UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Statistical Models and Bayesian Inference

2627-2-F8205B011

Obiettivi formativi

L'insegnamento si articola in due moduli appartenenti all'ambito delle scienze statistiche, dell'informatica e delle scienze sociali, con l'obiettivo di fornire agli studenti conoscenze teoriche e competenze operative per l'analisi statistica avanzata, con particolare attenzione alla riproducibilità e alla replicabilità delle analisi, nonché alla comunicazione efficace dei risultati.

Il primo modulo introduce i principali metodi avanzati per l'inferenza statistica, tra cui il bootstrap, i modelli miscuglio e i metodi di previsione. Le attività didattiche prevedono l'utilizzo del software R e dell'ambiente RMarkdown per l'analisi di dati reali e simulati, consentendo allo studente di sviluppare autonomia nel ragionamento statistico, capacità di problem solving e competenze nella produzione di report riproducibili.

Il secondo modulo approfondisce l'approccio bayesiano all'inferenza statistica, integrandolo con i metodi classici presentati nel primo modulo. Lo studente acquisisce le basi teoriche dei modelli bayesiani, dei metodi computazionali Monte Carlo tramite catene di Markov (MCMC) e dell'impiego di strumenti software quali R e SAS per la stima, la valutazione e il confronto dei modelli. Anche in questo modulo viene posta particolare attenzione alla riproducibilità delle analisi e alla produzione di documenti integrati contenenti codice, risultati e relativa interpretazione.

Per la descrizione dettagliata dei contenuti si rimanda al syllabus pubblicato nelle pagine dei due rispettivi insegnamenti.

Nel suo complesso, l'insegnamento consente allo studente di acquisire solide basi teoriche e competenze applicative per affrontare problemi di analisi statistica avanzata in diversi ambiti, tra cui la biostatistica, l'epidemiologia, la genetica, la salute pubblica e le scienze sociali. Al termine del percorso, grazie all'approccio orientato alla pratica, alla riproducibilità e alla comunicazione scientifica, lo studente sarà in grado di approfondire autonomamente gli argomenti trattati e di applicare le conoscenze acquisite in contesti di ricerca e professionali.

Risultati di apprendimento attesi (Descrittori di Dublino)

DdD 1 – Conoscenza e capacità di comprensione

Al termine dell'insegnamento lo studente sarà in grado di:
• descrivere i fondamenti teorici dei principali modelli statistici avanzati, inclusi i modelli miscuglio e i modelli bayesiani;
• comprendere e spiegare le assunzioni alla base dei modelli, le procedure di stima (massima verosimiglianza con l'algoritmo expectation-maximization e inferenza bayesiana) e i principali criteri di valutazione dei modelli;
• riconoscere il ruolo della riproducibilità e della replicabilità nella ricerca scientifica attraverso l'utilizzo integrato di R, RStudio, RMarkdown e SAS.

DdD 2 – Capacità di applicare conoscenza e comprensione

Al termine dell'insegnamento lo studente sarà in grado di:
• sviluppare simulazioni da distribuzioni univariate e multivariate;
• analizzare dati reali provenienti da differenti ambiti applicativi mediante R, RMarkdown e SAS;
• applicare tecniche statistiche avanzate, tra cui bootstrap, modelli a miscuglio, inferenza bayesiana, diagnostica, selezione e validazione dei modelli, nonché metodi di previsione;
• produrre report riproducibili che integrino codice, risultati, interpretazioni e discussione critica delle analisi.

DdD 3 – Autonomia di giudizio

Al termine dell'insegnamento lo studente sarà in grado di:
• valutare criticamente la validità delle assunzioni dei modelli e l'adeguatezza dei metodi statistici rispetto ai quesiti di ricerca;
• confrontare modelli alternativi utilizzando appropriati criteri diagnostici e misure di accuratezza predittiva;
• interpretare criticamente i risultati ottenuti, discutendone punti di forza, limiti e implicazioni nei diversi contesti applicativi.

DdD 4 – Abilità comunicative

Al termine dell'insegnamento lo studente sarà in grado di:
• comunicare efficacemente i risultati delle analisi statistiche mediante report strutturati, documentati e riproducibili;
• presentare in forma scritta analisi, interpretazioni e conclusioni con un linguaggio scientifico chiaro e rigoroso, motivando le scelte metodologiche adottate.

DdD 5 – Capacità di apprendimento

Al termine dell'insegnamento lo studente sarà in grado di:
• proseguire autonomamente lo studio di metodi statistici avanzati utilizzando la bibliografia, i materiali didattici e gli esempi di codice forniti;
• aggiornare e ampliare le proprie competenze metodologiche e computazionali, applicando strumenti di data science e inferenza statistica in contesti interdisciplinari e di ricerca.

Contenuti sintetici

Per la descrizione dettagliata si rimanda al syllabus pubblicato nelle pagine dei due rispettivi insegnamenti.

Programma esteso

Per la descrizione dettagliata si rimanda al syllabus pubblicato nelle pagine dei due rispettivi insegnamenti.

Prerequisiti

Per la descrizione dettagliata si rimanda al syllabus pubblicato nelle pagine dei due rispettivi insegnamenti.

Metodi didattici

Per la descrizione dettagliata si rimanda al syllabus pubblicato nelle pagine dei due rispettivi insegnamenti.

Modalità di verifica dell'apprendimento

Per la descrizione dettagliata si rimanda al syllabus pubblicato nelle pagine dei due rispettivi insegnamenti.

Testi di riferimento

Albert, J. (2009). Bayesian computation with R (2nd ed.). Springer.

Albert, J., & Hu, J. (2019). Probability and Bayesian modeling. Chapman and Hall/CRC.

Bartolucci, F., Farcomeni, A., & Pennoni, F. (2013). Latent Markov models for longitudinal data. Chapman and Hall/CRC.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (2007). Discrete multivariate analysis: Theory and practice. Springer.

Blitzstein, J. K., & Hwang, J. (2014). Introduction to probability. Chapman and Hall/CRC.

Dey, D. K., Ghosh, S. K., & Mallick, B. K. (Eds.). (2000). Generalized linear models: A Bayesian perspective. CRC Press.

Gentle, J. E., Härdle, W. K., & Mori, Y. (Eds.). (2004). Handbook of computational statistics: Concepts and methods. Springer.

Lange, K. (2010). Numerical analysis for statisticians (2nd ed.). Springer.

Migon, H. S., Gamerman, D., & Louzada, F. (2014). Statistical inference: An integrated approach (2nd ed.). Chapman and Hall/CRC.

Pennoni, F. (2026). Dispensa di Inferenza Bayesiana: Teoria e applicazioni con R e SAS. Dipartimento di Statistica e Metodi Quantitativi, Università degli Studi di Milano-Bicocca.

Pennoni, F. (2026). Dispensa di Modelli Statistici II: Teoria e applicazioni con R. Dipartimento di Statistica e Metodi

Quantitativi, Università degli Studi di Milano-Bicocca.

R Core Team. (2026). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Robert, C. P., & Casella, G. (2004). Monte Carlo statistical methods (2nd ed.). Springer.
SAS Institute Inc. (2012). SAS/STAT 12.1 user's guide: The MCMC procedure. SAS Institute Inc.

Periodo di erogazione dell'insegnamento

Per la descrizione dettagliata si rimanda al syllabus pubblicato nelle pagine dei due rispettivi insegnamenti.

Lingua di insegnamento

Per la descrizione dettagliata si rimanda al syllabus pubblicato nelle pagine dei due rispettivi insegnamenti.

Sustainable Development Goals

PARTNERSHIP PER GLI OBIETTIVI