



UNIVERSITÀ DEGLI STUDI DI PALERMO

DIPARTIMENTO	Scienze Economiche, Aziendali e Statistiche		
ANNO ACCADEMICO OFFERTA	2020/2021		
ANNO ACCADEMICO EROGAZIONE	2021/2022		
CORSO DILAUREA MAGISTRALE	STATISTICA E DATA SCIENCE		
INSEGNAMENTO	ADVANCED STATISTICAL METHODS C.I.		
CODICE INSEGNAMENTO	21226		
MODULI	Si		
NUMERO DI MODULI	2		
SETTORI SCIENTIFICO-DISCIPLINARI	SECS-S/01		
DOCENTE RESPONSABILE	CHIODI MARCELLO	Professore Ordinario	Univ. di PALERMO
ALTRI DOCENTI	ABBRUZZO ANTONINO	Professore Associato	Univ. di PALERMO
	CHIODI MARCELLO	Professore Ordinario	Univ. di PALERMO
CFU	9		
PROPEDEUTICITA'			
MUTUAZIONI			
ANNO DI CORSO	2		
PERIODO DELLE LEZIONI	1° semestre		
MODALITA' DI FREQUENZA	Facoltativa		
TIPO DI VALUTAZIONE	Voto in trentesimi		
ORARIO DI RICEVIMENTO DEGLI STUDENTI	ABBRUZZO ANTONINO Lunedì 15:00 17:00 DSEAS secondo piano stanza 222 CHIODI MARCELLO Martedì 15:00 17:00 stanza del docente (edificio 13); eccezionalmente su teams Venerdì 12:00 13:00 stanza del docente (edificio 13); eccezionalmente su teams		

<p>PREREQUISITI</p>	<p>Elementi di calcolo delle probabilità e di inferenza statistica. Fondamenti della teoria delle ipotesi statistiche. Modelli lineari generalizzati.</p>
<p>RISULTATI DI APPRENDIMENTO ATTESI</p>	<p>Conoscenza e capacità di comprensione. Conoscenza delle metodologie della statistica bayesiana e non parametrica. Acquisizione del linguaggio e della terminologia propri della disciplina. Capacità di comprensione delle derivazioni, delle proprietà teoriche e dei nessi interni dei metodi presentati.</p> <p>Capacità di applicare conoscenza e comprensione Capacità di affrontare problemi concreti con i metodi acquisiti durante le lezioni frontali. Capacità di utilizzare l'ambiente statistico R per applicare i metodi acquisiti durante le lezioni frontali e per verificare per via simulativi i risultati teorici.</p> <p>Autonomia di giudizio Essere in grado di comprendere criticamente caratteristiche, potenzialità e limiti dei metodi bayesiani e non parametrici. Essere in grado di inquadrare uno specifico problema in termini bayesiani e non parametrici.</p> <p>Abilità comunicative Essere in grado di discutere le caratteristiche di un dato problema. Essere in grado di usare la terminologia statistica e la formalizzazione dei problemi in un'esposizione scritta.</p> <p>Capacità di apprendimento Essere in grado di consultare la letteratura scientifica sull'argomento; capacità di apprendere le estensioni dei modelli studiati a lezione; capacità di apprendimento di software statistico specialistico anche diverso da quello impiegato in aula.</p>
<p>VALUTAZIONE DELL'APPRENDIMENTO</p>	<p>La verifica dell'apprendimento avviene attraverso un colloquio orale, subordinato al superamento della prova scritta.</p> <p>Prova scritta La prova scritta mira a rilevare le conoscenze e le abilità possedute dallo studente e la sua capacità di riportarle in un elaborato scritto che tenga conto anche della proprietà di linguaggio statistico posseduto. La prova scritta riguarderà la verifica dei concetti di statistica bayesiana, non-parametrica e semiparametrica, acquisiti durante il corso e sarà svolta anche con l'ausilio del software statistico R, attraverso lo studio di un dataset reale. La prova avrà una durata massima di 3 ore. La soglia di sufficienza (pari a un punteggio di 18 su una scala 18-30), propedeutica al superamento della prova, è raggiunta se lo studente mostra un utilizzo adeguato dei termini relativi ai concetti in oggetto d'esame.</p> <p>Prova orale La prova orale mira ad approfondire l'elaborato scritto e alla valutazione dell'apprendimento dello studente. Questa consisterà in almeno due domande finalizzate a graduare meglio la valutazione delle conoscenze, competenze e abilità possedute dallo studente, e la sua capacità di trasmetterle con idoneo linguaggio statistico. La soglia della sufficienza, della prova orale, sarà raggiunta quando lo studente abbia mostrato conoscenza e comprensione degli argomenti almeno nelle linee generali (definizione dei concetti). Quanto più, invece, l'esaminando abbia brillantemente superato la prova scritta e abbia dato evidenza, nella prova orale, delle sue capacità argomentative ed espositive, nonché di proprietà di linguaggio statistico e di collegamento con le altre discipline, tanto più la valutazione sarà positiva.</p> <p>METODO DI VALUTAZIONE FINALE La valutazione finale dell'esame prenderà in considerazione due aspetti: i) la padronanza degli argomenti; ii) la proprietà di linguaggio, valutati nel complesso della prova scritta e orale. Il docente avrà anche la possibilità di tenere conto dei fattori di contesto dell'esame (come ad esempio la partecipazione attiva durante le lezioni e le esercitazioni, oppure la presenza di qualche disabilità) ai fini della determinazione dell'esito della prova.</p> <p>Entrambe le prove (scritta e orale) sono valutate in trentesimi e si ritengono superate con una votazione minima di 18/30. Il voto finale è dato dalla media aritmetica semplice delle due prove.</p>
<p>ORGANIZZAZIONE DELLA DIDATTICA</p>	<p>Il corso sarà suddiviso in lezioni frontali ed esercitazioni. Tutti gli argomenti teorici sviluppati nelle lezioni frontali verranno affrontati in termini applicativi mediante attività di laboratorio informatico-statistico con l'utilizzo dell'ambiente di programmazione R.</p>

**MODULO
BAYESIAN STATISTICS**

Prof. ANTONINO ABBRUZZO

TESTI CONSIGLIATI

A first course in Bayesian statistical methods, Hoff, Peter D, 2009, Springer Science & Business Media. Capitoli da 5 a 11.

Bayesian Modelling using WinBUGS, Ntzoufras Ioannis, 2009, Wiley. Capitoli da 3 a 8.

Appunti del docente.

TIPO DI ATTIVITA'	C
AMBITO	21031-Attività formative affini o integrative
NUMERO DI ORE RISERVATE ALLO STUDIO PERSONALE	54
NUMERO DI ORE RISERVATE ALLE ATTIVITA' DIDATTICHE ASSISTITE	21

OBIETTIVI FORMATIVI DEL MODULO

Il corso guida lo studente verso la conoscenza delle metodologie della statistica bayesiana e l'acquisizione della capacità di applicare tali metodologie a casi pratici reali. Alla fine del corso lo studente dovrà essere in grado di riconoscere i pregi e difetti delle tecniche Bayesiane rispetto a quelle classiche, e descrivere insiemi di dati reali complessi sfruttando le tecniche apprese.

PROGRAMMA

ORE	Lezioni
3	Inferenza Bayesiana con il campionamento di Gibbs
3	Inferenza Bayesiana con l'algoritmo di Metropolis-Hastings
3	Metodi basati sull'inferenza Bayesiana approssimata
3	Inferenza Bayesiana basata sull'approssimazione integrata di Laplace

ORE	Esercitazioni
2	Applicazione dell'algoritmo di campionamento di Gibbs
3	Applicazione dell'algoritmo di Metropolis-Hastings
2	Applicazione dell'algoritmo approssimato Bayesiano
2	Applicazione dell'algoritmo INLA

**MODULO
NON PARAMETRIC STATISTICS**

Prof. MARCELLO CHIODI

TESTI CONSIGLIATI

Density Estimation for Statistics and Data Analysis - B.W. Silverman. Statistics and Applied Probability, London: Chapman and Hall, 1986. (Capitoli: 1, 3, 4, 5 - Disponibile presso la Bib.Sc.Econ.Aziend.Statistiche).

Applied smoothing techniques for data analysis: The kernel approach with splines illustrations. - ADRIAN W. BOWMAN, ADELCHI AZZALINI. CLARENDON PRESS • OXFORD 1997. (Capitoli 1, 2, 3, 4, 5, 6 - Disponibile presso la Bib.Sc.Econ.Aziend.Statistiche).

Generalized additive models. An Introduction with R. S.N. Wood. Statistical Science: Chapman & Hall, 2006. (Capitoli: 2, 3, 4, 5 - Disponibile presso la Bib.Sc.Econ.Aziend.Statistiche).

Materiale didattico (dispense e lucidi) forniti dal docente.

TIPO DI ATTIVITA'	C
AMBITO	21031-Attività formative affini o integrative
NUMERO DI ORE RISERVATE ALLO STUDIO PERSONALE	108
NUMERO DI ORE RISERVATE ALLE ATTIVITA' DIDATTICHE ASSISTITE	42

OBIETTIVI FORMATIVI DEL MODULO

Il modulo si propone di guidare lo studente alla conoscenza delle metodologie della statistica non parametrica e all'acquisizione della capacità di applicare tali metodologie a casi pratici reali. Alla fine del corso lo studente dovrà essere in grado di riconoscere i pregi e difetti delle tecniche nonparametriche e semiparametriche rispetto a quelle parametriche, e descrivere insiemi di dati reali complessi sfruttando e integrando le tecniche non parametriche e semiparametriche con le tecniche parametriche

PROGRAMMA

ORE	Lezioni
2	Introduzione alla statistica non parametrica. Test d'ipotesi non parametrici Test non parametrici, Ipotesi funzionali, stime non parametriche
6	Densità kernel univariate: Istogramma, Stima della densità attraverso la funzione kernel, proprietà e scelta di h. Kernel variabili e proprietà. Cross Validation e k-fold cross-validation. Minimizzazione del MISE. Proprietà teoriche e simulazioni
8	Densità kernel multivariate: Stimatore kernel multivariato e forme di H. Proprietà dello stimatore kernel multivariato: AMISE. Analisi discriminante e cluster analysis non parametrica. Kth nearest neighbour
8	Non parametric regression: Local Polinomial regression, Cubic splines e Regressione spline penalizzata. Regressione quantilica. Regressione penalizzata e funzione obiettivo aggiunta. Regressione penalizzata e scelta del parametro di smoothing. Modelli additivi generalizzati e metodo di stima dei parametri. Cenno ai metodi non parametric bayesiani
ORE	Esercitazioni
3	Applicazioni a casi reali dei test di ipotesi non parametrici.
3	Uso delle densità kernel univariate per casi reali. esercitazioni in R. uso del software didattico MLANP
6	Uso delle densità kernel multivariate per casi reali.
6	Applicazione a casi reali dei modelli di regressione non parametrica e semiparametrica.