

FACOLTÀ	Economia
ANNO ACCADEMICO	2011-2012
CORSO DI LAUREA	Statistica per l'Analisi dei Dati L41
INSEGNAMENTO	Statistica 3
TIPO DI ATTIVITÀ	Caratterizzante
AMBITO DISCIPLINARE	Statistico-probabilistico
CODICE INSEGNAMENTO	6650
ARTICOLAZIONE IN MODULI	NO
SETTORI SCIENTIFICO DISCIPLINARI	SECS-S/01
DOCENTE RESPONSABILE (MODULO 1)	Marcello Chiodi Professore Ordinario Università di Palermo
CFU	10
NUMERO DI ORE RISERVATE ALLO STUDIO PERSONALE	154
NUMERO DI ORE RISERVATE ALLE ATTIVITÀ DIDATTICHE ASSISTITE	96
PROPEDEUTICITÀ	Matematica 1, Statistica 2, Calcolo delle Probabilità
ANNO DI CORSO	2
SEDE DI SVOLGIMENTO DELLE LEZIONI	Aula 5 (ed.13 primo piano) (soggetto a variazioni)
ORGANIZZAZIONE DELLA DIDATTICA	Lezioni frontali, Esercitazioni in aula di informatica.
MODALITÀ DI FREQUENZA	Facoltativa
METODI DI VALUTAZIONE	Esame scritto con PC ed esame orale
TIPO DI VALUTAZIONE	Voto in trentesimi
PERIODO DELLE LEZIONI	Terzo quadrimestre
CALENDARIO DELLE ATTIVITÀ DIDATTICHE	Lun. Giov. Ven. 12-14 Mar. Mer. 10-12
ORARIO DI RICEVIMENTO DEGLI STUDENTI	Lun. 10-12 Mer. 12-14

RISULTATI DI APPRENDIMENTO ATTESI

Conoscenza e capacità di comprensione

Conoscenza del linguaggio e della terminologia dei modelli di dipendenza
Comprensione del modello normale multivariato (ad un livello base)
Conoscenza del linguaggio e della terminologia dei modelli lineari
Comprensione della unicità di approccio all'analisi della dipendenza da variabili quantitative e qualitative
Comprensione dell'analisi dei modelli lineari e della regressione logistica nell'ambito dei GLM
Comprensione dell'analisi della verosimiglianza per la costruzione di test e stimatori
Comprensione dei diversi tipi di scomposizione della devianza
Comprensione del concetto di varianza "spiegata"
Conoscenza delle tecniche di selezione dei modelli
Comprensione dei modelli di regressione e dell'analisi della varianza a più vie come casi particolari dei modelli lineari.
Conoscenza dell'analisi della collinearità
Conoscenza delle tecniche di analisi dei residui

Capacità di applicare conoscenza e comprensione

Essere in grado di applicare la metodologia dei modelli di dipendenza lineari e non a diversi casi di dati reali, provenienti da contesti diversi
Capacità di discriminare modalità diverse di acquisizione dei dati
Capacità di distinzione fra diverse tipologie di matrici del disegno in funzione della natura dei dati e delle variabili.
Capacità di costruzione della matrice del disegno più opportuna in funzione della natura dei dati e delle variabili.
Capacità di impiego del software specialistico (R) per l'analisi dei modelli lineari e non
Capacità di interpretare il significato dei parametri
Capacità di applicare l'analisi dei residui e di comprenderne i risultati
Capacità di utilizzare le tecniche di selezione di un modello statistico di dipendenza

Autonomia di giudizio

Essere in grado di produrre risultati in relazione e al fenomeno sostantivo
Essere in grado di giudicare l'utilità dei modelli impiegati per analizzare relazioni tra variabili
Essere in grado di analizzare criticamente i risultati utilizzando tecniche di analisi dei residui.
Essere in grado di sintetizzare i risultati e le conclusioni.
Capacità di effettuare un processo di selezione di un modello statistico di dipendenza, integrando con le informazioni provenienti dall'analisi dei residui
Capacità di giudicare il modello di dipendenza più adeguato a descrivere un fenomeno
Capacità di giudicare il software da utilizzare per l'implementazione di particolari tecniche

Abilità comunicative

Essere in grado di redigere un rapporto tecnico con livelli di approfondimento a seconda dell'ambito applicativo e del destinatario;
scegliere le rappresentazioni grafiche più opportune in relazione ai problemi;
esporre verbalmente il contenuto del rapporto sia in forma sintetica che in forma estesa, in funzione delle conoscenze statistiche del destinatario

Capacità d'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.

Obiettivi formativi del corso

Il corso ha come obiettivo fondamentale di offrire allo studente gli strumenti fondamentali per la costruzione e l'analisi di modelli statistici di dipendenza.. Alla fine del corso lo studente dovrà essere in grado di scrivere relazioni di tipo statistico, utilizzando l'opportuno software e scegliendo i modelli adeguati.

Conoscenze propedeutiche

Algebra delle matrici

Variabili aleatorie discrete e continue: distribuzione di probabilità. Principi e metodi della stima e della verifica di ipotesi. Funzione di verosimiglianza

Materiale didattico

Il materiale didattico consisterà anche in dispense e lucidi forniti dal docente insieme con codice R per alcuni esercizi. Per maggiori approfondimenti verranno indicate le fonti bibliografiche più utili durante il corso.

Gli appunti: Modelli lineari; appunti di marcello chiodi;

potranno essere scaricate dal sito internet del Corso di Laurea dal sito <http://dssm.unipa.it/chiodi>.

In particolare i files più aggiornati si trovano seguendo il link della materia statistica 3 all'indirizzo

<http://dssm.unipa.it/chiodi/teaching/files>

Statistica 3		
ORE LEZIONI	ORE ESERCIZI - TAZIONI	ARGOMENTI
2		Introduzione al corso e problemi reali introduttivi;
2	1	□ variabili statistiche multiple; definizione della matrice dei dati; momenti primi e secondi (multivariati) di una variabile statistica multipla e di combinazioni lineari di variabili statistiche multipla
2	1	□ la distribuzione normale bivariata; funzioni di regressione; distribuzione normale multivariata (cenni); combinazioni lineari di variabili normali; risultati principali sulle distribuzioni condizionate;
2	1	Funzioni di regressione e dipendenza in media: modello lineare come approssimazione del problema generale della dipendenza in media; regressione non parametrica
2	1	□ Modelli lineari: versatilità del modello lineare generale; componente sistematica e accidentale; la matrice delle x; osservazioni ripetute; disegni fattoriali; regressione multipla; regressione polinomiale, regressori del tipo 0/1, modelli con interazioni;
8	4	□ assunzioni di base nel modello lineare; la verosimiglianza del modello lineare: costruzione del test LR e degli stimatori ML; minimi quadrati ordinari; distribuzione campionaria di b, scomposizione della devianza nel modello lineare e verifica di ipotesi. test F per la verifica di ipotesi nel modello lineare; configurazioni della matrice x e di x'x; modello lineare: verifica di ipotesi generali; intervalli di confidenza per E(Y); errori di previsione.
6	4	□ regressione multipla; scomposizione della devianza empirica e coefficiente di determinazione lineare multipla r ² ; prova di ipotesi particolari; test per l'eliminazione di q regressori; la multicollinearità
4	2	Scelta delle variabili; confronto fra modelli; AIC; errori di previsione e cross validation.
4	4	Violazione delle assunzioni di base nella regressione e nel modello lineare: analisi dei residui; proprietà dei residui empirici; grafici impiegati per l'analisi dei residui.
8	6	□ stima dei parametri del modello lineare con fattori qualitativi; analisi della varianza ad una via; assunzioni di base; divergenza dalla linearità per fattori quantitativi, differenza fra i test di omogeneità; analisi della varianza a due vie; significato delle interazioni; analisi della varianza con variabili concomitanti: l'analisi della covarianza. allontanamento dalle assunzioni di base nel modello lineare e nell'analisi della varianza; analisi dei residui nell'analisi della varianza.
16	10	Modelli lineari generalizzati: funzioni legame e famiglia di distribuzioni esponenziale. Regressione logistica, Regressione di Poisson. Stima dei parametri e scelta del modello
4	2	Cenni a tecniche esplorative di analisi multivariata: analisi delle componenti principali; definizione di distanza per dati multivariati; cluster analysis (cenni)