



UNIVERSITÀ
DEGLI STUDI
DI PADOVA



Bollettino Notiziario - A.A. 2021/2022

LAUREA MAGISTRALE IN SCIENZE STATISTICHE

Curriculum: Corsi comuni

ANALISI DEI DATI (DATA MINING)

Titolare: Prof. BRUNO SCARPA

Periodo: I anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 56A+18L; 9,00

Prerequisiti:

Sostanziali ma non formali: Modelli statistici II Analisi dei dati Multidimensionali, Elementi programmazione, Elementi di Algebra Lineare

Conoscenze e abilità da acquisire:

Lo scopo del corso è di introdurre gli studenti alla comprensione e alla reale capacità di utilizzo di strumenti di data mining e di metodi statistici per l'analisi dei dati.

Attività di apprendimento previste e metodologie di insegnamento:

Lezioni frontali. Esercitazioni in laboratorio

Contenuti:

- Nozioni generali: motivazioni e contesto, contrasto tra aderenza ai dati e complessità del modello ovvero contrasto tra distorsione e varianza, tecniche generali per la selezione del modello (AIC, BIC, convalida incrociata, oltre ai test statistici classici), suddivisione dei dati in un insieme di lavoro e uno di verifica. - Metodi di regressione: richiami sui modelli lineari e sui glm; regressione non parametrica mediante il metodo della regressione locale, splines di regressione, splines di lisciamiento, modelli additivi, alberi, mars, projection pursuit, reti neurali (cenni). - Metodi di classificazione: mediante la regressione lineare, richiami sulla regressione logistica e multilogit, modelli additivi, alberi, polymars, reti neurali, combinazione di classificatori (bagging, boosting, foreste casuali), support vector machines. - Metodi di analisi interna: Introduzione ai metodi di analisi di dati su reti e alle reti sociali. Nozioni sui metodi di raggruppamento: dissimilarità, metodo delle k-medie, metodi gerarchici. - Miscellanea: Introduzione alla sentiment analysis, tecniche di visualizzazione dei dati, cenni ad aspetti computazionali.

Modalità di esame:

La prova d'esame consta di tre parti: una parte "teorica", una "pratica" ed una "orale".

Criteri di valutazione:

Correttezza e qualità delle prove d'esame

Testi di riferimento:

Azzalini, A. e Scarpa, B., Data analysis and data mining: an introduction. New York: Oxford University Press, 2012 Azzalini, A. e Scarpa, B., Analisi dei dati e data mining. Milano: Springer-Verlag Italia, 2004 Hastie, Trevor J.; Tibshirani, Robert, The elements of statistical learning: data mining, inference, and prediction. New York: Springer, 2009

Eventuali indicazioni sui materiali di studio:

Strumenti informatici Lo strumento di calcolo primario adottato per questo corso è l'ambiente di programmazione R; questo può essere prelevato, assieme alla relativa documentazione, da una postazione CRAN.

ANALISI DEI DATI IN FINANZA

Titolare: Prof. MASSIMILIANO CAPORIN

Periodo: Il anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 42A+22L; 9,00

Prerequisiti:

Anche se non strettamente necessario, è fortemente consigliato avere acquisito i contenuti dei corsi Serie storiche finanziarie ed Economia dei Mercati Finanziari.

Conoscenze e abilità da acquisire:

Lo scopo del corso è di introdurre gli studenti alla comprensione e alla reale capacità di utilizzo di metodi statistici per l'estrazione delle informazioni contenute in dati finanziari grezzi, l'analisi degli stessi, e la modellazione di fenomeni finanziari. Partendo dal problema finanziario, verranno introdotte diverse procedure e metodologie comunemente utilizzate in finanza. Il corso sarà sviluppato su alcune problematiche attuali della finanza, in particolare la misurazione e l'analisi dei rischi finanziari, la misurazione del rischio sistemico e delle reti finanziarie, il trattamento di dati finanziari in alta frequenza. Data la particolare natura degli argomenti trattati, e considerando che il corso è prevalentemente applicativo, svolto quasi integralmente con attività che richiedono un utilizzo intenso dell'elaboratore, è caldamente consigliata la frequenza.

Attività di apprendimento previste e metodologie di insegnamento:

Le lezioni prevedono sia elementi teorici sia analisi empiriche. Il docente fornirà esempi di applicazione delle tecniche e metodologie che gli studenti dovranno replicare in modo autonomo o in piccoli gruppi utilizzando dati reali forniti dal docente. Gli studenti potranno utilizzare un software di loro scelta per sviluppare le analisi richieste. I risultati e le scelte di implementazione saranno oggetto di confronto e valutazione durante le ore di lezione.

Contenuti:

1. Metodi e modelli per l'analisi dei dati finanziari in alta frequenza: stima della volatilità giornaliera; individuazione dei salti nei prezzi; modelli per variabili realizzate; analisi tecnica e trading algoritmico con dati in alta frequenza. 2. Metodi e modelli per l'analisi del rischio di mercato: il Value-at-Risk con la simulazione storica, metodi parametrici o semi-parametrici; Expected Shortfall; tecniche di back-testing. 3. Metodi e modelli per l'analisi del rischio sistemico: approcci basati su modelli VAR; approcci basati su modelli per i quantili o le varianze e le correlazioni; dal rischio sistemico alle reti finanziarie.

Modalità di esame:

Per studenti frequentanti: 30% sulla base della partecipazione alle esercitazioni ed altre attività svolte durante il corso; 70% con discussione durante le sessioni d'esame di un report o di una presentazione o di analisi empiriche sviluppati su un argomento assegnato dal docente a partire da un articolo scientifico o da un working paper o da un capitolo di libro. Per studenti non frequentanti: 50% prova pratica da svolgere a casa; 50% prova scritta.

Criteri di valutazione:

La valutazione della preparazione dello studente si baserà sulla comprensione degli argomenti svolti, sull'acquisizione dei concetti e delle metodologie proposte e sulla capacità di applicarli e di implementarli in modo autonomo e consapevole.

Testi di riferimento:

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio:

Trattandosi di un corso composto di vari moduli non è possibile indicare un solo testo. All'inizio di ogni modulo verranno forniti riferimenti bibliografici (in inglese) ed il materiale delle lezioni (lucidi e dati).

APPRENDIMENTO AUTOMATICO

Titolare: Prof. FABIO AIOLLI

Mutuato da: Laurea magistrale in Computer Science (Ord. 2021)

Periodo: I anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 40A+8L; 6,00

Prerequisiti:

È opportuno avere familiarità con le conoscenze matematiche relative al Calcolo delle Probabilità e all'Analisi di funzioni multivariate. Inoltre è consigliabile avere conoscenze di base relative alla Programmazione e all'Intelligenza Artificiale. L'insegnamento non prevede propedeuticità.

Conoscenze e abilità da acquisire:

In questo insegnamento si presentano alcuni dei concetti fondamentali che caratterizzano l'Apprendimento Automatico, cioè quella classe di tecniche ed algoritmi che a partire da dati empirici permettono di acquisire nuova conoscenza, oppure di correggere e/o raffinare conoscenza già disponibile. Tali tecniche sono particolarmente utili per problemi per cui è impossibile o molto difficile pervenire ad una formalizzazione utilizzabile per la definizione di una soluzione algoritmica ad-hoc. Esempi di tali problemi sono compiti percettivi, come il riconoscimento visivo di cifre manoscritte, e problemi in cui i dati sono corrotti dal rumore e/o incompleti. L'insegnamento tratta principalmente metodi numerici. Sono previste esercitazioni in laboratorio informatico che consentono allo studente di sperimentare le conoscenze acquisite mediante l'applicazione a piccoli esempi pratici.

Attività di apprendimento previste e metodologie di insegnamento:

L'insegnamento prevede lezioni frontali ed esercitazioni in laboratorio informatico. Le esercitazioni in laboratorio informatico consistono nella sperimentazione da parte degli studenti delle tecniche viste a lezione sotto vari scenari operativi. In questo modo gli studenti possono verificare sperimentalmente i concetti appresi e acquisire sia capacità di applicazione dei concetti appresi che di giudizio critico.

Contenuti:

La struttura e le tematiche dell'insegnamento saranno le seguenti: - Introduzione: Quando Applicare le Tecniche Proprie dell'Apprendimento Automatico; Paradigmi di Apprendimento Automatico; Gli ingredienti Fondamentali dell'Apprendimento Automatico. - Apprendimento supervisionato: Complessità dello Spazio delle Ipotesi; Misure di Complessità; Esempi di Algoritmi di Apprendimento Supervisionato; - Alberi di Decisione: Apprendimento di Alberi di Decisione; Trattamento di Dati Numerici, di Dati Mancanti, di Costi; Tecniche di Pruning e Derivazione di Regole di Decisione. - Apprendimento Probabilistico: Apprendimento Bayesiano; Esempi di Applicazione al Paradigma Supervisionato e al Paradigma Non-Supervisionato (clustering); Classificatore Ottimo di Bayes; EM. - Reti Neurali e Support Vector Machines: Cenni di Reti Neurali; Margine di Classificazione; Support Vector Machines per Classificazione e Regressione; Funzioni Kernel. - Aspetti Applicativi: Pipeline di Classificazione; Rappresentazione dei dati e Selezione di Variabili;

Model Selection; Clustering; Comitati; Sistemi di Raccomandazione.

Modalità di esame:

Lo studente deve superare un esame scritto e, se ritenuto necessario dal docente, un esame orale.

Criteri di valutazione:

Il testo dell'esame scritto contiene alcune domande che consentono di valutare il livello di apprendimento delle nozioni impartite durante l'insegnamento e la capacità dello studente nell'analizzarle criticamente. Sono poi presenti domande in cui si richiede allo studente di mostrare di aver compreso gli aspetti applicativi trattati all'interno delle attività svolte in laboratorio informatico. Tali domande hanno lo scopo di valutare se lo studente ha sviluppato la capacità di applicare le nozioni apprese durante l'insegnamento. Nel caso in cui la valutazione dello scritto non risulti soddisfacente per lo studente, il docente può integrare l'esame scritto con un esame orale per meglio verificare la preparazione dello studente.

Testi di riferimento:

Alpaydin, Ethem, Introduction to machine learning. Cambridge: The MIT press, 2010 Mitchell, Tom M., Machine learning. New York: McGraw-Hill, 1998

Eventuali indicazioni sui materiali di studio:

Vengono rese disponibili, come riferimento, i lucidi utilizzati a lezione.

BIOSTATISTICA COMPUTAZIONALE E BIOINFORMATICA

Titolare: Prof. DAVIDE RISSO

Periodo: Il anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 42A+22L; 9,00

Prerequisiti:

Statistica progredito, Calcolo delle Probabilità e Modelli Statistici II.

Conoscenze e abilità da acquisire:

Introduzione ai problemi biologici legati alla complessità dei dati provenienti dalle nuove tecniche di sequenziamento. Introduzione ai modelli statistici per dati genomici e trascrittomici. Capacità di affrontare l'analisi completa dei dati: dal dato grezzo all'interpretazione del risultato. Capacità di scrivere una breve tesina su un dataset opportunamente assegnato dal docente.

Attività di apprendimento previste e metodologie di insegnamento:

Lezioni frontali e laboratori informatici

Contenuti:

Il completamento del progetto genoma umano e con esso l'inizio di una serie di progetti di sequenziamento sistematico di molti organismi complessi ha aumentato enormemente la quantità di informazioni disponibili riguardanti sequenze geniche e proteiche. Questa grande disponibilità di dati biologici ha quindi rivoluzionato e rivoluzionerà ulteriormente la ricerca genetica e la comprensione di molti aspetti biologici quali la regolazione genica, l'interazione fra proteine e l'attivazione e la soppressione di vie metaboliche. In questo contesto quindi, la quantità di dati congiuntamente alla natura complessa degli stessi hanno reso l'analisi statistica un passo obbligato per la loro comprensione. Il corso tratterà i seguenti argomenti: - Introduzione alla Genomica, Trascrittomica e Epigenomica. - Allineamento di sequenze. Algoritmi di allineamento, allineamenti globali e locali. Applicazione alla quantificazione dell'espressione dell'RNA. - Analisi di dati di espressione derivanti da esperimenti di RNA-seq. Normalizzazione dei dati, metodi globali e locali (lowess), trasformazioni per la stabilizzazione della varianza. Applicazione di analisi cluster e analisi discriminante. Verifica d'ipotesi per l'identificazione di geni differenzialmente espressi, test moderati, approcci permutazionali. Problema dei confronti multipli, controllo del False Discovery Rate (FDR). Metodi di classificazione, Gene set Analysis. - Introduzione all'analisi di altri dati genomici, quali sequenziamento del DNA, accessibilità della cromatina, interazione proteine-RNA (immunoprecipitazione).

Modalità di esame:

Prova scritta

Criteri di valutazione:

Saranno criteri di valutazione: la capacità espositiva della tesina, la congruenza dei metodi usati per l'analisi dei dati assegnati, e la completezza delle risposte nell'esame scritto. Ulteriori criteri saranno l'analisi critica dei risultati e l'indipendenza nell'affrontare i temi proposti.

Testi di riferimento:

Irizarry, Rafael A., Data analysis for the life sciences with R. Boca Raton: Chapman and Hall CRC, 2017

Eventuali indicazioni sui materiali di studio:

Materiale didattico predisposto dal docente. NOTA: si tratta di informazioni di massima che potrebbero subire variazioni al momento della presa in carico dell'insegnamento da parte del docente titolare

CALCOLO DELLE PROBABILITA'

Titolare: Prof.ssa ALESSANDRA BIANCHI

Periodo: I anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 50A+32E; 9,00

Prerequisiti:

Solide basi di Analisi Matematica e Algebra lineare.

Conoscenze e abilità da acquisire:

Al termine del corso, lo studente: - possiede nozioni approfondite della teoria classica del Calcolo delle Probabilità, incluso la convergenza di variabili

aleatorie e i principali teoremi limite; - è in grado di risolvere autonomamente problemi di probabilità relativi alle applicazioni, ed in particolare alle applicazioni statistiche.

Attività di apprendimento previste e metodologie di insegnamento:

82 ore di lezioni frontali, delle quali 56 di teoria e 26 di esercitazioni

Contenuti:

Richiami: somme infinite, operazioni insiemistiche, convergenza in spazi metrici. Statistica descrittiva ed esempi di motivazione. Modelli probabilistici discreti: spazi di probabilità discreti, densità discreta; spazi uniformi e calcolo combinatorio; probabilità condizionale ed indipendenza di eventi; distribuzioni discrete notevoli (di Bernoulli, binomiale, ipergeometrica, di Poisson, geometrica, binomiale negativa); variabili aleatorie discrete e loro leggi; valor medio e sue proprietà. Spazi di probabilità generali e sigma-algebre: generatori e sigma-algebra dei boreliani, teorema di unicità per misure; indipendenza di sigma-algebre e lemma di Borel-Cantelli. Costruzione di misure, misura prodotto, funzioni di ripartizione, misura di Lebesgue. Variabili aleatorie generali e loro leggi. Variabili aleatorie reali, valor medio e sue proprietà; integrale di Lebesgue e teorema di trasformazione. Distribuzioni assolutamente continue, esempi (uniforme continua, di Cauchy, esponenziale, Gamma, beta, gaussiana); trasformazioni di variabili aleatorie e loro leggi. Statistiche d'ordine. Leggi condizionali. Concetti di convergenza per variabili aleatorie. Legge debole e legge forte dei grandi numeri; applicazioni. Funzioni caratteristiche, distribuzioni gaussiane multivariate e teorema del limite centrale.

Modalità di esame:

Prova scritta

Criteri di valutazione:

Esame finale (prova scritta).

Testi di riferimento:

Pra, Paolo Dai; Caravenna, Francesco, Probabilità: Un'introduzione attraverso modelli e applicazioni. : Springer, 20130710 Klenke, Achim., Probability Theory [electronic resource] : A Comprehensive Course / by Achim Klenke.. Cham: Springer International Publishing, 2020 Ross, Sheldon M., Probabilità e statistica per l'ingegneria e le scienze. Santarcangelo di Romagna: Maggioli, 2014 Resnick, Sidney I., >probability path. Boston [etc: Birkhauser, 1999

Eventuali indicazioni sui materiali di studio:

Materiali di studio saranno resi disponibili attraverso la piattaforma moodle. NOTA: si tratta di informazioni di massima che potrebbero subire variazioni al momento della presa in carico dell'insegnamento da parte del docente titolare

DATI MULTI-FONTE E ANALISI TERRITORIALI

Titolare: Prof. MARCO TOSI

Periodo: Il anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 32A+32L; 9,00

Prerequisiti:

-

Conoscenze e abilità da acquisire:

Gli studenti acquisiranno le competenze necessarie per costruire e gestire basi dati multi-fonte, integrando fonti "tradizionali" (es. censimenti, indagini campionarie) con quelle "moderne" (es. web, social network). Gli studenti saranno in grado di gestire, analizzare e rappresentare i dati territoriali, grazie ai Sistemi Informativi Geografici (GIS).

Attività di apprendimento previste e metodologie di insegnamento:

Insegnamento sarà sviluppato mediante un'intensa attività laboratoriale, in cui gli studenti utilizzeranno oltre ai programmi statistici conosciuti, il software QGIS per gestire ed analizzare dati da fonti diverse.

Contenuti:

Il corso è diviso in tre moduli: Il primo modulo riguarda la creazione e gestione di basi dati multi-fonte tramite procedure di (1) record linkage, (2) di identificazione e correzione degli errori, e (3) di post-stratificazione. Verranno esposte le tecniche deterministiche e non-deterministiche di linkage, i metodi di imputazione dei valori mancanti, e le tecniche di ponderazione e creazione dei pesi. Le aree di applicazione di questi metodi sono molteplici, quali, ad esempio, il settore sanitario (es. Oxford Record Linkage Study), la sicurezza nazionale (es. Multistate Anti-terrorism System), e prevenzione del crimine (es. frodi dei conti correnti). Il modulo di Analisi Territoriale prevede l'utilizzo dei sistemi informativi geografici (GIS) per la gestione e l'analisi di dati spaziali affrontando le principali problematiche della gestione del dato: dalle differenze dei modelli di dati raster e vettoriali, sistemi di riferimento, regole cartografiche e algoritmi di analisi. Il terzo modulo riguarda l'uso dei media digitali a scopi di ricerca in ambito demografico e sociale. Si vedrà come diversi tipi di dati digitali sono generati e quali potenzialità e limiti presentano in diversi ambiti di ricerca, con particolare riferimento all'indagine di fenomeni che includano la dimensione geografica e territoriale. Saranno discussi diversi casi di studio e saranno svolte delle esercitazioni in cui i dati digitali verranno elaborati attraverso l'uso di GIS.

Modalità di esame:

Saranno valutate e rappresenteranno voto d'esame diverse attività svolte durante le lezioni, quali esercitazioni e discussione di diversi case studies. Per la parte di GIS sarà valutata la produzione di una cartografia tematica.

Criteri di valutazione:

Partecipazione attiva alle lezioni e alle attività laboratoriali. Capacità di gestire dati multi-fonte. Capacità di produrre cartografie tematiche.

Testi di riferimento:

Christen, Peter, Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection. : Springer, 2012 Dempster AP, Laird NM and Rubin DB, Maximum likelihood from incomplete data via the EM algorithm. : Journal of the Royal Statistical Society, Series B, 1977 Fabrizio, Mauro, Esercizi svolti in QGIS guida pratica all'analisi vettoriale. Palermo: D. Flaccovio, 2018 Hine, C; Snee, H; Morey, Y; Roberts, S; Watson, H, Digital Methods for Social Science: An Interdisciplinary Guide to Research Innovation. : Palgrave Macmillan, 201510

Eventuali indicazioni sui materiali di studio:

I materiali (articoli, base dati, e slides) saranno condivisi sulla piattaforma Moodle.

FISICA DELLE PARTICELLE: FONDAMENTI, STRUMENTI, E METODI DI ANALISI

Titolare: Dott. TOMMASO DORIGO

Periodo: I anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00

Prerequisiti:

L'ammissione alla LM in Scienze statistiche permette di avere gli strumenti matematici per l'insegnamento

Conoscenze e abilità da acquisire:

Generale comprensione dello stato della ricerca in fisica delle particelle, dei modelli teorici alla base della nostra comprensione della materia a livello subnucleare, degli strumenti per la rivelazione delle particelle e il loro studio, e delle tecniche di analisi in uso per inferenza statistica in fisica fondamentale, incluse le metodologie statistiche per l'estrazione di intervalli di confidenza per parametri misurabili in presenza di errori sistematici, test delle ipotesi, e analisi multivariate.

Attività di apprendimento previste e metodologie di insegnamento:

Lezioni frontali, esercizi per casa e successiva discussione collettiva, esercizi con software di analisi dati al computer

Contenuti:

La nostra comprensione del mondo fisico subnucleare si basa su misure ottenute dalla collisione energetica di particelle subatomiche. Gli esperimenti al CERN di Ginevra, in particolare, studiano i processi di più alta energia mai ottenuti in moli enormi di dati, utilizzando tecniche avanzate di analisi e inferenza statistica. In questo corso verrà fornita una panoramica della nostra conoscenza di base della fisica fondamentale, senza presupporre competenze pregresse dell'argomento. Lo scopo del corso è di mettere lo studente in grado di comprendere gli strumenti, i metodi, e gli obiettivi della ricerca agli acceleratori di particelle, e aprire la possibilità di un suo futuro coinvolgimento in questa affascinante ricerca, grazie alla collaborazione con il dipartimento di Fisica e Astronomia di Padova.

Modalità di esame:

Esame orale

Criteri di valutazione:

Verifica delle competenze acquisite, dimostrazione della comprensione dei principali argomenti del corso

Testi di riferimento:

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio:

Materiale didattico predisposto dal docente NOTE: si tratta di informazioni di massima che potrebbero subire variazioni al momento della presa in carico dell'insegnamento da parte del docente titolare

GLI OBIETTIVI DELLA RICERCA CLINICA IN ONCOLOGIA

Titolare: Prof.ssa VALENTINA GUARNERI

Periodo: I anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00

Prerequisiti:

Nessuno

Conoscenze e abilità da acquisire:

La rapida evoluzione delle conoscenze sulla biologia tumorale e l'innovazione dei trattamenti sta influenzando in modo rilevante i diversi aspetti della sperimentazione clinica in ambito oncologico. Infatti, chi opera nel campo della ricerca clinica è sempre più spesso sollecitato a confrontarsi con le innovazioni della ricerca e con le metodologie relative sia alla pianificazione degli studi sia all'analisi dei dati. Sempre di più sono richieste figure con competenze statistiche/metodologiche specifiche. Il corso ha lo scopo di fornire elementi base di oncologia, di esplorare il ruolo della sperimentazione clinica in questo ambito e di come dall'avanzamento scientifico scaturisca la necessità di concepire studi clinici dal disegno innovativo. Il corso sottolineerà la necessità di continui feedback tra statistico e clinico nell'identificazione degli elementi chiave che entrano in gioco nel calcolo della dimensione campionaria, nella definizione degli obiettivi e nell'analisi dei risultati di uno studio clinico. Questo scambio è infatti di primaria importanza per condurre studi clinici che possano fornire risultati clinicamente utili.

Attività di apprendimento previste e metodologie di insegnamento:

Lezioni frontali

Contenuti:

–Nozioni generali su epidemiologia ed eziologia, basi biologiche dei tumori –Fondamenti della terapia medica e nuovi farmaci –Fattori prognostici e predittivi. –Ruolo della sperimentazione clinica in oncologia: –Prospettiva storica –Le fasi della sperimentazione clinica; ruolo degli studi di fase I in oncologia –La dimensione campionaria nell'era della medicina di precisione –Disegno di studi clinici classico e trials innovativi –La sperimentazione clinica nelle diverse fasi di malattia oncologica –Endpoints clinici e Endpoints statistici –Interpretazione dei risultati dal punto di vista del clinico: significatività statistica vs rilevanza clinica –Ruolo delle metanalisi –Biomarcatori –Validità analitica, validità clinica, utilità clinica –Integrazione di biomarcatori nel disegno di studi clinici

Modalità di esame:

Esame orale e/o scritto

Criteri di valutazione:

Correttezza e qualità delle prove d'esame.

Testi di riferimento:

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio:

Materiale didattico messo a disposizione dal docente

INTRODUCTION TO STOCHASTIC PROCESSES

Titolare: Prof. MARCO FORMENTIN

Periodo: I anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00

Prerequisiti:

Un corso base di Calcolo delle Probabilità.

Conoscenze e abilità da acquisire:

Conoscenza approfondita di modelli Markoviani a tempo discreto e tempo continuo, con capacità di risolvere autonomamente esercizi e problemi anche di livello avanzato.

Attività di apprendimento previste e metodologie di insegnamento:

64 ore di lezioni frontali (34 teoria e 30 esercitazioni)

Contenuti:

Definizione di processo stocastico. Probabilità condizionata e valore atteso condizionato. Indipendenza condizionata. Catene di Markov a tempo discreto: definizione. Matrice di transizione, leggi congiunte e proprietà di Markov. Random Walk e sue proprietà. Catene ricorrenti, positivamente ricorrenti e transienti. Criterio della matrice potenziale. Tempi di arresto e proprietà di Markov forte. Probabilità e tempo medio di assorbimento. Classificazione degli stati. Distribuzioni invarianti. Teorema di Markov. Periodicità. Teorema ergodico. Long-Run behavior. Coupling fra catene di Markov. Metodi Monte Carlo: l'algoritmo di Metropolis-Hastings. Processo di Poisson: costruzione del processo e definizioni equivalenti. Principali proprietà ed alcune importanti applicazioni. Catene di Markov a tempo continuo: definizione. Generatore infinitesimo. Equazioni di Kolmogorov. Distribuzioni invarianti. Competition theorem e rappresentazione grafica. Applicazioni.

Modalità di esame:

Esame scritto con esercizi simili a quelli svolti in classe. Potranno essere richieste enunciati e dimostrazioni di teoremi.

Criteri di valutazione:

Lo studente dovrà dimostrare di saper acquisito le conoscenze teoriche sui processi di Markov viste nel corso e di saperle applicare correttamente in esercizi sui processi stocastici di congrua difficoltà.

Testi di riferimento:

Pierre Brouaud, Markov Chains, Gibbs Fields, Monte Carlo Simulation and queues. : Springer, 1998 John Robert Norris, Markov Chains. : Cambridge, 1997

Eventuali indicazioni sui materiali di studio:

Tutti gli argomenti d'esame vengono illustrati a lezione. Materiale addizionale (esercizi, appunti del docente) saranno disponibile su Moodle.

INTRODUZIONE ALLA BIOLOGIA

Titolare: Prof. LIBERO VITIELLO

Periodo: I anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00

Prerequisiti:

Nessuno

Conoscenze e abilità da acquisire:

Obiettivo del corso è trasmettere le conoscenze di base di biologia e genetica necessarie ai laureati in scienze statistiche per poter in seguito applicare le loro competenze a ricerche nel campo delle scienze della vita.

Attività di apprendimento previste e metodologie di insegnamento:

Lezioni frontali, a fine corso una o più lezioni seminariali su uso della statistica in vari campi della biologia (a seconda degli interessi espressi dagli studenti durante il corso). A seconda degli interessi manifestati dagli studenti durante il corso, nelle ultime lezioni verranno inoltre scelti e discussi in aula uno o più esempi di lavori scientifici.

Contenuti:

INTRODUZIONE Atomi e molecole: cenni sulla chimica dei viventi L'acqua come solvente delle reazioni biologiche Le principali classi di molecole biologiche La teoria cellulare L'organizzazione dei viventi, cenni di sistematica Le caratteristiche principali di batteri, virus e cellule eucarioti. STRUTTURA E FUNZIONE DELLA CELLULA La membrana plasmatica, proprietà e funzioni Gli apparati membranosi I mitocondri, struttura in rapporto alla funzione; il metabolismo energetico Il citoscheletro Il compartimento nucleare I processi di endocitosi e secrezione La trasduzione del segnale recettoriale La divisione

cellulare BASI MOLECOLARI DELL'INFORMAZIONE EREDITARIA Composizione e struttura chimica del DNA e degli RNA Il codice genetico e sue proprietà La replicazione del DNA L'organizzazione del genoma negli eucarioti e nei procarioti I cromosomi umani Trascrizione e maturazione del RNA La regolazione dell'espressione genica La sintesi proteica GENETICA E GENOMICA UMANA I differenti tipi di trasmissione dei caratteri ereditari (AD, AR, XD, XR, eccezioni alla trasmissione mendeliana). I caratteri quantitativi e multifattoriali; la variabilità genetica Le conseguenze patologiche delle mutazioni Le principali tecniche di analisi del DNA (PCR, sequenziamento) Il progetto genoma umano e l'era della genomica Definizione dei diversi progetti "omici" (trascrittoma, proteoma, metaboloma)

Modalità di esame:

Esame scritto, a risposta multipla (con possibilità di dover aggiungere brevi risposte per esteso)

Criteri di valutazione:

Alla fine del corso gli studenti dovranno essere in grado di: Descrivere le basi chimico-fisiche che regolano il funzionamento delle macromolecole biologiche; Descrivere le caratteristiche generali delle macromolecole biologiche (carboidrati, lipidi, proteine, acidi nucleici); Descrivere le caratteristiche principali dei vari tipi di organismi viventi e le relazioni evolutive che li collegano; Illustrare l'organizzazione della cellula ed in particolare riconoscere le diverse strutture cellulari mettendole in relazione con la loro funzione. Illustrare la funzione della cellula e descriverne i seguenti processi fondamentali (replicazione, trascrizione e traduzione del materiale genetico; mitosi e meiosi); Descrivere i vari modi nei quali l'informazione genetica si riflette sulle caratteristiche dei singoli individui e sulla insorgenza di patologie. Descrivere le differenze tra analisi genetica ed analisi genomica ed avere familiarità con le metodologie di base di entrambe.

Testi di riferimento:

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio:

I file delle lezioni ed eventuali altri materiali didattici utilizzati in aula saranno messi a disposizione degli studenti tramite moodle. Il corso non ha un testo di riferimento specifico; durante la prima lezione saranno illustrate le varie possibilità tra le quali scegliere.

MARKETING PROGREDITO

Titolare: Prof. ROBERTO GRANDINETTI

Periodo: I anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00

Prerequisiti:

Conoscenze di base di Economia Aziendale e Marketing, acquisite nei relativi insegnamenti. Strumentazione basilare di analisi statistica acquisita nel corso della triennale.

Conoscenze e abilità da acquisire:

Approfondimento di alcune tematiche relative alla strategia dell'impresa, al suo approccio al mercato, ai processi di innovazione, alle forme dell'internazionalizzazione, alla crescita aziendale e alla gestione delle relazioni inter-organizzative, al rapporto tra impresa e territorio. Sviluppo di un caso aziendale e confronti con uno o più campioni di riferimento utilizzando la banca dati AIDA. Capacità di lavorare in gruppo. Capacità di comunicare in pubblico.

Attività di apprendimento previste e metodologie di insegnamento:

Lezioni frontali del docente. Testimonianze aziendali. Presentazione dei case studies sviluppati dai gruppi e loro discussione. Lavori di gruppo. Laboratorio per l'utilizzo della banca dati AIDA.

Contenuti:

Impresa, strategia e dati di bilancio. I processi di innovazione nelle imprese. I knowledge-intensive business services. L'internazionalizzazione dell'impresa e le sue forme. Globalizzazione e global value chains. I processi di crescita aziendale. La gestione delle relazioni inter-organizzative come leva di vantaggio competitivo. La natalità aziendale. I distretti industriali, loro caratteristiche ed evoluzione nel tempo. Le imprese all'interno dei distretti industriali. La banca-dati AIDA.

Modalità di esame:

--

Criteri di valutazione:

Il voto ottenuto dalla valutazione dell'elaborato collettivo viene pesato con il voto conseguito nella prova scritta individuale.

Testi di riferimento:

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio:

Slides relative alle lezioni frontali, articoli e documenti, resi disponibili nel sito (Moodle) all'inizio e durante il corso. NOTA: si tratta di informazioni di massima che potrebbero subire variazioni al momento della presa in carico dell'insegnamento da parte del docente titolare

METODI E MODELLI STATISTICI PER LA FINANZA

Titolare: Prof. FRANCESCO LISI

Periodo: Il anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 42A+22L; 9,00

Prerequisiti:

Serie storiche economiche. Pur non essendo un prerequisito stringente, è fortemente consigliato Serie storiche finanziarie.

Conoscenze e abilità da acquisire:

Lo scopo del corso è di fornire degli strumenti avanzati ed aggiornati che consentano allo studente di stimare ed utilizzare modelli - anche non standard - che tengano conto delle principali caratteristiche delle serie storiche finanziarie. La presentazione delle tecniche e dei modelli appropriati sarà illustrata tramite l'uso di serie reali. I pacchetti software utilizzati saranno R ed S+Finmetrics.

Attività di apprendimento previste e metodologie di insegnamento:

Tutte le metodologie proposte verranno implementate con un opportuno software e applicate a dati reali durante le esercitazioni in aula computer.

Contenuti:

Programma: - Introduzione: richiami alle principali caratteristiche delle serie finanziarie e ai modelli che le descrivono. - La stima dei modelli della classe GARCH: verosimiglianza dei modelli garch, stime MLE, stime QML. - Costruzione di un software per la stima di un modello GARCH. - Modelli multivariati per l'analisi e la previsione della volatilità. - Modelli GARCH multivariati: la funzione di autocorrelazione incrociata, problematiche generali, il modello VECH, il modello VECH diagonale, il modello BEKK, il modello CCC, il modello DCC, il modello PC-GARCH. - Dati ad alta frequenza: introduzione e principali caratteristiche. - Modelli di decomposizione per le variazioni di prezzo (ADS); Modelli per l'analisi e la previsione delle durate (modelli ACD). - Modelli a volatilità stocastica

Modalità di esame:

Prova scritta + esercitazione per casa

Criteri di valutazione:

La valutazione della preparazione dello studente si baserà sulla comprensione degli argomenti svolti sull'acquisizione dei concetti e delle metodologie proposte e sulla capacità di applicarli e di implementarli in modo autonomo e consapevole.

Testi di riferimento:

Tsay R., Analysis of Financial Time Series. : Wiley, 2010

Eventuali indicazioni sui materiali di studio:

Lucidi delle lezioni che verranno forniti di volta in volta prima della lezione stessa.

METODI INFORMATICI PER LA STATISTICA E IL DATA SCIENCE

Titolare: Prof. EMANUELE DI BUCCIO

Periodo: I anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00

Prerequisiti:

I prerequisiti sono relativamente semplici, ma necessari: elementi di strutture di dati (variabile, file, vettore, matrice), algoritmi elementari, sistemi di elaborazione, e sistemi di gestione delle basi di dati. La conoscenza di un linguaggio di programmazione è utile, ma non strettamente necessaria.

Conoscenze e abilità da acquisire:

S'intende fornire le conoscenze concrete di metodi e strumenti informatici affinché uno studente posseda maggiore competenza in Statistica di un informatico e maggiore competenza in Informatica di uno statistico. Particolare enfasi sarà posta su programmazione e gestione dei dati e sul superamento del modo di scrivere software indotto da linguaggi come R e da pacchetti di software preconfezionato.

Attività di apprendimento previste e metodologie di insegnamento:

I contenuti saranno trattati in forma prevalentemente laboratoriale mediante lo sviluppo di programmi e l'utilizzo di librerie software in Python. Gli elementi metodologici saranno introdotti allo scopo di conoscere le problematiche sottostanti, di progettare e realizzare progetti, e di utilizzare gli strumenti in modo consapevole.

Contenuti:

1. Introduzione a Python: ambiente, costrutti, primi esempi. 2. Raccolta, organizzazione e gestione di grandi masse di dati: pattern matching, parsing. 3. Strutture di dati fondamentali: liste, hash, grafi, alberi. 4. Algoritmi fondamentali: ricorsione, ricerca, ordinamento. 5. Architetture distribuite con MapReduce. 6. Sistemi di raccomandazione 7. Decomposizione e riduzione della dimensionalità. 8. Rappresentazione e reperimento. 9. Reti, link e click-through: WWW, Link Analysis, Pagerank.

Modalità di esame:

L'esame si potrà svolgere in due modalità che prevedono entrambe l'esecuzione di un mini-progetto ed un esame orale che include la discussione del mini-progetto. Il mini-progetto è un progetto di una applicazione di data science. Durante la discussione del mini-progetto gli studenti dovranno essere in grado di spiegare le problematiche, le metodologie, gli strumenti e i risultati ottenuti. La tematica deve essere scelta tra le seguenti: - Flussi di dati (Stream Processing) - Insiemi frequenti (Frequent Itemset) - Raggruppamento sovrapposto - Raggruppamento di immagini / musica / film - Raccomandazione di immagini / musica / film - Percettori e SVM - Analisi spettrale dei grafi - Simrank - Motore di ricerca di argomenti / tematiche - Decomposizione CUR - PageRank sensibile agli argomenti - Link Spam e Link Farm - Pubblicità sul Web Le due modalità sono: [Modalità A] Mini-progetto di gruppo con consegna di una relazione scritta, software, e presentazione finale di gruppo in presenza degli altri gruppi; la presentazione dovrà coinvolgere tutti i membri del gruppo. Questa modalità prevede la consegna di alcuni documenti durante il corso e una presentazione intermedia. Dopo la presentazione finale, verrà somministrato un questionario con domande sul proprio mini-progetto e su quello degli altri gruppi che hanno presentato. Nel caso in cui le presentazioni finali non si svolgessero in presenza o il numero di gruppi non consentisse di effettuare il questionario, a ciascun membro del gruppo saranno poste delle domande relative alle lezioni svolte in laboratorio. [Modalità B] Mini-progetto individuale con consegna della relazione e del software via moodle alcuni giorni prima dell'esame. L'esame orale consisterà nella discussione individuale del mini-progetto e potrà includere domande relative al programma svolto a lezione; le domande potranno includere la stesura o la modifica di codice Python.

Criteri di valutazione:

Si valuterà la comprensione delle problematiche e la capacità di trovare e di progettare soluzioni automatizzate per l'organizzazione, la gestione e l'analisi di dati allo scopo di svolgere i compiti illustrati nei contenuti e previsti dal progetto della prova orale.

Testi di riferimento:

Aho, Alfred, V.;Ullman, Jeffrey D., Foundations of Computer Science. : W H Freeman & Co, 1994 Leskovec, Jurij; Rajaraman, Anand; Ullman, Jeffrey D., Mining of Massive Datasets. USA: Cambridge University Press, 2020

Eventuali indicazioni sui materiali di studio:

Materiale didattico (e.g. slides e codice) sarà distribuito durante le lezioni in aggiunta ai testi di riferimento. Alcuni testi, specialmente quelli per la programmazione e la gestione dei dati, saranno indicati all'inizio delle lezioni.

METODI STATISTICI PER DATI CON ELEVATA DIMENSIONALITÀ

Titolare: Prof.ssa MANUELA CATTELAN

Periodo: Il anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 44A+20L; 9,00

Prerequisiti:

Calcolo delle Probabilità Statistica Progredito Analisi dei Dati (Data Mining)

Conoscenze e abilità da acquisire:

Lo studente apprenderà alcuni concetti fondamentali dell'analisi statistica con data sets di grandi dimensioni e saprà applicare diverse metodologie statistiche a tali data sets.

Attività di apprendimento previste e metodologie di insegnamento:

Lezioni frontali in aula e lezioni in laboratorio informatico.

Contenuti:

Nel corso verranno affrontate le seguenti tematiche: - verifica di ipotesi su larga scala e test multipli: Bonferroni, FWER e FDR e loro controllo, metodi bayesiani empirici e FDR locale; - modelli di regressione per dati con elevata dimensionalità: a) dati con numerosità elevata: algoritmi incrementali a memoria limitata, discesa stocastica del gradiente, inferenza; b) dati con un numero elevato di variabili: sparsità, penalizzazioni che inducono sparsità, Lasso e sue generalizzazioni ed estensioni; - stima di matrici di varianza e covarianza ad elevata dimensionalità: a) modelli fattoriali, modelli che inducono sparsità nella matrice di precisione/concentrazione; b) grafi di concentrazione, algoritmi di apprendimento basati su calcoli locali.

Modalità di esame:

Esame scritto con quesiti inerenti sia la parte teorica sia la parte pratica.

Criteri di valutazione:

La valutazione sarà basata sulla capacità di rispondere correttamente e con completezza ai quesiti posti e sulla capacità di applicare le diverse metodologie di analisi apprese durante il corso.

Testi di riferimento:

Bühlmann, Peter; van de Geer, Sara, Statistics for High-Dimensional Data : Methods, Theory and Applications. : Springer Berlin Heidelberg, 2011 Efron, Bradley, Large-scale inference empirical Bayes methods for estimation, testing, and prediction. Cambridge: Cambridge University Press, 2013 Hastie, Trevor; Tibshirani, Robert; Wainwright, Martin, Statistical Learning with Sparsity. Philadelphia, PA: Chapman and Hall/CRC, 2015

Eventuali indicazioni sui materiali di studio:

Materiale fornito dal docente e libri indicati all'inizio del corso.

METODI STATISTICI PER DATI FUNZIONALI

Titolare: Prof. BRUNO SCARPA

Periodo: Il anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 44A+20L; 9,00

Prerequisiti:

--

Conoscenze e abilità da acquisire:

--

Attività di apprendimento previste e metodologie di insegnamento:

--

Contenuti:

Programma di massima che potrebbe subire qualche variazione. 1) Metodi esplorativi 2) Modelli frequentisti - liscio di funzioni e funzioni vincolate - registrazione di curve - metodi di riduzione della dimensionalità (componenti principali di funzioni e relative regolarizzazioni) - modelli con esplicative funzionali - modelli per risposta funzionale ed esplicative multivariate - modelli per risposta funzionale ed esplicative funzionali e modelli concorrenti - modelli per equazioni differenziali 3) Modelli bayesiani - modelli parametrici - modelli non parametrici (BNP)

Modalità di esame:

--

Criteri di valutazione:

--

Testi di riferimento:

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio:

--

METODI STATISTICI PER IL MARKETING**Titolare:** Prof. OMAR PACCAGNELLA**Periodo:** Il anno, 1 semestre**Indirizzo formativo:** Corsi comuni**Tipologie didattiche:** 58A+6L; 9,00**Prerequisiti:**

Calcolo delle probabilità, Statistica progredito

Conoscenze e abilità da acquisire:

Il corso introduce metodi e modelli statistici a supporto delle decisioni di marketing

Attività di apprendimento previste e metodologie di insegnamento:

Le lezioni sono frontali tenute dal docente, eventualmente integrate da esercitazioni in aula informatica. Sono previste testimonianze aziendali di esperti provenienti dal mondo del lavoro.

Contenuti:

1. Analisi dei comportamenti di acquisto: la fedeltà alla marca
2. La misura di fenomeni non direttamente osservabili rilevanti nel marketing: la misura della soddisfazione del consumatore
3. Misure dell'efficacia della comunicazione pubblicitaria
4. La segmentazione del mercato: tecniche e applicazioni
5. Il posizionamento di prodotti e marche: strategie e tecniche
Studi di caso

Modalità di esame:

Homework più esame finale orale

Criteri di valutazione:

Si vuole valutare se lo studente abbia appreso i concetti teorici introdotti durante le lezioni e sia in grado di rispondere, utilizzando opportuni metodi e modelli statistici, a domande che emergono nel disegnare strategie di marketing.

Testi di riferimento:

Hair J.F., Black W.C., Babin B.J., Anderson R.E., Multivariate data analysis. : Pearson, 2010

Eventuali indicazioni sui materiali di studio:

Materiale distribuito durante il corso. East R. "Comportamento del consumatore", Apogeo, 2003. Fabbris L. "Statistica multivariata", McGraw-Hill, Milano, 1997. Molteni L., Troilo G. "Ricerche di marketing", McGraw-Hill, Milano, 2003. Leeflang P.S.H., Wittink D.R., Wedel M., Naert P.A. Building Models for Marketing Decisions, Boston, Kluwer Academic Publishers, 2000. Wedel M., Kamakura W.A. Market Segmentation, Boston, Kluwer Academic Publishers, 2000. Chakrapani C. Statistics in Marketing Research, Londra, Arnold Publishers, 2004. Hanssens D.M., Parsons L.J., Schultz R.L. Market Response Models, Boston, Kluwer Academic Publishers, 2001. Franses P.H., Paap R. Quantitative Models in Marketing Research, Cambridge, Cambridge University Press, 2001. Eliashberg J., Lilien G.L. (a cura di) (1993) Handbooks in Operations Research and Management Science. Marketing, Vol. 5, North Holland. Brasini S., Freo M., Tassinari F., Tassinari G. Marketing e pubblicità. Bologna: Il Mulino, 2010. Dolnicar, S., Grün, B., Leisch, F. Market segmentation analysis: Understanding it, doing it, and making it useful (p. 324). Springer Nature, 2018.

METODI STATISTICI PER L'ANALISI DEI DATI GERARCHICI**Titolare:** Prof. LUIGI GROSSI**Periodo:** Il anno, 1 semestre**Indirizzo formativo:** Corsi comuni**Tipologie didattiche:** 42A+22L; 9,00**Prerequisiti:**

L'insegnamento di "Metodi statistici per l'analisi di dati gerarchici" ha come prerequisito l'insegnamento di "Calcolo delle Probabilità" e "Statistica Progredito".

Conoscenze e abilità da acquisire:

Il corso si propone di introdurre metodologie e tecniche avanzate per il trattamento, l'interpretazione e l'analisi di dati aventi una struttura gerarchica o longitudinale.

Attività di apprendimento previste e metodologie di insegnamento:

Il corso ha una connotazione pratica molto importante, per cui viene richiesta una partecipazione attiva da parte dello studente. L'approccio didattico prevede infatti attività di lavoro, di discussione e di approfondimento sui temi del corso, sia in autonomia sia in gruppo.

Contenuti:

Introduzione: - Richiami dei principali concetti e strumenti per lo studio della variabilità intra e tra gruppi. I modelli multilivello. Approfondimenti pratici: - Definizione e caratterizzazione della struttura gerarchica in funzione dell'obiettivo dell'analisi. - Centrazione delle variabili in un contesto multilivello (definizioni, effetti, criticità). Approfondimenti metodologici: - La stima dei parametri nei modelli multilivello e gli effetti delle numerosità campionarie dei diversi livelli sull'accuratezza delle stime. - La struttura di covarianza dei dati nei modelli multilivello. - Estensione in ambito multilivello di modelli e metodi statistici classici, quali GLM e modelli a classi latenti. Tutte le diverse sezioni prevedono la visione e la discussione collegiale di applicazioni per favorire la comprensione e l'approfondimento dei vari temi trattati.

Modalità di esame:

La prova d'esame si articola in due parti: - lavori di gruppo su casi-studio - discussione orale finale

Criteri di valutazione:

La valutazione della preparazione dello studente si baserà sulla comprensione e la padronanza degli argomenti trattati e sulla capacità di applicare le soluzioni più appropriate per lo studio del dataset d'interesse. Assumeranno una particolare rilevanza nel processo di valutazione il grado di autonomia nella gestione del problema e lo spirito critico sia nell'applicazione delle soluzioni più idonee, sia nell'interpretazione dei risultati raggiunti.

Testi di riferimento:

Paul Roback, Julie Legler, Beyond Multiple Linear Regression. Applied Generalized Linear Models And Multilevel Models in R. Boca Raton (FL, USA): Chapman and Hall/CRC, 2021

Eventuali indicazioni sui materiali di studio:

Materiale didattico predisposto dal docente Testi di riferimento Roback P., Legler J. (2021), Beyond Multiple Linear Regression Applied Generalized Linear Models And Multilevel Models in R, 1st Edition, Chapman and Hall. Bartholomew, D. J., and Knott, M., and Moustaki, I. (2011) Latent variable models and factor analysis: a unified approach, 3rd ed., John Wiley & Sons, London, UK. ISBN 9780470971925. Hox J. (2010). Multilevel analysis: techniques and applications, 2a ed., New York: Routledge. Muthèn B (2004) Latent variable analysis: growth mixture modeling and related techniques for longitudinal data. In: Kaplan D (ed) Handbook of quantitative methodology for the social sciences. Sage, Newbury Park, pp 345–368 Skrondal A. e Rabe-Hesketh, S. (2004). Generalized latent variable modeling: Multilevel, longitudinal and structural equation models, Boca Raton, FL: Chapman & Hall/CRC. Snijders T. e R. Bosker (2012). Multilevel analysis. An introduction to basic and advanced multilevel modelling, 2a ed., Los Angeles: Sage Publications. Vermunt, J.K. (2003). Multilevel latent class models. Sociological Methodology, 33, 213-239. Vermunt, J.K., Tran, B., and Magidson, J. (2008). Latent class models in longitudinal research. In: Menard, S. (ed.), Handbook of Longitudinal Research: Design, Measurement, and Analysis, pp. 373-385. Burlington, MA: Elsevier.

METODI STATISTICI PER L'INFERENZA CAUSALE

Titolare: Prof. ADRIANO PAGGIARO

Periodo: Il anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 42A+22L; 9,00

Prerequisiti:

Il corso prevede le conoscenze di: Statistica 2, Modelli Statistici 1

Conoscenze e abilità da acquisire:

La necessità di inferire nessi causali da osservazioni empiriche è sempre più sentita in diversi campi applicativi, fra i quali l'epidemiologia, la biostatistica, l'economia e le scienze sociali. Il corso presenta diversi approcci e metodi per l'inferenza causale sia dal punto di vista metodologico che con numerosi studi di caso su dati reali. L'obiettivo è consentire allo studente di essere in grado di: 1. Definire nel dettaglio una domanda di ricerca riguardante una relazione causale fra variabili di interesse, alla quale si possa rispondere con un approccio empirico. 2. Individuare di volta in volta il metodo più adatto alla soluzione di un problema concreto e comprenderne le assunzioni necessarie per poter rispondere alla domanda causale di interesse. 3. Utilizzare il metodo in modo appropriato con l'ausilio del software Stata. 4. Interpretare i risultati in modo corretto alla luce della domanda iniziale, delle assunzioni fatte e dei dati disponibili.

Attività di apprendimento previste e metodologie di insegnamento:

Il corso è organizzato in lezioni frontali (42 ore), nelle quali si presentano i principali aspetti metodologici e numerose applicazioni empiriche in ambito causale, e studi di caso in aula informatica (22 ore) dove gli studenti possono applicare le conoscenze acquisite analizzando dati reali.

Contenuti:

1) Introduzione - inferenza statistica e causalità - dati sperimentali, osservazionali e quasi-sperimentali; selection bias - cenni ai diversi approcci per l'inferenza causale 2) Approccio con risultati potenziali: Rubin Causal Model - eventi fattuali e controfattuali - metodi di identificazione dell'impatto parametrici e non parametrici - selezione nelle osservabili: regressione, p-score matching, Regression Discontinuity Design - selezione nelle non osservabili: variabili strumentali, diff-in-diff, modelli per dati longitudinali 3) Modelli ad equazioni strutturali (SEM) - modelli strutturali e causalità - specificazione, identificazione e stima di modelli complessi - effetti diretti e indiretti: path analysis, mediazione e moderazione - analisi fattoriale esplorativa e confermativa - modelli strutturali con variabili latenti e/o errori di misura

Modalità di esame:

La verifica delle conoscenze e competenze acquisite avviene attraverso una prova pratica in aula informatica e una successiva discussione orale dei risultati.

Criteri di valutazione:

La valutazione della preparazione dello studente si basa su: - Comprensione di una domanda di ricerca e capacità di dare una risposta coerente su basi empiriche a partire dai dati disponibili. - Autonomia e spirito critico nello scegliere e applicare le metodologie acquisite nel corso per la soluzione di specifici casi reali.

Testi di riferimento:

Kline, Rex B., Principles and practice of structural equation modeling. New York: Guilford press, 2011 Acock, Alan C., Discovering structural equation modeling using Stata. Texas: Stata Press, 2013 Angrist, Joshua D.; Pischke, Jörn-Steffen, Mastering Metrics: The path from cause to effect. Princeton: Princeton University Press, 2015

Eventuali indicazioni sui materiali di studio:

Libri di testo, slide e dataset forniti dal docente, materiale didattico integrativo disponibile sulla piattaforma Moodle.

MODELLI STATISTICI PER DATI ECONOMICI

Titolare: Prof. MAURO BERNARDI

Periodo: Il anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 44A+20L; 9,00

Prerequisiti:

Quelli previsti dall'ordinamento del corso di laurea.

Conoscenze e abilità da acquisire:

Lo scopo del corso è di fornire allo studente la preparazione necessaria per costruire ed impiegare modelli statistici, prevalentemente dinamici, in presenza di dati di natura economica. La presentazione e lo studio delle caratteristiche principali delle varie classi di modelli saranno costantemente accompagnati dall'interpretazione economica dell'applicazione dei modelli attraverso esempi basati su dati reali.

Attività di apprendimento previste e metodologie di insegnamento:

Lezioni frontali teoriche. Utilizzo di esempi con dati reali per l'interpretazione economica dei risultati di un modello. Sessioni di esercizi analitici.

Contenuti:

Introduzione: classificazione e caratteristiche dei dati economici; principali classi di modelli utili per la modellazione di relazioni economiche. Modelli per serie storiche multivariate: modelli VAR, SVAR e VECM, specificazione, identificazione, stima e loro impiego. Modelli per dati di panel: specificazione e stima di modelli con componente d'errore a una via, modelli dinamici per dati di panel. Modelli ad equazioni simultanee: forma strutturale, forma ridotta, identificazione e stima (cenni). Modelli state-space: esempi, specificazione, stima e loro impiego (cenni).

Modalità di esame:

L'esame consiste di una prova scritta composta da tre esercizi di cui due teorici e uno pratico che consiste nello svolgimento di un caso di studio con l'uso del software R.

Criteri di valutazione:

La valutazione mira a stabilire se e in quale misura lo studente è in grado di costruire modelli adeguati per dati economici e la sua capacità di interpretarne i risultati da un punto di vista economico.

Testi di riferimento:

Baltagi, Badi H., *Econometric analysis of panel data*. Chichester: Wiley and Sons, 2008 Tsay, Ruey S., *Multivariate time series analysis with R and financial applications*. Hoboken: Wiley, 2014 Durbin, James, *Time series analysis by state space methods*. Oxford: Oxford University Press, 2012 Petris, Giovanni, *Dynamic linear models with R*. Dordrecht [etc: Springer, 2009 Lutkepohl, Helmut, *New introduction to multiple time series analysis*. Berlin: Springer, 2006 Greene, William H., *Econometric analysis*. Upper Saddle River: Prentice Hall, 2012

Eventuali indicazioni sui materiali di studio:

Testi principali di riferimento: Lutkepohl e Baltagi. Slides del docente.

MODELLI STATISTICI PER DATI SOCIALI

Titolare: Prof.ssa FAUSTA ONGARO

Periodo: l'anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 42A+22L; 9,00

Prerequisiti:

Conoscenze di contenuti impartiti nel corso di Modelli statistici 2.

Conoscenze e abilità da acquisire:

1. Conoscere e utilizzare correttamente alcune tecniche di EHA 2. Conoscere e usare correttamente modelli multilevel/gerarchici e frailty 3. Saper usare procedure di SAS per effettuare le analisi statistiche proposte nel corso

Attività di apprendimento previste e metodologie di insegnamento:

Lezioni frontali, esercitazioni in aula informatica durante le quali si userà il software SAS

Contenuti:

1. Metodi per l'analisi di dati di durata - Richiami di nozioni base per analisi dati durata: concetti base; funzioni base nel continuo e nel discreto; tipologie di metodi - Metodi non parametrici: Kaplan Meier e life table - Modelli a tempo continuo: classi di modelli; il modello semiparametrico a rischi proporzionali e sue estensioni oltre il modello base (variabili tempo dipendenti, effetti non proporzionali, rischi competitivi); modelli parametrici a rischi proporzionali e a tempi accelerati (esponenziale, weibull, esponenziale a tratti) - Modelli a tempo discreto: person period e modello a odds proporzionali (cenni). 2. Modelli frailty (nel continuo) - Introduzione al concetto di frailty - Frailty univariati: modelli semiparametrici e parametrici; metodi di stima - Shared frailty models 3. Modelli multilevel - Il modello multilevel come pooling parziale dei dati - Modelli a risposta continua: modelli a intercetta casuale; modelli a effetti casuali; inferenza per effetti fissi e per effetti casuali; - Modelli a risposta discreta: modelli logistici a effetti casuali; modelli a tempo discreto con effetti casuali; inferenza per effetti fissi ed effetti casuali.

Modalità di esame:

La prova d'esame è costituita da un test di teoria, seguito da una prova pratica al computer in aula ASID con uso di software SAS. Alla prova pratica sono ammessi gli studenti che hanno superato il test teorico.

Criteri di valutazione:

La valutazione mira a stabilire se lo studente ha appreso sia gli aspetti formali che interpretativi della modellistica proposta, nonché la sua capacità di applicarli correttamente in situazioni reali. Lo studente deve inoltre essere in grado di utilizzare correttamente il software SAS per applicare le tecniche e i metodi proposti durante il corso.

Testi di riferimento:

Singer J.D., Willet J. B., *Applied longitudinal data analysis. Modeling change and event occurrence*. Oxford: Oxford University Press, 2003 Allison P.D., *Survival analysis using SAS. A practical guide*, 2nd edition.. : SAS Institute Corporation, 2010 Snijders T.A.B., Bosker R.J., *Multilevel Analysis. An Introduction to Basic and Advanced Multilevel Modeling*. : Sage, 2012

Eventuali indicazioni sui materiali di studio:

Materiali di studio sono: i testi di riferimento, le slide delle lezioni/esercitazioni a computer, le basi di dati utilizzate per le esercitazioni a computer. Eventuali materiali aggiuntivi saranno distribuiti dal docente durante il corso. NB. Il corso prevede alcune ore di lezioni di SAS base da impartire agli studenti in concomitanza con l'inizio delle lezioni.

OTTIMIZZAZIONE STOCASTICA

Titolare: Prof. LUIGI DE GIOVANNI

Periodo: I anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 50A+14L; 9,00

Prerequisiti:

Conoscenze di base di calcolo delle probabilità

Conoscenze e abilità da acquisire:

Fornire una panoramica degli strumenti che aiutano a prendere le migliori decisioni anche quando le informazioni utili non sono completamente disponibili in modo certo ma sono date in condizioni di incertezza.

Attività di apprendimento previste e metodologie di insegnamento:

Lezioni teoriche in aula. Esercitazioni in laboratorio. Lavoro di gruppo. Le attività di apprendimento sono supportate da mezzi informatici come presentazioni elettroniche, software di ottimizzazione, fogli elettronici, software di simulazione.

Contenuti:

Il programma del corso verte sui seguenti argomenti: - Teoria delle Code - Analisi decisionale - Ottimizzazione robusta - Processi decisionali markoviani - Ottimizzazione stocastica - Revenue Management - Simulazione a eventi discreti - Simulazione di tipo continuo - Uso di software specifico per la Simulazione

Modalità di esame:

L'esame consiste in una prova scritta individuale, eventualmente integrata da una prova orale, e nell'analisi di un progetto (lavoro di squadra, tipicamente 2-3 persone) su un argomento da concordare con il docente.

Criteri di valutazione:

L'esame verifica il livello di apprendimento degli argomenti svolti e la capacità di applicarli per la soluzione di problemi decisionali in condizioni di incertezza.

Testi di riferimento:

G. Ghiani e R. Musmanno, Modelli e metodi decisionali in condizioni di incertezza e rischio. Milano: McGraw Hill, 2009

Eventuali indicazioni sui materiali di studio:

Oltre al libro di testo, ulteriore materiale sarà messo a disposizione nel sito dedicato al corso.

PROVA FINALE

Titolare: da definire

Periodo: Il anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: ; 20,00

PSICOMETRIA PER LE NEUROSCIENZE COGNITIVE

Titolare: Prof. LIVIO FINOS

Periodo: I anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00

Prerequisiti:

Modelli Statistici e tecniche di riduzione della molteplicità.

Conoscenze e abilità da acquisire:

Il corso offrirà la possibilità di lavorare con dati EEG e fMRI reali. Lo studente imparerà come gestire i dati, dal pre-processing dei dati, ai risultati statistici finali. Verranno inoltre introdotti problemi pratici della vita reale e lo studente imparerà come ragionare criticamente sui dati e come risolvere i problemi della vita reale (ad es. Dati mancanti, fonti di variabilità dei dati, ecc).

Attività di apprendimento previste e metodologie di insegnamento:

Si alterneranno lezioni frontali ed esercitazioni. Le lezioni frontali illustreranno gli aspetti teorici della materia e la loro applicazione. Le lezioni pratiche (laboratoriali) daranno allo studente l'occasione di applicare sui dati le nozioni acquisite. Saranno inoltre proposte attività di gruppo nei quali gli studenti si confronteranno con problemi e dati reali.

Contenuti:

Presentazione delle principali tecniche di rilevazione Introduzione, Misure dirette e indirette della attività neurologica e fisiologica Presentazione delle principali tecniche: Tecniche trattate: - Elettroencefalografia (EEG) - Risonanza magnetica strutturale (sMRI) - Risonanza Magnetica funzionale (fMRI) Dal

segnale neurofisiologico al dato: Preprocessing del segnale - MRI: correzione del movimento (rigid body transformation), Slice?Timing correction, Filtri Spaziali e Temporali, Global intensity normalization, Rototraslazioni dei dati. - EEG: FiltraggioSpazio-temporale del segnale, Estrazione delle sorgenti del segnale (blind source separation). Definizione delle features di interesse per l'analisi. Dal dato all'informazione - Mass-univariate Analysis of fMRI (1st and 2nd level) - Mixed models per l'analisi di dati EEG - Test di permutazione - Test Multipli (da Bonferroni a Random field theory) - Multi-Voxel Pattern Analysis Approfondimenti - Diffusion Tensor Imaging (DTI) - Hyperalignment (Procrustes analysis) - Distanze Riemanniane - Independent Component Analysis, - Brain-computer Interface

Modalità di esame:

Prova scritta e orale

Criteri di valutazione:

Si valutano la comprensione dei contenuti proposti e la capacità di applicare in contesti pratici le conoscenze e le metodologie acquisite.

Testi di riferimento:

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio:

Il materiale didattico verrà fornito dai docenti all'inizio del corso. Inoltre lo studente avrà a disposizione le slide del corso, suddivise per argomento. Il materiale integrativo può essere scaricato dal sito Moodle dell'insegnamento.

SOCIAL MEDIA

Titolare: Dott. STEFANO SBALCHIERO

Periodo: I anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00

Prerequisiti:

Nessuno

Conoscenze e abilità da acquisire:

Il corso intende delineare un approccio metodologico allo studio dei social media con particolare riferimento all'analisi delle reti sociali e dei contenuti da esse veicolati. Di conseguenza, si dovrà acquisire la capacità di valutare criticamente le potenzialità e i limiti delle informazioni ricavabili, specie dal materiale di tipo testuale. Lo studente sperimenterà l'impostazione di una ricerca in questo ambito, attraverso l'analisi di esempi concreti di applicazione di metodi statistici ai contenuti della rete.

Attività di apprendimento previste e metodologie di insegnamento:

Lezioni frontali ed eventuali laboratori

Contenuti:

- Introduzione ai social media Definizioni Problematiche tecniche, etiche e sociali Finalità dell'analisi (esempi) - Inquadramento sociologico Cenni ontologico/epistemologici e conseguenze sulla conoscibilità della realtà sociale Approcci quantitativi/qualitativi/misti Approfondimento sul ruolo del ricercatore - rapporto osservatore/osservato: da "manipolativo" a "naturalistico" - Social Network Analysis Concetti Reti complete ed ego-centrate Tecniche di rilevazione Grafi, indici e affiliazioni Esempi e utilizzo negli ambiti sociali, aziendali e di marketing - Analisi automatica dei testi Definizioni Unità di analisi, dati e meta-dati Parsing, lemmatizzazione, stemming Analisi statistica dei testi e text mining Analisi lessicale, analisi testuale e relativi indici Tipi di matrici Esempi di applicazione e risorse informatiche - Sentiment Analysis Lo studio dei sentimenti e delle emozioni nel web Scoring, tagging automatico e umano Classificazione individuale e aggregata Integrated Sentiment Analysis Strumenti, potenzialità e limiti Esempi applicativi

Modalità di esame:

Prova pratica in aula informatica + prova orale

Criteri di valutazione:

Qualità delle prove di esame

Testi di riferimento:

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio:

Materiale didattico distribuito dal docente NOTA:si tratta di informazioni di massima che potrebbero subire variazioni al momento della presa in carico dell'insegnamento da parte del docente titolare

STATISTICA COMPUTAZIONALE PROGREDITO

Titolare: Prof. NICOLA SARTORI

Periodo: Il anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 42A+22L; 9,00

Prerequisiti:

Calcolo delle Probabilità e Statistica progredito. Utile, anche se non prerequisito, Statistica Computazionale.

Conoscenze e abilità da acquisire:

- Lo sviluppo di nuove tecniche computazionali per l'inferenza in modelli statistici. - Utilizzo di R per l'implementazione di tali tecniche.

Attività di apprendimento previste e metodologie di insegnamento:

Contenuti:

- Simulazione: algoritmi di accettazione e rifiuto; integrazione Monte Carlo; importance sampling e altri metodi di riduzione della varianza. - Metodi numerici e grafici per l'analisi di verosimiglianza e per l'inferenza Bayesiana. - L'algoritmo EM. - Metodi di ricampionamento: bootstrap e jackknife. - Catene di Markov e algoritmi Markov Chain Monte Carlo (MCMC): la teoria delle catene Markoviane; algoritmi MCMC; applicazioni all'inferenza Bayesiana.

Modalità di esame:

L'esame consiste in una prova pratica in aula ASID. La prova comprende sia parti teoriche che analisi empiriche.

Criteri di valutazione:

La valutazione si basa sulla correttezza e completezza della prova pratica.

Testi di riferimento:

Robert, Christian P.; Casella, George, *Introducing Monte Carlo methods with R*. New York: Springer, 2010 Davison, Anthony Christopher; Hinkley, David V., *Bootstrap methods and their application*. Cambridge [etc.]: Cambridge university press, 1997 Albert, Jim, *Bayesian computation with R*. Dordrecht: Springer Verlag, 2009 Gelman, Andrew; Meng, Xiao-Li; Brooks, Steve; Jones, Galin L., *Handbook of Markov Chain Monte Carlo*. Boca Raton: CRC Press (Taylor & Francis Group), 2011 Efron, Bradley; Tibshirani, Robert, *Introduction to the bootstrap* Bradley Efron and Robert J. Tibshirani. New York: London, Chapman & Hall, 0

Eventuali indicazioni sui materiali di studio:

Materiale didattico disponibile sulla pagina web del corso (su Moodle)

STATISTICA ITERAZIONE

Titolare: Prof.ssa GIOVANNA MENARDI

Periodo: Il anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: +64L; 9,00

Prerequisiti:

Vincolanti: Calcolo delle probabilità Statistica progredito Sostanziali (oltre ai vincolanti): Modelli statistici per dati sociali Analisi dei dati (Data mining)

Conoscenze e abilità da acquisire:

Il corso si prefigge di sviluppare le capacità operative degli studenti nell'analisi statistica di dati.

Attività di apprendimento previste e metodologie di insegnamento:

Presentazione e discussione critica dei problemi che si incontrano nell'analisi di dati reali in aula informatica

Contenuti:

Il corso si prefigge di offrire agli studenti l'opportunità di sviluppare capacità operative nell'analisi statistica di dati. A tale scopo alternerà momenti di presentazione e discussione critica dei problemi che si incontrano nell'analisi di dati reali con momenti di lavoro autonomo e di gruppo da parte degli studenti. L'attività prevista si basa sulla discussione e sul trattamento di problemi e dati reali, volti ad illustrare i concetti e ad utilizzare le tecniche presentate nei corsi di statistica, al fine di sviluppare una abilità operativa degli studenti in questo ambito. Per raggiungere questo obiettivo nella prima parte del corso sarà il docente che introdurrà alcuni problemi reali e suggerirà le tecniche più appropriate per l'analisi. Nella seconda parte saranno gli studenti a dover affrontare autonomamente alcuni problemi reali proposti dal docente e ad analizzare i dati con le tecniche che riterranno più adeguate. NB: Data la particolare natura dell'insegnamento, che richiede forte interazione tra studenti e docenti, è necessario che gli studenti vi si preiscrivano (secondo modalità che verranno comunicate nel mese di dicembre 2019), non potendosi escludere una limitazione negli accessi. In ogni caso saranno ammessi a partecipare all'insegnamento solo studenti del secondo anno della Laurea magistrale che abbiano superato entrambi gli esami di Calcolo delle probabilità e di Statistica progredito. Le eventuali limitazioni saranno basate su superamento e voto oltre che di questi due esami anche di Modelli statistici per dati sociali e di Analisi dei dati (Data mining).

Modalità di esame:

L'esame di accertamento consiste in tre prove: - partecipazione globale all'attività in aula (sia quantitativa che qualitativa) - lavoro di gruppo di analisi di un insieme di dati complesso, con presentazione di un rapporto finale - prova pratica individuale di analisi di un insieme di dati.

Criteri di valutazione:

- partecipazione all'attività in aula (sia quantitativa che qualitativa) - capacità di analizzare un insieme di dati complesso, con presentazione di un rapporto finale - capacità di analizzare un insieme di dati in breve tempo

Testi di riferimento:

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio:

Il materiale usato a lezione dai docenti sarà messo a disposizione degli studenti.

STATISTICA MEDICA ED EPIDEMIOLOGIA PROGREDITO

Titolare: Prof.ssa GIOVANNA BOCCUZZO

Periodo: Il anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 32A+32L; 9,00

Prerequisiti:

Statistica progredito Trattandosi di un corso progredito, gli studenti devono avere conoscenze di base di epidemiologia e di statistica medica. Conoscenza

di base di R e di SAS.

Conoscenze e abilità da acquisire:

Il corso fornisce metodi e strumenti avanzati per la ricerca clinica ed epidemiologica. In particolare, al termine del corso lo studente avrà le competenze necessarie per coordinare gli aspetti statistici di studi epidemiologici, dall'impostazione dello studio alla stesura del rapporto finale.

Attività di apprendimento previste e metodologie di insegnamento:

Sia le metodologie di insegnamento che le modalità di esame sono improntate sull'acquisizione sia di conoscenze sia di competenze specialistiche, tecniche e trasversali. Il corso è organizzato con lezioni frontali, laboratori in aula informatica, discussione di casi studio e articoli scientifici, e conferenze didattiche. Durante il corso gli studenti svolgono dei lavori in gruppi o a coppie. Per l'esame predispongono un report completo di studio di un problema reale, e fanno una presentazione ai loro colleghi su un tema concordato col docente.

Contenuti:

1. Richiami. Epidemiologia e causalità; tipi di studio in epidemiologia; misure di occorrenza di malattia e di associazione fra malattia e fattore di rischio; concetti di confondente e modificatore di effetto; stratificazione, appaiamento e standardizzazione; validità e fonti di distorsione. 2. Epidemiologia sociale. Metodi di controllo dei confondenti in studi osservazionali: regressione logistica per dati appaiati, propensity score matching, analisi con risposta ordinale; misure di salute della popolazione e di livello socio-economico; diseguglianze di salute: misura e decomposizione. Applicazioni in aula informatica con SAS e R. Lettura di articoli scientifici e discussione. 3. La Meta-Analisi. Introduzione alla meta-analisi, come preparare una meta-analisi, il calcolo dei singoli effetti, combinare i singoli effetti, valutare l'eterogeneità? e il publication bias, riportare i risultati della meta-analisi. Applicazioni in aula informatica con SAS e R. Lettura di articoli scientifici e discussione. 4. Metodologia per un protocollo di ricerca clinica. Il disegno dello studio, l'allocazione del trattamento, randomizzazione e standardizzazione, determinazione della numerosità campionaria, aspetti etici della ricerca, interpretazione dei risultati. Applicazioni in aula informatica con SAS e R. Lettura di articoli scientifici e discussione. 5. Epidemiologia delle malattie infettive. Sorveglianza sanitaria: notifica e sistema informativo delle malattie infettive. Modelli matematici di diffusione delle epidemie. Applicazioni in aula informatica con SAS e R. Lettura di articoli scientifici e discussione. 6. Elementi di farmacoepidemiologia e farmacovigilanza. Sulla base del tempo disponibile, potranno essere approfonditi altri argomenti, quali: Studi clinici randomizzati per la valutazione dei farmaci, approccio sequenziale al disegno delle prove cliniche; disegno sequenziale e disegno adattivo di prove cliniche confermate; metodi adattivi per la Fase I, II e III delle prove cliniche, vantaggi e svantaggi rispetto al disegno standard basato su ampiezze campionarie fisse; regole d'arresto per l'individuazione di trattamenti inefficaci.

Modalità di esame:

Esame orale con discussione di un'esercitazione finale, e presentazione in gruppi di un tema di natura metodologico/epidemiologica. L'esercitazione finale, sviluppata preferibilmente a coppie, verte sullo studio di un problema basato sull'analisi di dati reali. I temi sono concordati col docente.

Criteri di valutazione:

La valutazione si basa sia sulle conoscenze manifestate durante l'esame orale sia sull'esercitazione finale e la presentazione. Il voto finale è una media pesata dei voti riportati nelle tre parti, con pesi pari a 2/5 per la prova orale, 2/5 per l'esercitazione finale, e 1/5 per la presentazione in aula. Le valutazioni ottenute nell'esercitazione finale e nella presentazione rimangono valide per un anno, anche se la prova orale non dovesse essere sufficiente.

Testi di riferimento:

Vineis P., Duca P. e Pasquini, Manuale di metodologia epidemiologica.. : Numero speciale di Epidemiologia e prevenzione n.3, 1987 Rothman, Kenneth J.; Greenland, Sander, Modern epidemiology Kenneth J. Rothman, Sander Greenland, Timothy L. Lash. Philadelphia [etc.]: Wolters Kluwer Health: Lippincott Williams & Wilkins, 0 D.W. Hosmer, S. Lemeshow, Applied Logistic Regression. New York: Wiley, 2000 M. Borenstein et al, Introduction to Meta-Analysis. Chichester, UK: Wiley, 2009 Oakes, J. Michael; Kaufman, Jay S., Methods in social epidemiology edited by Michael Oakes, Jay S. Kaufman. San Francisco: Jossey-Bass, 2017 Edited by I. Dos Santos Silva, Cancer Epidemiology: Principles and Methods. Lyon, France: International Agency for Research on Cancer, 1999

Eventuali indicazioni sui materiali di studio:

Tutto il materiale, a parte i testi di riferimento coperti da copyright (disponibili in biblioteca), è reso disponibile su moodle. Riguarda le dispense delle lezioni, testi non coperti da copyright, programmi e dettagli sulle esercitazioni in aula informatica (alla conclusione dell'esercitazione), materiali per le presentazioni, letture di interesse e articoli scientifici.

STATISTICA PER LA TECNOLOGIA E L'INDUSTRIA

Titolare: Prof.ssa GIOVANNA CAPIZZI

Periodo: Il anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 48A+16L; 9,00

Prerequisiti:

Elementi di base del controllo statistico della qualità e inferenza (univariata e multivariata)

Conoscenze e abilità da acquisire:

- Conoscere e utilizzare correttamente alcune tecniche avanzate del controllo statistico dei processi produttivi di beni e servizi, dei sistemi tecnologici. - Conoscere e utilizzare correttamente alcune tecniche avanzate di sorveglianza on-line e off-line di dati complessi (dati multivariati, funzionali, dati multistadio, ecc.) - Conoscere e utilizzare correttamente metodi di sorveglianza specifici degli ambiti tecnologici ed industriali a contesti applicativi diversi (ad esempio a contesti bio-medici).

Attività di apprendimento previste e metodologie di insegnamento:

Il corso prevede sia lezioni frontali che lezioni in laboratorio informatico. Una parte delle lezioni frontali saranno dedicate all'illustrazione di alcuni articoli tratti dalla letteratura scientifica recente sul controllo statistico di processo (Journal Club).

Contenuti:

- Introduzione dei concetti di basi del controllo statistico della qualità e panoramica degli ambiti applicativi moderni di applicazione (individuazione di anomalie nei processi produttivi, nei flussi di comunicazione, di realizzazioni di eventi indesiderati in ambito bio-medico, ecc.) - Fase I e Fase II nel disegno di schemi di sorveglianza (criticità della sorveglianza off-line e on-line) - Controllo statistico di processo "univariato": caso parametrico e nonparametrico. - Carte di tipo Shewhart, carte con memoria markoviana, carte basate sulla statistica GLR (Generalized Likelihood Ratio), carte combinate, carte self-starting (short runs in ambito industriale e non industriale) - Controllo statistico di processo "multivariato": caso parametrico e non parametrico. - Carte di controllo multivariate (T2 di Hotelling, MEWMA, MCUSUM). - Misure di distanza euclidea e non euclidea per individuare osservazioni anomale. - Carte di controllo multivariate invarianti rispetto alla direzione dei cambiamenti. - Generalizzazione del concetto di distanza ed ordinamento al caso multidimensionale - Tecniche di riduzione della dimensionalità: - Metodo di Hawkins (regressione delle variabili). - Metodo delle proiezioni (PCA, PLS ,

ecc.) - Decomposizione MYT della carta di controllo di Hotelling. - Fault diagnosis: metodo di calcolo del contributo delle variabili e metodi principali per determinare le cause che generano un "fuori controllo" del processo. - Sorveglianza di dati funzionali. - Sorveglianza di processi multi stadio. - Metriche di performance degli schemi di sorveglianza - Metodi per il disegno degli schemi di sorveglianza (metodi esatti, approssimati, per permutazione, per simulazione, ecc.)

Modalità di esame:

- Presentazione di gruppo su un articolo recente nella letteratura su controllo statistico di processo. - Esame orale sulle diverse tematiche affrontate nel corso, partendo da uno degli articoli studiati durante le attività del Journal Club).

Criteri di valutazione:

La valutazione della preparazione dello studente si baserà sulla comprensione degli argomenti svolti, sull'acquisizione dei concetti e delle metodologie proposte e sulla capacità di applicarli in modo autonomo e consapevole.

Testi di riferimento:

Qiu, Peihua, Introduction to statistical process control Peihua Qiu. Boca Raton [etc.]: CRC press, 2013

Eventuali indicazioni sui materiali di studio:

- Il materiale usato a lezione (slide, articoli, codici in R, ecc.) dal docente sarà a disposizione nella pagina del corso nel Moodle del Dipartimento di Scienze Statistiche - I libri di testo saranno precisati all'inizio del corso, anche per tenere conto della preparazione pregressa degli studenti

STATISTICA PROGREDITO

Titolare: Prof. NICOLA SARTORI

Periodo: I anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 56A+26E; 9,00

Prerequisiti:

Calcolo delle Probabilità. Solide basi di Analisi matematica e Algebra lineare.

Conoscenze e abilità da acquisire:

Conoscenze: concetti e strumenti statistico-matematici essenziali per poter comprendere e affrontare problemi di Statistica metodologica. Abilità: capacità di riconoscere concetti e strumenti matematici utilizzati in altri contesti statistici; capacità di risolvere semplici problemi di carattere metodologico nonché di comprendere o formulare eventuali modelli atti a descriverli.

Attività di apprendimento previste e metodologie di insegnamento:

Didattica frontale: teoria (75%), esercitazioni (25%). Supporto: Servizio tutorato.

Contenuti:

- Specificazione del modello e valutazione dell'incertezza nell'inferenza. - Paradigmi dell'inferenza: frequentista e bayesiano. - Specificazione del modello: livelli di specificazione e schemi di osservazione; famiglie di distribuzioni notevoli; specificazione della distribuzione a priori; scambiabilità. - Richiami sugli elementi di base dell'inferenza statistica frequentista e bayesiana: problemi di stima puntuale, di stima intervallare, di verifica d'ipotesi, di previsione. Esempi notevoli. - Riduzione di dati e statistiche sufficienti. Criterio di fattorizzazione di Neyman-Pearson. Statistiche sufficienti minimali. Famiglie esponenziali e statistiche sufficienti. Statistiche costanti in distribuzione. Statistiche complete e teorema di Basu. - La funzione di verosimiglianza e quantità collegate (funzione punteggio, stima di massima verosimiglianza, informazione osservata). Verosimiglianza e sufficienza. Principi forte e debole di verosimiglianza. Parametri di disturbo e verosimiglianza profilo. Proprietà di invarianza della verosimiglianza. Cenni agli aspetti computazionali. - Inferenza basata sulla verosimiglianza: proprietà campionarie esatte. Lemma di Neyman-Pearson. Test uniformemente più potenti. Test uniformemente più potenti tra i non distorti. Diseguaglianza di Wald. Informazione attesa e diseguaglianza di Cramer-Rao. Teorema di Rao-Blackwell e Lehmann-Scheffé. Informazione attesa ed inferenza bayesiana: la distribuzione a priori di Jeffreys. Cenni ad altre distribuzioni a priori non informative. - Inferenza basata sulla verosimiglianza: proprietà asintotiche e distribuzioni approssimate. Consistenza. Rapporto di verosimiglianza e distribuzione asintotica. Forme asintoticamente equivalenti. Verosimiglianza profilo. Cenni ai modelli non regolari. - Inferenza bayesiana: approssimazioni asintotiche. Cenni ai metodi numerici. Distribuzioni marginali per parametri di interesse. - Errata specificazione del modello statistico ed equazioni di stima.

Modalità di esame:

Prova scritta a libro chiuso. Il dettaglio delle regole d'esame sarà disponibile alla pagina Moodle dell'insegnamento (accessibile da <https://elearning.unipd.it>)

Criteri di valutazione:

Il testo d'esame sarà composto da esercizi, suddivisi in quesiti, con grado di difficoltà comparabile. Si valuteranno la preparazione dello studente sui contenuti oggetto del corso e la sua capacità di risolvere semplici problemi di carattere metodologico.

Testi di riferimento:

Adelchi Azzalini, Inferenza statistica: una presentazione basata sul concetto di v. Milano: Springer-Verlag Italia, 2001 Luigi Pace & Alessandra Salvan, Introduzione alla statistica - II - Inferenza, verosimiglianza, modelli. Padova: Cedam, 2001

Eventuali indicazioni sui materiali di studio:

Testi di consultazione: - Adimari, G., Pauli, F. (2012). Esercizi di Statistica (corso progredito). (disponibile gratuitamente in formato pdf all'indirizzo https://homes.stat.unipd.it/sites/homes.stat.unipd.it/gianfrancoadimari/files/libro_eserc.pdf) - Casella, J., Berger, R.L. (2001). Statistical Inference (2nd edition). Duxbury Press. - Evans, M.J., Rosenthal, J.S. (2003). Probability and Statistics: The Science of Uncertainty. W.H. Freeman and Co. (disponibile gratuitamente in formato pdf all'indirizzo <http://www.utstat.toronto.edu/mikeevans/jeffrosenthal/>) - Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B. (2014). Bayesian Data Analysis (3rd edition). CRC Press.

STRUMENTI STATISTICI PER L'ANALISI DI DATI AZIENDALI

Titolare: Prof.ssa MARIANGELA GUIDOLIN

Periodo: Il anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 34A+30L; 9,00

Prerequisiti:

Formale: Calcolo delle Probabilità e Statistica Progredito Sostanziale: Analisi dei dati (data mining)

Conoscenze e abilità da acquisire:

Il corso si propone di fornire strumenti statistici avanzati a supporto del processo di decisione aziendale. Il corso ha una forte connotazione applicata, con una partecipazione attiva dello studente durante le settimane di lezione. E' prevista un'intensa attività in aula informatica, con l'analisi ed il commento di diversi studi di caso.

Attività di apprendimento previste e metodologie di insegnamento:

Il corso verrà erogato per mezzo di lezioni frontali, esercitazioni e studi di caso in aula informatica. Considerato il taglio fortemente applicativo del corso, la frequenza alle lezioni ed esercitazioni, seppure non obbligatoria, è vivamente consigliata.

Contenuti:

- Le decisioni in azienda - Modelli per le decisioni aziendali --- Modelli per dati ordinali --- Modelli gerarchici --- Modelli non parametrici --- Modelli per l'associazione di variabili (market basket analysis) - Analisi della customer base ---Modelli per dati di rete - Applicazioni e studi di caso: business intelligence - Tecniche di visualizzazione - Problematiche di privacy dei dati, General Data Protection Regulation, etica professionale nella pratica statistica.

Modalità di esame:

La prova d'esame si articola in due parti: - Prova orale relativa al programma svolto a lezione. - Lavoro di gruppo su un caso di studio prescelto e preventivamente concordato con il docente.

Criteri di valutazione:

La valutazione della preparazione dello studente si baserà sulla comprensione degli argomenti svolti e sulla capacità di adottare le soluzioni statistiche più idonee per l'analisi di dati reali. Autonomia e spirito critico nell'applicazione delle soluzioni sono criteri fondamentali considerati nel processo di valutazione.

Testi di riferimento:

Gelman, Andrew; Hill, Jennifer, Data analysis using regression and multilevel/hierarchical models. Cambridge: Cambridge University press,, 2006 Agresti, Alan, Analysis of Ordinal Categorical Data. Hoboken: John Wiley & Sons, 2012 Hastie, Trevor J.; Tibshirani, Robert; Friedman, Jerome, The elements of statistical learning data mining, inference, and prediction. New York: Springer, 2009 Azzalini, Adelchi; Scarpa, Bruno, Data analysis and data mining: an introduction. New York: Oxford University Press, 2012

Eventuali indicazioni sui materiali di studio:

Materiale didattico predisposto dal docente

TEMI E METODI DI POPOLAZIONE E SOCIETA'

Titolare: Prof. STEFANO MAZZUCO

Periodo: Il anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 38A+26L; 9,00

Prerequisiti:

Statistica progredito Calcolo delle probabilità Modelli statistici per dati sociali

Conoscenze e abilità da acquisire:

Gli studenti dovranno acquisire da una parte alcuni aspetti dell'analisi demografica e sociale non trattati altrove (analisi delle sequenze, agent-based models, costruzione di indicatori complessi) dall'altra una mentalità che permetta loro di collegare aspetti sostantivi e metodologici.

Attività di apprendimento previste e metodologie di insegnamento:

Il corso si compone di moduli tematici, in cui vengono esposte (scegliendo tra i temi e i metodi affrontati negli studi più recenti) problematiche rilevanti nello studio dei fenomeni demografici e sociali e alcuni metodi per affrontarne l'analisi empirica. Si tratta di temi e metodi non trattati in insegnamenti precedenti. Per ogni modulo è prevista un'esercitazione (individuale o a gruppi) con assegnazione di problemi concreti che verranno trattati ed esposti dagli studenti.

Contenuti:

Tema 1: La seconda transizione demografica e la de standardizzazione del corso di vita Metodo: Analisi delle sequenze Tema 2: La misura di concetti complessi Metodo: Indicatori sintetici e analisi multicriterio Tema 3: Interazione Sociale e fecondità Metodo: Agent-Based Computational Demography

Modalità di esame:

L'esame finale è orale. Durante le lezioni gli studenti saranno assegnati a lavori individuali e/o di gruppo la cui valutazione inciderà sul voto finale.

Testi di riferimento:

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio:

Il materiale di studio è costituito principalmente da articoli (verranno forniti dal docente in aula) relativi al tema e al metodo trattati

TEORIE E MODELLI DEMOGRAFICI

Titolare: Prof.ssa MARIA LETIZIA TANTURRI

Mutuato da: Laurea magistrale in Scienze Statistiche

Periodo: l'anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 48A+16L; 9,00

Prerequisiti:

Per seguire il corso con profitto è necessaria una conoscenza di base della demografia. Gli studenti che non hanno mai seguito un corso di demografia, dovranno studiare i concetti e le misure demografiche fondamentali sul volume di: - Rowland D. T. (2003), Demographic methods and concepts, Oxford University Press (ch. 1, 2,3,4,6,7) O alternativamente: - Livi Bacci M. (ultima edizione) , "Introduzione alla demografia", Loescher Editore.

Conoscenze e abilità da acquisire:

Il corso ha un duplice scopo. In primo luogo si propone di far acquisire agli studenti una conoscenza approfondita e rigorosa dei metodi dell'analisi demografica e dei modelli di popolazione. In particolare si dà ampio spazio a: a) ai modelli che enfatizzano come la popolazione cambi in funzione dell'interazione di più processi demografici; b) a come i fenomeni demografici possano essere modellati nel caso di statistiche carenti (es. Paesi in via di Sviluppo o piccole aree); c) a strumenti che permettono di prevedere gli sviluppi futuri di una popolazione. In secondo luogo si vuole introdurre gli studenti ad una conoscenza critica delle principali teorie di popolazione che tentano di spiegare il cambiamento demografico così come è avvenuto sia nei Paesi sviluppati che nei Paesi in via di sviluppo. Gli studenti conosceranno le principali determinanti della fecondità, mortalità, nuzialità e migratorietà. Il corso è basato sul pieno coinvolgimento degli studenti in varie attività che saranno proposte attraverso la piattaforma Moodle. L'attività di laboratorio e lo studio di casi accompagnerà le lezioni frontali così che gli studenti sappiano applicare le conoscenze acquisite a problemi concreti di ricerca e sappiano orientarsi nella letteratura demografica.

Attività di apprendimento previste e metodologie di insegnamento:

- Lezioni frontali - Esempi di casi di studio - Discussioni in classe di articoli scientifici che gli studenti leggeranno preventivamente - Esercitazioni in aula informatica - Esercitazioni da svolgere a casa

Contenuti:

1. Richiami di demografia di base. Misure della crescita demografica ed equazione della popolazione. Indicatori sintetici e tavole di mortalità, indicatori sintetici della fecondità, misurazione dei movimenti migratori. Misure e metodi per l'analisi demografica per generazioni e per contemporanei. Le ipotesi fondamentali dell'analisi demografica. Principali funzioni matematiche interpolanti la mortalità. Il modello di Bongaarts delle variabili intermedie della fecondità. 2. Modelli e teorie della mortalità. Le tavole tipo (approcci relazionale; metodo di Halley puro e modificato); decomposizione delle differenze di speranza di vita. Teorie del declino della mortalità (fasi della transizione epidemiologica e sanitaria). Prospettive sul futuro della longevità. 3. Modelli e teorie della fecondità, della nuzialità e delle variabili intermedie della fecondità. Tempo e quantum per coorte e per periodo, la traslazione. Metodo di Hajnal; i modelli analitici della nuzialità; gli indici I_f , I_g e I_m di Coale; il modello di Coale e Trussel; le misure di contraccezione e abortività. Teorie del declino della fecondità (teoria della Nuova economia della famiglia, teoria di Easterlin, approccio socio-culturale, seconda transizione demografica, teorie di genere, teoria delle preferenze, teoria dell'happiness). 4. Modelli e teorie delle migrazioni. Stime della migratorietà con fonti di vario livello di completezza. Stime delle migrazioni con l'equazione della popolazione. Teorie delle migrazioni: teoria demografica, teoria micro-economica classica, la teoria della "nuova economia della migrazione", teoria del mercato dualistico, teoria del sistema-mondo. 5. I modelli di popolazione. Popolazioni esponenziali, logistiche, malthusiane, stabili e stabili generalizzate. Relazioni sintetiche fra mortalità, fecondità, migrazioni, relazioni analitiche fra struttura e crescita demografica. Rimpiazzo delle generazioni. Teorie dell'equilibrio demografico: da Malthus alla transizione demografica. 6. Stime indirette con dati scarsi. Mortalità adulta con il metodo dei figli orfani; mortalità infantile con il metodo dei figli sopravvissuti; fecondità con il metodo dei figli propri. Stime della dinamica demografica comparando due rilevazioni di stato (per sesso ed età) a tempi differenziati. 7. Le previsioni demografiche. Modello a componenti di coorte. Previsioni stocastiche: le previsioni Onu. Valutazione della bontà e degli errori delle previsioni

Modalità di esame:

2 esercitazioni pratiche da svolgere durante il corso (40% del voto finale) - lettura di un articolo e sintesi in classe durante il corso (20% del voto finale) - esame orale (40% del voto finale) Per i frequentanti l'esame finale orale può essere sostituito dalla realizzazione di un project work, con lettura della bibliografia essenziale sul tema, analisi di dati demografici e la stesura di un rapporto di ricerca a commento dei risultati ottenuti.

Criteri di valutazione:

Si valuterà la capacità dello studente di: - acquisire una conoscenza approfondita dei metodi di analisi demografica - utilizzare criticamente tali conoscenze nell'analisi dei dati demografici - leggere criticamente la letteratura scientifica su alcuni temi demografici

Testi di riferimento:

Preston, S.H., P. Heuveline and M. Guillot, Demography, Measuring and Modeling Population Processes. Oxford: Blackwell, 2001., Livi Bacci M, Storia minima della popolazione del mondo. Bologna: Il Mulino, 2011

Eventuali indicazioni sui materiali di studio:

Oltre allo studio dei testi di riferimento, durante il corso sarà proposta la lettura critica di alcuni articoli della letteratura demografica su temi specifici. I riferimenti saranno indicati sulla piattaforma Moodle. NOTA: si tratta di informazioni di massima che potrebbero subire variazioni al momento della presa in carico dell'insegnamento da parte del docente titolare

TEORIE E MODELLI DEMOGRAFICI

Titolare: Prof.ssa MARIA LETIZIA TANTURRI

Periodo: l'anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 48A+16L; 9,00

Prerequisiti:

Per seguire il corso con profitto è necessaria una conoscenza di base della demografia. Gli studenti che non hanno mai seguito un corso di demografia, dovranno studiare i concetti e le misure demografiche fondamentali sul volume di: - Rowland D. T. (2003), Demographic methods and concepts, Oxford University Press (ch. 1, 2,3,4,6,7) O alternativamente: - Livi Bacci M. (ultima edizione) , "Introduzione alla demografia", Loescher Editore.

Conoscenze e abilità da acquisire:

Il corso ha un duplice scopo. In primo luogo si propone di far acquisire agli studenti una conoscenza approfondita e rigorosa dei metodi dell'analisi demografica e dei modelli di popolazione. In particolare si dà ampio spazio a: a) ai modelli che enfatizzano come la popolazione cambi in funzione dell'interazione di più processi demografici; b) a come i fenomeni demografici possano essere modellati nel caso di statistiche carenti (es. Paesi in via di

Sviluppo o piccole aree); c) a strumenti che permettono di prevedere gli sviluppi futuri di una popolazione. In secondo luogo si vuole introdurre gli studenti ad una conoscenza critica delle principali teorie di popolazione che tentano di spiegare il cambiamento demografico così come è avvenuto sia nei Paesi sviluppati che nei Paesi in via di sviluppo. Gli studenti conosceranno le principali determinanti della fecondità, mortalità, nuzialità e migratorietà. Il corso è basato sul pieno coinvolgimento degli studenti in varie attività che saranno proposte attraverso la piattaforma Moodle. L'attività di laboratorio e lo studio di casi accompagnerà le lezioni frontali così che gli studenti sappiano applicare le conoscenze acquisite a problemi concreti di ricerca e sappiano orientarsi nella letteratura demografica.

Attività di apprendimento previste e metodologie di insegnamento:

- Lezioni frontali - Esempi di casi di studio - Discussioni in classe di articoli scientifici che gli studenti leggeranno preventivamente - Esercitazioni in aula informatica - Esercitazioni da svolgere a casa

Contenuti:

1. Richiami di demografia di base. Misure della crescita demografica ed equazione della popolazione. Indicatori sintetici e tavole di mortalità, indicatori sintetici della fecondità, misurazione dei movimenti migratori. Misure e metodi per l'analisi demografica per generazioni e per contemporanei. Le ipotesi fondamentali dell'analisi demografica. Principali funzioni matematiche interpolanti la mortalità. Il modello di Bongaarts delle variabili intermedie della fecondità. 2. Modelli e teorie della mortalità. Le tavole tipo (approcci relazionale; metodo di Halley puro e modificato); decomposizione delle differenze di speranza di vita. Teorie del declino della mortalità (fasi della transizione epidemiologica e sanitaria). Prospettive sul futuro della longevità. 3. Modelli e teorie della fecondità, della nuzialità e delle variabili intermedie della fecondità. Tempo e quantum per coorte e per periodo, la traslazione. Metodo di Hajnal; i modelli analitici della nuzialità; gli indici I_f , I_g e I_m di Coale; il modello di Coale e Trussel; le misure di contraccezione e abortività. Teorie del declino della fecondità (teoria della Nuova economia della famiglia, teoria di Easterlin, approccio socio-culturale, seconda transizione demografica, teorie di genere, teoria delle preferenze, teoria dell'happiness). 4. Modelli e teorie delle migrazioni. Stime della migratorietà con fonti di vario livello di completezza. Stime delle migrazioni con l'equazione della popolazione. Teorie delle migrazioni: teoria demografica, teoria micro-economica classica, la teoria della "nuova economia della migrazione", teoria del mercato dualistico, teoria del sistema-mondo. 5. I modelli di popolazione. Popolazioni esponenziali, logistiche, malthusiane, stabili e stabili generalizzate. Relazioni sintetiche fra mortalità, fecondità, migrazioni, relazioni analitiche fra struttura e crescita demografica. Rimpiazzo delle generazioni. Teorie dell'equilibrio demografico: da Malthus alla transizione demografica. 6. Stime indirette con dati scarsi. Mortalità adulta con il metodo dei figli orfani; mortalità infantile con il metodo dei figli sopravvissuti; fecondità con il metodo dei figli propri. Stime della dinamica demografica comparando due rilevazioni di stato (per sesso ed età) a tempi differenziati. 7. Le previsioni demografiche. Modello a componenti di coorte. Previsioni stocastiche: le previsioni Onu. Valutazione della bontà e degli errori delle previsioni

Modalità di esame:

2 esercitazioni pratiche da svolgere durante il corso (40% del voto finale) - lettura di un articolo e sintesi in classe durante il corso (20% del voto finale) - esame orale (40% del voto finale) Per i frequentanti l'esame finale orale può essere sostituito dalla realizzazione di un project work, con lettura della bibliografia essenziale sul tema, analisi di dati demografici e la stesura di un rapporto di ricerca a commento dei risultati ottenuti.

Criteri di valutazione:

Si valuterà la capacità dello studente di: - acquisire una conoscenza approfondita dei metodi di analisi demografica - utilizzare criticamente tali conoscenze nell'analisi dei dati demografici - leggere criticamente la letteratura scientifica su alcuni temi demografici

Testi di riferimento:

Preston, S.H., P. Heuveline and M. Guillot, Demography, Measuring and Modeling Population Processes. Oxford: Blackwell, 2001., Livi Bacci M, Storia minima della popolazione del mondo. Bologna: Il Mulino, 2011

Eventuali indicazioni sui materiali di studio:

Oltre allo studio dei testi di riferimento, durante il corso sarà proposta la lettura critica di alcuni articoli della letteratura demografica su temi specifici. I riferimenti saranno indicati sulla piattaforma Moodle. NOTA: si tratta di informazioni di massima che potrebbero subire variazioni al momento della presa in carico dell'insegnamento da parte del docente titolare

THEORY AND METHODS FOR INFERENCE

Titolare: Prof.ssa ALESSANDRA SALVAN

Periodo: Il anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00

Prerequisiti:

Insegnamenti del primo anno della Laurea Magistrale in Scienze Statistiche, in particolare Calcolo delle Probabilità e Statistica Progredito.

Conoscenze e abilità da acquisire:

Il corso mira a fornire agli studenti una comprensione approfondita della teoria dell'inferenza statistica, secondo diversi approcci, in particolare bayesiano e frequentista, evidenziando aspetti unificanti e specificità. Il corso ha l'obiettivo di sviluppare abilità avanzate di "problem solving" attraverso esercizi assegnati settimanalmente a gruppi e discussi in classe. Si preparano inoltre gli studenti alla consultazione di diverse fonti bibliografiche, sia testi, sia articoli su riviste specializzate. Agli studenti è proposto anche un lavoro individuale di preparazione alla stesura di un articolo. Sono inoltre previsti piccoli seminari degli studenti per sviluppare abilità di presentazione orale.

Attività di apprendimento previste e metodologie di insegnamento:

Lezioni, homework, presentazioni scritte e orali degli studenti

Contenuti:

- Modelli statistici e valutazione dell'incertezza nell'inferenza. Inferenza bayesiana e frequentista, specificazione del modello in ambito bayesiano e frequentista, valutazione frequentista dell'incertezza e problemi di distribuzione. Approssimazioni asintotiche e metodo delta. - Funzioni generatrici, approssimazioni di momenti, trasformazioni. Momenti, cumulanti e loro funzioni generatrici. Sviluppi di Edgeworth e Cornish-Fisher. Notazioni $Op(\cdot)$ e $op(\cdot)$. Approssimazioni di momenti e trasformazioni. Approssimazione di Laplace. - Riduzione dei dati e del modello. Modelli statistici dominati e fattorizzazioni della densità. Statistiche sufficienti e costanti in distribuzione, condizionamento, completezza. Riduzione dei dati e del modello in presenza di parametri di disturbo. - Famiglie esponenziali e di gruppo. Famiglie esponenziali naturali, applicazione del valore atteso, funzione di varianza e caratterizzazione. Sufficienza e completezza nelle famiglie esponenziali. Distribuzioni marginali e condizionate in famiglie esponenziali multiparametriche. Famiglie di dispersione esponenziali e modelli lineari generalizzati. Gruppi di trasformazioni, orbite e invarianti massimali, famiglie di gruppo semplici e composite. - Inferenza di verosimiglianza. Quantità di verosimiglianza osservate e attese, loro proprietà in famiglie esponenziali e di gruppo. Esempi con diversi schemi di osservazione e modelli: dati raggruppati e censurati, campionamento sequenziale, processi stocastici. Invarianza, sufficienza e proprietà distributive esatte. Diseguaglianze universali. Parametri ortogonali e parametrizzazione mista in famiglie esponenziali. Riparametrizzazioni. Statistiche

ancillari. Inferenza condizionata in famiglie di scala e posizione. Consistenza e teoria asintotica del primo ordine, procedure inferenziali collegate. Verosimiglianza profilo e relative quantità pivotali approssimate. Teoria asintotica nelle famiglie esponenziali. Modelli non regolari. Inferenza condizionata approssimata e teoria asintotica di ordine superiore. - Inferenza bayesiana. Approssimazioni asintotiche. A priori non informative. Metodi bayesiani empirici. Inferenza basata sulla distribuzione a posteriori: stima puntuale, regioni di credibilità, previsione. Modelli lineari. - Aspetti computazionali in R dell'inferenza di verosimiglianza e bayesiana. Esempi con parametro scalare e vettoriale. Parametri di interesse e verosimiglianza profilo. Verosimiglianza e bootstrap parametrico. Algoritmo EM. Inferenza bayesiana (elementi di base). - Equazioni di stima e pseudo-verosimiglianze. Effetti una non corretta specificazione del modello. Equazioni di stima. Quasi-verosimiglianza. Pseudo-verosimiglianze. Verosimiglianza composita. Verosimiglianza empirica. Verosimiglianza condizionata e verosimiglianza marginale in famiglie esponenziali e di gruppo. Verosimiglianze profilo e integrata. - Inferenza secondo i paradigmi decisionali negli approcci frequentista e bayesiano. Problemi di decisione statistica. Ottimalità nella stima puntuale, verifica di ipotesi e regioni di confidenza. Procedure ottimali in famiglie esponenziali e di gruppo.

Modalità di esame:

1/3 homework, 1/3 esame scritto finale, 1/3 presentazione scritta e orale di un lavoro individuale di rassegna basato su un paio di articoli scientifici recenti.

Criteri di valutazione:

La valutazione terrà conto di come gli argomenti presentati siano padroneggiati in applicazioni e problemi, della capacità di valutazione critica delle metodologie presentate, della abilità dimostrata nell'interazione con argomenti di ricerca recenti.

Testi di riferimento:

Davison, Anthony Christopher, Statistical Models. New York: Cambridge University Press, 2003 Pace, Luigi; Salvan, Alessandra, Principles of Statistical Inference, from a Neo-Fisherian Perspective. Singapore: World Scientific, 1997 Severini, Thomas A., Likelihood Methods in Statistics. Oxford: Oxford University Press, 2000 Severini, Thomas A., Elements of Distribution Theory. Cambridge: Cambridge University press, 2005 Young, G. A.; Smith, R. L., Essentials of Statistical Inference. Cambridge: Cambridge University Press, 2005

Eventuali indicazioni sui materiali di studio:

Materiale didattico sarà reso disponibile sulla pagina Moodle del corso.

TIROCINIO FORMATIVO

Titolare: Prof. FRANCESCO LISI

Periodo: Il anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: ; 1,00