

**CORSO DI STUDIO: SCIENZE STATISTICHE (L41)**
**ANNO ACCADEMICO: 2023 - 2024**
**DENOMINAZIONE DELL'INSEGNAMENTO:**
**Inferenza Statistica e Modelli Lineari (Statistical Inference and Linear Models)**

Principali informazioni sull'insegnamento	
Anno di corso	Secondo anno
Periodo di erogazione	Secondo semestre (19 febbraio – 7 giugno 2024)
Crediti formativi universitari (CFU/ETCS):	10
SSD	SECS-S/01 (Statistica)
Lingua di erogazione	italiano
Modalità di frequenza	La frequenza non è obbligatoria ma fortemente raccomandata

Docente	
Nome e cognome	Nunziata Ribecco
Indirizzo mail	nunziata.ribecco@uniba.it
Telefono	080.5049251
Sede	Largo Abbazia Santa Scolastica, 53 (stanza n. 17)
Sede virtuale	Piattaforma Microsoft Teams: Ricevimento Studenti (codice teams: <b>rfl3ft0</b> ) Tesisti inferenza (codice teams: <b>xiz8pwr</b> )
Ricevimento	martedì: 9,30 - 11,30; mercoledì: 9,30 - 11,30. Durante il periodo delle lezioni l'orario di ricevimento si modifica come segue: martedì: 9,00 - 11,00 (primo e secondo semestre); mercoledì: 9,00 - 11,00 (primo semestre); mercoledì: 11,30 - 13,30 (secondo semestre).

Organizzazione della didattica			
Ore			
Totali	Didattica in presenza	Pratica (laboratorio, campo, esercitazione, altro)	Studio individuale
250	49	21	180
CFU/ETCS			
10	7	3	

Obiettivi formativi
<p>Il corso si propone di preparare professionisti in grado di analizzare i fenomeni reali mediante gli strumenti di inferenza statistica classica che vengono presentati sia sotto l'aspetto teorico che applicativo. Vuole, inoltre, porre le basi per l'impostazione bayesiana e decisionale, per le analisi multivariate e le tecniche non parametriche. Gli argomenti in programma sono quelli dell'Inferenza statistica classica e vengono trattati con la finalità di approfondire sia la parte statistico-matematica che quella applicativa, sviluppando l'abilità a identificare un problema e a risolverlo con un approccio di tipo professionale. Si affrontano i problemi inferenziali, inizialmente, sotto l'aspetto generale dei metodi (stima dei parametri e verifica delle ipotesi) che vedono, prevalentemente, la loro applicazione ad una o più popolazioni normali con uno sguardo, anche, alle popolazioni dicotomiche, il tutto viene presentato facendo riferimento sia a piccoli che a grandi campioni. Inoltre, vengono analizzati alcuni problemi di statistica non parametrica, alcune tecniche di ricampionamento e di teoria delle decisioni. La finalità con cui si presentano questi temi è quella di rendere lo studente padrone dell'aspetto statistico matematico e capace di essere autonomo nell'analisi inferenziale di dati campionari riguardanti fenomeni che si realizzano in svariati contesti applicativi (sperimentale, economico, sociale, ecc. ...).</p>

<b>Prerequisiti</b>	Lo studente deve aver sostenuto l'esame di Istituzioni di Statistica, inoltre deve avere conoscenze di Istituzioni di analisi matematica, Analisi matematica ed elementi di algebra lineare e di Calcolo delle probabilità
<b>Metodi didattici</b>	La didattica in presenza riguarda argomenti teorici ed esercitazioni durante le quali vengono esposti dei problemi atti a sviluppare la capacità dello studente ad applicare la teoria nel contesto dei fenomeni reali. Vengono svolte, inoltre, attività di laboratorio con il software R, per un numero di ore pari circa ad un terzo dell'intera durata del corso, con applicazioni dei metodi di inferenza statistica a fenomeni reali.
<b>Risultati di apprendimento previsti</b>  <b>DD1 Conoscenza e capacità di comprensione</b>  <b>D2 Conoscenza e capacità di comprensione applicate</b>  <b>DD3-5 Competenze trasversali</b>	<p>I risultati di apprendimento attesi sono racchiusi in questi punti:</p> <ul style="list-style-type: none"> <li>- conoscere i metodi per stimare i parametri ed essere in grado di applicarli (prevalentemente stimatori di massima verosimiglianza) sia per i parametri di popolazioni normali che per quelli di popolazioni dicotomiche;</li> <li>- conoscere ed essere in grado di verificare le proprietà di uno stimatore ed affrontare il problema della determinazione della sua precisione (stima intervallare);</li> <li>- essere in grado di affrontare il problema della verifica delle ipotesi sia parametriche che non parametriche con riferimento ad una o più popolazioni;</li> <li>- saper interpretare i risultati ottenuti ed essere capaci di prendere delle decisioni ai fini operativi;</li> <li>- avere dimestichezza con i metodi inferenziali ed acquisire la padronanza degli strumenti al fine di poter operare in totale autonomia allorché vengono affrontati dei problemi concreti;</li> <li>- acquisire la logica inferenziale sia in termini metodologici che applicativi.</li> </ul> <p>Al termine dell'insegnamento lo studente deve conoscere la teoria dell'inferenza classica e le sue applicazioni:</p> <ul style="list-style-type: none"> <li>- Teoria della stima: Determinazione dei principali stimatori mediante il metodo della massima verosimiglianza. Conoscere le proprietà di questi stimatori ed essere capaci di verificare le proprietà di un generico stimatore lineare. Sapere cosa si intende per precisione di una stima ed essere capace di determinarla.</li> <li>- Teoria della verifica d'ipotesi parametrica: Conoscere cosa è un'ipotesi statistica parametrica e tutti i problemi ad essa collegati. Saper applicare gli strumenti di verifica d'ipotesi per una o più popolazioni normali e dicotomiche.</li> <li>- Modello lineare: Conoscere cosa si intende per relazione fra due variabili e come affrontare i problemi inferenziali ad essa connessi sia nel caso di variabile antecedente non quantitativa (confronto fra più medie di popolazioni normali) che nel caso di variabile antecedente quantitativa (regressione lineare).</li> <li>- Inferenza non parametrica: Conoscere cosa si intende per inferenza non parametrica, sapere quando farne ricorso e con quali strumenti.</li> </ul> <p>Lo studente deve essere in grado, al termine del corso di insegnamento, di applicare i metodi dell'inferenza statistica presentati a fenomeni concreti e, in particolare, saper interpretare i risultati ottenuti dall'applicazione di detti metodi. L'insegnamento è una delle tre metodologie previste dal percorso formativo. Pertanto, al termine del corso lo studente ha acquisito gli strumenti necessari ad affrontare l'analisi di dati campionari in tutti i contesti applicativi ed è, anche, in grado di interpretare i risultati e giustificare le decisioni a cui dovranno condurre i risultati ottenuti.</p> <p>Gli strumenti metodologici di analisi inferenziale vengono presentati costantemente con riferimento a fenomeni reali. Pertanto data la trasversalità della disciplina, attraverso esercitazioni ed attività di laboratorio, lo studente analizza fenomeni quali quelli sociali, economici, biomedici, ecc., mediante</p>

	<p>l'utilizzo degli strumenti inferenziali e commenta i risultati dell'analisi. Pertanto, lo studente non solo acquisisce le competenze metodologiche ma è in grado, anche, di acquisire l'elasticità mentale per applicare gli strumenti in maniera trasversale. Le prove di laboratorio e le esercitazioni vengono svolte sistematicamente in aula e vedono coinvolti gli studenti.</p>
<p><b>Contenuti di insegnamento (Programma)</b></p>	<ul style="list-style-type: none"> <li>• Introduzione al corso: scopi e logica dell'inferenza statistica.</li> <li>• Alcuni elementi di calcolo delle probabilità; variabili casuali; principali variabili casuali discrete e continue; alcuni teoremi limite.</li> <li>• Introduzione all'inferenza statistica: concetti fondamentali; popolazione e campione.</li> <li>• Tecniche di inferenza statistica: stima dei parametri, stimatori e proprietà, alcuni metodi di stima, intervalli di confidenza; verifica di ipotesi parametriche e non parametriche, lemma di Neyman-Pearson, rapporto di verosimiglianza.</li> <li>• Dimensione campionaria.</li> <li>• Problemi di inferenza sui parametri di una popolazione normale: inferenza sulla media e sulla varianza.</li> <li>• Problemi di inferenza sui parametri di due e più popolazioni normali.</li> <li>• Confronto fra le medie di due popolazioni: campioni indipendenti e campioni dipendenti (dati appaiati).</li> <li>• Confronto fra le varianze di due popolazioni.</li> <li>• Confronto fra le medie di più popolazioni (ANOVA).</li> <li>• Il modello lineare: regressione lineare.</li> <li>• Altri modelli lineari e non.</li> <li>• L'ANOVA in termini di modello lineare.</li> <li>• Inferenza sul coefficiente di correlazione.</li> <li>• Problemi di inferenza sul parametro di una popolazione dicotomica.</li> <li>• Problemi di inferenza sui parametri di due popolazioni dicotomiche.</li> <li>• Problemi di inferenza non parametrica.</li> <li>- <b>Approfondimenti</b></li> <li>• Robustezza e stimatori robusti.</li> <li>• Metodi di ricampionamento: Jackknife e Bootstrap.</li> <li>• Concetti introduttivi di teoria delle decisioni.</li> </ul>
<p><b>Testi di riferimento</b></p>	<p><u>Per la parte teorica</u>          PICCOLO D., Statistica per le decisioni, Terza edizione, Il Mulino, 2020 (capp. 8 – 19)          NEWBOLD P., CARLSON W. L., THORNE B., Statistica, seconda edizione, 2015          AGRETI A., FINLAY B., Statistical Methods for the Social Sciences, Prentice Hall, New Jersey 2007          CHIEPPA M., RIBECCO N., VITALE C., Teoria e metodi statistici, Edizioni scientifiche italiane, Napoli 1994          G. CICCHITELLI, Statistica: Principi e Metodi- seconda edizione, Pearson, 2012 (capp.12 - 21 e Appendici B e C).          FRED N., JONES S., BERGQUIST T., BONNINI S., Statistica per le scienze economiche e aziendali, Isedi, Torino, 2019 (capitoli 4 - 9)          FREUND R.J., WILSON W.J., Metodi statistici, Piccin, Padova 2000          PIERACCINI L., Fondamenti di inferenza statistica, G. Giappichelli, Torino 2007</p> <p><u>Per la parte esercitativa relativa alle lezioni frontali</u>          MURRAY R. SPIEGEL, Statistica, Collana Schaum, McGraw Hill Italia, Milano 2003          BERNSTEIN S., BERNSTEIN R., Statistica Inferenziale, Collana Schaum, McGraw Hill Italia, Milano 2003.</p> <p><u>Per la parte esercitativa relativa al laboratorio</u>          IACUS S.M., MASAROTTO G., Laboratorio di Statistica con R, MacGraw-Hill Informatica, 2021          COCCARDA R., FRASCATI F., Manuale interattivo di statistica con R con MyLab, Pearson 2015</p>

	RACUGNO W., VENTURA L., 2017, Biostatistica, casi di studio in R, EGEA, Milano.
<b>Note ai testi di riferimento</b>	Lo studente è libero di scegliere uno qualunque dei testi indicati sia per la parte teorica che per quella pratica
<b>Materiali didattici</b>	Il materiale didattico relativo agli approfondimenti è disponibile sulla piattaforma e-learning ( <a href="http://dief.osel.it">dief.osel.it</a> )

<b>Valutazione</b>	
Modalità di verifica dell'apprendimento	<ul style="list-style-type: none"> <li>• Durante il corso vengono svolte delle prove che non hanno alcuna validità ai fini dell'esame di profitto ma hanno lo scopo di valutare e verificare le competenze acquisite dagli studenti sugli argomenti trattati sino al momento della prova. Tali prove possono prevedere lavori di gruppo su argomenti di statistica-matematica, a cui segue la presentazione in power point da parte degli studenti, e prove individuali in cui vengono posti dei problemi che richiedono l'applicazione di tecniche inferenziali idonee al contesto.</li> <li>• L'esame di profitto prevede una prova scritta, a fine corso, in cui si richiede allo studente di risolvere, tramite metodi statistici idonei, problemi riguardanti casi reali. A tale prova segue una discussione orale che può essere sostenuta in una delle date previste dal calendario di esame.</li> <li>• L'eventuale esito negativo della prova sostenuta a fine corso non pregiudica l'ammissione all'esame di profitto, che verrà svolto secondo le seguenti modalità: una prova scritta in cui sono contenuti quesiti da risolvere tramite l'applicazione dei metodi inferenziali ed una prova orale.</li> <li>• La prova orale prevede la discussione dei risultati ottenuti e la verifica delle conoscenze su ulteriori argomenti che non sono oggetto della prova scritta.</li> </ul>
Criteri di valutazione	Alla valutazione finale contribuiscono sia la prova scritta che quella orale. In particolare, quello che contribuisce alla valutazione finale, oltre la corretta applicazione dei metodi ai problemi concreti che vengono sottoposti durante la prova scritta, l'elasticità mentale e la capacità logica dello studente nell'interpretare i risultati e, soprattutto, nella presentazione dei metodi.
Criteri di misurazione dell'apprendimento e di attribuzione del voto finale	La valutazione è in trentesimi e a determinare il voto finale contribuiscono in ugual misura la valutazione della prova scritta e quella della prova orale. L'esame è superato allorché lo studente ha conseguito un minimo di 18/30.
<b>Altro</b>	<p><b>Laboratorio Statistico (3 CFU)</b></p> <p>E' possibile acquisire i 3 CFU previsti per il Laboratorio Statistico nell'ambito del corso di Inferenza Statistica e Modelli Lineari, previa richiesta al Coordinatore del Consiglio Interclasse di Scienze Statistiche. A tal fine gli studenti che lo richiedono dovranno svolgere ulteriori 21 ore (pari a 7 ore per CFU) di lezioni frontali di Laboratorio in R. I materiali del laboratorio sono pubblicati sulla piattaforma e-learning (<a href="http://dief.osel.it">dief.osel.it</a>) a cui è possibile accedere dopo la registrazione.</p> <p>Per acquisire l'idoneità è necessario:</p> <ol style="list-style-type: none"> <li>1. Frequentare il laboratorio.</li> </ol> <p>Coloro che avranno accumulato un numero di assenze superiore a tre non potranno partecipare alla valutazione finale.</p> <ol style="list-style-type: none"> <li>2. Aver conseguito un voto medio (domande a risposta multipla e elaborato in R) non inferiore a 18/30.</li> <li>3. Superare la verifica finale di R (voto minimo 18/30) a cui sarà possibile partecipare solo se sono soddisfatti i requisiti previsti al punto 1. e 2.</li> </ol>