UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

# COURSE SYLLABUS

# Generalized Linear Model

**2627-1-F8205B003-F8205B003-1**

## Obiettivi formativi

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

modelli lineari
modelli GLS
modelli multivariati
modelli multilevel

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

Diagnosticare i risolvere le violazioni delle ipotesi del modello OLS
Applicare modelli per dati con struttura gerarchica
Applicare modelli multiviariati (con piu variabili dipendenti)
Utilizzare le principali procedure di SAS ed R dedicate ai modelli lineari, GLS, multivariati e multilevel

## Contenuti sintetici

Il corso ha quale obiettivo lo studio di modelli più avanzati del modello lineare classico. Si presentano perciò

- modelli lineari generalizzati,
- modelli lineari multivariati
- modelli multilevel

## Programma esteso

Il corso ha quale obiettivo l'introduzione alla specificazione, stima e verifica di modelli interpretativi dei dati di tipo lineare più avanzati del modello lineare classico. Si presentano perciò

· Modelli lineari che non rispettano le ipotesi del modello lineare classico: modelli con errori esteroschedastici e correlati, modelli non lineari, trattamento di outlier, modelli GLS

· Modelli lineari multivariati: dal modello classico multivariato al modello seemingly unrelated

· Modelli multilevel per dati gerarchici: la natura dei dati gerarchici, anova ad effetti fissi, modelli mixed (random slope, random intercept)

Ciascun ambito sarà l'oggetto specifico di un modulo del corso. L'attività formativa è svolta attraverso lezioni teoriche e lezioni pratiche in laboratorio statistico-informatico nelle quali si affronteranno analisi su casi empirici mediante l'uso dei software R e SAS. Il materiale del corso (sia delle lezioni teoriche sia delle lezioni pratiche) e ulteriori informazioni verranno riportate sulla pagina web dedicata nella piattaforma e-learning unimib: http://elearning.unimib.it/.

## Prerequisiti

Si richiede una buona conoscenza della

Statistica descrittiva univariata : indici di posizione; indici di variabilità:a; indici di simmetria e di curtosi.

Statistica descrittiva bivariata: connessione, dipendenza in media, correlazione lineare, regressione lineare bivariata, multipla, multivariata, polinomiale, non lineare.

Teoria della probabilità: popolazione e campione; significato di probabilità nella versione classica ; elementi di calcolo combinatorio; tipi di campionamento; distribuzioni di variabili casuali univariate; variabili casuali Normale , t di Student, F d Snedecor ; distribuzioni casuali campionarie

Inferenza: teoria della stima, proprietà dello stimatore puntuale; stima intervallare; verifica di ipotesi, test di ipotesi di Neyman Pearson; test di ipotesi sulle medie basati su Normale , t di Student; test d ipotesi sulla varianza.

Modello lineare classico: ipotesi; stima dei parametri del modello nel campione e nella popolazione; proprietà degli stimatori dei minimi quadrati; test di ipotesi sui parametri basati su Normale , t di Student, ; test di ipotesi sul modello e su gruppi di parametri , su un parametro basata F di Snedecor

Algebra delle matrici

Si suggerisce a chi non provenga da corsi triennali di statistica o economia di seguire preventivamente i corsi introduttivi del corso di laurea di biostatistica calcolo delle probabilità, introduzione all' inferenza stastistica, introduzione ai modelli statistici,modelli statistici per dati categoriali e di conoscere i pacchetti statistici R e SAS.

## Metodi didattici

Le lezioni si distinguono in parte teorica e parte applicata. Durante la parte teorica vengono presentate i framework metodologici relativi al corso, che vengono poi applicati durante le lezioni pratiche in laboratorio. In laboratorio si

utilizzano i software SAS ed R con attenzione ai codici e alla lettura dei degli output dei modelli. Lezioni ed esercitazioni saranno registrate sulla piattaforma e-learning.

## Modalità di verifica dell'apprendimento

L'esame si svolge in laboratorio. Lo studente dovrà rispondere a due quesiti teorici tra un insieme di domande predeterminate che conoscerà già all'inizio del corso. Occorre argomentare la risposta in termini comprensibili ed esaurienti riportando le dimostrazioni richieste. Il punto di riferimento per le risposte sono le slides e la dispensa: ovviamente si possono riportare le conoscenze acquisite dai libri consigliati. Si devono riportare formule e grafici: se risulta difficoltoso si possono scrivere su foglio con penna e poi scannerizzarlo. La lunghezza richiesta delle risposte dipenderà dalla domanda: si suggeriscono risposte che non superino i quattro fogli dattiloscritti in calibri 12 interlinea 1.5 (12000 battute spazi inclusi).
La seconda parte dell'esame consterà in un esercizio pratico su dati reali o simulati forniti dal docente mediante l'uso di pacchetti statistici. Gli strumenti statistici che dovrà utilizzare saranno quelli appresi al corso. Nell'elaborato tutti i grafici e gli output dovranno essere opportunamente commentati, sia da un punto di vista teorico, sia rispetto all'applicazione in esame. Lo svolgimento avviene tramite l'ambiente R. Lo studente potrà utilizzare i codici delle esercitazioni durante l'esame. Tali codici verranno forniti il giorno della prova.

## Testi di riferimento

Il principale testo di riferimento è la dispensa del corso, resa disponibile in formato digitale sulla piattaforma elearning. La dispensa contiene sia la parte teorica che esempi pratici.

Testi consigliati:

- Wooldridge, J. M. (2015). Introductory econometrics: A modern approach. Cengage learning.
- Freund, R. J., Wilson, W. J., and Sa, P. (2006), Regression Analysis: Statistical Modeling of a Response Variable, 2nd edition, Academic Press
- Baltagi B. H. (2008), Econometrics, fourth Edition, Springer Berlin

-Rencher , A. C. ,Methods of Multivariate Analysis, Wiley

-Tom Snijders, T., Bosker, R., Multilevel Analysis: An Introduction To Basic And Advanced Multilevel Modeling, SAGE Publictaions Inc

- Littell, R. C., Freund, R. J., and Spector, P. C. (2002), SAS for Linear Models, 4th Edition, Cary, NC: SAS Institute Inc.
- Manual SAS/STAT 15.1
- Faraway, J. J. (2004). Linear models with R. Chapman and Hall/CRC.

## Periodo di erogazione dell'insegnamento

II Semestre, III periodo

## Lingua di insegnamento

Italiano

## Sustainable Development Goals

ISTRUZIONE DI QUALITÁ