UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Statistical Models in Epidemiology (blended)

2122-2-F8203B008

Obiettivi formativi

1. **Conoscenza e capacità di comprensione (knowledge and understanding):**

Il corso si propone di introdurre metodi e tecniche statistiche per lo studio (1) dei tempi di sopravvivenza: stima della distribuzione dei tempi di sopravvivenza, confronto tra curve di sopravvivenza, modelli semi-parametrici, modelli parametrici, rischi competitivi, modelli multivariati di sopravvivenza, regressione quantilica, disegno dello studio; (2) dei dati longitudinali di tipo continuo e discreto: modelli marginali, modelli a effetti misti, modelli GEE, dati mancanti.

2. **Capacità di applicare conoscenza e comprensione (applying knowledge and understanding):**

Sulla base delle conoscenze acquisite, lo studente sarà in grado di applicare le conoscenze acquisite per analizzare con strumenti appropriati i tempi di sopravvivenza e i dati longitudinali, utilizzando il software statistico Stata (StataCorp), di ampio uso in epidemiologia e biostatistica. Applicazioni in R and SAS saranno presentate brevemente in appositi seminari.

3. **Autonomia di giudizio (making judgements):**

Utilizzando le metodologie presentate a lezione, lo studente dovra' utilizzare le tecniche statistiche presentate per analizzare come variabili di esposizione, fattori di rischio, trattamenti, provenienti da studi epidemiologici o clinici possono avere un impatto sui tempi di sopravvivenza e sulle misure ripetute di dati longitudinali, e in modo critico valutare la performance del modello statistico utilizzato.

4. **Abilità comunicative (communication skills)**

Lo studente dovra' essere in grado di comprendere il materiale presentato in classe e di comunicare in modo autonomo le tecniche statistiche imparate in classe e esporre in modo chiaro come interpretare i risultati ottenuti.

5. **Capacità di apprendimento (learning skills):**

Lo studente acquisira' le capacita' e conoscenze necessarie per analizzare i dati di sopravvivenza e longitudinali provenienti da uno studio epidemiologico o da un studio clinico, attraverso un percorso composto da lezioni teoriche e da applicazioni in Stata.

Contenuti sintetici

1. Analisi dei Dati di Sopravvivenza: Introduzione all'analisi della sopravvivenza. Censura e troncamento. Funzioni di sopravvivenza, densità e rischio. Procedura “st” in Stata. Stima e confronto delle funzioni di sopravvivenza. Il modello di regressione di Cox. Analisi dei rischi competitivi. Modelli di sopravvivenza parametrici. Regressione quantilica. Disegno dello studio.
2. Analisi dei dati longitudinali: Introduzione all’analisi dei dati longitudinali: i modelli marginali e i modelli a effetti casuali per risposte continue e discrete.

Programma esteso

1. Analisi dei Dati di Sopravvivenza: Introduzione all'analisi della sopravvivenza. Definizione ed esempi. Censura e troncamento. Funzioni di sopravvivenza, densità e rischio. Procedura “st” in Stata. Stima e confronto delle funzioni di sopravvivenza di Kaplan-Meier e Nelson-Aalen. Tavole di sopravvivenza. Log-rank test, Wilcoxon-Breslow-Gehan, Tarone-Ware test. Il modello di regressione di Cox. Formulazione e ipotesi. Verosimiglianza Parziale. Interpretazione dei parametri del modello. Diagnostica: Martingale e Deviance dei residuals, valutazione delle ipotesi del modello. Analisi dei rischi competitivi: definizione di net e crude survival, la funzione di incidenza cumulativa, hazard causa specifica e la funzione subhazard. Il modello di Fine e Gray. Modelli parametrici: esponenziale e weibull. La regressione quantilica. La sopravvivenza multivariata. Introduzione al disegno dello studio. Applicazioni in R e SAS.

2. Analisi dei dati longitudinali: Introduzione all’analisi dei dati longitudinali: esempi e motivazioni. Modelli di risposta per lo studio dei dati longitudinali: estensione dei modelli standard. Struttura della matrici di varianza e covarianza. Analisi dei profili di risposta. I modelli a effetti casuali per risposte continue e discrete. Modelli a effetti fissi. Introduzione ai valori mancanti: definizione, trattamento e possibili metodi di stima. Introduzione ai dati mancanti.

Prerequisiti

Le lezioni vengono svolte in lingua inglese, anche il materiale distribuito e' in lingua inglese: dunque la

conoscenza della lingua inglese e' considerata un requisito importante per la partecipazione attiva in classe, per la comprensione delle slides utilizzate, e per la lettura di tutto il materiale del corso.

Metodi didattici

Blended/Elearning: Lezioni frontali e online , group work, seminars, utilizzo di Stata. Le lezione online saranno svolte con un duplice scopo: il primo e' quello di consentire a studenti non frequentanti di seguire alcune lezioni, che verranno preparate e registrate e dunque rese disponibili nella piattaforma elearning. Il secondo e' quello di discutere e rivedere i compiti che gli studenti hanno preparato, evidenziando le problematiche presenti, proponendo un modello di soluzione al compito.

A causa delle emergenza Covid-19 le lezioni in presenza verranno trasmesse live e le lezioni registrateverranno rese disponibili sulla piattaforma di ateneo.

Modalità di verifica dell'apprendimento

1. Lavoro di gruppo (30% votazione finale) (max 5 - 6 studenti) che si basa sulla consegna di 9 compiti, uno a settimana, conta per il 30% della votazione finale.

2. Esame Finale (70% votazione finale).

3. Prova orale-facoltativa : La discussione orale dell'esame scritto serve per valutare la conoscenza dello studente laddove le risposte all'esame non siano state chiare e poste in modo impreciso.

Sia i compiti che si basano sul lavoro di gruppo, sia l'esame finale, hanno come fine quello di valutare il grado di apprendimento dello studente e se gli obiettivi del corso sono stati raggiunti in termine di formulazione delle ipotesi dello studio analizzato utilizzando Stata, analisi descrittive, inferenza statistica, interpretazione dei parametri, valutazione delle ipotesi dei modelli fittati.

L’orale può comportare sia l’aumento che la diminuzione della valutazione della prova scritta.

Per gli studenti che non possono partecipare al lavoro di gruppo, la valutazione si baserà sulla prova scritta e orale soltanto, con un peso del 100%.

L'esame scritto finale e' lo stesso per studenti frequentanti e studenti non frequentanti

Nel periodo di emergenza Covid-19 gli esami saranno solo telematici. Verranno svolti utilizzando la piattaforma Zoom e WebEx e nella pagina e-learning dell'insegnamento verrà riportato un link pubblico per l'accesso all'esame di possibili spettatori virtuali.

Testi di riferimento

1. Kleinbaum, D.G. and Klein, M. Survival Analysis, A self-learning text. (2013). Springer.

2. Fitzmaurice, G. M., Laird, N. M., and Ware, J. H. (2013). Applied Longitudinal Analysis, Chapman & Hall CRC.

3. Jewell, N.P. Statistics for Epidemiology. (2004). Chapman & Hall CRC.

Materiale integrativo sara' fornito durante le lezioni e messo a disposizione tramite la pagina web del corso.

Periodo di erogazione dell'insegnamento

II Semestre, III ciclo

Lingua di insegnamento

Inglese