UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Introduction To Statistical Models

2526-1-F8205B005

Obiettivi formativi

Il corso si propone di richiamare gli aspetti metodologici e applicativi di base dei modelli statistici secondo i punti di vista descrittivo e inferenziale. Argomenti principali sono il modello di regressione lineare sia semplice che multipla per dati di popolazione e per dati campionari, e la loro analisi con il software SAS. Il corso si rivolge ai laureati con formazione pregressa non quantitativa o non assimilabile a quella tipicamente acquisita nelle lauree triennali in Statistica.

Conoscenza e comprensione. Questo insegnamento fornirà conoscenze e capacità di comprensione relativamente a:

- Principali metodi per lo studio del legame di dipendenza di una variabile quantitativa da una o più variabili quantitative, insieme eventualmente a variabili qualitative, e per la costruzione della relativa modellistica lineare in ambito sia descrittivo sia inferenziale
- Applicazione dei principali metodi di analisi bivariata mediante esercizi numerici svolti con la calcolatrice (ossia, senza l'ausilio del software statistico)
- Logica e funzionamento alla base del software SAS e suo utilizzo nell'ambito dell'analisi bivariata e della modellistica di regressione con le relative rappresentazioni grafiche
- Lettura e interpretazione degli output delle analisi prodotte con SAS.

Capacità di applicare conoscenza e comprensione. Alla fine dell'insegnamento gli studenti saranno in grado di:

- Applicare i metodi di base dell'analisi bivariata, stabilire nel modo più opportuno il ruolo da assegnare alle variabili (ossia se dipendenti o indipendenti), e costruire i modelli di regressione più adeguati in base agli scopi delle analisi e alla natura dei dati a disposizione
- Selezionare un modello di regressione che sia al contempo parsimonioso ma di buon adattamento in base a criteri statistici descrittivi o inferenziali
- Interpretare i risultati delle analisi in modo critico e individuare gli eventuali margini di miglioramento nella costruzione delle modellistiche di base applicate in ottica sia interpretativo-descrittiva, sia predittiva
- Importare in SAS file di dati esterni di varia provenienza e formato e utilizzare in modo autonomo la sintassi

di base di SAS.

L'insegnamento consente allo studente di acquisire le principali basi teoriche e applicative relativamente all'analisi bivariata e alla costruzione dei modelli di regressione lineare necessarie in qualsiasi contesto lavorativo in cui si utilizzino file di dati e che rappresentano una base imprescindibile per il proseguimento del percorso universitario.

Contenuti sintetici

Dipendenze in media e lineare, interpolazione fra punti e per punti, funzione di regressione. Regressione lineare semplice e multipla per dati di popolazione e per dati campionari. Selezione del modello di regressione in base a criteri descrittivi e inferenziali. Applicazioni a dati reali con il software SAS.

Programma esteso

- Richiami sulle forme principali di dipendenza su dati di popolazione quantitativi: dipendenze in media e lineare. Interpolazione fra punti e per punti. Funzione di regressione. Criterio di accostamento dei minimi quadrati
- Regressione lineare: retta dei minimi quadrati, bontà di adattamento, estensione a più variabili esplicative, correlazioni multipla e parziale, regressori qualitativi
- Modello di regressione lineare semplice e multipla per dati campionari: distribuzione normale multivariata e sue proprietà, specificazione del modello, ipotesi, metodi di stima dei minimi quadrati e di massima verosimiglianza, intervalli di confidenza, verifica di ipotesi
- Selezione del modello di regressione lineare multipla: criteri descrittivi e inferenziali
- Analisi di casi empirici con SAS

Prerequisiti

Per questa attività formativa è consigliata la conoscenza degli argomenti trattati nei corsi di base di Statistica e di Inferenza Statistica.

Metodi didattici

Lezioni teoriche in aula ed esercitazioni pratiche in laboratorio statistico-informatico con il software SAS.
Le lezioni saranno tutte erogate in presenza in modalità erogativa. Verranno comunque messe a disposizione le video-registrazioni in asincrono sia delle lezioni di teoria sia delle esercitazioni in laboratorio.

Modalità di verifica dell'apprendimento

L'esame di Introduzione ai Modelli Statistici consiste in una prova scritta (durata complessiva: 2 ore) con tre quesiti (articolati in più punti) che riguardano gli aspetti sia teorici sia applicativi delle modellistiche trattate durante il corso.

Le domande teoriche riguardano gli aspetti metodologici degli argomenti trattati e consentono di verificare le conoscenze teoriche acquisite in merito alla logica e agli aspetti di base dell'analisi affrontata su dati di popolazione o su dati campionari, dell'analisi bivariata, della specificazione formale dei modelli di regressione, dei problemi legati alla determinazione o alla stima dei parametri e alla valutazione della bontà di adattamento del modello ai dati, e della conduzione dell'inferenza statistica in un'ottica di modellistica. Permettono inoltre di verificare la capacità di utilizzare in autonomia il linguaggio simbolico-formale statistico, di fornire in modo appropriato le definizioni e di dimostrare analiticamente i principali risultati teorici. Le domande applicative riguardano la trasposizione nella pratica della teoria, e consentono di verificare le capacità di comprensione e di applicazione della teoria, e di lettura e interpretazione dei risultati delle analisi ottenute con il software SAS. Inoltre, l'esame in forma scritta permette complessivamente di verificare la capacità di espressione mediante utilizzo adeguato del linguaggio tecnico statistico.

La prova orale è facoltativa (su richiesta del docente o dello studente) e riguarda argomenti sia teorici sia pratici. L’accesso alla prova orale è subordinato al superamento della prova scritta con un esito di almeno 18/30. Si fa presente che la prova orale può comportare sia l'aumento, sia il mantenimento, che la diminuzione della valutazione conseguita alla prova scritta.

Considerata l'abbondanza di materiale didattico messo a disposizione dalla docente sulla piattaforma e-learning del corso, non si prevede alcuna distinzione fra esami per studenti frequentanti ed esami per studenti non frequentanti. Infine non si prevedono prove in itinere.

Testi di riferimento

- Materiale didattico della docente pubblicato sul sito e-learning del corso (ad accesso riservato)
- Delwiche, L.D., Slaughter, S.J. (2012), The Little SAS Book: A Primer, SAS Institute
- Freund, R. J., Wilson, W. J., and Sa, P. (2006), Regression Analysis: Statistical Modeling of a Response Variable, 2nd edition, Academic Press
- Johnston, J. (1993), Econometrica, 3rd edition, Franco Angeli, Milano
- Littell, R. C., Freund, R. J., and Spector, P. C. (2002), SAS for Linear Models, 4th Edition, Cary, NC: SAS Institute Inc.
- Piccolo, D. (2010), Statistica, Il Mulino, Bologna
- Spencer N. (2004), SAS Programming - The One-Day Course, 1st Edition, Chapman and Hall/CRC, New York
- Zelterman, D. (2010), Applied Linear Models with SAS, Cambridge University Press, New York
- Zenga, M. (2014), Lezioni di Statistica Descrittiva, Giappichelli, Torino

Periodo di erogazione dell'insegnamento

I Semestre, I periodo

Lingua di insegnamento

Italiano

Sustainable Development Goals

ISTRUZIONE DI QUALITÁ