

# Statistica

A.A. 2013/2014

CREDITI (CFU): 9

CORSO DI LAUREA IN ECONOMIA E COMMERCIO (Verona)

DOCENTE: Marco Minozzo ([marco.minozzo@univr.it](mailto:marco.minozzo@univr.it))

TELEFONO: 045-8028234

ORARIO DI RICEVIMENTO: martedì 12:00 – 13:00

DOCENTE ESERCITAZIONI: Mauro Mussini

## Piattaforma E-learning di Ateneo

Il materiale didattico relativo al corso (programma definitivo, esercitazioni, temi d'esame, ecc.) si trova sulla piattaforma E-learning di Ateneo. In caso di problemi di accesso al servizio di E-learning, si prega di contattare i responsabili dell'area informatica.

## Calendario delle lezioni

Il corso è articolato sulle 14 settimane del primo semestre. Le lezioni si terranno dal 23 settembre 2013 al 10 gennaio 2014 per un totale di 80 ore (56 ore di lezione (7 CFU) e 24 ore di esercitazioni (2 CFU)), secondo il seguente calendario:

lunedì	ore 10:10 – 11:50	Aula A
martedì	ore 10:10 – 11:50	Aula A
giovedì	ore 14:00 – 15:40	Aula A
venerdì	ore 08:30 – 10:10	Aula A

In linea di massima, le esercitazioni si terranno il giovedì. Le lezioni sono sospese durante le festività e durante la pausa didattica che va dal 1 novembre 2013 al 15 novembre 2013.

## Obiettivi formativi

Il corso si propone di fornire le tecniche base della statistica descrittiva, del calcolo delle probabilità e della statistica inferenziale a studenti di corsi di laurea in discipline economiche ed aziendali che abbiano già acquisito le indispensabili nozioni di matematica fornite con gli insegnamenti di base. Nel loro insieme queste tecniche forniscono una metodologia di analisi quantitativa utile a fini descrittivi, interpretativi e decisionali, fondata sulla osservazione, rilevazione ed elaborazione dei fenomeni collettivi. Da un punto di vista applicativo, queste tecniche sono indispensabili nell'interpretazione delle informazioni statistiche ufficiali nonché nella realizzazione di indagini statistiche di fenomeni economici e sociali. Oltre a fornire la necessaria strumentazione statistico-matematica, il corso si pone l'obiettivo di fornire anche gli strumenti concettuali necessari per una valutazione critica delle metodologie proposte.

## **Programma**

(il programma definitivo sarà distribuito a fine corso)

### Statistica Descrittiva

Concetti introduttivi; fenomeni collettivi; popolazione e unità statistiche; indagini censuarie e campionarie; il questionario; la raccolta, lo spoglio e la classificazione dei dati; caratteri qualitativi ordinali e non ordinali; caratteri quantitativi discreti e continui; caratteri ciclici; caratteri trasferibili; fenomeni di movimento e fenomeni di stato; fonti statistiche.

Tipi di dati statistici; la matrice dei dati; distribuzioni unitarie semplici, doppie e multiple; distribuzioni di frequenza (assolute) semplici, doppie e multiple; distribuzioni di frequenza relative e percentuali; distribuzioni in classi; densità di frequenza; distribuzioni pesate; distribuzioni di quantità.

Frequenze cumulate e retrocumulate; funzione di ripartizione a gradini per distribuzioni di frequenza; funzione di ripartizione continua per dati in classi.

Rappresentazioni grafiche; grafici a barre e a nastri; areogrammi; cartogrammi; pictogrammi; diagrammi radar; istogramma per distribuzioni in classi.

Sommatorie semplici e doppie; produttorie; operazioni con le sommatorie.

Gli indici di localizzazione; la media aritmetica; la media geometrica; la media armonica; proprietà della media aritmetica: media aritmetica di una trasformazione lineare, somma degli scarti, somma del quadrato degli scarti, media aritmetica del miscuglio; la media quadratica; la media cubica; la media potenziata di quarto ordine e le altre medie potenziate; le medie lasche; la mediana per distribuzioni unitarie, di frequenza e in classi; la mediana come centro di grado 1; quartili, decili, percentili e quantili per distribuzioni unitarie, di frequenza e in classi; moda e classe modale.

I numeri indici a base fissa; cambiamento di base; i numeri indici a base mobile; passaggio dai numeri indici a base fissa a quelli a base mobile; le variazioni relative e la variazione media relativa; i numeri indici complessi di Laspeyres e di Paasche.

La variabilità e gli indici di variabilità; il campo di variazione; la differenza interquartile; gli scostamenti semplici medi dalla media; lo scarto quadratico medio (deviazione standard); la varianza; la varianza di una trasformazione lineare e del miscuglio; la standardizzazione; le differenze medie; gli indici relativi di variabilità: il coefficiente di variazione.

L'asimmetria e gli indici di asimmetria; la curtosi e le misure di curtosi; i momenti dall'origine e i momenti centrali.

Distribuzioni doppie e multiple, unitarie e di frequenza; media aritmetica della somma di più variabili; media aritmetica del prodotto di due variabili; codevarianza e covarianza; varianza della somma di due o più variabili; distribuzioni condizionate; media aritmetica e varianza condizionata; indipendenza; indice di dipendenza  $\chi^2$ ; coefficiente di contingenza  $C$ ; paradosso di Simpson (cenni).

Interpolazione statistica; il metodo dei minimi quadrati per funzioni lineari nei parametri; la retta dei minimi quadrati (per distribuzioni doppie unitarie); il metodo dei minimi quadrati per funzioni riconducibili ad una retta tramite trasformazione delle variabili; il coefficiente di correlazione lineare  $r$  di Bravais; la disuguaglianza di Cauchy-Schwarz; il coefficiente di determinazione  $R^2$ ; decomposizione della devianza totale in devianza spiegata più devianza residua; la retta dei minimi quadrati per distribuzioni di frequenza in tabella a doppia entrata.

## Calcolo delle Probabilità

Modelli deterministici e modelli probabilistici; eventi elementari e spazio campionario; alberi degli eventi; eventi aleatori e operazioni tra eventi.

Elementi di calcolo combinatorio: permutazioni semplici, disposizioni semplici, combinazioni semplici, disposizioni con ripetizione, permutazioni tra elementi non tutti distinti (cenni).

Algebre e sigma-algebre; definizione assiomatica della probabilità; funzione di probabilità; spazi di probabilità; diverse interpretazioni della probabilità; primi teoremi sulla probabilità; legge della somma.

Probabilità condizionata; proprietà della probabilità condizionata; legge del prodotto; indipendenza stocastica tra eventi; partizioni di eventi; formula delle probabilità totali; teorema di Bayes.

Variabili aleatorie; condizione di misurabilità; funzione di ripartizione; proprietà della funzione di ripartizione; variabili aleatorie discrete e continue; distribuzione di probabilità e funzione di densità di probabilità; trasformata di una variabile aleatoria  $Y=g(X)$ ; valore atteso  $E(X)$ , varianza  $Var(X)$ , moda, momenti non centrati e momenti centrati;  $E(X)$  e  $Var(X)$  di trasformate lineari e non di una variabile aleatoria; disuguaglianza di Markov e disuguaglianza di Tchebycheff.

Particolari distribuzioni discrete: uniforme (con  $E(X)$  e  $Var(X)$ ), Bernoulli (con  $E(X)$  e  $Var(X)$ ), binomiale (con  $E(X)$ ), Poisson (cenni), geometrica (cenni).

Particolari distribuzioni continue: rettangolare (con  $E(X)$  e  $Var(X)$ ), normale (standardizzazione, tavole della normale standard, calcolo di probabilità e quantili), esponenziale negativa (cenni); distribuzione lognormale (cenni).

Variabili aleatorie doppie discrete; funzione di ripartizione congiunta; distribuzione di probabilità congiunta; distribuzioni di probabilità marginali e condizionate; indipendenza tra variabili aleatorie; variabili aleatorie multiple (cenni); valore atteso di una funzione di due variabili aleatorie; covarianza  $Cov(X,Y)$ ; coefficiente di correlazione di Bravais  $\rho(X,Y)$ ; valore atteso condizionato  $E(X|Y)$  e varianza condizionata  $Var(X|Y)$ .

Valore atteso e varianza di combinazioni lineari di variabili aleatorie; valore atteso e varianza della media campionaria di variabili aleatorie indipendenti.

Legge (debole) dei grandi numeri (con dimostrazione); legge dei grandi numeri di Bernoulli per frequenze relative.

Teorema del limite centrale per variabili aleatorie indipendenti e identicamente distribuite; approssimazione della distribuzione binomiale alla distribuzione normale.

## Statistica Inferenziale

Introduzione all'inferenza statistica; campioni casuali (probabilistici); variabilità campionaria; statistiche campionarie; media campionaria; varianza campionaria e varianza campionaria corretta; frequenza relativa campionaria; distribuzioni campionarie; distribuzioni chi-quadrato,  $t$  di Student,  $F$  di Fisher.

Stima puntuale; stimatori; proprietà degli stimatori: correttezza, efficienza, consistenza; errore quadratico medio; stima della media di una popolazione normale; stima della varianza di una popolazione normale; stima di una proporzione di una popolazione dicotomica.

Stima per intervallo; intervallo di confidenza per la media di una popolazione normale (con  $\sigma^2$  noto e  $\sigma^2$  incognito); intervallo di confidenza per la media di una popolazione qualsiasi (grandi campioni); intervallo di confidenza per la varianza di una popolazione normale (con media nota e

media incognita); intervallo di confidenza per la proporzione di una popolazione dicotomica (grandi campioni).

Verifica delle ipotesi; ipotesi nulla e ipotesi alternativa; ipotesi semplici e ipotesi composte; errori di I e di II tipo; livello di significatività e potenza; test unilaterali e bilaterali; verifica di ipotesi sulla media di una popolazione normale (con  $\sigma^2$  noto e  $\sigma^2$  incognito); verifica di ipotesi per la varianza di una popolazione normale; verifica di ipotesi per la proporzione di una popolazione dicotomica (grandi campioni); verifica di ipotesi su due proporzioni di popolazioni dicotomiche (grandi campioni); verifica di ipotesi su due varianze di popolazioni normali; verifica di ipotesi su due medie di popolazioni normali (con varianze note, e con varianze incognite ma uguali).

### **Libro di testo**

- G. CICCHITELLI (2012), Statistica: principi e metodi, Seconda edizione, Pearson Italia, Milano.

### **Testi di approfondimento**

- S. BERNSTEIN, R. BERNSTEIN (2003), Statistica descrittiva, Collana Schaum's, numero 109. McGraw-Hill, Milano.
- S. BERNSTEIN, R. BERNSTEIN (2003), Calcolo delle Probabilità, Collana Schaum's, numero 110. McGraw-Hill, Milano.
- S. BERNSTEIN, R. BERNSTEIN (2003), Statistica inferenziale, Collana Schaum's, numero 111. McGraw-Hill, Milano.
- D. PICCOLO (1998), Statistica, Seconda edizione 2000. Il Mulino, Bologna.
- D. PICCOLO (2010), Statistica per le decisioni, Nuova edizione. Il Mulino, Bologna.
- E. BATTISTINI (2004), Probabilità e statistica: Un approccio interattivo con Excel. McGraw-Hill, Milano.
- F. P. BORAZZO, P. PERCHINUNNO (2007), Analisi statistiche con Excel. Pearson Education.
- M. R. MIDDLETON (2004), Analisi statistica con Excel. Apogeo.

### **Guida allo studio del libro di testo**

(la guida definitiva sarà distribuita a fine corso)

Le seguenti parti del libro di testo sono escluse dal programma:

- paragrafi 5.6, 5.7, 9.4, 10.3, 11.2.3 e 16.3.1;
- capitoli 20 e 21;
- appendici A.4, A.6, A.7 e B.6;
- testo a pag. 446 ed esempio 19.5 alle pagg. 446-447;
- combinazioni con ripetizione a pag. 522;
- dimostrazione del valore atteso e della varianza della distribuzione di Poisson alle pagg. 526-527.

Le seguenti parti del libro di testo sono da considerarsi come “lettura”:

- intervallo di confidenza a pag. 441;
- intervallo di confidenza a pag. 444;
- intervallo di confidenza a pag. 447;
- intervallo di confidenza a pag. 449.

L'utilizzo del software Excel non è in nessun modo influente ai fini dell'esame. Allo stesso modo, l'utilizzo del software R è escluso dal programma.

## **Guida allo studio**

Durante lo svolgimento del corso sarà indicato, per ogni specifico argomento, quali parti del libro di testo sono da studiare e quali parti sono da escludere. Un programma definitivo ed una guida definitiva allo studio del libro di testo saranno resi disponibili a fine corso. Si consiglia di seguire le lezioni e le esercitazioni e di prendere regolarmente gli appunti.

## **Conoscenze preliminari**

Per affrontare in modo adeguato gli argomenti trattati nell'insegnamento di Statistica è fondamentale aver già seguito e superato l'esame di Matematica. Nel caso non si fosse già superato l'esame di Matematica (che per questo anno accademico 2013/2014 non è propedeutico all'esame di Statistica), e si volesse ugualmente seguire l'insegnamento di Statistica, si consiglia di sottoporsi (anche con l'aiuto di appunti o libri di testo) alla prova di autovalutazione presente sulla piattaforma E-learning di Ateneo. Se la soluzione degli esercizi di questa prova risultasse particolarmente difficile (o addirittura impossibile), si consiglia di desistere dall'impegnarsi con l'insegnamento di Statistica fino a quando non si siano maturate le adeguate competenze matematiche.

## **Modalità di svolgimento delle lezioni ed esercitazioni**

La maggior parte del corso si svolgerà attraverso una serie di lezioni frontali alle quali gli studenti sono fortemente invitati a partecipare prendendo regolarmente gli appunti. Fanno parte integrante del corso una serie di esercitazioni. Alcune delle esercitazioni, da svolgere a casa individualmente, saranno successivamente corrette in aula. Tutte le esercitazioni sono indispensabili per una adeguata comprensione degli argomenti del corso.

## **Attività di tutorato in aula**

Al termine delle lezioni, in corrispondenza degli appelli di gennaio e febbraio 2014, sono previste alcune esercitazioni aggiuntive in aula (tutoraggio). Altre esercitazioni aggiuntive potrebbero tenersi in corrispondenza della sessione di appelli estiva e della sessione di appelli autunnale. Informazioni più dettagliate a riguardo saranno fornite a tempo debito. Le ore complessive per queste attività di tutorato sono 35.

## **Laboratorio Excel**

Il laboratorio di Excel è un ciclo di lezioni che si tengono in aula informatica sull'utilizzo avanzato del software Excel, rivolto agli studenti dell'insegnamento di Statistica (CdL in Economia e Commercio, sede di Verona). Questa attività verrà erogata attraverso una lezione/esercitazione frontale di 2 ore alla settimana, per 10 settimane del primo semestre, per un totale di 20 ore. L'attività (facoltativa) è rivolta in particolare agli studenti del secondo anno, e riservata a chi abbia raggiunto sufficienti risultati da stabilirsi in base alla carriera (superamento degli esami del primo anno, numero di CFU acquisiti, media dei voti, ecc.). Il programma dell'attività prevede alcune lezioni di introduzione al software ed alle sue principali funzioni. Verranno quindi presi in considerazione alcuni degli argomenti trattati negli insegnamenti di matematica e di statistica quali, ma non solo, l'algebra matriciale, l'ottimizzazione, l'interpolazione e le simulazioni con il metodo Monte Carlo. Nel suo complesso, l'attività si pone l'obiettivo di innalzare il livello delle abilità quantitative ed informatiche degli studenti che vorranno parteciparvi, fornendo delle competenze

che potrebbero essere utili sia per la preparazione del lavoro di tesi sia nel mondo del lavoro. Si sottolinea che questo ciclo di lezioni in aula informatica è assolutamente facoltativo e che non è in nessun modo influente ai fini dell'esame.

### **Corso intensivo estivo di Statistica ad Alba di Canazei**

Come per gli anni scorsi, anche per quest'anno accademico è probabile che venga attivato un corso intensivo estivo di Statistica nella sede di Alba di Canazei dell'Università degli Studi di Verona. Appena possibile si daranno informazioni più dettagliate a riguardo.

### **Orario di ricevimento**

Nel caso di sovrapposizioni (con altre lezioni, ecc.) delle ore previste per il ricevimento studenti, si prega di contattare direttamente il docente.

### **Modalità di esame**

La prova di esame consiste di una prova scritta (di circa 2 ore e 30 minuti) composta da una selezione di esercizi e da una serie di domande a risposta chiusa. Per la prova scritta si potrà usare solamente una calcolatrice e non sarà consentito utilizzare nessun altro materiale (libri, appunti, ecc.). Saranno ammessi alla prova orale (facoltativa) soltanto gli studenti che avranno riportato un voto maggiore od uguale a 15/30 sia negli esercizi che nelle domande a risposta chiusa. Per sostenere le prove lo studente deve presentarsi munito di tessera universitaria, ovvero di libretto universitario, o di idoneo documento di riconoscimento.

Gli studenti che lo vorranno potranno sottoporsi, la seconda settimana del mese di novembre 2013, ad una prova intermedia sulla prima parte del programma relativa alla statistica descrittiva. Questa prova intermedia (della durata di circa 1 ora) consisterà in una serie di domande a risposta chiusa. Dell'esito positivo di tale prova se ne terrà conto in sede di esame e solo per i due appelli della sessione invernale. Il superamento della prova intermedia potrà comportare un innalzamento del voto conseguito in sede di esame fino ad un massimo di tre punti.