

ANALISI MULTIVARIATA

Settore SECS-S/01, codice B003956

6 CFU

Corso di laurea in Statistica
semestre II, anno II

ORARIO: 2 incontri su 4 ore settimanali

DOCENTE: Carla Rampichini

Dipartimento di Statistica Informatica, Applicazioni 'Giuseppe Parenti'

Tel: 055 4237246

e-mail: rampichini@disia.unifi.it

RICEVIMENTO: inviare e-mail per appuntamento

OBIETTIVO: Il corso fornisce un'introduzione all'analisi statistica multivariata, con particolare riferimento a metodi esplorativi. Obiettivo principale del corso è introdurre gli studenti alle tecniche statistiche che permettono lo studio congiunto di più caratteri, al fine di rappresentare le unità statistiche e/o le variabili in uno spazio di dimensione inferiore a quello originale, studiare la relazione tra caratteri o la somiglianza tra unità statistiche. Saranno trattati in particolare i seguenti argomenti: analisi in componenti principali, analisi delle corrispondenze; misure di similarità tra unità e tra variabili, analisi dei gruppi.

Le lezioni teoriche saranno compendiate da esercitazioni, utilizzando il software SAS. Saranno assegnati compiti da svolgere a casa. Gli studenti potranno svolgere i compiti assegnati individualmente utilizzando SAS o altro software statistico appropriato (STATA, R, SPSS, o altro).

PREREQUISITI: Statistica, Algebra Lineare e Geometria Analitica

PROVA D'ESAME. L'esame consiste in una prova pratica e in una prova orale. Per essere ammesso alla prova orale lo studente deve aver superato la prova pratica.

COMPITI A CASA: ogni settimana saranno assegnati dei compiti da svolgere a casa e da restituire la settimana successiva utilizzando la piattaforma Moodle. I compiti consegnati in ritardo non saranno accettati salvo deroghe del docente. È possibile svolgere i compiti in gruppo, tuttavia ogni studente deve scrivere e consegnare il suo compito individualmente.

e-learning

Il calendario delle lezioni, i lucidi, i compiti a casa assegnati settimanalmente, avvisi e altro materiale utilizzato a lezione sono disponibili sulla piattaforma Moodle di Ateneo:

<http://e-l.unifi.it/>

Per accedere alla piattaforma Moodle occorre essere registrati, quindi la prima volta autenticarsi inserendo Username e Password corrispondenti all'Autenticazione unica servizi CSIAF e compilare il form successivo. Per altre informazioni vedere la guida on-line: http://e-l.unifi.it/file.php/1/Guida_rapida_Moodle_Unifi_studenti.pdf

La pagina WEB http://local.disia.unifi.it/rampichini/statistica_multivariata.htm contiene link a materiale per lo studio individuale di R, SAS o STATA.

Programma dettagliato

1. Introduzione all'analisi multivariata. La matrice dei dati (cap. 1)
2. Richiami di algebra lineare. Operazioni di base su vettori e matrici.
3. Introduzione a SAS.
4. La matrice di varianza e covarianza e la matrice di correlazione (cap. 2)
5. Rappresentazioni grafiche dei dati multivariati (cap. 5)
6. Analisi in componenti principali (cap. 6)
7. Analisi delle corrispondenze (cap. 7)
8. Distanze e indici di similarità (cap. 8)
9. Analisi dei gruppi (cap. 9)

COMPETENZE ACQUISITE. Al termine del corso, lo studente sarà in grado di:

1. Identificare e comprendere la struttura dei dati multivariati e di formulare i quesiti di ricerca appropriati nei termini dei parametri di interesse.
2. Comprendere le assunzioni necessarie per l'utilizzo delle metodologie di analisi multivariata studiate nel corso.
3. Implementare l'analisi dei metodi studiati per mezzo di un software statistico.
4. Leggere articoli scientifici e comprendere l'uso delle metodologie di analisi statistica multivariata utilizzate nel corso.

Testo consigliato per l'esame

Zani S. e Cerioli A. (2007) *Analisi dei dati e data mining per le decisioni aziendali*, Milano: Giuffrè, capp. 1, 2, 5, 6, 7, 8, 9

Testi di utile consultazione

Bartholomew, D. J., Steele, F., Moustaki, I. & Galbraith, J. I. (2008). *Analysis of Multivariate Social Science Data, second edition*, Texts in Statistical Science, Chapman and Hall/CRC, Boca Raton, FL.

Everitt B. (2007) *An R and S-plus companion to Multivariate analysis*, London: Springer
Dati e codice in R scaricabili da: <http://biostatistics.iop.kcl.ac.uk/publications/everitt/>

Hair J. F., Anderson R.E., Tatham R.L. & W.C. Black. (2009), *Multivariate data analysis*, 7th edition Upper Saddle River, Prentice-Hall.

Härdle W, Simar L. (2007). *Applied Multivariate Statistical Analysis*, Second Edition. Springer-Verlag Berlin Heidelberg.

Johnson R.A. e Wichern D.W. (2007). *Applied Multivariate Statistical Analysis*. Sixth Edition. Pearson Education International. Capp.1-4, 8,9,11,12.1-12-5

Marchetti G. (1997), *Introduzione all'analisi statistica dei dati multivariati*, dispense ad uso degli studenti, Dipartimento di Statistica "G. Parenti", Firenze.

Krzanowski, W. J. (1990), *Principles of multivariate analysis: a user's perspective*, Oxford, Clarendon press.

Seber G. A. F. (2004). *Multivariate Observations*. Wiley.

Regole di comportamento

Il docente si aspetta dagli studenti un elevato livello di responsabilità e rettitudine accademica. Siccome la validità di un corso dipende in toto dall'integrità del lavoro fatto dagli studenti, è imprescindibile che questi ultimi si comportino in maniera impeccabile nella confezione dei lavori a casa e nel comportamento in classe. Ciò significa che ci si aspetta che gli studenti non arrivino in ritardo, dimostrino rispetto nei confronti del docente e dei colleghi, non disturbino la lezione. In classe non è consentito l'uso di alcuna apparecchiatura elettronica quali telefonini, ipad, ipod e quant'altro. La copiatura e il plagio possono comportare l'abbassamento del voto finale.