



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

SCUOLA DI SCIENZE

Bollettino Notiziario

Anno Accademico 2017/2018

**Laurea magistrale in Scienze Statistiche
(Ord. 2014)**

Curriculum: Corsi comuni

ANALISI DEI DATI (DATA MINING)

(Titolare: Prof. BRUNO SCARPA)

Periodo: I anno, 2 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 34A+30L; 9,00 CFU

Prerequisiti :

Sostanziali ma non formali:

Modelli statistici II

Analisi dei dati Multidimensionali,

un primo corso di Programmazione,

un primo corso di Algebra Lineare

Conoscenze e abilita' da acquisire :

Lo scopo del corso Ã di introdurre gli studenti alla comprensione e alla reale capacitÃ di utilizzo di strumenti di data mining e di metodi statistici per lâ€™analisi dei dati.

AttivitÃ di apprendimento previste e metodologie di insegnamento :

Lezioni frontali. Esercitazioni in laboratorio

Contenuti :

- Nozioni generali: motivazioni e contesto, contrasto tra aderenza ai dati e complessitÃ 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 liscio, 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: nozioni sui metodi di raggruppamento: dissimilaritÃ , metodo delle k-medie, metodi gerarchici. Analisi delle associazioni tra variabili, algoritmo Apriori. Reti sociali (cenni).
- Miscellanea: sentiment analysis (cenni), tecniche di visualizzazione dei dati, cenni ad aspetti computazionali.

ModalitÃ di esame :

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

Criteri di valutazione :

Correttezza e qualitÃ 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

Eventuali indicazioni sui materiali di studio :

Strumenti informatici

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

ANALISI DEI DATI IN FINANZA

(Titolare: Prof. FRANCESCO LISI)

Periodo: Il anno, 2 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 34A+30L; 9,00 CFU

Prerequisiti :

Anche se non strettamente necessario, Ã fortemente consigliato avere acquisito i contenuti di Serie storiche finanziarie.

Conoscenze e abilita' da acquisire :

Lo scopo del corso Ã di introdurre gli studenti alla comprensione e alla reale capacitÃ di utilizzo di metodi statistici per lâ€™analisi e la modellazione di fenomeni finanziari.

Partendo dal problema finanziario, verranno presentate varie procedure computazionali basate su tecniche parametriche e nonparametriche e di ricampionamento.

Il corso sarÃ sviluppato su alcune problematiche attuali della finanza, quali, ad esempio, stima e controllo del rischio, la costruzione e valutazione di strategie di trading e la misurazione della performance di un portafoglio.

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 :

1. Tecniche statistiche per l'analisi del rischio finanziario. Modelli per il calcolo del Valore a Rischio (VaR)
2. Prezzaggio di opzioni mediante modelli GARCH
3. Tecniche statistiche di stima della volatilità
4. Modelli di regressione non parametrica e loro applicazioni finanziarie.
5. Analisi tecnica dei mercati finanziari
6. Introduzione alla borsa elettrica.
7. Misure e metodi di valutazione della performance di un portafoglio.

Modalità di esame :

Prova pratica in aula pc + 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 :

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio :

Lucidi forniti di volta in volta prima delle lezioni.

Trattandosi di un corso composto di vari moduli non è possibile indicare un solo testo. All'inizio di ogni modulo verranno forniti riferimenti bibliografici sia in italiano che in inglese.

APPRENDIMENTO AUTOMATICO

(Titolare: Prof. FABIO AIOLLI) - Mutuato da: Laurea magistrale in Informatica (Ord. 2014)

Periodo: I anno, 1 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 40A+8L; 6,00 CFU

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 o sono 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 di Concetti:

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 e Selezione di Variabili Categoriche; Model Selection, Holdout, CrossValidation, LeaveOneOut CV; Criteri Esterni e Interni per Valutare un Sistema di Clustering; Sistemi di Raccomandazione: Tipologie, Approcci, Misure di Valutazione.

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 :

Mitchell, Tom M., *Machine learning*. New York: McGraw-Hill, 1998

Alpaydin, Ethem, *Introduction to machine learning*. Cambridge: The MIT press, 2010

Eventuali indicazioni sui materiali di studio :

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

BIOSTATISTICA COMPUTAZIONALE E BIOINFORMATICA

(Titolare: Prof.ssa CHIARA ROMUALDI)

Periodo: Il anno, 1 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 40A+24L; 9,00 CFU

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 Proteomica. Database di dati genomici di riferimento mondiale disponibili al National Center of Biotechnology Information (NCBI), GeneBank, GEO, EntrezGene, OMIM, e in altri centri di riferimento, SwissProt, UniProt, Pfam. Sistemi integrati di interazione con questi database (interfacce web).

- Allineamento di sequenze. Algoritmi di allineamento, allineamenti globali e locali. Programmazione Dinamica, algoritmi euristici (BLAST, FASTA). Significatività dello score di un allineamento.

- Analisi di dati di espressione derivanti da esperimenti di microarray. 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.

Modalità di esame :

Esame Scritto

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 :

Parmigiani G, Garrett ES, Irizarry R, and Zeger SL., *The analysis of gene expression data: methods and software*. New York: Springer, 2003

Gentleman R. Carey V.J. Huber, *Bioinformatics and computational biology solutions using R and Bioconductor*,. New York: Springer, 2005

Ewens, Warren J., Grant, Gregory R., *Statistical Methods in Bioinformatics. An introduction*. New York: Springer, 2005

Eventuali indicazioni sui materiali di studio :

Materiale didattico predisposto dal docente.

CALCOLO DELLE PROBABILITÀ

(Titolare: Prof. PAOLO DAI PRA)

Periodo: I anno, 1 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 56A+26E; 9,00 CFU

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 :

Spazio di probabilità : Spazio campionario, eventi, algebre e sigma-algebre, probabilità e sue proprietà fondamentali. Elementi di calcolo combinatorio e spazi discreti uniforme. Probabilità condizionata ad un evento e sue proprietà . Formula delle probabilità totali, formula di Bayes. Indipendenza di eventi e di sigma-algebre. Prove ripetute indipendenti: caso finito e caso infinito . Definizione di liminf e limsup di eventi e lemma di Borel-Cantelli.

Distribuzioni discrete unidimensionali: Variabili aleatorie, distribuzione e sue proprietà , densità di probabilità discrete, media, varianza e momenti di ordine k . Distribuzioni discrete notevoli: distribuzione uniforme, di Bernoulli, binomiale, geometrica, ipergeometrica, binomiale negativa, di Poisson. Distribuzioni discrete multidimensionali. Vettori aleatori multidimensionali. Distribuzione congiunta e distribuzione marginale di variabili aleatorie discrete. Indipendenza e non-correlazione. Covarianza e coefficiente di correlazione. Casi notevoli: distribuzione multinomiale e ipergeometrica multinomiale. Somma di variabili aleatorie e alcuni casi notevoli. Distribuzioni discrete condizionate e valore atteso condizionato.

Distribuzioni assolutamente continue unidimensionali. Funzione di ripartizione, caso singolare, caso assolutamente continuo e densità di probabilità , media, varianza e momenti di ordine k . Distribuzioni assolutamente continue notevoli: distribuzione uniforme, esponenziale, Gaussiana, Gamma, di Cauchy, Chi-quadro, Beta. Variabili aleatorie con distribuzione mista. Stima delle code per variabili normali.

Distribuzioni assolutamente continue multidimensionali. funzione di ripartizione congiunta e marginale, densità di probabilità congiunta e marginale, Indipendenza e noncorrelazione. Covarianza e coefficiente di correlazione. Densità della somma di variabili aleatorie e alcuni casi notevoli. Esempi: distribuzione di Dirichlet, distribuzione Gaussiana n -dimensionale, distribuzione uniforme multidimensionale.

Riordinamento di numeri aleatori indipendenti con distribuzione uniforme in $[0, 1]$. Distribuzioni assolutamente continue condizionate e valore atteso condizionato.

Funzione generatrice di probabilità , funzione generatrice dei momenti, funzione caratteristica. Disuguaglianza di Markov e di Chebicev, disuguaglianza di Jensen. Convergenza di variabili aleatorie e teoremi limite. Convergenza quasi certa, in probabilità , in media di ordine p , in distribuzione: esempi, proprietà e relazioni tra i diversi tipi di convergenza. Casi particolari: convergenza della distribuzione geometrica a quella esponenziale, convergenza della distribuzione binomiale a quella di Poisson, il teorema di De Moivre Laplace. Legge dei grandi numeri (in senso debole e in senso forte). Teorema del limite centrale e applicazioni.

Modalità di esame :

Prova scritta

Criteri di valutazione :

Esame finale (100%)

Testi di riferimento :

A. DasGupta, *Probability for Statistics and Machine Learning: Fundamentals and Advanced Topics*. : Springer, 2011

S. Ross, *Calcolo delle Probabilità* . Milano: Apogeo, 2013

Eventuali indicazioni sui materiali di studio :

Le lezioni del corso saranno disponibili agli studenti sul sito di moodle

COMPUTATIONAL FINANCE

(Titolare: Prof. MASSIMILIANO CAPORIN)

Periodo:

1 anno, 1 semestre

Indirizzo formativo:

Corsi comuni

Tipologie didattiche:

18A+46L; 9,00 CFU

Prerequisiti :

Elements of Economics and Mathematics of Financial Markets, elements of Statistics and Econometrics.

Conoscenze e abilità da acquisire :

The course, based on two modules, aims at providing to the students the ability to address computational problems and issues in the broad area of finance. Emphasis will be given to the asset allocation framework. At the end of the course students will become advanced users of a statistical software enabling them to formalize and solve the computational problem related to an empirical finance question. The main module of the course will cover the formalization of computational problems into a statistical package.

The minor introductory module (first 10-12 lectures) will focus on an introduction to the financial economic theories and models needed to understand the main quantitative module.

Attività di apprendimento previste e metodologie di insegnamento :

Theoretical lectures and empirical computer sessions.

Contenuti :

Introduction (minor module)

- Introduction to financial instruments and markets;
- Investment choices under uncertainty and the approach of Markowitz;
- Market equilibrium, CAPM and APT, and market efficiency.

Main module:

1. The formalization of computational problems into a statistical package
2. Asset Allocation: from the approach of Markowitz to Risk Budgeting
3. Backtesting and performance evaluation
4. Introduction to Market Risk Management

The program might be subject to changes depending on a number of elements including: the interest of the students and their ability to solve computational problems with the statistical software; the occurrence of particular events in the financial markets. Changes to the program content will affect the list of tasks included in the team work.

Modalità di esame :

The exam will be given in the form of a group homework. Each group (a team), will receive, at a beginning of the course (groups will be formed within the first two weeks of lectures), a list of tasks pointing at computational finance questions. Each team will have to coordinate activities, inducing team members to interact. During the exam session, each team will show results in the form of a presentation (PowerPoint-like). Each team member must have full knowledge of the presentation and of the analyses performed by the

team and of the main findings.

Criteri di valutazione :

The evaluation of the group homework will be based on the following criterias:

- presence of appropriate answers to the various tasks assigned to the team;
- appropriateness of the quantitative tools adopted by the team;
- interpretation/economic intuition of the results obtained;
- interaction across team members.

Testi di riferimento :

Hull, J.C., *Options, Futures and other derivatives.* : Prentice Hall,
Roncalli, T., *Introduction to risk parity and budgeting.* : Chapman & Hall,
Bodie, Z., Kane, A. and Marcus, A.J., *Investments.* : McGraw Hill,
Hull, J.C., *Risk management and financial institutions.* : Wiley Finance,
Barucci, E., Marsala, C., Nencini, M., and Sgarra, C., *Ingegneria finanziaria.* : Egea,
Elton, E.J., Gruber, M.J., Brown, S.J., and Goetzmann, W.N., *Modern Portfolio Theory and Investment Analysis.* : Wiley,

Eventuali indicazioni sui materiali di studio :

Lecture notes will be distributed to students

Computer sessions and example codes will also be made available as well as the data sets used.

GLI OBIETTIVI DELLA RICERCA CLINICA IN ONCOLOGIA

(Titolare: Prof. PIERFRANCO CONTE)

Periodo: I anno, 2 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 64A; 9,00 CFU

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

INTRODUZIONE ALLA BIOLOGIA

(Titolare: Dott. LIBERO VITIELLO)

Periodo: I anno, 1 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 64A; 9,00 CFU

Prerequisiti :

Nessuno

Conoscenze e abilità da acquisire :

Obiettivo del corso Ã 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 piÃ¹ lezioni seminariali su uso della statistica in vari campi della biologia (a seconda degli interessi espressi dagli studenti durante il corso)

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;

Illustrare la funzione della cellula e descriverne i 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: 1 anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00 CFU

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 pi¹ campioni di riferimento utilizzando la banca dati AIDA.

Attività di apprendimento previste e metodologie di insegnamento :

Lezioni frontali del docente.

Testimonianze aziendali.

Discussioni in aula.

Lavori di gruppo.

Laboratorio per l'utilizzo della banca dati.

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 :

Prova scritta individuale a domande con risposte multiple. Elaborato relativo al caso aziendale. L'elaborato, risultato di un lavoro di squadra, viene prodotto in formato presentazione e come testo scritto, e discusso in aula prima della valutazione finale.

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 e documenti, resi disponibili nel sito all'inizio e durante il corso.

MARKETING QUANTITATIVO

(Titolare: Dott. CLAUDIO MOSCARDINI)

Periodo: I anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00 CFU

Prerequisiti :

Non sono richiesti prerequisiti

Conoscenze e abilità da acquisire :

L'insegnamento vuole immergere lo studente in un ufficio analisi di marketing che supporta per tutti gli aspetti quantitativi i marketing managers nella ideazione, implementazione e valutazione delle azioni di marketing.

Nell'insegnamento verranno presentati gli strumenti di analisi quantitativa pi¹ rilevanti e usati e verranno sottolineate le tematiche di marketing per cui i metodi statistici e gli strumenti di analisi statistica sono indispensabili e quelli in cui sono utili.

Attività di apprendimento previste e metodologie di insegnamento :

Lezioni frontali. Case studies in laboratorio

Contenuti :

1) Il marketing analitico, strategico e operativo

2) Il Piano di marketing

- Il posizionamento di brand e di prodotto

- La valutazione quantitativa del mercato

- Le ricerche di mercato obiettivi, metodologia, utilizzo dati

- Analisi comportamenti dei clienti

- Pianificazione domanda

- Pricing: redditività e sostenibilità economica

- La customer base. Come e perché segmentare

- Marketing della customer base: loyalty, upselling, cross selling and pricing customization

3) La valutazione delle azioni di marketing: sistemi di misurazione e controllo

- Reporting: scelta dei kpi e rolling forecast

- Misurazione del valore generato dall'offerta: ex post analysis

- One2one: selezione, targeting e definizione campagne

- Misurazione del valore generato da One2one Marketing: ex post target e controllo

4) Impatto di internet e nuovi tools

- Il ruolo e l'utilizzo dei social media

- Big data e gli strumenti di analisi

Modalità di esame :

Prova orale

Criteri di valutazione :

Correttezza e qualità delle prove d'esame.

Testi di riferimento :

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio :
Materiale didattico disponibile sulla pagina web del corso.

METODI E MODELLI STATISTICI PER LA FINANZA

(Titolare: Prof. FRANCESCO LISI)

Periodo: Il anno, 1 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 34A+30L; 9,00 CFU

Prerequisiti :

Serie storiche economiche.

Pur non essendo un prerequisito stringente, \hat{A} fortemente consigliato Serie storiche finanziarie.

Conoscenze e abilità da acquisire :

Lo scopo del corso \hat{A} 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 sar \hat{A} illustrata tramite lâ€™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 lâ€™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 baser \hat{A} sulla comprensione degli argomenti svolti sull'acquisizione dei concetti e delle metodologie proposte e sulla capacit \hat{A} 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. MASSIMO MELUCCI)

Periodo: I anno, 2 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 64A; 9,00 CFU

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 e' utile, ma non strettamente necessaria.

Conoscenze e abilità da acquisire :

Sâ€™intende fornire le conoscenze concrete di metodi e strumenti informatici affinché uno studente possieda maggiore competenza in Statistica di un informatico e maggiore competenza in Informatica di uno statistico. Particolare enfasi sar \hat{A} posta su programmazione e gestione dei dati.

Attività di apprendimento previste e metodologie di insegnamento :

I contenuti saranno trattati in forma prevalentemente laboratoriale mediante lo sviluppo di programmi e lâ€™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, SQL, XML, JSON.
3. Strutture di dati fondamentali: liste, hash, grafi, alberi.
4. Algoritmi fondamentali: ricorsione, ricerca, ordinamento.
5. Architetture distribuite, parallele e cloud: Hadoop, MapReduce.
6. Motori di ricerca: crawling, indexing, ranking, Relevance Feedback.
7. Reti, link e click-through: WWW, Link Analysis, HITS, Pagerank, Learning to Rank.
8. Latent Semantic Analysis, decomposizione.
9. Topic Modeling e dati in streaming: Latent Dirichlet Allocation, Probabilistic Latent Semantic Analysis, Streaming (Twitter), Spark.

Modalità di esame :

Data la natura e le modalità d'insegnamento, la prova d'esame sar \hat{A} orale e si baser \hat{A} sulla discussione di un progetto d'approfondimento.

delle tematiche affrontate. Il progetto preveder  la programmazione e la gestione dei dati.

Criteri di valutazione :

Si valuter  la comprensione delle problematiche e la capacit  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 :

Leskovec, Juri; Rajaraman, Anand; Ullman, Jeffrey D., *Mining Massive Datasets.* : Cambridge University Press, 2014

Melucci, Massimo, *Information Retrieval.* : Franco Angeli, 2013

Aho, Alfred; Ullmann, Jeffrey D., *Fondamenti di informatica.* : Zanichelli,

Atzeni, et al, *Basi di dati.* : McGraw-Hill, 2014

Eventuali indicazioni sui materiali di studio :

Materiale didattico sar  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 IL MARKETING

(Titolare: Prof.ssa FRANCESCA BASSI)

Periodo: Il anno, 1 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 58A+6L; 9,00 CFU

Prerequisiti :

Calcolo delle probabilit , Statistica progredito

Conoscenze e abilita' da acquisire :

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

Attivit  di apprendimento previste e metodologie di insegnamento :

Le lezioni sono frontali tenute dal docente. Sono previste esercitazioni in aula informatica e testimonianze aziendali di esperti provenienti

Contenuti :

1. Analisi dei comportamenti di acquisto

La misura della fedelt  alla marca.

Modelli di scelta tra marche.

Indicatori della frequenza di acquisto.

2. La misura dei fenomeni non direttamente osservabili rilevanti nel marketing

La misura della soddisfazione del consumatore.

I modelli a classi a latenti.

Studi di caso.

3. Misure dell'efficacia della comunicazione pubblicitaria

Tipologie di risposta del consumatore all'azione pubblicitaria.

Percezione e memorizzazione del messaggio.

Modelli di risposta delle vendite e delle quote di mercato.

Studi di caso.

4. La segmentazione del mercato

Le fasi operative.

Schema a priori e tecniche statistiche di segmentazione binaria e multipla, modelli loglineari.

Impiego della cluster analysis nella segmentazione a posteriori.

La conjoint analysis nella segmentazione flessibile.

Studi di caso.

5. Il posizionamento di prodotti e marche

Strategie di posizionamento.

L'analisi fattoriale, analisi della correlazione canonica, analisi delle corrispondenze per lo studio delle preferenze dei consumatori.

Tecniche di multidimensional scaling per la formazione di mappe di percezione dei consumatori.

Studi di caso.

Modalita' di esame :

L'esame   orale con eventuale homework

Criteri di valutazione :

Con l'esame e gli homework si valuta se lo studente ha appreso i concetti teorici introdotti durante le lezioni e se   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.

Testi consigliati

East R. *Comportamento del consumatore*, Apogeo, 2003 (cap. 1, 2, 3, 4, 10).

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 (cap 9, modelli di risposta delle vendite e delle quote di mercato; cap. 12, modelli stocastici).

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 (cap. 3 modelli di risposta delle vendite e delle quote di mercato).

Franses P.H., Paap R. *Quantitative Models in Marketing Research*, Cambridge, Cambridge University Press, 2001 (cap 5, modelli razionali).

Eliashberg J., Lilien G.L. (a cura di) (1993) *Handbooks in Operations Research and Management Science. Marketing*, Vol. 5, North Holland, cap 6.

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 CFU

Prerequisiti :

Modelli Statistici 1, Statistica 2

Conoscenze e abilita' da acquisire :

La necessit  di inferire nessi causali da osservazioni empiriche   sempre pi  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 e' consentire allo studente di individuare di volta in volta il metodo pi  adatto alla soluzione di un problema concreto e di utilizzarlo ed interpretarne i risultati in modo appropriato.

Attivit  di apprendimento previste e metodologie di insegnamento :

Lezioni teoriche frontali e studi di caso con software statistici in aula ASID

Contenuti :

1) Introduzione

- inferenza statistica e causalit 
- 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

3) Modelli ad equazioni strutturali (SEM)

- modelli strutturali e causalit 
- specificazione, identificazione e stima di modelli complessi
- effetti diretti e indiretti: path analysis, mediazione e moderazione
- modelli con variabili latenti e/o errori di misura

Modalit  di esame :

Prova pratica in aula informatica e discussione dei risultati.

Criteri di valutazione :

Lo studente deve dimostrare autonomia e spirito critico nell'applicare le metodologie acquisite nel corso per la soluzione di 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, J rn-Steffen, *Mastering Metrics: The path from cause to effect*. Princeton: Princeton University Press, 2015

Eventuali indicazioni sui materiali di studio :

Materiale didattico distribuito dal docente

MODELLI STATISTICI 2

(Titolare: Prof.ssa ALESSANDRA SALVAN) - Mutuato da: Laurea in Statistica per le Tecnologie e le Scienze (Ord. 2014)

Periodo: I anno, 1 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 48A+16L; 9,00 CFU

Prerequisiti :

Istituzioni di Analisi Matematica

Algebra Lineare

Istituzioni di Probabilit 

Statistica 1 e 2

Modelli Statistici 1

Conoscenze e abilita' da acquisire :

Il corso mira ad approfondire teoria e applicazioni dei modelli di regressione con particolare riferimento ai modelli lineari generalizzati.

Sono trattati modelli per dati continui, binari, categoriali e di conteggio. Sono forniti alcuni elementi introduttivi ai modelli per dati correlati. Il corso tratta inoltre gli strumenti necessari per l'analisi dei dati utilizzando modelli di regressione, tramite il software statistico R.

Attivit  di apprendimento previste e metodologie di insegnamento :

Il corso verr  erogato per mezzo di lezioni di teoria ed esercitazioni in laboratorio informatico.

Contenuti :

I modelli lineari generalizzati (Generalized Linear Models, GLM)

- Famiglie esponenziali, di dispersione esponenziali e GLM: modelli, momenti, funzione di legame e verosimiglianza.
- Inferenza sui parametri di un GLM (stima puntuale, verifica d'ipotesi e regioni di confidenza)
- Adeguatezza dei modelli: devianza e residui. Selezione del modello.
- Modelli per dati binari.
- Modelli per risposte politomiche non ordinali e ordinali.
- Modelli per dati di conteggio: schemi di campionamento, modelli di regressione Poisson, tabelle di contingenza e modelli log-lineari.
- Sovradispersione con dati binari e di conteggio: diagnosi e modelli mistura; regressione beta-binomiale e binomiale negativa.
- Modelli per dati di conteggio con inflazione di zeri.
- Inferenza basata su equazioni di stima e quasi-verosimiglianza.
- Modelli per risposte correlate: modelli marginali, risposte normali multivariate, equazioni di stima generalizzate, MLG misti (con effetti fissi e casuali).

Modalità di esame :

Esame scritto in aula informatica (si richiede l'uso di R per lo svolgimento di alcuni calcoli).

Criteri di valutazione :

Si valuteranno la preparazione dello studente sui contenuti oggetto del corso, la sua capacità di analizzare le caratteristiche dei modelli e di interpretare e valutare criticamente i risultati delle analisi svolte.

Testi di riferimento :

- Agresti, A., *Foundations of Linear and Generalized Linear Models*. Hoboken: John Wiley & Sons Inc, 2015
Dobson, A. and Barnett, A., *An Introduction to Generalized Linear Models, Third Edition*. Boca Raton, FL: Chapman and Hall/CRC, 2008
Madsen, H. and Thyregod, P., *Introduction to General and Generalized Linear Models*. Boca Raton, FL: Chapman and Hall/CRC, 2010
Azzalini, A., *Inferenza Statistica: una Presentazione basata sul Concetto di Verosimiglianza*. Milano: Springer-Italia, 2001
Pace, L., Salvan, A., *Introduzione alla Statistica - II. Inferenza, Verosimiglianza, Modelli*. Padova: Cedam, 2001
Bortot, P., Ventura, L., Salvan, A., *Inferenza Statistica: Applicazioni con S-Plus e R*. Padova: Cedam, 2000
Pace, L., Salvan, A., *Teoria della Statistica: Metodi, Modelli, Approssimazioni Asintotiche*. Padova: Cedam, 1996

Eventuali indicazioni sui materiali di studio :

Materiale didattico sarà reso disponibile sulla pagina web del corso.

MODELLI STATISTICI PER DATI ECONOMICI

(Titolare: Prof. MASSIMILIANO CAPORIN)

Periodo: Il anno, 1 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 42A+22L; 9,00 CFU

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, di interpretazione economica dei risultati di un modello e 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 $d\hat{\epsilon}^{TM}$ 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 :

Esame scritto.

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.

Testi di riferimento :

- Greene W. H., *Econometric Analysis (7th edition)*. : Prentice Hall, 2012
Baltagi, B. H., *Econometric Analysis of Panel Data (4th edition)*. : Wiley, 2008
Tsay R. S., *Multivariate Time Series Analysis With R and Financial Applications*. : Wiley, 2014
Durbin J. and Koopman S.J., *Time Series Analysis by State Space Methods (2nd edition)*. : Oxford University Press, 2012
Petris G., Petrone S., Campagnoli P., *Dynamic Linear Models with R*. : Springer, 2009
Lutkepohl, H., *New Introduction to Multiple Time Series Analysis*. : Springer, 2007

Eventuali indicazioni sui materiali di studio :

Testi principali di riferimento, testi ed articoli di consultazione.

Materiale didattico integrativo reso disponibile durante il corso.

MODELLI STATISTICI PER DATI SOCIALI

(Titolare: Prof.ssa FAUSTA ONGARO)

Periodo: I anno, 1 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 42A+22L; 9,00 CFU

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 (programmazione e analisi) nelle quali si userà il software SAS

Contenuti :

1. Modelli per dati di durata

- Richiami di nozioni base per analisi dati durata: concetti base; funzioni base nel continuo e nel discreto; metodi non parametrici
- 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, cenni ad altri modelli a rischio non monotono)
- Modelli a tempo discreto: person period e modello a odds proporzionali (base e estensioni).

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 :

Prova scritta composta da test su teoria e da esercitazione SAS al computer.

Eventuale prova orale successiva potrebbe essere richiesta ad integrazione della valutazione della prova scritta.

Criteri di valutazione :

La valutazione mira a stabilire se e in quale misura lo studente ha appreso gli aspetti non solo formali ma anche interpretativi della modellistica proposta e qual è la sua capacità di applicarli correttamente in situazioni reali.

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.

OTTIMIZZAZIONE STOCASTICA

(Titolare: Prof. GIOVANNI ANDREATTA)

Periodo: I anno, 2 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 50A+14L; 9,00 CFU

Prerequisiti :

Una buona conoscenza e comprensione della lingua inglese sia scritta che parlata.

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 solo in termini probabilistici.

Attività di apprendimento previste e metodologie di insegnamento :

Lezioni teoriche in aula. Esercitazioni in laboratorio.

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

Modalita' 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 3-4 persone) su un argomento da concordare con il docente.

Testi di riferimento :

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

Eventuali indicazioni sui materiali di studio :

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

PROCESSI STOCASTICI

(Titolare: Prof. MARCO FERRANTE)

Periodo: I anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00 CFU

Prerequisiti :

Un corso base di Calcolo delle Probabilit 

Conoscenze e abilita' da acquisire :

Conoscenza approfondita delle catene di Markov a tempo discreto e tempo continuo, con capacita di risolvere autonomamente esercizi e problemi anche di livello avanzato.

Attivita' di apprendimento previste e metodologie di insegnamento :

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

Contenuti :

Definizione di processo stocastico. Probabilit  condizionata e valore atteso condizionato. Indipendenza condizionata.

Catene di Markov a tempo discreto: definizione. Matrice di transizione, leggi congiunte e propriet  di Markov. Random Walk e sue propriet . Tempi di arresto e propriet  di Markov forte. Probabilit  e tempo medio di assorbimento. Classificazione degli stati. Distribuzioni invarianti. Teorema di Markov. Periodicit . Teorema ergodico.

Processo di Poisson: costruzione del processo e definizioni equivalenti. Principali propriet  ed alcune importanti applicazioni.

Catene di Markov a tempo continuo: definizione. Matrice generatrice. Principali propriet , classificazione degli stati, probabilit  e tempo medio di assorbimento, distribuzioni invarianti. Teorema ergodico.

Applicazioni: Processi di nascita e morte. Teoria delle code. Modelli legati allo sport e all'Information Retrieval

Modalita' di esame :

Esame scritto

Criteri di valutazione :

Homeworks (10%) - Esame finale (90%)

Testi di riferimento :

J.Norris, *Markov Chains*. Cambridge: Cambridge University Press, 1996

Paolo Baldi, *Calcolo delle probabilit * (2 ed.). Milano: McGraw-Hill, 2011

PROVA FINALE

(Titolare: da definire)

Periodo: Il anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: ; 20,00 CFU

SISTEMI INFORMATIVI

(Titolare: Prof.ssa MARISTELLA AGOSTI) - Mutuato da:

Periodo: I anno, 1 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 48A; 6,00 CFU

Conoscenze e abilita' da acquisire :

Competenze critiche e avanzate di ideazione, progettazione, realizzazione e valutazione di sistemi di reperimento dell'informazione (Information Retrieval: IR) e motori di ricerca (search engines).

Attivita' di apprendimento previste e metodologie di insegnamento :

L'insegnamento si basa su lezioni frontali e attivit  sperimentali.

La frequenza alle lezioni   fortemente consigliata, perch  il corso   progettato per studenti che frequentano con costanza le lezioni, seguono le attivit  sperimentali del corso e realizzano il progetto.

Informazioni relative al diario delle lezioni, al materiale didattico, alle scadenze e modalit  d'esame sono rese disponibili dal docente agli studenti in Moodle e/o nel Web.

Contenuti :

I sistemi e i servizi di reperimento dell'informazione sono sistemi pervasivi che vengono utilizzati ogni giorno in attivit  di acquisizione di informazioni, di gestione amministrativa, di acquisizione di competenze, di intrattenimento e comunicazione da parte di milioni di utenti

esperti e finali. I motori di ricerca e i sistemi di reperimento dell'informazione gestiscono grandi quantità di informazioni, documenti e risorse digitali (big data).

I motori di ricerca (search engines) sono fra gli strumenti di reperimento dell'informazione più conosciuti e utilizzati per acquisire informazioni su persone, eventi e risultati scientifici per prendere decisioni e per svolgere attività di lavoro e personali da pagine Web e da vari altri tipi di documenti disponibili in formato digitale (ad esempio: quotidiani, notizie, riviste scientifiche, documenti d'archivio e di biblioteche generali e specialistiche).

Sono strumenti di reperimento dell'informazione anche quelli incorporati nei diversi media sociali (social media) che vengono utilizzati ogni giorno per scopi professionali e personali da utenti di tutto il mondo e tutti quei sistemi che permettono di recuperare informazioni dai messaggi di posta elettronica, dai documenti e dagli archivi presenti sui dispositivi di memorizzazione permanente utilizzati nei diversi strumenti di elaborazione delle informazioni che le persone oggi hanno in uso personale.

I contenuti del corso permettono di acquisire competenze per la ideazione, progettazione, realizzazione e valutazione di sistemi di reperimento dell'informazione e motori di ricerca:

- Elementi introduttivi (quadro d'insieme) per la rappresentazione, gestione e reperimento automatico dell'informazione in formato digitale
- Indicizzazione: strutture dati idonee al reperimento dell'informazione
- Modelli e sistemi per il reperimento dell'informazione
- Valutazione: collezioni sperimentali, misure di efficacia e di efficienza
- Web search: la struttura del Web, la stima della dimensione del Web, interrogazioni e utenti, agenti Web (Web Crawlers, Crawler), Search Engine Optimization (SEO).

Modalità di esame :

1. Progetto su un argomento scelto in accordo con il docente e con presentazione seminariale.
2. Esame scritto con domande sui contenuti del corso.
3. Eventuale integrazione orale.

Criteri di valutazione :

La prova finale serve a verificare l'acquisizione di competenze relative alla capacità di valutare, ideare, progettare, ristrutturare, realizzare e gestire sistemi di reperimento dell'informazione e motori di ricerca.

Il progetto serve a verificare in modo sperimentale la capacità di applicare le competenze acquisite.

Testi di riferimento :

Croft, W. Bruce; Metzler, Donald; Strohm, Trevor, Search engines: information retrieval in practice. Boston: Pearson, 2010

Eventuali indicazioni sui materiali di studio :

Indicazioni utili verranno date agli studenti all'inizio delle lezioni.

Il docente fornisce materiali di studio in Moodle e/o nel Web.

SOCIAL MEDIA

(Titolare: Dott. ANDREA SCIANDRA)

Periodo: I anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00 CFU

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

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"

- I dati di testo

Lo stemming

Lo scoring

Il tagging automatico e umano

Classificazione individuale e aggregata

Cenni al text mining e all'analisi statistica dei testi

Esempi con acquisizione testi da link esterni

- Sentiment analysis

Dati provenienti da "social networking sites"

Lo studio dei sentimenti e delle emozioni nel web

Strumenti, potenzialità e limiti
Esempi applicativi

- L'analisi delle reti sociali

Concetti

Tecniche di rilevazione

Reti personali

Indici e misure

Esempi e utilizzo negli ambiti sociali, aziendali e di marketing

Modalità di esame :

Elaborato individuale/di gruppo e/o prova orale

Criteri di valutazione :

Qualità delle prove di esame

Testi di riferimento :

Ceron, Andrea; Curini, Luigi, *Social Media e Sentiment Analysis*. Milano: Springer, 2014

Chiesi, Antonio M., *L'analisi dei reticoli*. Milano: F. Angeli, 2016

Bolasco, Sergio; De_Mauro, Tullio, *L'analisi automatica dei testifare ricerca con il text mining*. Roma: Carocci Editore, 2013

Eventuali indicazioni sui materiali di studio :

Materiale didattico distribuito dal docente

STATISTICA COMPUTAZIONALE PROGREDITO

(Titolare: Prof. NICOLA SARTORI)

Periodo: Il anno, 1 semestre

Indirizzo formativo: Corsi comuni

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

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 :

Lezioni frontali e laboratori informatici

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 prova scritta e su una eventuale discussione del relativo elaborato.

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

Eventuali indicazioni sui materiali di studio :

Materiale didattico disponibile sulla pagina web del corso

STATISTICA ITERAZIONE

(Titolare: Prof.ssa GIOVANNA MENARDI)

Periodo: Il anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00 CFU

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 offrire agli studenti l'opportunità di sviluppare capacità operative 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 alterneranno 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.

N.B 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 2017), 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 (progredito) 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 :

Data la particolare natura dell'insegnamento, la valutazione non può svolgersi nelle forme tradizionali.

L'accertamento consisterà quindi in

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

Criteri di valutazione :

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

Prerequisiti :

Statistica progredito

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 :

Il corso si avvale di lezioni frontali, esercitazioni e laboratori in aula informatica (SAS e R). Sono inoltre previste delle conferenze didattiche tenute da esperti della materia.

Contenuti :

1. Richiami ai tipi di studio in epidemiologia e approfondimenti. Studi caso-controllo: selezione dei casi e dei controlli, caso-coorte, caso-controllo cumulati, caso-controllo con casi prevalenti. Studi di mortalità proporzionale.
2. Misure di occorrenza di malattia e mortalità: incidenza puntuale, cumulata, prevalenza. Relazione fra incidenza e prevalenza. Analisi del rischio di malattia in funzione di un fattore di esposizione: Rischio relativo, rischio attribuibile, odds-ratio. Relazione fra rischio relativo e odds ratio. Inferenza su incidenza, prevalenza, rischio relativo e odds-ratio. Errore di primo e secondo tipo, calcolo della numerosità campionaria.
3. Tabelle di contingenza e modelli log-lineari (richiami) e loro applicazioni in epidemiologia.
4. Applicazione della regressione logistica a studi caso-controllo: appaiamento 1:1, appaiamento 1:m. Propensity score matching.
5. La meta-analisi. Introduzione alla meta-analisi, come preparare una meta-analisi, il calcolo degli effetti, combinare gli effetti, valutare l'eterogeneità e ricerca dei moderatori, gestire database complessi, il publication bias, riportare i risultati della meta-analisi.

Come ultima parte del corso, sarà approfondito uno dei seguenti temi:

6. Epidemiologia genetica. L'interazione fra fattori ambientali e genetici all'origine della malattia. Natura del coinvolgimento genetico nella malattia. Principali concetti di genetica e tipi di studio in epidemiologia genetica. Approcci epidemiologici allo studio dell'aggregazione familiare, misure di aggregazione familiare. Applicazione di modelli di regressione per dati familiari: equazioni di stima generalizzate e modelli marginali. Approcci genetici allo studio dell'aggregazione familiare: analisi dell'ereditarietà mediante i modelli a componenti de varianza e la path analysis.
7. Epidemiologia ambientale. Analisi descrittiva: analisi temporale e mappatura. Analisi dei tassi di malattia mediante regressione di Poisson o binomiale, effetti di confondimento (tipicamente stato socio-economico). Metodi di analisi dei cluster di malattia: conta nelle celle, adiacenza, metodo delle distanze. Sistemi di sorveglianza delle malattie.
8. Epidemiologia delle malattie infettive: rapporti ospite-parassita, sorgenti e serbatoi d' infezione, veicoli e vettori d'infezione, modalità di

trasmissione. Sorveglianza sanitaria: notifica e sistema informativo delle malattie infettive. Modelli matematici di diffusione delle epidemie.

Modalità di esame :

Esame orale con discussione di un'attività esercitazione finale. L'attività esercitazione finale, sviluppata singolarmente o a coppie, verte sullo studio di un problema basato sull'analisi di dati reali. Il tema è concordato col docente.

Criteri di valutazione :

La valutazione si basa sulle conoscenze manifestate durante l'esame orale e sull'attività esercitazione finale. Il giudizio finale è una media del voto assegnato alle due prove.

Testi di riferimento :

Vineis P., Duca P. e Pasquini, Manuale di metodologia epidemiologica.. : Numero speciale di Epidemiologia e prevenzione n.3, 1987

Agresti A., Categorical Data Analysis.. New Jersey: Wiley, 2002

D.W. Hosmer, S. Lemeshow, Applied Logistic Regression. New York: Wiley, 2000

M. Borenstein et al, Introduction to Meta-Analysis. Chichester, UK: Wiley, 2009

Joseph L. Fleiss, Bruce Levin, Myunghee Cho Paik, Statistical Methods for Rates and Proportions, 3rd Edition. : Wiley, 2003

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 :

Dispense fornite dal docente (che non sostituiscono i testi di riferimento).

Materiali per le esercitazioni.

STATISTICA PER LA TECNOLOGIA E L'INDUSTRIA

(Titolare: Prof.ssa GIOVANNA CAPIZZI)

Periodo: Il anno, 2 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 48A+16L; 9,00 CFU

Prerequisiti :

Elementi di base del controllo statistico della qualità e di programmazione degli esperimenti

Conoscenze e abilità da acquisire :

- Conoscere e utilizzare correttamente alcune tecniche avanzate del controllo statistico dei processi produttivi e dei sistemi tecnologici.

- Conoscere e utilizzare correttamente alcune tecniche avanzate di programmazione degli esperimenti per il miglioramento della performance dei processi.

- Conoscere ed utilizzare correttamente le conoscenze di base della teoria dell'affidabilità dei sistemi

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.

Contenuti :

1. Il controllo statistico di processo

- Introduzione al controllo statistico della qualità : richiami dei concetti di base.

- Controllo statistico di processo "univariato": caso nonparametrico.

- Controllo statistico di processo "multivariato":

- Carte di controllo multivariate (T2 di Hotelling, MEWMA, MCUSUM).

- Tecniche di riduzione della dimensionalità :

- Metodo di Hawkins (regressione delle variabili), metodo basato sulla t di Doganakay e Tucker.

- 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 guasto del processo.

- Sorveglianza dei profili.

2. Utilizzo di alcuni piani sperimentali per il controllo della qualità .

- Randomizzazione, modelli misti, dispositivi a blocchi. Piani fattoriali.

- Applicazioni in presenza di componenti di variazione gerarchica).

3. Affidabilità

- Concetti generali.

- Misure dell'affidabilità ed esempi di calcolo.

Modalità di esame :

Orale

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 :

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio :

- Il materiale usato a lezione 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.ssa ALESSANDRA ROSALBA BRAZZALE)

Periodo: I anno, 2 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 56A+26E; 9,00 CFU

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 :

- Richiami sugli elementi di base dell'inferenza statistica: problemi di stima puntuale, di stima intervallare, di verifica d'ipotesi.
- La funzione di verosimiglianza e sue proprietà (invarianza, disuguaglianza di Wald). Quantità collegate alla verosimiglianza (funzione di punteggio, informazione osservata e attesa) e loro proprietà .
- Famiglie esponenziali.
- Statistiche sufficienti.
- Stimatori di massima verosimiglianza: definizione, esempi, proprietà (equivarianza, consistenza, normalità asintotica).
- Disuguaglianza di Cramer-Rao. Stimatori ottimi tra i non distorti.
- Test del rapporto di verosimiglianza: definizione, esempi; distribuzione asintotica, forme asintoticamente equivalenti; regioni di confidenza collegate.
- Verosimiglianza profilo.
- Lemma di Neyman-Pearson. Test uniformemente più potenti.
- Quantità pivotali e equazioni di stima.
- Effetti di errata specificazione del modello statistico e metodi robusti.
- Inferenza bayesiana parametrica: teorema di Bayes, famiglie coniugate; casi particolari (modello normale-normale, beta-binomiale, pareto-uniforme); intervalli di credibilità e verifica d'ipotesi.

Modalità di esame :

Prova scritta a libro chiuso. Materiale ammesso: penna (blu/nero), calcolatrice, formulario individuale (1 facciata A4).

Criteri di valutazione :

Testo d'esame costituito da 3 esercizi, suddivisi in 4-6 quesiti, con grado di difficoltà comparabile.

Testi di riferimento :

Luigi Pace & Alessandra Salvan, *Introduzione alla statistica - II - Inferenza, verosimiglianza, modelli*. Padova: Cedam, 2001

Adelchi Azzalini, *Inferenza statistica: una presentazione basata sul concetto di verosimiglianza* (2a ed.). Milano: Springer-Verlag Italia, 2001

Eventuali indicazioni sui materiali di studio :

Testi di consultazione:

(in italiano)

- Cifarelli, D.M. e Muliere, P. (1989). *Statistica bayesiana. Appunti ad uso degli studenti*. Gianni Luculano Editore, Pavia.

(in inglese)

- Beaumont, G.P. (1980). *Intermediate Mathematical Statistics*. Chapman & Hall, London.

- Welsh, A.H. (1996). *Aspects of Statistical Inference*. Wiley, New York.

Eserciziari (in italiano):

- Andreatta, G. e Runggaldier, W.J. (1983). *Statistica matematica: problemi ed esercizi risolti*. Liguori Editore, Napoli.

STATISTICAL MODELS

(Titolare: Prof.ssa LUISA BISAGLIA)

Periodo: Il anno, 2 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 64A; 9,00 CFU

Prerequisiti :

Calcolo delle probabilità , Statistica progredito

Conoscenze e abilità da acquisire :

The objective of the whole course is to get students acquainted with the fundamentals, basic properties and use of the most important recent modeling techniques, to gain experience in model building and to get some hands-on experience by analysing some real data by using R, Bugs and other up-to-date statistical software.

Attività di apprendimento previste e metodologie di insegnamento :

Lectures and Laboratories

Contenuti :

Generalized linear mixed models

o Introduction to the course: basic ideas

o Generalized linear models: structure and inference

o Extending GLMs: First instances of models for hierarchical data

o Generalized linear mixed models

- o Introduction to hierarchical models and to GLMMs
- o Likelihood inference in GLMMs
- o Bayesian Hierarchical Models
- o Practical sessions with R and R-Bugs

Time series analysis

- o Introduction. Linear time series models.
- o Linear time series models: model specification.
- o Linear time series models: parameter estimation and forecasting.
- o Introduction to spectral analysis
- o Nonlinear models: an introduction
- o Nonlinear models: Markov-Switching Models and Threshold Autoregression Models
- o Long-memory models. Integer AutoRegressive models

Spatial statistics

1. Introduction to spatial statistics:
2. Estimation and modeling of spatial correlations:
3. Prediction and Interpolation (kriging):
4. Spatio-temporal modeling:
5. Second order spatial models for network data:
6. Gibbs-Markov random fields on networks:
7. Simulation and estimation of a Markov random field on a network:
8. Hierarchical spatial models and Bayesian statistics:

Modalita' di esame :

A written exam for each parts of the course.
Each exam will be marked independently by the corresponding instructor.
At the end of the course, students will receive a final mark based on all 3 exams results.

Criteri di valutazione :

At the end of the course, students will receive only a final mark based on all 3 exams results.

Testi di riferimento :

CONTENUTO NON PRESENTE

Eventuali indicazioni sui materiali di studio :

Mc Cullagh, P & Nelder J.A., *Generalized Linear Models*. New York: Chapman & Hall, 1989.
Gelman, A. & Hill J., *Data Analysis Using Regression and Multilevel/Hierarchical Models*. --: Cambridge University Press, 2007.
Fahrmeir L., Tutz, G., *Multivariate Statistical Modelling Based on Generalized Linear Models*. --: Springe, 2001. chapter 6
McCulloch, C.E., Searle, S.R., *Generalized, Linear and Mixed Models*. --: Wiley, 2001.
Brockwell P.J., Davis R.A., *Introduction to Time Series and Forecasting*. --: Springer, 1996.
Fan J., Yao Q., *Nonlinear time series*. --: Springer-Verlag, 2003.
Tsay R.S., *Analysis of Financial Time Series*. - -: Wiley-Interscience, 2005.
Wei W., *Time Series*
Banerjee, S. ,Carlin, B.P. and Gelfand. A.E (2014) *Hierarchical Modeling and Analysis for Spatial Data*, CRC Press, New York (second edition)
Gaetan, C. and Guyon, X. (2010) *Spatial Statistics and Modeling*, Springer, New York.

STRUMENTI STATISTICI PER L'ANALISI DI DATI AZIENDALI

(Titolare: Dott.ssa MARIANGELA GUIDOLIN)

Periodo: Il anno, 2 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 32A+32L; 9,00 CFU

Prerequisiti :

Formali:
Calcolo delle Probabilit  e Statistica Progredito
Sostanziale:
Analisi dei dati (data mining)

Conoscenze e abilita' 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 attivit  in aula informatica, con lâ€™analisi ed il commento di diversi studi di caso.

Attivit  di apprendimento previste e metodologie di insegnamento :

Il corso verr  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,   vivamente consigliata.

Contenuti :

- Le decisioni in azienda
- Modelli per le decisioni
- Analisi della customer base
- Applicazioni e studi di caso: business intelligence

Modalita' di esame :

Esame scritto

Criteri di valutazione :

La valutazione della preparazione dello studente si baser  sulla comprensione degli argomenti svolti e sulla capacit  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 :

CONTENUTO NON PRESENTE

TEMI E METODI DI POPOLAZIONE E SOCIETÀ

(Titolare: Prof.ssa FAUSTA ONGARO)

Periodo: Il anno, 2 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 38A+26L; 9,00 CFU

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 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)

Periodo: Il anno, 1 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: 48A+16L; 9,00 CFU

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) ai modelli che enfatizzano come la popolazione cambi in funzione dell'interazione di 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 come avvenuto sia nei Paesi sviluppati che nei Paesi in via di sviluppo. Gli studenti conosceranno le principali determinanti della fecondità, mortalità, nuzialità e migratorietà.

L'attività di laboratorio e lo studio di casi accompagneranno le lezioni frontali 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 migrazione 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.

THEORY AND METHODS FOR INFERENCE

(Titolare: Prof.ssa ALESSANDRA SALVAN)

Periodo: Il anno, 2 semestre

Indirizzo formativo: Corsi comuni

Tipologie didattiche: 64A; 9,00 CFU

Prerequisiti :

Insegnamenti del primo anno della Laurea Magistrale, 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 frequentista basata sulla verosimiglianza. Agli studenti è proposto anche un lavoro personale su argomenti di ricerca recenti in tale ambito. Il corso fornisce inoltre un'introduzione all'inferenza bayesiana.

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.
- Funzioni generatrici, approssimazioni di momenti, trasformazioni.

- Verosimiglianza: quantità osservate e attese. Proprietà esatte e riparametrizzazioni.
- Inferenza basata sulla verosimiglianza: teoria asintotica del primo ordine.
- Inferenza bayesiana: approssimazioni asintotiche, distribuzioni a priori non informative, metodi bayesiani empirici, stima e verifica d'ipotesi bayesiane, inferenza bayesiana per il modello lineare.
- Equazioni di stima e pseudo-verosimiglianze: effetti una non corretta specificazione del modello, quasi verosimiglianza, verosimiglianza composita, verosimiglianza empirica.
- Aspetti computazionali in R dell'inferenza di verosimiglianza e bayesiana. Algoritmo EM.
- Riduzione dei dati e del modello: statistiche costanti in distribuzione, ancillari e condizionamento, completezza, pseudo-verosimiglianze per l'inferenza con parametri di disturbo.
- Inferenza secondo il paradigma decisionale.
- Famiglie esponenziali: modelli e inferenza.
- Famiglie di dispersione esponenziale e modelli lineari generalizzati.
- Famiglie di gruppo: modelli e inferenza.

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. Gli studenti di dottorato possono sostenere l'esame solo nell'appello fissato alla fine del corso.

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 web del corso.

TIROCINIO FORMATIVO

(Titolare: Prof. BRUNO SCARPA)

Periodo: Il anno, 2 semestre
Indirizzo formativo: Corsi comuni
Tipologie didattiche: ; 1,00 CFU