UNIVERSITA' DEGLI STUDI DI PERUGIA
DIPARTIMENTO DI FILOSOFIA SCIENZE SOCIALI UMANE E DELLA FORMAZIONE
Corso di Laurea in Scienze per l'Investigazione e la Sicurezza

3. INDICATORI STATISTICI

Prof. Maurizio Pertichetti

Statistica sociale

3. INDICATORI STATISTICI

Gli indicatori statistici sono strumenti con i quali si può descrivere in modo sintetico un fenomeno statistico. Un indicatore statistico è sostanzialmente un numero che si ottiene in modo opportuno operando mediante differenze e rapporti tra i dati statistici raccolti e moltiplicando eventualmente i rapporti per convenienti potenze di 10. Esso fornisce informazioni sul comportamento reciproco dei dati medesimi e, di conseguenza, sul fenomeno statistico che si vuole studiare.
In considerazione della loro natura, gli indicatori (o indici) si distinguono in:
- **indici assoluti** quando sono espressi nella stessa unità di misura del fenomeno;
- **indici relativi** quando non dipendono dall'unità di misura del fenomeno e si ottengono rapportando due misure assolute oppure un indice assoluto al suo massimo.

Infine, gli **indici normalizzati** sono indici relativi che assumono valori in un intervallo finito quasi sempre [0,1] oppure [-1,+1].

ESEMPIO. La tabella sottostante riassume l'andamento delle iscrizioni alla prima classe di un istituto di scuola media superiore in un periodo di 5 anni, distinguendo maschi e femmine.

Anno scolastico	**0**	**1**	**2**	**3**	**4**
Maschi	82	94	94	85	90
Femmine	65	72	69	76	76
Totale	**147**	**166**	**163**	**161**	**166**

Può essere interessante confrontare i numeri degli alunni iscritti in un dato anno scolastico rispetto a quelli dell'anno precedente o il numero dei maschi rispetto a quello delle femmine o altro ancora. Il confronto può avvenire operando per differenza o per rapporto tra i dati considerati. In ogni caso si ottengono degli indicatori statistici.

Indicatori ottenuti per differenza

Gli indicatori ottenuti per differenza sono di due tipi: assoluti e relativi.

- Un indicatore ottenuto per differenza si dice **assoluto** quando è ricavato eseguendo la differenza pura e semplice tra la dimensione o consistenza di un fenomeno ad una certa data o ad un certo periodo e la dimensione o consistenza dello stesso fenomeno ad un tempo antecedente a quello considerato. In formula:

  $$d = x_t - x_0$$

  Con riferimento all'esempio della tabella, un indicatore assoluto è quello che si ottiene effettuando la differenza tra il numero degli alunni iscritti nell'anno 1 e quello degli iscritti nell'anno 0:

  $$d = 166 - 147 = 19$$

  Questo indicatore ci dice banalmente che il numero degli iscritti nell'anno 1 è cresciuto di 19 unità rispetto all'anno 0.

- Un indicatore ottenuto per differenza è **relativo** quando è ricavato eseguendo il rapporto tra un indicatore assoluto e la dimensione o consistenza che il fenomeno osservato aveva al tempo antecedente. In formula:

  $$d = (x_t - x_0) / x_0$$

  Riprendendo, con riferimento alla stessa tabella, il numero degli alunni iscritti nell'anno 1 (=166) e quello degli iscritti nell'anno 0 (=147), si ha:

  $$d = (166 - 147) / 147 = 0{,}1293$$

  In genere questi indicatori vengono moltiplicati per 100 e il risultato viene a rappresentare una percentuale:

  $$d = (x_t - x_0) * 100 / x_0$$

  nell'esempio precedente:

  $$d = (166 - 147) * 100 / 147 = 12{,}93$$

Indicatori ottenuti per rapporto: I rapporti statistici

I rapporti statistici sono misure statistiche elementari finalizzate al confronto tra i dati stessi. In un rapporto statistico si mettono a confronto due termini, frequenze o quantità, di cui uno almeno è di natura statistica e tale che tra i due termini sussiste un qualche legame logico, cioè richiedono l'esistenza di una relazione o nesso logico tra le due quantità poste al numeratore e al denominatore. I rapporti così costruiti permettono di confrontare l'intensità di un fenomeno misurato su un collettivo, in tempi o luoghi diversi, e sono largamente impiegati nella descrizione di fenomeni di tipo socio-economico. Più precisamente sono utili per offrire una immediata ed efficace rappresentazione sintetica del fenomeno di interesse.

Assumono un valore sempre positivo e non dipendono dall'unità di misura consentendo così un confronto tra fenomeni diversi, ovvero vengono calcolati per eliminare l'influenza di circostanze che altrimenti non renderebbero confrontabili i dati stessi. Indicano il numero di unità della quantità posta al numeratore che corrispondono in media a una unità (o a 100, 1000, …, se il rapporto è moltiplicato per comodità di lettura per 100, 1000, …) della quantità posta al denominatore.
Se ne possono costruire diversi. Ne prendiamo in considerazione quattro, in base ad altrettanti tipi di rapporti:

- Rapporto di composizione
- Rapporto di coesistenza
- Rapporto di derivazione
- Rapporto di densità

i primi tre si ottengono dividendo grandezze omogenee e sono pertanto numeri puri. L'ultimo è invece il rapporto di due grandezze eterogenee ed è pertanto caratterizzato da un misura.

Rapporto di composizione

Il rapporto di composizione è una misura ottenuta rapportando il valore di una parte di un determinato collettivo o popolazione a quello dell'intero collettivo o popolazione. Ovvero quando l'ammontare complessivo di una quantità (frequenza o intensità) viene classificato in più modalità o classi, il rapporto tra la quantità (parziale) corrispondente ad una modalità e la quantità totale costituisce un rapporto di composizione. Il rapporto di composizione è detto anche rapporto di parte al tutto.

Rapporto di composizione = parte del fenomeno / totale fenomeno

Riprendendo la tabella e ponendo l'attenzione sull'anno 4, si può constatare che i maschi (=90) costituiscono la seguente frazione del totale degli iscritti (=166): **R=90/166=0,54216** .
È questo, così calcolato, un rapporto di composizione. Il rapporto di composizione può assumere valori compresi tra 0 e 1 (oppure tra 0 e 100, 0 e 1000, …, se il rapporto è moltiplicato per comodità di lettura rispettivamente per 100, 1000, …). Per cui tornando all'esempio: **R=90*100/166=54,22**
Sono esempi di rapporti di composizione:

- maschi residenti sul totale dei residenti di una comunità;
- residenti di una determinata fascia di età sul totale dei residenti;
- pezzi di scarto sul totale dei pezzi prodotti da un'azienda;
- operai di un'azienda sul totale dei dipendenti dell'azienda;
- fumatori sul totale dei componenti di una comunità.

Alcuni rapporti di composizione noti:

- Tasso di occupazione: su una popolazione di persone che hanno compiuto 15 anni, è la percentuale di coloro che hanno un lavoro;
- Tasso di disoccupazione: su una popolazione di persone che hanno compiuto 15 anni, è la percentuale di coloro che non hanno un lavoro;
- Tasso di scolarità: sugli alunni di una determinata fascia di età, è la percentuale di coloro che frequentano la scuola. Per esempio, in Italia, il tasso di scolarità della secondaria di 2° grado è la percentuale di coloro che frequentano quest'ordine di scuola rispetto al totale dei giovani della fascia di età 15-18 anni.

Le frequenze relative semplici (o percentuali) ottenute dividendo la frequenza di ciascuna classe o modalità in cui è classificato un carattere statistico per il numero totale di unità statistiche esaminate, sono esempi di rapporti di composizione: $\mathbf{f_i = n_i / n}$.

Anno scolastico	4	Rapp comp = f_i	Rapp comp = f_i * 100
Maschi	90	0,5422	54,22
Femmine	76	0,4578	45,78
Totale	**166**	**1,0000**	**100,00**

Nell'impiego per sintesi, confronti spazio-temporali e come base per indicatori, il rapporto di composizione evidenzia il contributo relativo di singole modalità o classi rispetto al totale.

Rapporto di coesistenza

Se l'ammontare complessivo di una quantità (frequenza o intensità) viene classificato in più modalità o classi, il rapporto di coesistenza è il rapporto tra la frequenza di una modalità rispetto alla frequenza corrispondente di un'altra modalità. Si può anche definire come un rapporto fra le intensità o frequenze riferite a fenomeni fra loro antitetici ma che coesistono. Evidenziano l'eventuale squilibrio fra le due grandezze coesistenti.

Rapporto di coesistenza = frequenza modalità A / frequenza modalità B

Questo rapporto, come si può capire facilmente, è un numero puro non negativo. Generalmente, i rapporti di coesistenza vengono moltiplicati per 100 o per 1000.
Prendendo ancora la tabella e considerando sempre l'anno 4, si può calcolare il rapporto $R_{M/F}$ fra il numero dei maschi e quello delle femmine o, viceversa, il rapporto $R_{F/M}$ tra il numero delle femmine e quello dei maschi:

Anno scolastico	4
Maschi	90
Femmine	76
Totale	**166**

Rapp coesist	$I_{M/F}$
M/F	**M*100/F**
1,18	**118,42**

Rapp coesist	$I_{F/M}$
F/M	**F*100/M**
0,84	**84,44**

I due rapporti maschi/femmine e femmine/maschi, moltiplicati per 100, si chiamano più propriamente **indice di mascolinità** e **indice di femminilità**.
L'**indice di mascolinità** ci dice che, nella comunità presa in esame, ci sono **118,4** maschi ogni 100 femmine, mentre l'**indice di femminilità** ci dice che ci sono **84,4** femmine ogni 100 maschi.
Può capitare che il confronto riguardi i pezzi di scarto (perché difettosi) della produzione di certo materiale rispetto ai pezzi non difettosi (perché buoni). Il rapporto $R_{D/B}$ fra il numero dei pezzi difettosi (pezzi D) e quello dei pezzi buoni (pezzi B) è un rapporto di coesistenza che, moltiplicato per 100, va sotto il nome di indicatore di qualità. Ad esempio, considerata un'azienda che produce lampadine, se su 100 lampadine prodotte ve ne sono 5 difettose, l'indice di qualità della produzione è il numero I_q tale che: $\mathbf{I_q=5*100/95=5,26}$. Vale a dire che vi sono 5,26 lampadine difettose ogni 100 buone.
Un altro rapporto di coesistenza è l'indice di vecchiaia, il quale indica il grado di invecchiamento di una popolazione. Lo si calcola moltiplicando per 100 il rapporto fra il numero delle persone "anziane" e quello dei "giovani" di una certa comunità.

RESIDENTI PER GRANDI CLASSI DI ETA'					I_v indice di vecchiaia
anno	**0-14**	**15-64**	**65 e oltre**	**totale**	**65 ed oltre * 100 / 0-14**
1990	13.899	74.922	21.009	109.830	**151,2**
1995	12.072	72.904	23.255	108.231	**192,6**
2000	11.955	71.203	24.581	107.739	**205,6**
2005	12.737	70.352	26.480	109.569	**207,9**
2010	13.702	71.938	27.684	113.324	**202,0**
2013	13.828	69.865	28.535	112.228	**206,4**

Questo indice mostra che per l'anno 2013 ad esempio ci sono 206,4 anziani ogni 100 giovani.

Rapporto di derivazione
Un rapporto di derivazione è ottenuto dividendo la frequenza o intensità di un fenomeno per la frequenza o intensità di un altro fenomeno che, sul piano logico o temporale, ne costituisce l'antecedente, la causa o il presupposto.
Rapporto di derivazione = modalità susseguente / modalità antecedente
Anche in questo caso assume valori >0 (oppure >100, >1000, …, se il rapporto è moltiplicato per comodità di lettura rispettivamente per 100, 1000, …).
Alcuni esempi di rapporti di derivazione:

- In un determinato periodo di tempo si sono verificati in una data regione **N** incidenti stradali, a causa dei quali sono morte **n** persone. Il numero **n/N** è un rapporto di derivazione. Solitamente moltiplicato per 100, è chiamato più propriamente **indice di pericolosità stradale** di quella regione.
- Se si prende in esame una popolazione e si considera il rapporto fra il numero **n** dei nati vivi in un dato periodo e la numerosità **N** della popolazione in quel periodo, si ottiene un rapporto di derivazione. Se si moltiplica tale rapporto per 1000, si ottiene il cosiddetto **quoziente di natalità**: $\mathbf{Q_N=n*1000/N}$.

- Un altro rapporto di derivazione è il tasso di fecondità di una data popolazione in un determinato periodo (di solito, un anno). È il valore relativo di nati in quel periodo rispetto al numero delle donne in età feconda (fra i 15 e i 49 anni, secondo il parere di alcuni demografi, 15-44 anni secondo quello di altri).

- Analogamente il **tasso di mortalità** $\mathbf{T_M}$ è dato da: $\mathbf{T_M=m*1000/N}$ dove **m** indica il numero dei morti della popolazione N nel periodo.

Movimento naturale

Ripartizione geografica	nati vivi	morti	popolazione media	quoziente di natalità	tasso di mortalità
nord	249.677	259.406	26.211.512	9,53	9,90
centro	104.740	112.846	11.124.059	9,42	10,14
mezzogiorno	208.182	174.406	20.663.632	10,07	8,44
Totale	562.599	546.658	58.175.310	9,67	9,40

$$\text{tasso di natalità} = \frac{\text{nati vivi}}{\text{popolazione media}} \times 1.000 \qquad \text{morti ogni 1.000 abitanti}$$

$$\text{tasso di mortalità} = \frac{\text{morti}}{\text{popolazione media}} \times 1.000 \qquad \text{morti ogni 1.000 abitanti}$$

popolazione media = (popol inizio anno + popolaz fine anno) / 2

Va detto che la "popolazione" di cui si parla va intesa in senso statistico, potendo essere un insieme di persone, di aziende , di beni in senso lato, eccetera.

Rapporto di densità
Un rapporto di densità è definito mediante il confronto tra la dimensione globale di una popolazione o di un fenomeno e la dimensione spaziale o temporale cui il fenomeno stesso fa riferimento.
Rapporto di densità = dimensione globale / dimensione temporale o spaziale
Il rapporto di densità può assumere valori >0 (oppure >100, >1000, …, se il rapporto è moltiplicato per comodità di lettura rispettivamente per 100, 1000, …).
Le quantità al numeratore e al denominatore sono grandezze eterogenee. L'inverso del rapporto di densità può dar luogo ad un **rapporto di estensione**.
Alcuni esempi di rapporti di densità:

- Il rapporto fra il numero degli abitanti di una data regione (città, provincia, stato, eccetera) e la misura della sua superficie è un rapporto di densità. Se la superficie è misurata, in chilometri quadrati, tale rapporto darà il numero medio di abitanti per chilometro quadrato.
- Il rapporto fra il numero degli alunni che frequentano una data scuola ed il numero delle classi di quella scuola è un altro rapporto di densità: dà il numero medio di alunni per classe.

Alcuni rapporti di densità noti:

- Prodotto interno lordo pro capite: rapporto tra il prodotto interno lordo e la popolazione;
- Numero medio componenti per famiglia: rapporto tra popolazione e numero di famiglie residenti nello stesso territorio;
- Numero medio di alunni per insegnante: rapporto tra il numero di alunni e di insegnati per livello di istruzione (medie, superiori, etc.);
- Indice di dotazione di posti letto negli istituti di cura: rapporto tra il numero di posti letto degli istituti di cura e la popolazione;
- Indice di diffusione TV: rapporto tra il numero di abbonamenti TV e la popolazione.

Tassi

Nel caso della popolazioni, le statistiche ad essa riferite possono essere distinte in:

- statistiche di stato o dati di consistenza, se riferite ad un dato istante, come ad esempio nel caso dell'ammontare della popolazione residente per sesso, età, ecc, alla data di un censimento, o al 31.12 di un dato anno;
- statistiche di flusso, se riferite alla dinamica o alle variazioni della popolazione intervenute nel corso di un periodo di tempo (es. nascite, morti, migrazioni nell'anno 2015).

I dati relativi ai flussi dipendono, a parità di altre condizioni, dall'ammontare della popolazione. I confronti nel tempo e nello spazio dei fenomeni relativi a tali flussi, quali ad esempio la natalità, la mortalità, ecc, vengono effettuati mediante particolari rapporti statistici denominati **tassi**, dove il dato di flusso viene rapportato all'**ammontare medio della popolazione**.
I tassi consentono di studiare il manifestarsi di un fenomeno di interesse nel tempo e di procedere a confronti temporali e spaziali anche con livelli di approfondimento diversi (a riguardo si parla di tassi generici o grezzi, tassi specifici e tassi standardizzati).

Numeri indice

I numeri indice, di grande importanza soprattutto in campo economico (i più noti sono quelli utilizzati per determinare l'aumento del costo della vita "inflazione"), sono particolari rapporti statistici che misurano la variazione dell'intensità di un fenomeno in situazioni temporali o spaziali diverse. Assumono un valore sempre positivo, si configurano come "puri numeri", in quanto non dipendono dall'unità di misura nella quale è espresso il fenomeno originario, e servono per facilitare la comprensione delle variazioni intervenute nel tempo o nello spazio nei dati statistici.
In generale si distingue tra:

- **numeri indici semplici**: qualora l'obiettivo è il confronto tra due singole "situazioni" differenti;
- **numeri indici complessi**: se invece l'obiettivo è quello di descrivere in modo sintetico la variazione, ad esempio, di un gruppo di n beni e/o di n servizi simultaneamente, elencati in una serie storica multipla, in due "situazioni" differenti.

Data una serie storica, i numeri indice si calcolano rapportando due valori della serie storica del fenomeno riferiti a due tempi, luoghi o condizioni differenti. La grandezza posta a denominatore, che viene assunta come termine di confronto, prende il nome di **base** e può essere:

- **fissa**, quando si esegue il rapporto fra due dati statistici, uno dei quali, quello posto al denominatore è sempre lo stesso e si assume quindi come base, termine di confronto, di riferimento, mentre quello posto al numeratore varia nel tempo;
- **mobile**, quando si vogliono mettere in risalto le variazioni del fenomeno fra istanti temporali consecutivi e pertanto si sceglie come base, ad esempio, il periodo che precede immediatamente quello relativo al valore posto al numeratore.

La scelta della base dipende dalla tipologia di confronto che si intende effettuare e conseguentemente dalle variazioni che si vogliono evidenziare:

- nel caso di base fissa sarà possibile confrontare tra loro tutte le diverse situazioni presentate;
- nel caso di base mobile, invece, potrà essere rilevata solamente la variazione relativa tra la situazione di un dato anno e l'anno immediatamente precedente.

L'indice relativo alla situazione base è generalmente posto uguale a 100 (e per questo tutti i rapporti sono moltiplicati per 100). In tal modo si usano meno cifre decimali e basta sottrarre 100 per ottenere la variazione percentuale. A rigore, tuttavia, il numero indice è il semplice rapporto tra le due grandezze considerate.

In questo corso limiteremo la nostra attenzione ai numeri indici semplici.

Un numero indice semplice l'abbiamo definito come il rapporto tra due numeri riferiti alle intensità di un fenomeno in tempi o luoghi diversi. Per costruirli è necessario considerare una serie temporale o territoriale di un dato fenomeno X, di termine generale x_t , con t=0,1,2,….,n. E quindi:

$$x_0 \quad x_1 \quad x_2 \quad \ldots\ldots \quad x_n$$

- i numeri indice saranno a **base fissa** se ciascuna intensità della serie viene rapportata ad una unica intensità che resta costante. Supponendo come base fissa t=0, i numeri indice sono del tipo:

$$_0I_0 = \frac{x_0}{x_0} \qquad _0I_1 = \frac{x_1}{x_0} \qquad _0I_2 = \frac{x_2}{x_0} \quad \ldots\ldots\ldots\ldots \quad _0I_n = \frac{x_n}{x_0}$$

- i numeri indice saranno a **base mobile** (o concatenati) se ciascuna intensità viene rapportata a quella del termine precedente. Per cui:

$$_{0\text{-}1}I_0 = \frac{x_0}{x_{0-1}} \qquad _0I_1 = \frac{x_1}{x_0} \qquad _1I_2 = \frac{x_2}{x_1} \quad \ldots\ldots\ldots\ldots \quad _{n-1}I_n = \frac{x_n}{x_{n-1}}$$

Questo rapporto non si calcola perché manca il termine x_{0-1}

Esempio di calcolo di numero indice temporale

anno	2010	2011	2012	2013	2014	2015
retribuzione annua	17.166	17.853	18.818	19.552	19.884	20.242

indice a base fissa (2010 = 100,00)	17.166/ 17.166	17.853/ 17.166	18.818/ 17.166	19.552/ 17.166	19.884/ 17.166	20.242/ 17.166
	1,000	1,040	1,096	1,139	1,158	1,179
	17.166x100/ 17.166	17.853x100/ 17.166	18.818x100/ 17.166	19.552x100/ 17.166	19.884x100/ 17.166	20.242x100/ 17.166
	100,000	**104,002**	**109,624**	**113,900**	**115,834**	**117,919**

indice a base mobile		17.853/ 17.166	18.818/ 17.853	19.552/ 18.818	19.884/ 19.552	20.242/ 19.884
		1,040	1,054	1,039	1,017	1,018
		17.853x100/ 17.166	18.818x100/ 17.853	19.552x100/ 18.818	19.884x100/ 19.552	20.242x100/ 19.884
		104,002	**105,405**	**103,901**	**101,698**	**101,800**

Nel particolare la costruzione di un numero indice si ottiene ricorrendo alle proporzioni.

Data la serie di valori:

anno	2010	2011	2012	2013	2014	2015
retribuzione annua	17.166	17.853	18.818	19.552	19.884	20.242

dei quali il primo, quello corrispondente all'anno 2010 viene identificato = a 100,000, i numeri indice si ottengono:

per il 2011:	17.166	:	100,000	=	17.853	:	X	;	X	=	17.853	*	100,000	/	17.166	=	**104,002**
per il 2012:	17.166	:	100,000	=	18.818	:	X	;	X	=	18.818	*	100,000	/	17.166	=	**109,624**
per il 2013:	17.166	:	100,000	=	19.552	:	X	;	X	=	19.552	*	100,000	/	17.166	=	**113,900**
per il 2014:	17.166	:	100,000	=	19.884	:	X	;	X	=	19.884	*	100,000	/	17.166	=	**115,834**
per il 2015:	17.166	:	100,000	=	20.242	:	X	;	X	=	20.242	*	100,000	/	17.166	=	**117,919**

e se a ciascuno dei numeri indice togliamo 100,000 otteniamo esattamente la variazione percentuale:

per il 2011:	**4,002**	%	ovvero	17.853	-	17.166	*	100,000	/	17.166	=	**4,002**
per il 2012:	**9,624**	%	ovvero	18.818	-	17.166	*	100,000	/	17.166	=	**9,624**
per il 2013:	**13,900**	%	ovvero	19.552	-	17.166	*	100,000	/	17.166	=	**13,900**
per il 2014:	**15,834**	%	ovvero	19.884	-	17.166	*	100,000	/	17.166	=	**15,834**
per il 2015:	**17,919**	%	ovvero	20.242	-	17.166	*	100,000	/	17.166	=	**17,919**

analoga considerazione vale nel caso dei numeri indice temporali a base mobile.

Esempio di calcolo di numero indice spaziale

riparzione geografica	nord	centro	sud e isole	italia
spesa media mensile anno 2000	4.726	4.162	3.507	4.217

indice a base fissa (nord = 100,00)	4.726/ 4.726	4.162/ 4.726	3.507/ 4.726	4.217/ 4.726
	1,000	0,881	0,742	0,892
	4.726x100/ 4.726	4.162x100/ 4.726	3.507x100/ 4.726	4.217x100/ 4.726
	100,000	**88,066**	**74,207**	**89,230**

Facendo riferimento al Sud e Isole, il risultato si legge così: in media nell'anno 2000 una famiglia dell'italia meridionale ha speso mensilmente 74,2 lire contro le 100,0 spese da una famiglia del nord, ovvero ha speso il 25,8 % in meno.

Come osservazione conclusiva va detto che un indicatore perché sia valido deve essere:

- misurabile: che sia ottenuto per differenza o per rapporto, ogni indicatore è un numero (puro o corredato di dimensione) che fornisce una data misura;
- comprensibile: deve essere chiara la caratteristica del fenomeno che l’indicatore vuole mettere in risalto;
- efficace: la conoscenza dell’indicatore permette di ottenere frequenti informazioni sul fenomeno studiato;
- efficiente: le informazioni ottenute mediante l’indicatore non sono dispendiose né in termini di tempo né in termini di denaro.

Un buon indicatore, infine, permette sia di tenere sotto controllo il fenomeno che si studia sia di migliorare la situazione se questo si rende necessario.