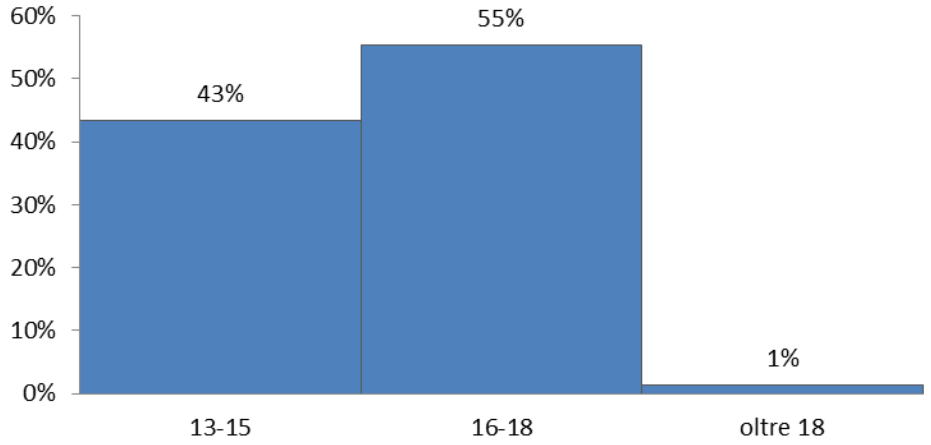

Si è inoltre analizzata la domanda n. 1 riguardante il comune di residenza dell'intervistato (Figura 2). Si osserva che il maggior numero di intervistati risiede a Bari (l'88%), seguono Modugno (6%), Palo del Colle e Bitritto (l'1,2%) in misura inferiore gli studenti provengono da altri comuni limitrofi alla città.

**Figura 2.** *Distribuzione degli intervistati per Comune di residenza.*




Infine, si è analizzata l'età degli intervistati (domanda 4 del questionario) suddividendoli in classi di età come riportato in Figura 3. La maggior parte degli studenti (55%) appartiene alla fascia di età tra i 16 e 18 anni.

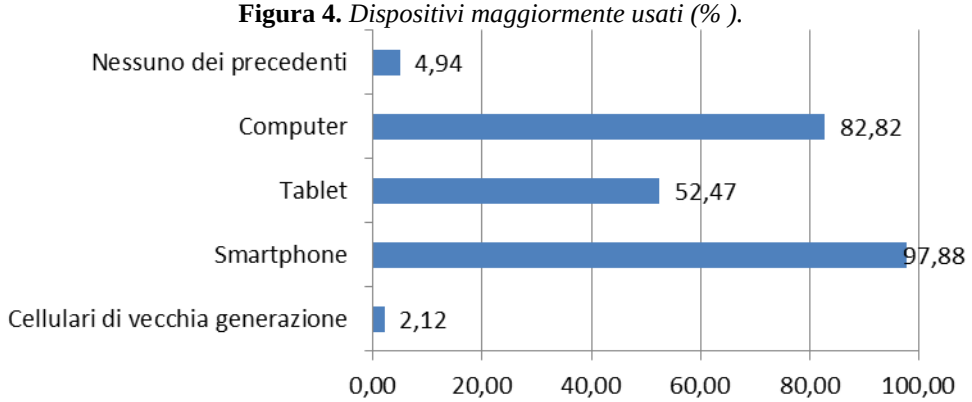
**Figura 3.** *Distribuzione degli intervistati per fascia d'età.*




### *3.2 Dispositivi tecnologici più utilizzati dagli adolescenti, social network "usati" e tempo dedicato*

Quasi tutti gli adolescenti intervistati possiedono dispositivi di nuova generazione con i quali risultano connessi alla rete, sia fuori che dentro casa. Come si evince dalla domanda 14, i cui risultati sono riportati in Figura 4, il dispositivo maggiormente utilizzato dagli intervistati è lo smartphone con circa il 98%, a cui segue il computer con l'83%.

**Figura 4.** *Dispositivi maggiormente usati (% ).*




Dalla domanda 17 del questionario emerge che i social network più utilizzati sono sicuramente le *App* di messaggistica istantanea (98%), seguite da Instagram e Youtube, mentre si registra una certa disaffezione a Facebook, utilizzato solo dal 78% degli intervistati (Figura 5).

**Figura 5.** *Social network più popolari.*

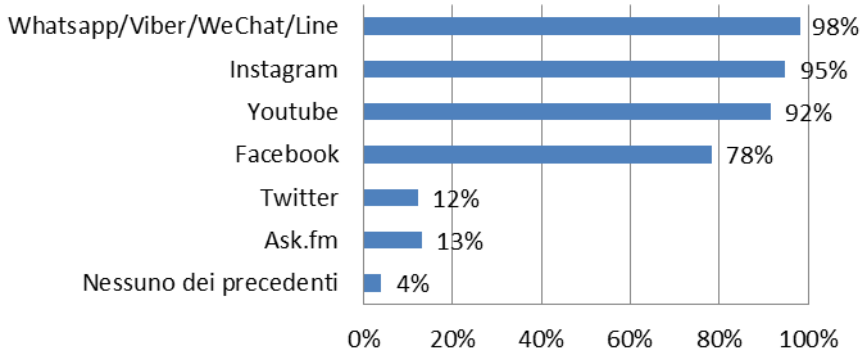
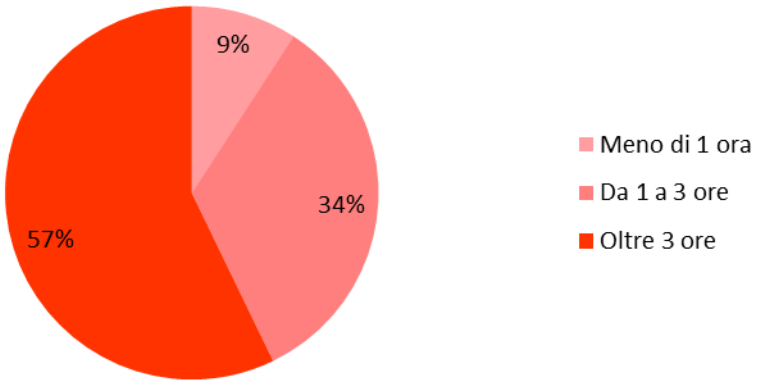



**Figura 6.** *Tempo dedicato ai social network.*



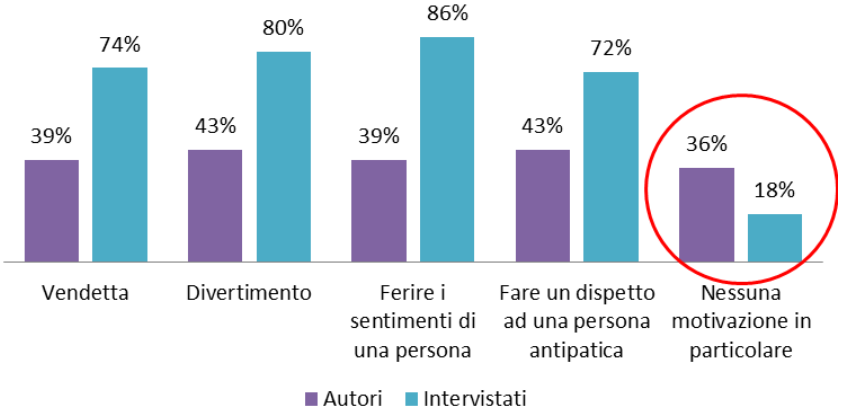

Inoltre, risulta che il tempo dedicato ai social per più del 50% degli intervistati superi le 3 ore al giorno (domanda n. 19) (Figura 6).

### *3.3 Motivazioni generiche del fenomeno cyber bullismo e motivazioni degli autori*

Quali sono le possibili motivazioni del cyberbullismo, in generale e in particolare per gli studenti che hanno dichiarato di essere stati autori di atti di cyberbullismo? Le domande 21 e 34 del questionario mirano a far emergere questo aspetto. Pertanto, i risultati ottenuti evidenziano che per tutti gli intervistati le più frequenti cause risultano il "Ferire i sentimenti di una persona" e il "Divertimento" con risposte superiori all'80% (Figura 7). Non si rileva, invece, una distinzione netta tra le motivazioni che inducono gli autori a compiere atti di cyberbullismo.

È interessante notare come se per il 18% degli intervistati possono verificarsi atti di cyberbullismo senza una motivazione precisa, questa percentuale si raddoppia nelle risposte degli autori.

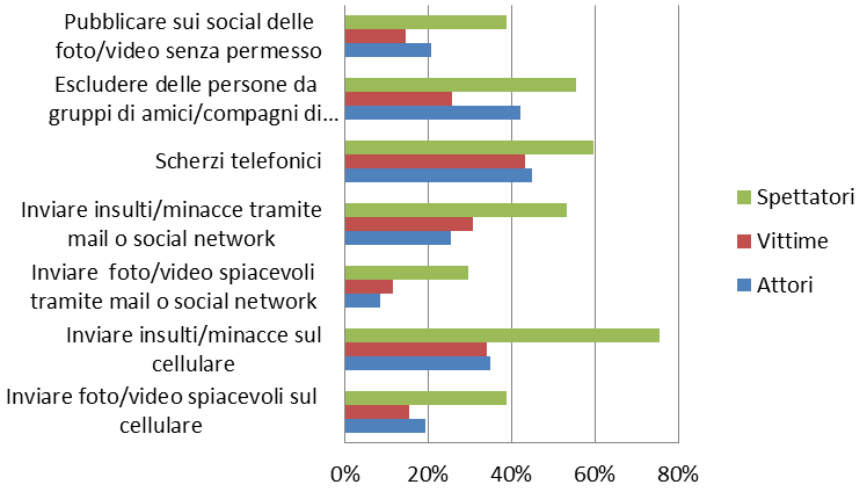
**Figura 7.** *Principali motivazioni del cyberbullismo.*




### *3.4 Tipologie di azioni compiute o subite o osservate*

Le domande 2, 29 e 33 del questionario mirano ad analizzare le tipologie di azioni subite dagli intervistati che hanno dichiarato di essere state vittime, quelle messe in campo da coloro che sono stati autori di atti di cyberbullismo e quelle rilevate da coloro che sono stati spettatori di queste azioni. Per le vittime gli atti più frequenti sono gli scherzi telefonici e le minacce ricevute sul cellulare; per gli autori sono gli scherzi telefonici e l'esclusione dai gruppi di amici; per gli spettatori sono gli insulti e le minacce inviate sul cellulare.

**Figura 8.** *Tipi di atti commessi.*




È da notare come la frequenza degli atti osservati assumano percentuali significativamente più alte rispetto a quelle dichiarate sia dalle vittime che dagli autori

(Figura 8). Ciò significa che gli studenti, se sono spettatori, sono più inclini a classificare come atti di cyberbullismo degli atti che non sono tali, sia per gli autori, che verosimilmente non danno peso ai gesti da loro compiuti, che per le vittime che subiscono queste azioni.

### *3.5 Suddivisione del campione di intervistati: la vittima*

Dalla domanda 22 risulta che l'11% degli alunni è stato vittima di cyberbullismo, in particolare il 6% al di fuori della scuola, il 4% all'interno della scuola e solo l'1% sia a scuola che fuori (Figura 9).

**Figura 9.** *Vittime e luoghi in cui avvengono gli atti di cyberbullismo.*

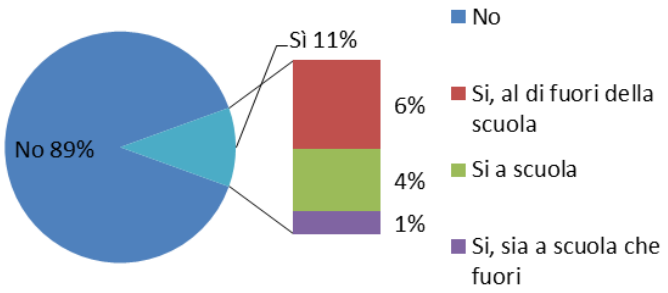
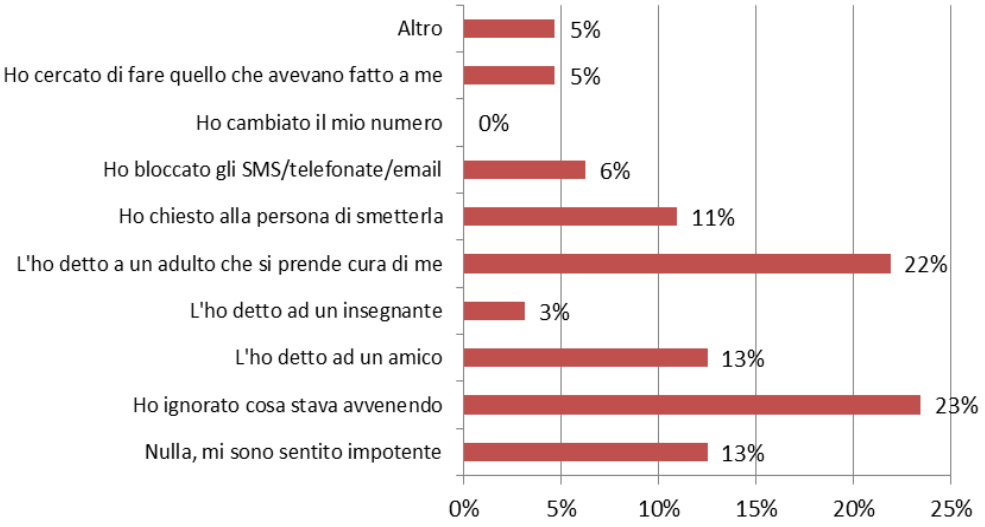



**Figura 10.** *Reazione delle vittime agli atti di cyberbullismo.*




Alla domanda 26 "Cosa hai fatto quando sei stato vittima di atti di cyberbullismo" ha risposto l'11% di studenti i quali avevano ammesso di essere stato vittima di cyberbullismo; gli studenti hanno reagito per lo più confidandosi con un adulto,

genitore o insegnante, o un amico (38%), oppure ignorando cosa stava accadendo e sperando che gli episodi non si ripetessero (23%). I risultati sono sintetizzati in Figura 10.

È da notare che solo il 3% delle vittime di cyberbullismo si rivolge ad un docente per avere supporto, questo dato mette in luce come è necessario all'interno della scuola individuare delle figure che si occupino di queste forme di disagio degli adolescenti e proprio come richiede il Ministero all'interno delle linee guida, sia individuato un docente referente con il compito di coordinare le iniziative di prevenzione e di contrasto del cyberbullismo.

Si è quindi proceduto ad analizzare le cause del fenomeno ipotizzando delle relazioni di associazione tra il carattere "Essere stato vittima di cyberbullismo" (domanda 22) e la "Soddisfazione per il proprio aspetto fisico" (domanda 9) o con "Credi di avere buoni rapporti con i tuoi genitori?" (domanda 13).

Come mostrato nella tabella seguente (Tabella 1), il 33% degli studenti è insoddisfatto del proprio aspetto fisico, ma non si osserva associazione tra le due variabili. Infatti, *l'indice di connessione di Pearson, chi-quadro*, è risultato pari a 18,93.

**Tabella 1.** *Relazione tra "Soddisfazione per il proprio aspetto fisico" e "l'essere stata vittima"*

| Quanto sei soddisfatto del tuo aspetto fisico? | Sei stato vittima di cyberbullsimo? | | | | Totale |
|---|---|---|---|---|---|
| | No | Si a scuola | Si, al di fuori della scuola | Si, sia a scuola che fuori | |
| Abbastanza | 215 | 7 | 9 | 1 | 232 |
| Molto | 51 | 1 | 2 | 0 | 54 |
| Per nulla | 22 | 2 | 1 | 1 | 26 |
| Poco | 90 | 7 | 12 | 4 | 113 |
| Totale | 378 | 17 | 24 | 6 | 425 |

Per una interpretazione più immediata, è stato usato l'*indice di contingenza quadratica media relativa di Cramer*, per il quale vale:

$$0 \leq \varphi^2 \leq 1,$$

che è risultato pari a 0,015. Pertanto, essendo il valore prossimo allo zero non è possibile ritenere che esista una evidente associazione tra i due caratteri.

Nella tabella sottostante (Tabella 2) sono state classificate l'"Essere stato vittima di cyber-bullismo" con la domanda "Credi di avere buoni rapporti con i tuoi genitori?" ed è stata analizzata la relazione tra dette variabili. Anche in questo caso i caratteri non risultano associati, infatti l'*indice di contingenza quadratica media relativa di Cramer* è pari a 0,024.

**Tabella 2.** *Relazione tra "La relazione con i genitori" e "L'essere stata vittima"*

| Credi di avere buoni rapporti con i tuoi genitori? | Sei mai stato vittima di cyberbullismo? | | | | Totale |
|---|---|---|---|---|---|
| | No | Si a scuola | Si, fuori della scuola | Si, sia a scuola che fuori | |
| No, con nessuno dei miei genitori | 15 | 2 | 1 | | 18 |
| Si con entrambi | 298 | 10 | 16 | 1 | 325 |
| Si, solo com mia madre | 51 | 4 | 7 | 3 | 65 |
| Si, solo con mio padre | 14 | 1 | 0 | 2 | 17 |
| Totale complessivo | 378 | 17 | 24 | 6 | 425 |

### *3.6 Suddivisione del campione degli intervistati: lo spettatore*

Dalla domanda 27 risulta che il 52% degli alunni è stato spettatore di atti di cyberbullismo; in particolare, il 17% al di fuori della scuola, il 16% all'interno della scuola e solo il 19% sia a scuola che fuori (Figura 11).

**Figura 11.** *Spettatori di atti di cyberbullismo.*

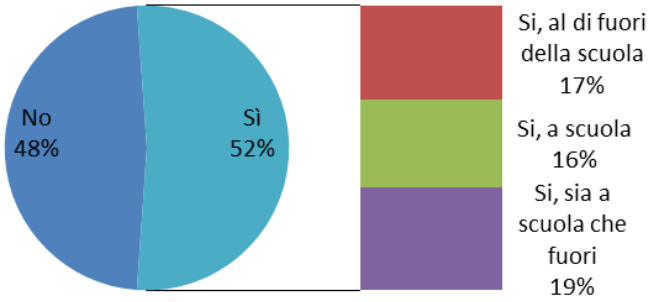
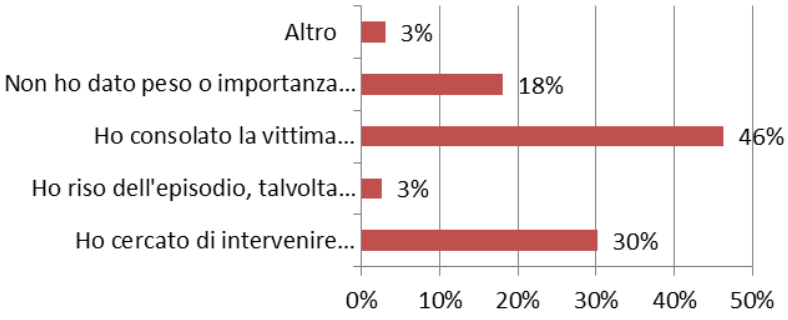



**Figura 12.** *Reazione degli spettatori.*




Alla domanda 30 "Se hai visto o saputo di qualcuno che è stato/a vittima di cyberbullismo, qual è stata la tua reazione", quasi la metà degli studenti dichiara di essere stata spettatrice di atti di cyberbullismo e dichiara di aver consolato la vitti-

ma, cercando di aiutarla (46%); solo il 30% degli studenti ha cercato di intervenire affinché gli episodi non si ripetessero. I risultati sono riportati in Figura 12.

### *3.7 Suddivisione del campione di intervistati: l'attore*

Alla domanda 31 "Hai mai preso parte ad episodi di cyberbullismo?", il 9% degli studenti ammette di aver preso parte ad episodi di cyberbullismo di cui, in particolare, il 2% al di fuori della scuola, il 4% all'interno della scuola e il 3% sia a scuola che fuori dalla scuola (Figura 13).

**Figura 13.** *Gli attori di atti di cyberbullismo.*

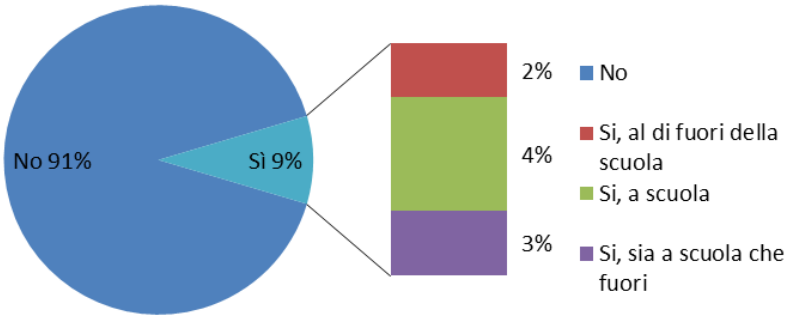
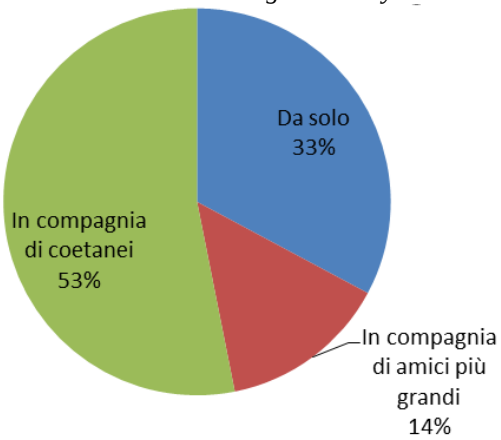



**Figura 14.** *Modalità di esecuzione degli atti di cyberbullismo.*




Come riportato in Figura 14, alla domanda 32 "Negli episodi ai quali hai preso parte, eri: in compagnia o da solo", il 67% degli intervistati ha dichiarato che le azioni sono avvenute in compagnia o di coetanei (53%) o di amici più grandi (14%). È da notare come chi ha compiuto cyberbullismo ha spesso avuto l'ausilio di terzi (collaboratori o semplici spettatori).

Si è quindi proceduto ad analizzare le cause del fenomeno ipotizzando la presenza di associazione tra il carattere "Essere stato attore di cyberbullismo" (domanda 31) e "Credi di avere buoni rapporti con i tuoi genitori?" (domanda 13).

**Tabella 3.** *Relazione tra "La relazione con i genitori" e "L'essere stato attore"*

| Credi di avere buoni rapporti con i tuoi genitori? | Hai mai preso parte ad episodi di cyberbullismo? | | | | Totale |
|---|---|---|---|---|---|
| | No | Si, a scuola | Si, al di fuori della scuola | Si, sia a scuola che fuori | |
| No, con nessuno dei miei genitori | 16 | 1 | 0 | 1 | 18 |
| Si con entrambi | 299 | 13 | 5 | 8 | 325 |
| Si, solo com mia madre | 56 | 4 | 4 | 1 | 65 |
| Si, solo con mio padre | 15 | 1 | 0 | 1 | 17 |
| Totale complessivo | 386 | 19 | 9 | 11 | 425 |

L'analisi di un legame tra la "relazione con i genitori" e l'essere stato "autore" di cyberbullismo non ha prodotto risultati significativi, l'*indice di contingenza quadratica media relativa (V di Cramer)* è pari a 0,007.

Infine, si è analizzato il legame tra "Essere stato attore di cyberbullismo" (domanda 31) e "Essere stato vittima di cyberbullismo" (domanda 21).

**Tabella 4.** *Relazione tra "L'essere stato vittima" e "L'essere stato attore"*

| Sei mai stato vittima di cyberbullismo? | Hai mai preso parte ad episodi di cyberbullismo? | | | | Totale |
|---|---|---|---|---|---|
| | No | Si, a scuola | Si, al di fuori della scuola | Si, sia a scuola che fuori | |
| No | 351 | 13 | 5 | 9 | 378 |
| Si a scuola | 13 | 2 | 1 | 1 | 17 |
| Si, al di fuori della scuola | 17 | 4 | 2 | 1 | 24 |
| Si, sia a scuola che fuori | 5 | 0 | 1 | 0 | 6 |
| Totale complessivo | 386 | 19 | 9 | 11 | 425 |

L'analisi di una relazione tra l'essere stato "vittima" e l'essere stato "autore" di cyberbullismo non ha prodotto risultati significativi, infatti il valore dell'*indice di contingenza quadratica media relativa (V di Cramer)* è pari a 0,021, sebbene ci siano 12 studenti su 425 cioè il 2,8% che hanno rivestito contemporaneamente i due ruoli. Questo indica che gli adolescenti hanno una notevole confusione e difficoltà nell'individuare gli atti di cyberbullismo e a dar loro la giusta gravità considerandoli a volte come scherzi innocenti e rivestendo alternativamente i due ruoli.

## 4. Confronto tra i dati locali e quelli su scala nazionale

I risultati relativi ad alcune domande del questionario somministrato nell'ambito dell'indagine fra la popolazione scolastica sono stati comparati con i risultati ottenuti da alcune domande simili relative al report dell'IPSOS sul cyberbullismo per "Save the Children" del 2013. L'indagine è stata condotta su scala nazionale su un campione di 405 maschi e di 405 femmine, tra i 12 e i 17 anni, di cui 317 del Sud e isole. Tali dati non sono perfettamente sovrapponibili con quelli dell'indagine scolastica, poiché provenienti da campioni con strutture differenti e con quesiti non identici, ma alcuni specifici item possono essere confrontati per comprendere l'andamento del fenomeno nel tempo. "Save the Children" ha, a sua volta, già confrontato i propri risultati con un'altra indagine del 2011.

Il primo punto di contatto tra le due ricerche è l'uso dei principali dispositivi di connessione alla rete. Dal 2013 al 2011 si impenna la crescita dell'uso degli strumenti più sofisticati, in particolare lo smartphone che passa dal 35% al 71% (Figura 15.a) con valori più alti, pari al 74%, per il Sud e le Isole (Figura 15.b).

**Figura 15.a** *Dispositivi maggiormente utilizzati (fonte ricerca IPSOS).*

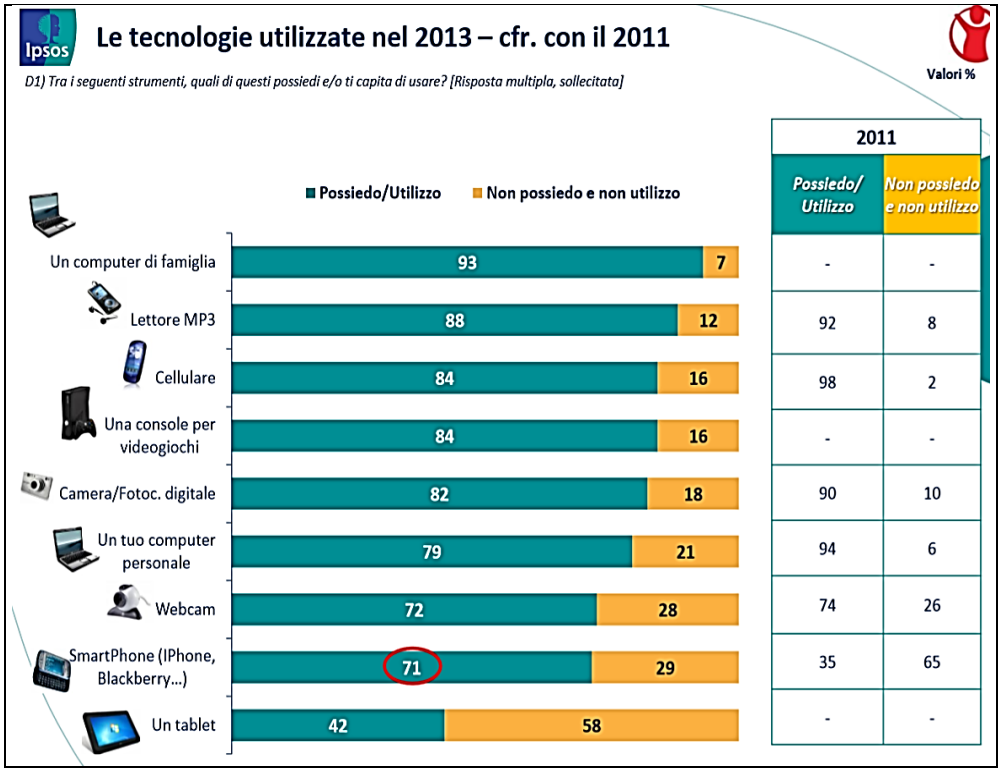


**Figura 15.b** *Dispositivi maggiormente utilizzati (fonte ricerca IPSOS).*

**Le tecnologie utilizzate nel 2013 – approfondimento età per sesso/area geografica**

*D1) Tra i seguenti strumenti, quali di questi possiedi e/o ti capita di usare? [Risposta multipla, sollecitata]*

Valori %

| Possiedo/Utilizzo | TOTALE | TOTALE Maschi | TOTALE Femm. | ETA' PER SESSO 12-14 anni Maschi | 12-14 anni Femm. | 15-17 anni Maschi | 15-17 anni Femm. | AREA GEOGRAFICA Nord Ovest | Nord Est | Centro | Sud + Isole |
|---|---|---|---|---|---|---|---|---|---|---|---|
| UN COMPUTER DI FAMIGLIA | 93 | 93 | 94 | 94 | 95 | 92 | 93 | 95 | 97 | 91 | 92 |
| LETTORE MP3 | 88 | 89 | 87 | 87 | 87 | 91 | 88 | 89 | 90 | 89 | 87 |
| CELLULARE | 84 | 82 | 85 | 83 | 85 | 81 | 86 | 82 | 86 | 84 | 84 |
| UNA CONSOLE PER VIDEOGIOCHI | 84 | 93 | 76 | 94 | 78 | 93 | 74 | 84 | 83 | 82 | 87 |
| TELECAMERA/ FOTOCAMERA DIGITALE | 82 | 78 | 86 | 77 | 86 | 80 | 86 | 81 | 80 | 82 | 85 |
| UN TUO COMPUTER PERSONALE | 79 | 79 | 79 | 73 | 75 | 84 | 82 | 72 | 79 | 86 | 81 |
| WEBCAM | 72 | 74 | 71 | 65 | 70 | 80 | 72 | 71 | 63 | 73 | 77 |
| SMARTPHONE (IPHONE, BLACKBERRY ...) | 71 | 71 | 71 | 64 | 67 | 76 | 74 | 72 | 62 | 72 | 74 |
| UN TABLET | 42 | 44 | 40 | 43 | 39 | 44 | 40 | 41 | 42 | 37 | 44 |

*Base: totale intervistati*

Ipsos Public Affairs

La crescita è confermata anche nell'indagine scolastica, in cui l'uso dello smartphone ha una frequenza del 98% circa.

Un altro punto di contatto tra le due indagini è il tempo trascorso on-line; mentre nelle due ricerche dell'IPSOS del 2011 e del 2013 le percentuali di coloro che sono connessi per più di tre ore al giorno sono circa il 50% (Figura 16), nell'indagine eseguita a scuola risulta che sempre più adolescenti (circa il 57%) passano più di tre ore al giorno connessi alla rete.

Si registra inoltre la rapidità dell'evoluzione di piattaforme sociali come Whatsapp; infatti nell'indagine del 2013 solo il 20% aveva un profilo Whatsapp (Figura 17) mentre la percentuale nell'indagine condotta a scuola è il 98%, non ci sono dati nel 2011 dato che l'applicazione non esisteva. I risultati della nostra indagine hanno sottolineato inoltre la disaffezione a Faceboook, infatti nel 2011 l'81% utilizzava la piattaforma, la percentuale si solleva nel 2013 (87% ) mentre nel 2018 si riduce (78%). Nella ricerca effettuata presso il nostro Istituto scolastico, si riscontra la presenza di Instagram, che ha fatto la sua comparsa nel 2013 e risulta assente nelle indagini anteriori.

**Figura 16.** *Tempo trascorso on line, in ore pro-die (fonte ricerca IPSOS).*

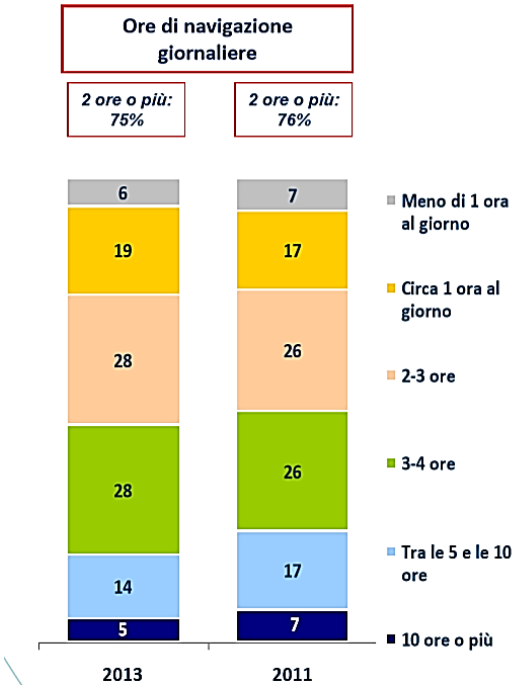
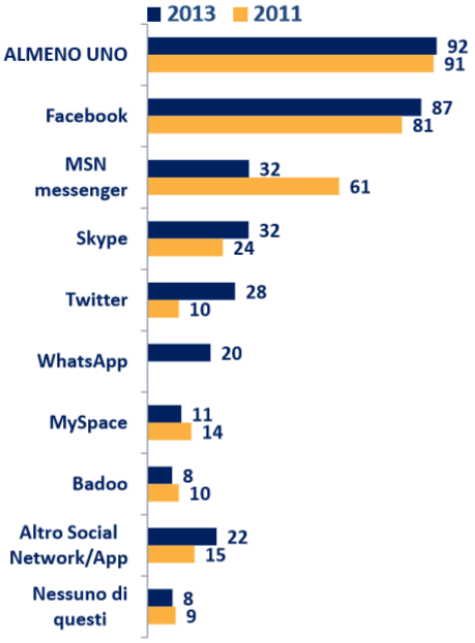
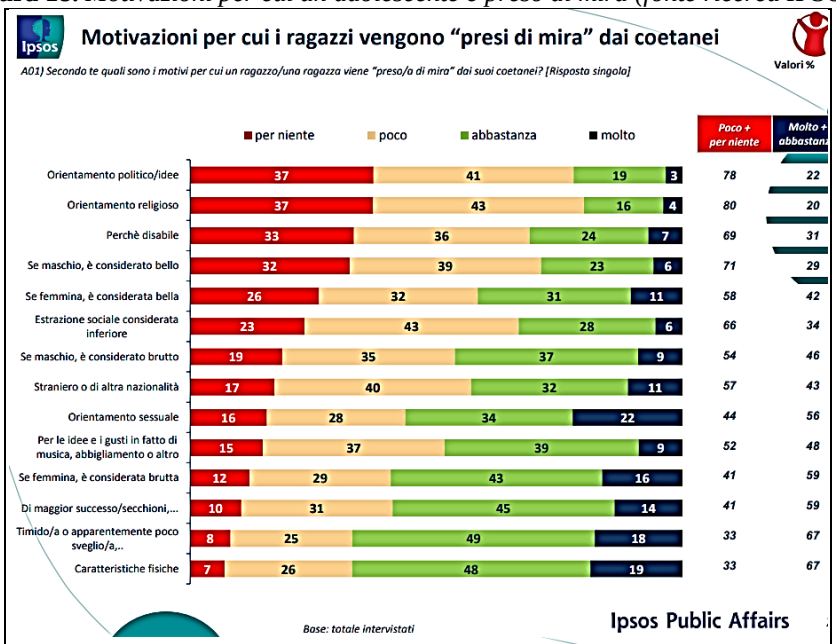



**Figura 17.** *Piattaforme social maggiormente usate (fonte ricerca IPSOS).*



Circa il cyberbullismo, dal rapporto IPSOS risulta che, secondo gli intervistati del Sud (Figura 18), i ragazzi maggiormente "presi di mira" sono quelli con particolari caratteristiche fisiche o quelli considerati timidi, femmine o maschi che siano, o ancora i cosiddetti "secchioni" o, infine, quelli con diverso orientamento sessuale.

**Figura 18.** *Motivazioni per cui un adolescente è preso di mira (fonte ricerca IPSOS).*




**Motivazioni per cui i ragazzi vengono "presi di mira" dai coetanei - approfondimento età per sesso/area geografica**

*A01) Secondo te quali sono i motivi per cui un ragazzo/una ragazza viene "preso/a di mira" dai suoi coetanei? [Risposta singola]*

Valori %

| Molto + abbastanza | TOTALE | | | ETA' PER SESSO | | | | AREA GEOGRAFICA | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | 12-14 anni | | 15-17 anni | | | | | |
| | TOTALE | Maschi | Femm. | Maschi | Femm. | Maschi | Femm. | Nord Ovest | Nord Est | Centro | Sud + Isole |
| Orientamento politico/idee | 22 | 21 | 23 | 19 | 23 | 23 | 23 | 19 | 19 | 24 | 25 |
| Orientamento religioso | 20 | 22 | 18 | 21 | 22 | 24 | 16 | 21 | 15 | 19 | 23 |
| Perché disabile | 31 | 30 | 32 | 31 | 36 | 30 | 30 | 30 | 30 | 27 | 35 |
| Se maschio, è considerato bello | 29 | 28 | 30 | 28 | 31 | 27 | 29 | 28 | 25 | 29 | 31 |
| Se femmina, è considerata bella | 42 | 41 | 42 | 41 | 45 | 41 | 40 | 40 | 44 | 39 | 43 |
| Estrazione sociale considerata inferiore | 34 | 32 | 34 | 34 | 35 | 31 | 33 | 25 | 30 | 29 | 41 |
| Se maschio, è considerato brutto | 46 | 46 | 46 | 52 | 49 | 42 | 44 | 47 | 40 | 45 | 49 |
| Straniero o di altra nazionalità | 43 | 43 | 43 | 41 | 45 | 45 | 41 | 42 | 40 | 45 | 44 |
| Orientamento sessuale | 56 | 58 | 55 | 62 | 55 | 55 | 55 | 58 | 50 | 54 | 59 |
| Per le idee e i gusti in fatto di musica, abbigliamento o altro | 48 | 47 | 50 | 54 | 55 | 42 | 45 | 49 | 43 | 38 | 55 |
| Se femmina, è considerata brutta | 59 | 57 | 62 | 65 | 63 | 51 | 61 | 62 | 53 | 59 | 60 |
| Di maggior successo/secchioni,... | 59 | 58 | 60 | 62 | 63 | 56 | 57 | 60 | 52 | 58 | 62 |
| Timido/a o apparentemente poco sveglio/a,.. | 67 | 64 | 70 | 68 | 71 | 61 | 70 | 64 | 70 | 73 | 65 |
| Caratteristiche fisiche | 67 | 62 | 71 | 66 | 77 | 60 | 67 | 71 | 58 | 67 | 68 |

*Base: totale intervistati*

Ipsos Public Affairs

Nell'indagine presso il nostro Istituto scolastico è emerso invece che non c'è una precisa motivazione per gli atti, inoltre è da sottolineare che il 36% degli autori ha dichiarato che ci possono essere atti di cyberbullismo anche senza una ragione particolare.

## 5. Conclusioni

I danni causati da cyberbullismo, cyber-minacce e sexting, ormai dilaganti su scala nazionale e internazionale, possono essere maggiori dei danni causati dal bullismo tradizionale, in ragione delle caratteristiche proprie della comunicazione mediata dalle nuove tecnologie. Infatti, le vittime di cyberbullismo esperiscono l'impos–sibilità di sottrarsi alle vessazioni. È già tristemente ampia la letteratura su questo fenomeno. La scuola potrebbe essere residenza elettiva del bullismo, poiché è luogo in cui gli adolescenti passano gran parte del loro tempo e primo luogo di socializzazione e di creazione e disfacimento dei legami significativi non famigliari; inoltre, il mondo virtuale rinforza le relazioni reali di bullismo: il carattere virtuale della rete sostiene la potenziale assenza di limiti, rende anonimi e quindi apparentemente non perseguibili, e consente di falsare i protagonisti e i contenuti.

Questa è la ragione per cui la scuola è il luogo ideale per analizzare questi temi e con gli studenti stessi interrogarsi sulle ragioni del fenomeno e sulle possibili vie d'uscita.

L'indagine conferma i dati di precedenti indagini: gli smartphone sono sempre più diffusi e circa il 60% degli adolescenti sono connessi ai vari social network per più di tre ore al giorno. Questa esposizione continua a questi strumenti e alla rete fa emergere una nuova esigenza educativa a cui sia la famiglia che la scuola devono dare una risposta. Gli adolescenti vanno educati all'uso di questi nuovi mezzi informandoli sulle conseguenze della diffusione dei loro dati privati e in particolare della loro immagine, spesso sono presenti dei ragazzi che con molta leggerezza e ignoranza concedono informazioni che diventano l'oggetto principale di atti di cyberbullismo. Inoltre, bisogna ribadire che l'uso di uno schermo e di una tastiera non assicura l'anonimato e l'impunità e che quindi le azioni compiute posso essere perseguibili.

Il problema è di grande portata, anche se vittime e autori ne parlano ancora poco. Solo l'11% degli intervistati dichiara di essere stata vittima di cyberbullismo e solo il 9% degli intervistati dichiara di esserne stato autore. Le informazioni sono più significative se a parlare è uno spettatore. Il 52% degli intervistati dichiara di aver assistito ad atti di cyberbullismo.

Le motivazioni per atti di cyberbullismo sono varie, ma per il 36% degli autori possono verificarsi degli atti senza alcuna ragione. Questo è il dato più preoccupante perché espone tutti gli adolescenti al pericolo di essere colpiti e denuncia come questi atti non hanno alla base necessariamente una volontà intenzionale di voler far male ma una assoluta superficialità e incapacità di attribuire valore alle proprie azioni e una assoluta incapacità a provare empatia per la vittima.

Gli atti più frequenti sono scherzi telefonici, minacce e insulti sul telefono e esclusione intenzionale delle vittime da gruppi di amici sui social, atti che sono percepiti dagli adolescenti vittime come estremamente gravi ma che vengono compiuti dagli autori con estrema superficialità. E' da notare che circa il 3% degli intervistati è stato sia vittima che autore di atti di cyberbullismo questo sta ad indicare come non c'è una netta distinzione tra i due ruoli e l'estrema confusione che avvertono sia le vittime che gli autori di questi atti.

## Riferimenti bibliografici

Borra S.; Di Ciaccio A. (2014). *Statistica (metodologie per le scienze economiche e sociali)*. Mc Graw-Hill Education.

Borrazzo F. P.; Perchinunno P.(2007). *Analisi statistiche con Excel*. Pearson Education.

Girone G. (2009). *Statistica*. Cacucci Editore, Bari.

Olivieri D. (2007). *Fondamenti di Statistica*. CEDAM Edizioni, Bologna.

Delvecchio F. (2015). *Statistica per lo studio dei fenomeni sociali*. CLEUP, Padova.

## Sitografia

http://www.altalex.com/documents/news/2016/09/21/bullismo-e-cyberbullismo (Legge cyberbullismo)

http://www.miur.gov.it/documents/20182/0/Linee+Guida+Bullismo+2017.pdf/4df7c320-e98f-4417-9c31-9100fd63e2be?version=1.0 (Linee guida del Ministero dell'Istruzione)

http://www.repubblica.it/cronaca/2018/02/06/news/sette_bambini_su_dieci_non_denunciano_agli_adulti_di_essere_vittima_di_cyberbullismo-188160004/

http://www.moige.it/indaginecyberbullismo

http://www.rollingstone.it/rolling-affairs/news-affairs/il-cyberbullismo-segreto-e-lo-squadrismo-su-facebook/2018-03-03/

http://www.generazioniconnesse.it/site/it/ilcyberbullismo/

http://www.e-abc.eu/it/ricerca/

http://questure.poliziadistato.it/it/Bari/articolo/15715a7c31175a9fb743273465

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
*Metodi e analisi statistiche 2018* ISBN: 978-88-6629-048-3
DIPARTIMENTO DI ECONOMIA E FINANZA pp. 297-309

# Atteggiamenti psicosociali sugli usi dei mezzi tecnologici

**Antonia Maria de Candia, Carmela Visaggi**
*I.I.S.S."Galileo Ferraris" - Molfetta (Ba)*

**Riassunto:** La legge n.71/2017 "*Disposizioni a tutela dei minori per la prevenzione ed il contrasto del fenomeno del cyber-bullismo*" assegna alle istituzioni scolastiche il compito di contrastare il fenomeno del cyber-bullismo in tutte le sue ma–nifestazioni con azioni di prevenzione e strategie di attenzione, tutela ed educazione dei minori coinvolti (vittime o responsabili di illeciti) assicurando l'attuazione degli interventi senza distinzione d'età. In tale ottica va letta l'indagine statistica svolta dagli studenti dell'I.I.S.S. "Galileo Ferraris" che hanno aderito al PLS Statistica 2017/18, i cui obiettivi prioritari sono la valutazione della percezione e della diffusione del fenomeno all'interno della comunità scolastica di appartenenza e la classificazione dei nuovi fenomeni della devianza online allo scopo di potenziare il sistema di governance (istituzione, società civile,adulti e minori).

## 1. Introduzione

Il gruppo di lavoro che ha partecipato al PLS Statistica 17/18 è composto da 13 studenti, di cui nove frequentanti la classe quarta e quattro la classe quinta dell'ITT degli indirizzi *Elettronica ed Elettrotecnica* ed *Informatica*, rispettivamente. Inoltre, parallelamente la classe quarta sez. C del *Liceo Scientifico* opzione *Scienze Applicate* ha seguito lo stesso percorso formativo nell'ambito delle attività di Alternanza Scuola Lavoro. Il tema oggetto di indagine è stato prontamente condiviso dal gruppo di lavoro in quanto il cyber-bullismo è un fenomeno di cui i giovani parlano ed è ritenuto molto grave perché in pochissimo tempo le vittime possono vedere la propria reputazione danneggiata in una comunità molto ampia dal momento che i contenuti, una volta pubblicati, possono riapparire a più riprese in luoghi diversi. Accade spesso che genitori e insegnanti ne rimangono a lungo all'oscuro, perché non hanno accesso alla comunicazione in rete degli adolescenti, per cui, prima che

un caso venga alla luce, può trascorrere molto tempo. Le vittime, gli autori e gli spettatori di atti di bullismo in rete hanno spesso remore a parlare. Di conseguenza per i familiari e gli insegnanti è difficile riconoscere il problema. Non è sempre possibile distinguere chiaramente tra autori e vittime, perché molti giovani rimasti vittime del mobbing lo hanno praticato anche in prima persona. Gli studenti hanno perciò riflettuto sugli indizi in ambito scolastico che possono indurre i docenti e il gruppo classe a nutrire fondati sospetti sull'esistenza del fenomeno. Il clima in classe che peggiora progressivamente, la mancanza di coesione tra i compagni di classe, la presenza di complicità solo nel prendersela con capri espiatori, comportamenti sgarbati, la rottura di amicizie, l'isolamento/esclusione di alcuni compagni di classe, il marinare spesso la scuola, il calo di rendimento, improvvisi atteggiamenti aggressivi, ecc. sono stati individuati come possibili sintomi dell'esistenza del problema o di possibili disagi.

La prevenzione del fenomeno è stata ritenuta perciò fondamentale e solo in una scuola che assume un atteggiamento chiaro gli studenti possono trovare il coraggio di segnalare un problema. Questa dunque la principale spinta a proporre un questionario on line, rivolto agli studenti di tutto l'istituto, per perseguire i seguenti obiettivi specifici:

- ✓ indagare sull'utilizzo critico e consapevole dei social network e dei media;
- ✓ valutare la percezione e la diffusione del cyber-bullismo da parte degli studenti;
- ✓ inquadrare i nuovi fenomeni della devianza online per potenziare il sistema di governance (istituzione, società civile, adulti e minori).

Gli studenti hanno elaborato i dati, dapprima procedendo con un'analisi univariata (raccolta, tabulazione e opportune rappresentazioni grafiche per ciascun item) per poi passare ad un'analisi bivariata al fine di studiare la presenza di eventuali relazioni tra alcuni caratteri, in particolare, in relazione al sesso oppure all'età dei rispondenti, a tale scopo sono stati utilizzati alcuni indici statistici, quali l'indice di connessione di Pearson $\chi^2$ e l'indice V di Cramer . A supporto delle attività la piattaforma http://economia.osel.it/ e i libri di testo in adozione [1] [2].

## 2. Materiali e metodi

Gli studenti hanno concordato le modalità di codifica dei dati attribuendo un codice a ciascuna domanda e un valore numerico relativo alla corrispondente opzione della risposta scelta dal rispondente. E' stata compilata in Excel la matrice dei dati

(database) dove per ogni riga è stato riportato l'identificativo del rispondente e ad ogni colonna è stata associata la corrispondente risposta; dal database sono state predisposte le tabelle pivot per procedere alla successiva analisi univariata dei dati raccolti.

Sono stati tabulati i dati per ciascun item del questionario, realizzate le distribuzioni di frequenza di ogni variabile (quasi tutte qualitative ad eccezione dell'età dei rispondenti e dei giudizi espressi con valutazione numerica) e scelta, di volta in volta, la rappresentazione grafica più idonea a descrivere le caratteristiche principali delle mutabili in esame, alcune delle quali sconnesse altre ordinali. La scelta del grafico è stata fatta in primis sulla base della tipologia dei dati e anche sulla base della maggior chiarezza nella lettura del grafico stesso.

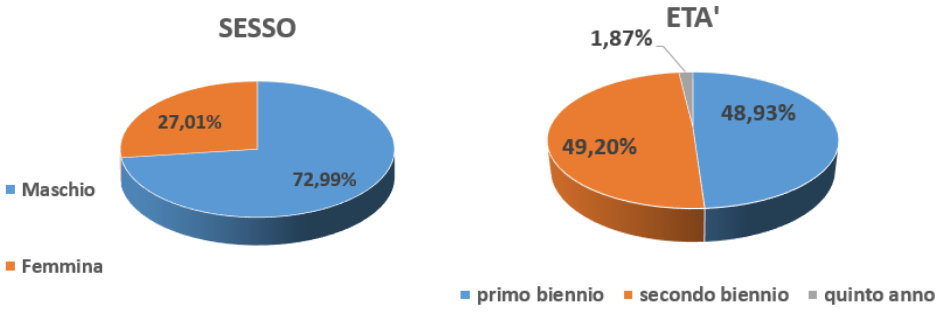
Successivamente l'analisi bivariata è stata finalizzata allo studio del grado di dipendenza statistica di due caratteri distinti; in particolare, per alcuni quesiti, si è cercato di capire la presenza o meno di un'eventuale connessione tra il sesso dei rispondenti e le risposte da essi fornite. Mediante la predisposizione di tabelle a doppia entrata sono state calcolate le distribuzioni marginali dei due caratteri in esame e le distribuzioni condizionate.

Gli studenti hanno confrontato le distribuzioni di frequenze relative condizionate e calcolato l'indice *Chi-quadrato di Pearson* e, successivamente, l'indice *V di Cramer* per stabilire l'eventuale associazione tra le mutabili analizzate.

## 3. Risultati

Nella Figura 1 sono riportati i risultati relativi alle caratteristiche del campione osservato rispetto al genere, all'età, alla classe frequentata dagli studenti (primo biennio/secondo biennio/quinto anno).

**Figura 1.** *Caratteristiche della popolazione statistica*

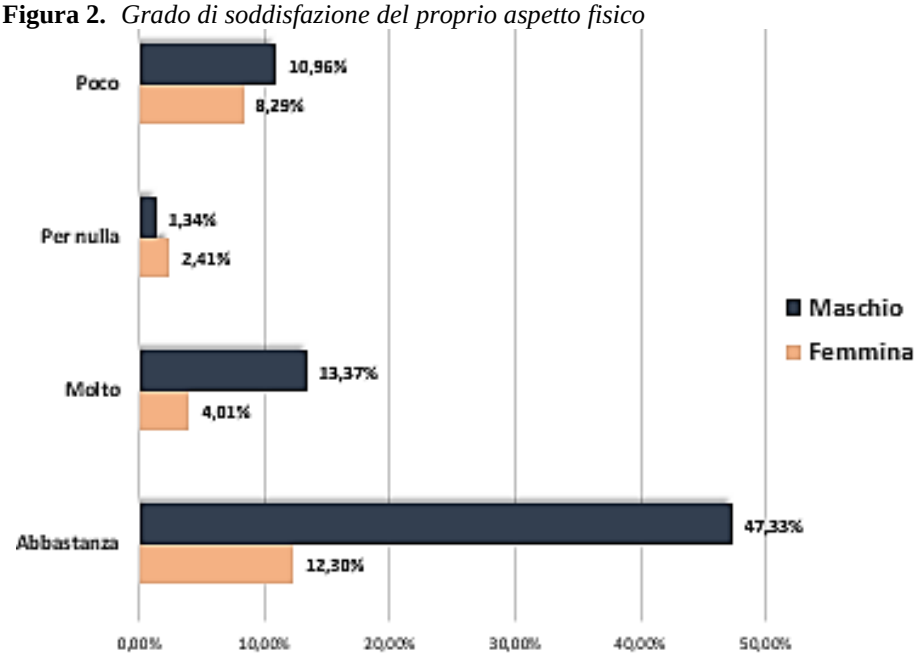


Come si può notare, il campione è prettamente maschile (72,99% dei rispondenti), dato coerente con la realtà scolastica; infatti l'utenza dell'istituto è prevalentemente maschile, soprattutto nelle classi dell'ITT. La quasi totalità delle risposte proviene dagli studenti frequentanti il primo e secondo biennio (equamente ripartite); solo un numero esiguo di studenti frequentanti la classe quinta ha risposto al questionario.

Nei sotto paragrafi successivi sono stati analizzati gli esiti in riferimento a ciascun obiettivo oggetto dell'indagine. In particolare, l'ambiente familiare sembra non avere influenza sul fenomeno del cyber-bullismo dal momento che la quasi totalità dei rispondenti vive in un contesto familiare "*normale*" (la composizione familiare del 96,99% dei rispondenti prevede la presenza di entrambi i genitori), i genitori hanno un'occupazione (il 90,37% dei padri e il 50% delle madri) ed hanno come titolo di studio un diploma di scuola secondaria di secondo grado (il 66,31% dei padri e il 62,03% delle madri) o un diploma di laurea (l'11,76% dei padri e il 14,71% delle madri).

Poiché il rapporto con il proprio corpo è un problema sentito in età adolescenziale per le ripercussioni che può avere da un punto di vista psicologico, emotivo e fisico e in considerazione del fatto che spesso gli atti di cyber-bullismo sono rivolti ai coetanei in sovrappeso o con qualche difetto fisico, è stato esaminato tale aspetto relativamente al campione di rispondenti (Figura 2).

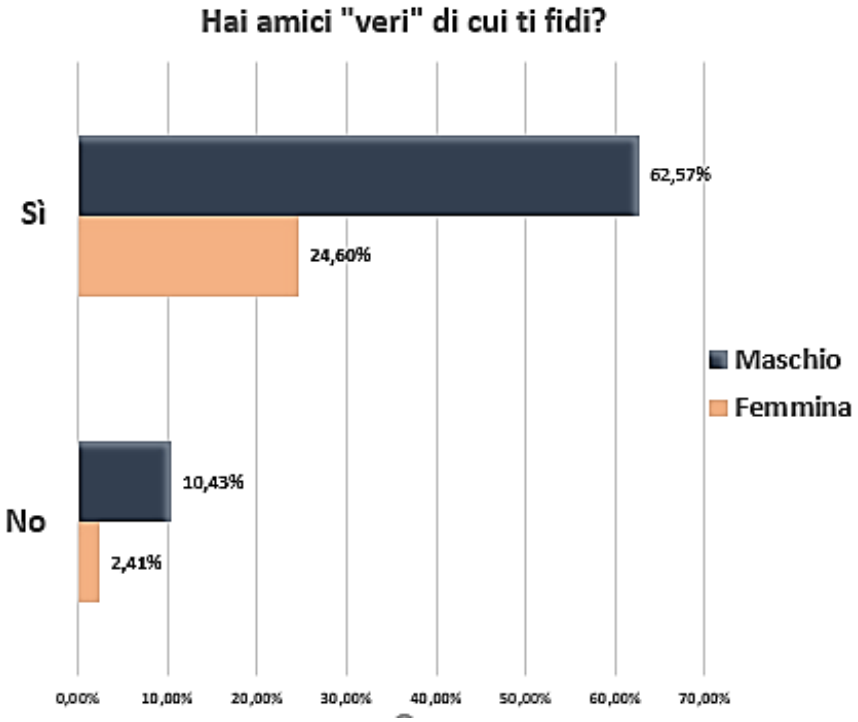
**Figura 2.** *Grado di soddisfazione del proprio aspetto fisico*



Indipendentemente dal sesso, la maggior parte dei rispondenti dichiara di essere abbastanza/molto soddisfatta del proprio aspetto fisico (l'indice V di Cramer ha infatti valori molto vicini allo zero) e di avere amici "veri" di cui fidarsi (Figura 3).

Inoltre, l'86,9% ritiene di avere buoni rapporti con entrambi i genitori e il 72,73% buoni rapporti con i compagni di classe.

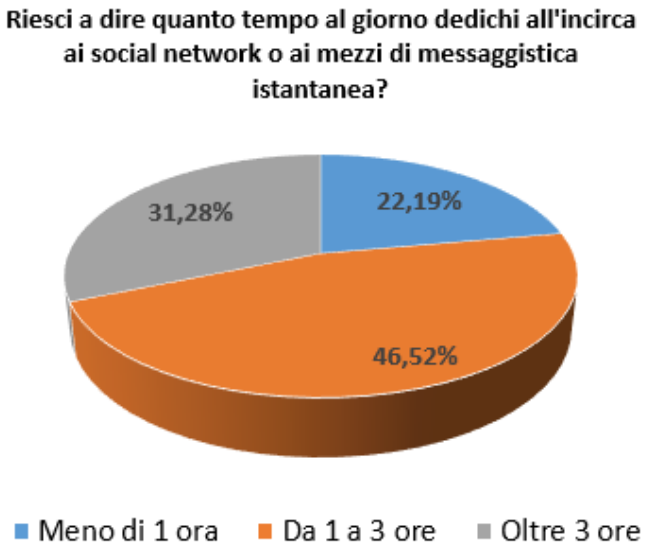
**Figura 3** *Rapporti di amicizia*




## 3.1 *Utilizzo critico e consapevole dei social network e dei media*

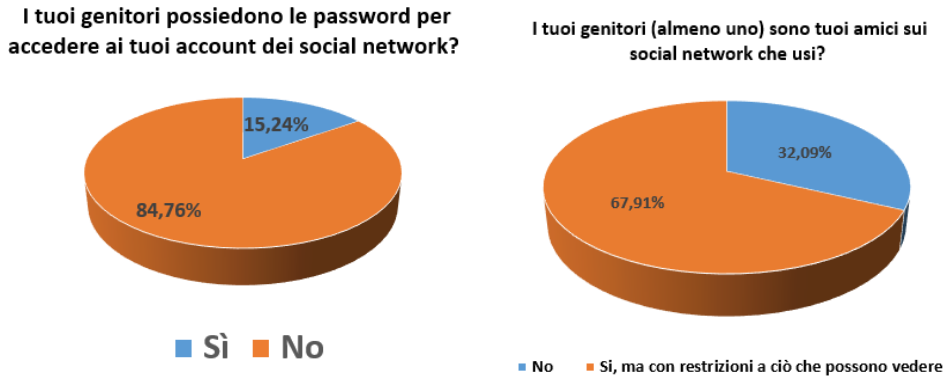
Sono stati esaminati gli item del questionario finalizzati a valutare l'utilizzo critico e consapevole dei nuovi strumenti di comunicazione che Internet mette a disposizione e dei mezzi tecnologici da parte dei rispondenti. In generale è emerso che quasi tutti i rispondenti hanno accesso ad Internet sia in casa (99,47%) che fuori casa (96,79%). Il 42,56% di essi quantifica da 1 a 3 ore il tempo dedicato quotidianamente ai social network o ai mezzi di messaggistica istantanea (figura 4), valori che risultano inferiori a quelli rilevati nell'ultima indagine MIUR curata da ***Skuola.net*** in collaborazione con l'Università di Firenze presentati il 6/02/2018 durante il Safer Internet Day 2018 [6].

**Figura 4** *Tempo dedicato ai media*



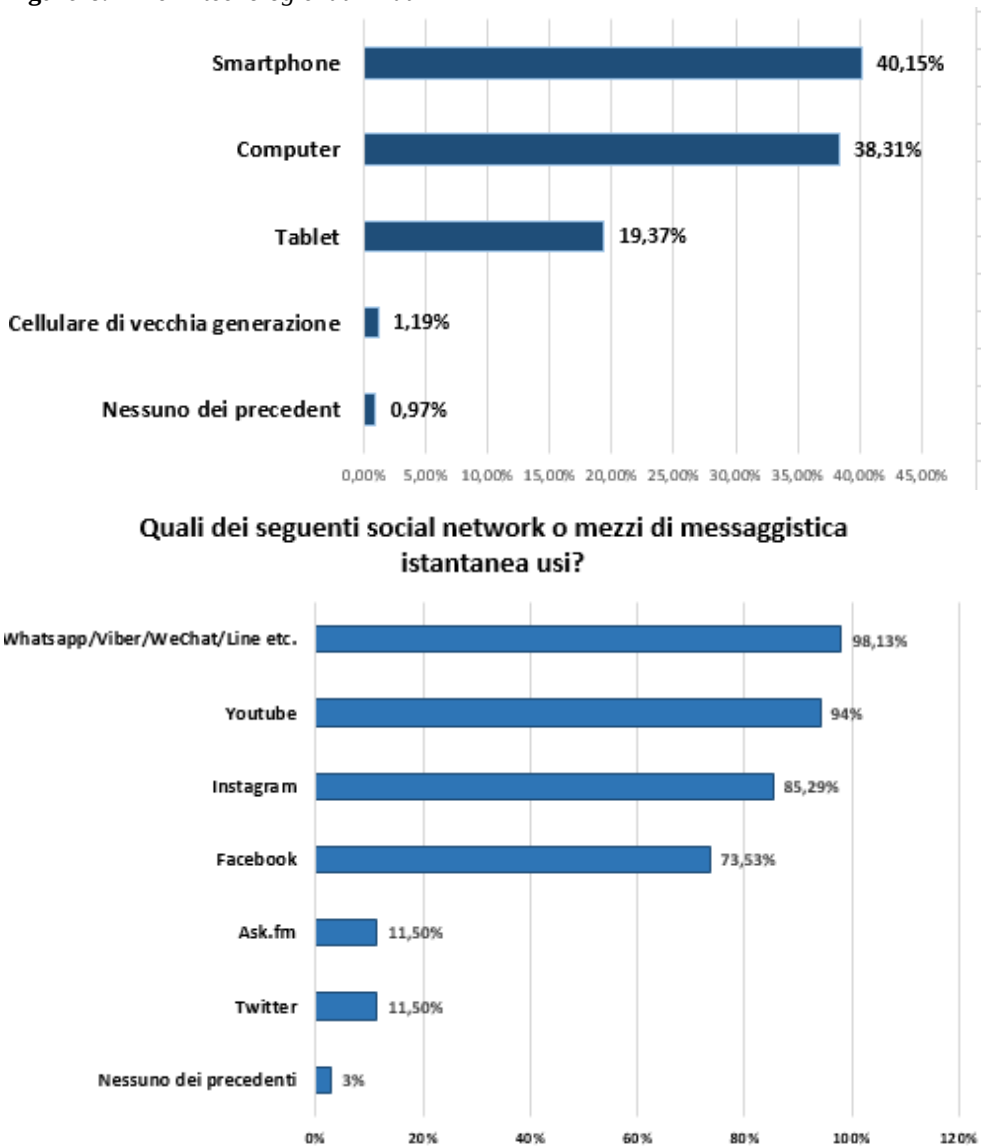

La stragrande maggioranza degli studenti (84,76%) non consente ai genitori l'accesso ai propri account dei social network ma, nello stesso tempo, il 67,91% dei genitori utilizza i social network dal momento che essi sono "*amici*" dei figli sui social network che questi utilizzano, anche se con restrizioni a ciò che possono vedere (Figura 5).

**Figura 5** *Relazioni tra genitori e figli sui social network*




Lo smartphone (40,15%), il computer (38,31%), il tablet (19,37%) sono i mezzi tecnologici usati solitamente mentre i social network o i mezzi di messaggistica più utilizzati sono Whatsapp/Viber/WeChat/Line etc., seguiti da Youtube, Instagram e Facebook, dati in linea con l'indagine MIUR (Figura 6).

**Figura 6.** *Mezzi tecnologici utilizzati*




## 3.2 *Percezione e diffusione del cyber-bullismo nella comunità scolastica*

Al fine di acquisire informazioni su quello che accade all'interno della comunità scolastica il questionario ha previsto una sezione che ha consentito di valutare la percezione e la diffusione del cyber-bullismo da parte dei rispondenti. Con riferimento a questi aspetti è risultato che il 40,91% ha dichiarato di aver assistito ad episodi in cui qualcuno è stato vittima di cyber-bullismo di cui il 18,98% sia a

scuola che fuori, il 15,24% al di fuori dell'ambiente scolastico e solo il 6,68% a scuola (Figura 7). Fra le motivazioni che inducono a mettere in atto episodi di cyber-bullismo i rispondenti evidenziano il divertimento (27,99%), la volontà di ferire i sentimenti della vittima (25,02%), fare un dispetto ad una persona antipatica (21,14%), il desiderio di vendetta (18,74%) e, per un'esigua minoranza, nessuna motivazione in particolare.

**Figura 7** *Testimoni di episodi di cyber-bullismo*

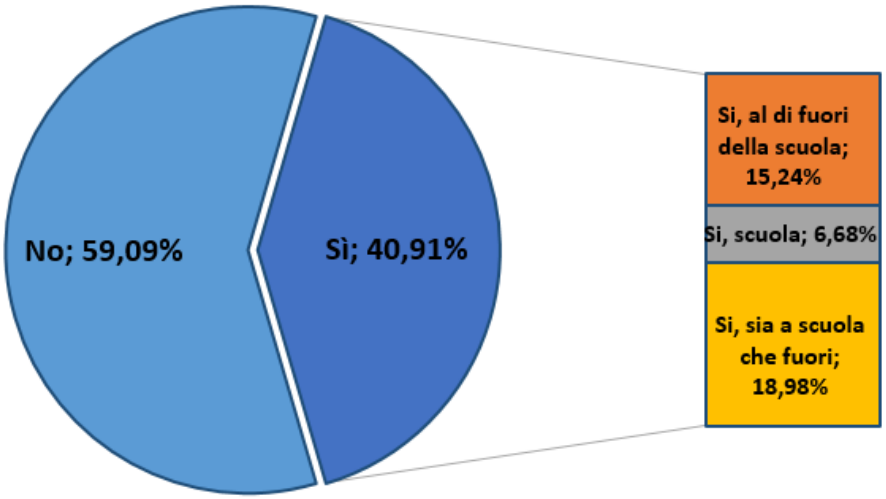
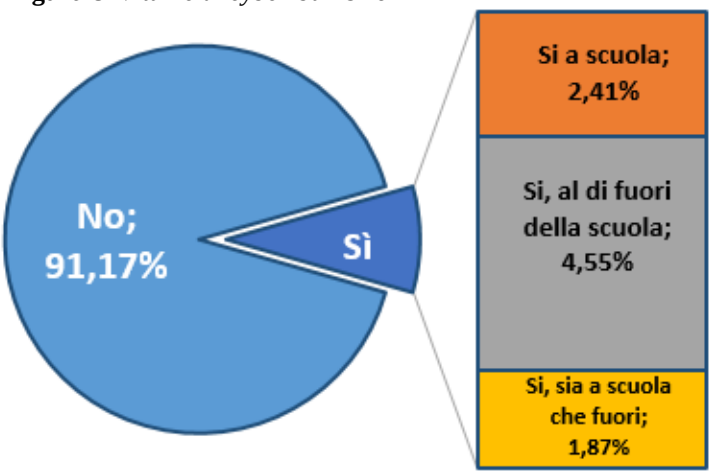



**Figura 8** *Vittime di cyber-bullismo*




Come mostrato in figura 8, solo l'8,83% dei rispondenti afferma di essere stato vittima di cyber-bullismo, con episodi verificatisi molto raramente (1 o 2 volte).

Dal grafico si evince infatti che il 91,17% del campione ha dichiarato di non esser mai stato vittima di cyber-bullismo. Mediante una verifica d'ipotesi sulla proporzione (parametro $\pi$) si è voluto confrontare il dato emerso dall'indagine nella Scuola con quello corrispondente dell'indagine MIUR che stabilisce che la percentuale di adolescenti che non hanno mai subito atti di cyber-bullismo sia pari al 94,4%. Posto un livello di significatività pari al 5%, è stato considerato il seguente sistema di ipotesi, mediante il quale si vuole verificare che $\pi$ sia pari alla percentuale dichiarata dal MIUR a partire dalla percentuale stimata mediante il campione di studenti dell'Istituto Ferraris:

$$H_0: \pi = 94{,}4$$
$$H_1: \pi < 94{,}4$$

Essendo la dimensione campionaria pari a $n = 374$, si è ritenuto di poter considerare che la distribuzione dello stimatore sia approssimativamente normale, pertanto si è fatto ricorso al test Z. :

$$z = \frac{p-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} = \frac{0{,}9118-0{,}944}{\sqrt{\frac{0{,}944(1-0{,}944)}{374}}} = -2{,}68$$

essendo $z_{0,05} = -1{,}64$ si rifiuta l'ipotesi formulata in quanto $-2{,}68 < -1{,}64$ , concludendo che la proporzione risultata dall'indagine svolta a scuola è differente da quella presentata dal MIUR.

L'esame delle risposte degli studenti (in totale 33) che hanno dichiarato di essere stati vittime di bullismo (pari all'8,82%) ha consentito, inoltre, di tracciare un profilo degli stessi evidenziando le principali caratteristiche della distribuzione delle loro età, come riportato nella Tabella 1:

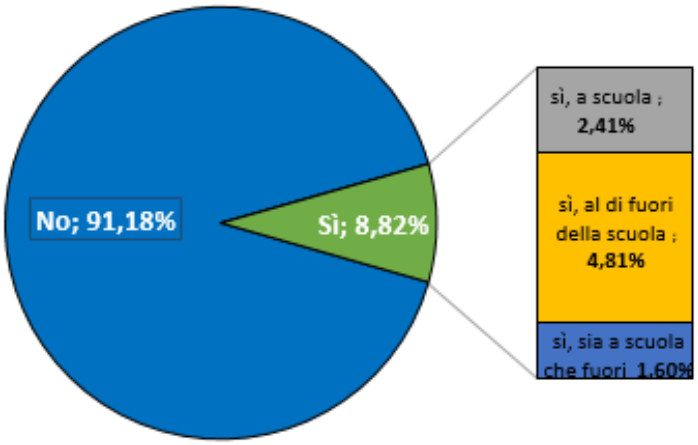
**Tabella 1**. *Principali valori caratteristici della distribuzione delle età delle vittime*

| | | | |
|---|---|---|---|
| Media | 15,57 | Varianza | 2,61 |
| Mediana | 15 | Deviazione standard | 1,61 |
| Moda | 14 | Coefficiente di variazione | 10,37% |
| Valore minimo | 13 | Indice di asimmetria di Fisher | 0,56 |
| Valore massimo | 19 | Indice di Curtosi | 2,05 |
| Range | 6 | Errore Standard | 0,28 |

Gli atti di cyber-bullismo si sono concretizzati in scherzi telefonici (27,94%), insulti/minacce sul cellulare (26,47%), esclusione da gruppi di amici/compagni di classe su Fb, Whatsapp (25%), ricezione di foto/video spiacevoli tramite email o social network (8,82%), pubblicazione di foto senza permesso sui social network e

ricezione di foto/video spiacevoli sul cellulare (5,88%). Le reazioni delle vittime di cyber-bullismo variano dalla volontà di ignorare la realtà dei fatti nella speranza che gli episodi non si ripetano (2,14%), il confidarsi con un genitore o un adulto (1,34%), comportarsi allo stesso modo del cyber-bullo (1,34%), nessuna reazione per la sensazione di impotenza avvertita (1,07%) ed infine, in minor misura, il confidarsi con un amico o il rivolgersi direttamente al cyberbullo chiedendogli di smetterla o bloccare gli SMS/telefonate/email. Successivamente si è indagato se i rispondenti avessero mai preso parte ad episodi di cyber-bullismo.

**Figura 9** *Partecipazione ad episodi di cyber-bullismo*




Dalla Figura 9 si evince che solo l'8,82% dei rispondenti ha partecipato ad episodi di cyber-bullismo, di cui il 4,81% al di fuori della scuola, il 2,41% a scuola e l'1,60% sia a scuola che fuori e per lo più in compagnia di coetanei (5,61%), da soli (2,08%) oppure in compagnia di amici più grandi.Tali episodi si sono concretizzati in scherzi telefonici (7,22%), insulti o minacce inviate sul cellulare della vittima (4,01%), esclusione della vittima da gruppi di amici/compagni di classe su Fb,Whatsapp, ecc. (3,74%), invio di foto/video spiacevoli sul cellulare (2,67%), invio di foto/video spiacevoli oppure insulti/minacce tramite email o social network (entambi con la stessa percentuale pari a 1,87%).

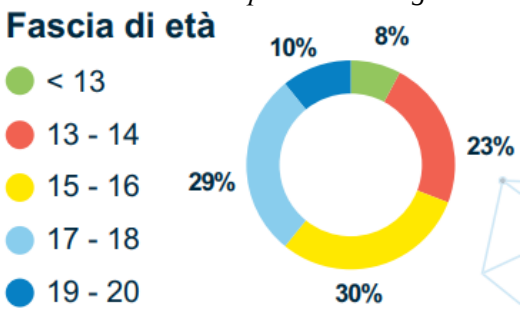
### *3.3 Potenziare il sistema di governance per i nuovi fenomeni della devianza on-line*

Gli esiti ottenuti sono stati confrontati con quelli dell'indagine condotta da*Skuola.netin collaborazione con l'Università di Firenze*, in occasione del *#SID2018 (Safer Internet Day)*, evento annuale, organizzato a livello internazionale con il

supporto della Commissione Europea nel mese di febbraio, al fine di promuovere un uso più sicuro e responsabile del web e delle nuove tecnologie, in particolare tra i bambini e i giovani di tutto il mondo.

Tale indagine ha coinvolto un campione di 3486 studenti la cui fascia d'età, riportata nella Figura 10, include anche un piccola percentuale (8%) di rispondenti di età inferiore a 13 anni.

**Figura 10** *Fasce d'età dei rispondenti indagine* ***Skuola.net***




Dal confronto emerge che gli studenti dell'Istituto "Ferraris" dedichino, quotidianamente, una minore quantità di tempo ai social network o ai mezzi di messaggistica istantanea rispetto ai rispondenti dell'indagine MIUR che invece dichiarano per il 31% di essere connessi da 3 a 4 ore al giorno, per il 22% da 5 a 10 ore e il 21% di essi è sempre connesso. Quest'ultima percentuale ("ragazzi sempre connessi") è passata dal 17% rilevato dalla stessa indagine nel 2016 al 21%. In realtà per il 31,28% dei rispondenti dell'Istituto "Ferraris", che hanno dichiarato di dedicare più di 3 ore ai social network, non è possibile quantificare esattamente il tempo di connessione. Infine, gli esiti delle due indagini relativamente ai social network utilizzati regolarmente, sono perfettamente in linea e vedono Whatsapp come strumento maggiormente usato seguito da Instagram, Youtube e Facebook in calo.

Fortunatamente a fronte del 12% dei rispondenti dell'indagine MIUR che dichiarano di comportarsi come cyberbulli, solo l'8,82% degli studenti dell'istituto ha preso parte ad atti di cyber-bullismo (di cui solo il 2,41% a scuola).

## 4. Conclusioni

I risultati dell'indagine, fortunatamente, denotano che nella comunità scolastica dell'Istituto non ci sono problematiche connesse all'uso non consapevole delle tecnologie digitali e che il fenomeno del cyber-bullismo è poco diffuso. Il sistema di

governance dell'Istituzione scolastica sembra andare dunque nella giusta direzione e la progettualità messa in atto nella comunità scolastica dell'IISS " Ferraris" volta alla prevenzione e al contrasto del fenomeno del cyber-bullismo e, in generale, al benessere e all'inclusione mediante interventi educativi e non punitivi, sembra produrre i suoi frutti.

L'approccio inclusivo scelto dall'Istituto è del resto in linea con il dettato normativo (legge n.71/2017) [3] che assegna un ruolo centrale alla scuola in cui il Dirigente Scolastico, nel definire le linee di indirizzo del Piano Triennale dell'Offerta Formativa (PTOF) e del Patto di Corresponsabilità (D.P.R. 235/07), deve aver cura che essi contemplino misure specificatamente dedicate alla prevenzione del cyber-bullismo.

Anche le misure di intervento immediato, qualora si venga a conoscenza di episodi di cyber-bullismo, devono essere integrate e previste nel Regolamento di Istituto e nel Patto di Corresponsabilità, al fine di meglio regolamentare l'insieme dei provvedimenti sia di natura disciplinare che di natura educativa e di prevenzione.

Ovviamente la scuola si sta attrezzando in tale direzione. "I pericoli del web e il cyber-bullismo" è stato, infatti, uno dei progetti POF svolti durante l'anno scolastico 2016/2017, rivolto agli studenti del primo biennio del Liceo OSA che ha coinvolto i docenti di diritto ed economia e di lingua italiana.

La scuola, inoltre, ritiene fondamentale assicurare la massima informazione alle famiglie di tutte le attività e iniziative intraprese, anche attraverso una sezione dedicata sul sito web dell'istituto.

La presenza di un docente referente contribuisce a raccogliere e diffondere le buone pratiche educative, organizzative e le azioni di monitoraggio, favorendo così l'elaborazione di un modello di e-policy d'istituto.

Anche se gli esiti dell'indagine possono ritenersi positivi ed in linea con analoghe indagini statistiche su tale fenomeno condotte a livello regionale nel novembre 2017 (http://www.ilnodoblunellescuolepugliesi.it/2017/11/15/prova-statistica) , tuttavia il dato comune è che le vittime raramente denunciano o si confidano con qualcuno, per cui sorge spontaneo il dubbio sulla sincerità delle risposte date, nonostante fossero protette dall'anonimato. I dati della Polizia postale parlano chiaro, soprattutto raccontando la poca voglia di denunciare dei ragazzi che rendono i numeri ufficiali solo la punta di un iceberg.

Per tale motivazione la scuola, per contrastare comportamenti dannosi online e, allo stesso tempo, accrescere la conoscenza del fenomeno potrebbe con la collaborazione dei docenti potenziatori di diritto portare a conoscenza degli studenti strumenti quali "iGloss@ 1.110, un Abc dei comportamenti devianti online", elaborato

dal Dipartimento per la Giustizia Minorile e di Comunità. Il glossario, nella ricognizione dei termini specialistici sui comportamenti online a rischio, offre infatti una sintetica spiegazione delle principali caratteristiche delle condotte devianti e dei risvolti socio-giuridici, con l'obiettivo non solo di descrivere e inquadrare i nuovi fenomeni della devianza online, ma favorire l'acquisizione di consapevolezza sulle conseguenze sociali e giudiziarie di queste specifiche trasgressioni.

Occorre puntare sulla prevenzione, non lasciare soli gli studenti, ma aiutarli a muoversi nel mondo virtuale; è un atto di responsabilità collettiva che dovrebbe essere condivisa e supportata dai genitori, dagli operatori economici e dalle Istituzioni.

## Riferimenti bibliografici

[1] M. BERGAMINI, A. TRIFONE, G. BAROZZI, *Matematica.verde, con Tutor* seconda edizione 2017, Ed. Zanichelli Vol. 4
[2] M. BERGAMINI, A. TRIFONE, G. BAROZZI, *Matematica.verde, con Tutor* seconda edizione 2017, Ed. Zanichelli Vol. 5
[3] LEGGE 29 maggio 2017, n. *71 Disposizioni a tutela dei minori per la prevenzione ed il contrasto del fenomeno del cyber-bullismo.* (17G00085), GU Serie Generale n.127 del 03-06-2017
[4] MIUR *Linee di orientamento per la prevenzione e il contrasto del cyber-bullismo*

## Sitografia

[5] http://economia.osel.it/
[6] http://www.miur.gov.it/-/il-6-febbraio-e-il-safer-internet-day
[7] http://www.iglossa.org/
[8] http://www.ilnodoblunellescuolepugliesi.it/2017/11/15/prova-statistica

UNIVERSITÀ DEGLI STUDI DI BARI ALDO MORO
DIPARTIMENTO DI ECONOMIA E FINANZA
PDF finito di comporre
il 14 dicembre 2018

ISBN 978-88-6629-048-3
9 788866 290483