**Metodi quantitativi per le decisioni aziendali in banca**

## Prof. Enrico Fabrizi

***OBIETTIVO DEL CORSO E RISULTATI DI APPRENDIMENTO ATTESI***

L’obiettivo del corso è fornire agli studenti un insieme di strumenti statistici e informatici di base per l’analisi e la modellazione dei dati. Gli studenti dovrebbero essere inoltre in grado di orientarsi nella letteratura scientifica laddove si fa ricorso a modelli di regressione e a metodi di classificazione supervisionata e non supervisionata

Al termine del corso gli studenti saranno in grado di:

* svolgere analisi statistiche di routine (statistiche descrittive, rappresentazioni grafiche, test di ipotesi) utilizzando il software R;
* effettuare operazioni elementari di gestione di data base in modo da prepararli per l’analisi statistica;
* stimare modelli di regressione lineare e logit/probit usando il software R, interpretare i risultati e le diagnostiche, mettere in atto strategie di stima alternative nel caso di fallimento delle ipotesi standard ;
* conoscere i fondamenti teorici e applicare con R tecniche di statistical learning legate alla classificazione supervisionata e non supervisionata (alberi di classificazione, random forests, Support Vector Machines, cluster analysis) e saperle applicare nell’analisi dei dati.

***PROGRAMMA DEL CORSO***

1. Richiami di statistica descrittiva e inferenziale: medie, misure di variabilità, intervalli di confidenza, test t per una media, test t per il confronto di due medie, ANOVA ad una via.
2. Introduzione all’ambiente R di analisi statistica dei dati. Elementi di analisi e rappresentazione grafica dei dati basata su R. Matrici dei dati e relazioni tra le variabili: misure di associazione e correlazione e relativi test;
3. Il modello di regressione lineare multipla per l’analisi di dati cross-section; inferenza sui parametri del modello; diagnostica. Inferenza robusta nel modello di regressione lineare.
4. Introduzione al problema della classificazione in statistica. Definizioni essenziali e strumenti per la valutazione dell’accuratezza nella classificazione.
5. Il modello di regressione logistica e il modello di regressione probit. Loro utilizzo come strumento di classificazione. Cenni alla regressione logistica multinomiale.
6. Altre tecniche statistiche di classificazione supervisionata: alberi di regressione, random forests, Support Vector Machines.
7. Cenni a tecniche statistiche di classificazione non supervisionata (cluster analysis)

***BIBLIOGRAFIA***

James G., Witten D., Hastie T., Tibshirani R., *An introducton to statistical learning with applications in R*, Springer Verlag, 2013.

In alternativa (stesso libro, ma tradotto in italiano):

James G., Witten D., Hastie T., Tibshirani R., *Introduzione all’apprendimento statistico, con applicazioni in R*, Piccin editore, 2020. (

Appunti e materiale integrativo fornito dal docente

***DIDATTICA DEL CORSO***

Le lezioni frontali in aula saranno integrate da materiale e video online postati dal docente sulla pagina blackboard del corso. E’ richiesta la partecipazione attiva degli studenti alle lezioni che potranno svolgersi anche nel laboratorio informatico.

***METODO E CRITERI DI VALUTAZIONE***

Prova in laboratorio a cui agli studenti sarà richiesto di compiere semplici analisi dei dati con il software R e commentare i risultati. Gli esercizi potranno essere integrati da domande di natura più teorica. Nello specifico ogni prova sarà articolata in 8 quesiti / problemi ognuno dei quali sarà valutato su una scala da 0 a 3.75 punti.

La prova orale è facoltativa; attraverso la prova orale il punteggio non può essere meodificato di oltre 2 punti (in più o in meno)

Durante il corso verranno assegnati esercizi da svolgere a casa (utilizzando il software R) e che permetteranno agli studenti di accumulare alcuni punti (max 4) che si sommeranno al voto ottenuto nella prova in laboratorio.

***AVVERTENZE E PREREQUISITI***

La frequenza alle lezioni, anche se non obbligatoria, è fortemente consigliata. Prerequisito del corso è la conoscenza delle principali tecniche di statistica descrittiva e inferenziale, ovvero i contenuti del volume Alan Agresti: Statistical Methods for the Social Sciences, 5th Edition (chapters 1-8).

***ORARIO E LUGO DI RICEVIMENTO STUDENTI***

# Gli orari di ricevimento sono disponibili on line nella pagina personale del docente, consultabile al sito <http://docenti.unicatt.it/>. In ogni caso è possibile contattare il docente via mail per ogni necessità.

Il ricevimento degli studenti si tiene normalmente presso lo studio del Docente (o in forma telematica via Teams/Skype), Facoltà di Economia e Giurisprudenza (palazzina di Economia, III piano).