

METODI MATEMATICI E STATISTICI PER LE APPLICAZIONI

Laurea Magistrale in Matematica
s.s.d. Mat/07

Il corso mira a fornire i principali strumenti di indagine statistica assieme ad argomenti avanzati che consentono di affrontare questioni di interesse fisico-matematico, economico, industriale e applicativo in genere. Il corso, sebbene inserito nel curriculum industriale, si configura come uno dei principali strumenti di analisi economica e, pertanto, consigliabile anche agli studenti del curriculum economico-finanziario. In ogni caso vengono fornite nozioni che rientrano nel bagaglio culturale di base per chi volesse tentare l'insegnamento della matematica applicata nelle scuole medie superiori.

Gli argomenti sono stati selezionati alla luce di svariati anni di insegnamento del corso di Metodi Matematici e Statistici 2 per la laurea magistrale in Ing. Gestionale, ove era particolarmente sentita la valenza applicativa dell'insegnamento in quanto essenziale per il controllo della qualità, l'analisi economica, la logistica, modelli di previsione stocastica e ottimizzazione.

Modalità di esame: svolgimento di un elaborato di corso, facoltativamente in ambiente MatLab, seguito da un colloquio orale.

Argomenti principali e loro finalità:

- 1) **Statistica descrittiva** che fornisce gli elementi di base per l'analisi di dati grezzi.
- 2) **Stimatori** che consentono di stimare le variabili di interesse a partire da un campionamento.
- 3) **Teoria delle decisioni statistiche** che fornisce i fondamenti scientifici della teoria delle scelte.
- 4) **Teoria della regressione** che analizza come formulare modelli analitici a partire da un insieme di dati. Essa storicamente nasce da studi di biologia e viene impiegata in tutti i settori delle scienze applicate, ad esempio fenomeni fisici, studio di processi economici, sociali, biologici.
- 5) **Analisi della varianza** che consente di analizzare l'influenza dei vari fattori in esperimenti in cui entrano più fattori.
- 6) **Test non parametrici** che si rivelano particolarmente utili nel caso in cui si hanno dati statistici con distribuzioni asimmetriche in cui è dubbia la validità del teorema del limite centrale. Sono particolarmente usati in statistica medica.
- 7) **Correlazione normale** che analizza il grado di dipendenza tra due variabili.
- 8) **Inferenza statistica Bayesiana** che tiene conto anche di valutazioni a priori dello sperimentatore. E' presente in un'ampia gamma di fenomeni.
- 9) **Metodi Monte Carlo.** Sono basati sulla generazione di numeri random con leggi assegnate e consentono di riprodurre fenomeni non deterministici in maniera relativamente semplice e realistica, ad esempio estrazioni da un'urna, moti browniani, diffusione di particelle, trasporto di particelle, dinamica delle popolazioni, dinamica di titoli.
- 10) **Processi stocastici ed equazioni differenziali stocastiche.** Costituiscono il modello matematico per fenomeni non deterministici. Particolare attenzione viene data alle catene di Markov (esempio tipico il problema della rovina del giocatore) e ottimizzazione stocastica nel caso discreto ove non sono applicabili i metodi analitici (tipico esempio il problema del commesso viaggiatore o la sequenza di prelievo di merci da un magazzino).
- 11) **Metodo della massima entropia.** Fornisce un metodo generale basato sulla teoria dell'informazione di Shannon che trova applicazioni in teoria dei segnali, fisica-matematica, riconoscimento di immagini, oltre che essere uno strumento per la risoluzione di equazioni integrali e approssimazione di dati e funzioni.

Tesi in atto correlate:

- 1) *Analisi della produzione di componenti elettronici*, presso la ST-Microelectronics
- 2) *Simulazione di celle solari*, presso l'Istituto di Microelettronica e Microsistemi del CNR, sede di Catania

Tesi di dottorato in atto correlate:

- 1) *Simulazione ed ottimizzazione di celle solari*, Dottorato di Ricerca in Matematica Applicata

Progetti di ricerca in atto correlati al corso

- 1) progetto PON denominato "AMBITION POWER", nel settore Energia e Risparmio Energetico "Sviluppo di tecnologie, prodotti e processi per le energie rinnovabili e/o per l'utilizzo razionale dell'energia e/o per l'efficienza energetica";
- 2) progetto ENIAC JU, denominato **ERG** nell'ambito della call ENIAC 2010 su "Energy for a green society: from sustainable harvesting to smart distribution, equipments, material, design solutions and their applications".

Programma dettagliato

Statistica descrittiva. Rappresentazione di dati e distribuzioni di frequenze. Quantili, moda, mediana, coefficiente di asimmetria e kurtosi. Spazi campionari. Distribuzione delle medie campionarie. Campionamenti con e senza ripetizione.

Stimatori. Stimatori non distorti per media, varianza e proporzione. Intervalli di confidenza per la media, sia nel caso in cui la varianza è nota che incognita, per la varianza e per le proporzioni. Distribuzioni t di Student e chi-quadro. Stima della differenza tra le medie di due popolazioni nel caso in cui sono note le varianze e nel caso di varianze incognite. Stimatori dei momenti e di massima verosimiglianza.

Teoria delle decisioni statistiche. Livello di significatività, errori di I e II tipo, potenza di un test. Inferenza sulla media, quando è nota la varianza. Inferenza sulla media quando la varianza non è nota. Inferenza sulla varianza. Inferenza sulla proporzione. Inferenza statistica sulla differenza delle medie di due popolazioni, quando sono note le varianze. Inferenza statistica sulla differenza delle medie di due popolazioni con le varianze incognite. Inferenza statistica sulla varianze di due popolazioni normali e distribuzione F di Fisher. Inferenza statistica sulle proporzioni in coppie di popolazioni.

Test del chi-quadro. Teorema di Pearson, bontà dell'adattamento ad una curva teorica di probabilità, test sulla indipendenza stocastica di due variabili aleatorie, relazione con il test sulla proporzione, modifica nel caso di stima di parametri.

Regressione lineare. Metodo dei minimi quadrati, retta di regressione, proprietà degli stimatori del coefficiente angolare e dell'ordinata all'origine, test sulla indipendenza del correttore dal predittore, analisi dei residui e coefficiente di determinazione, coefficiente di correlazione lineare.

Complementi di calcolo delle probabilità. Matrice di covarianza, funzioni caratteristiche e loro proprietà, funzioni caratteristiche delle principali distribuzioni di probabilità, criterio di indipendenza basato sulle funzioni caratteristiche, leggi normali multivariate e loro densità, operatori di proiezione ortogonale e loro caratterizzazione geometrica, teorema di Cochran.

Regressione lineare multipla. Stimatori dei coefficienti e loro proprietà, test sulla indipendenza parziale e globale dai predittori, predizione tramite la regressione lineare (valori attesi ed intervalli di confidenza), analisi del modello, caso di errore con varianza non costante, esempi di trasformazione di modelli non lineari in modelli lineari.

Analisi della varianza. Test sulla media per n campioni, formula di Huygens per la somma delle variazioni, tabella ANOVA per esperimenti ad un fattore, analisi della varianza per esperimenti a due fattori con e senza repliche e relative tabelle ANOVA, cenni sul caso generale.

Test non parametrici. Criterio del segno, trasformazione ai ranghi, criterio di Mann-Whitney, criterio di Kruskal-Wallis, cograduazione di Spearman, criterio dei cicli nella causalità.

Correlazione normale. Stimatore di massima verosimiglianza per il coefficiente di correlazione e trasformata di Fisher.

Inferenza statistica Bayesiana. Legge *a priori* e *a posteriori*, stimatore di Bayes, stimatore di massima verosimiglianza *a posteriori*.

Metodi Monte Carlo. Legge dei grandi numeri e teorema del limite centrale, velocità di convergenza, generazione di numeri casuali, moto browniano, limite diffusivo della passeggiata aleatoria, simulazioni MC nel trasporto di particelle e in problemi finanziari.

Processi stocastici. Funzione di transizione, principali processi stocastici continui e discreti, processo dei tempi di attesa.

Catene di Markov. Definizione, classificazione degli stati di una C.M., problemi di assorbimento, catene di nascita e morte, problema della rovina del giocatore, probabilità invarianti, teorema di Markov, criterio di regolarità, l'algoritmo di Metropolis, simulated annealing e ottimizzazione stocastica.

Equazioni differenziali stocastiche. Definizione, proprietà, formula di Ito.

Metodo della massima entropia. Teoria dell'informazione di Shannon, applicazione del metodo della massima entropia a equazioni integrali e applicazioni a modelli di interesse fisico-matematico e finanziario.

La presentazione del corso sarà online sul sito <http://www.dmi.unict.it/~romano/>