

# Corso di Statistica

Corso di Laurea in  
Ingegneria Edile  
ed  
Ingegneria Tessile

Docente: Ilia Negri

Esercitatore: Emilio Porcu

## Orario del corso:

Martedì: dalle 10.30 alle  
12.30

Giovedì: dalle 9.30 alle 11.30

## Ricevimento:

Giovedì: dalle 11.30 alle  
13.30

e-mail: [negri@unibg.it](mailto:negri@unibg.it)

## Programma del corso:

7/10	Presentazione, assiomi probabilità
9/10	Indipendenza e formula di Bayes
14/10	Campionamento, distribuzioni di frequenza e rappresentazioni grafiche
16/10	Moda, mediana, media, box-plot
21/10	Laboratorio con R - Presentazione
23/10	<b>Esercitazione 1</b> - Probabilità
28/10	Varianza e scarto quadratico medio
30/10	<b>Esercitazione 2</b> - Grafici, Media, Varianza
4/11	Analisi di più fenomeni, tabelle di contingenza
6/11	Analisi di regressione
11/11	Variabili casuali discrete, valore atteso
13/11	<b>Esercitazione 3</b> - Analisi più fenomeni
18/11	Variabili casuali continue, valore atteso
20/11	<b>Esercitazione 4</b> - Variabili Casuali
28/11	<b>Prima prova parziale</b>
2/12	Trasformazione di media e varianza
4/12	Inferenza, media e varianza campionarie
9/12	Analisi dati Sperimentali
11/12	<b>Esercitazione 5</b> - Inferenza
16/12	<b>Esercitazione 6</b> - Laboratorio con R
18/12	<b>Esercitazione 7</b> - Analisi dati sperimentali
8/1	Intervalli di confidenza
13/1	Verifica di ipotesi statistiche
15/1	<b>Esercitazione 8</b> - Intervalli di confidenza e verifica ipotesi

# Modalità d'esame:

- I prova parziale:  
2 esercizi (1 ora)  
1 domanda teoria (15 minuti)
- II Prova parziale:  
2 esercizi (1 ora)  
1 domanda teoria (15 minuti)
- Esame completo:  
2+2 esercizi (2 ore)  
2 domande di teoria (30 minuti)

Si accede alla seconda prova parziale con un voto superiore o uguale a 15 nella prima prova parziale

Se la media delle due prove parziali è superiore o uguale a 18 si registra il voto

La prima prova è valida un'anno accademico (fino a settembre 2004)

Nell'esame completo chi ha superato solo la prima prova può decidere di svolgere solo gli ultimi due esercizi e solo la seconda domanda di teoria.

# Libri di Testo:

- S.M. Iacus: *Statistica: 15 lezioni facili*, ed. cusl Milano
- Appunti della docente per le parti del programma non comprese nel testo.

# Altre letture:

- Gonick e Woolcott: *The Cartoon guide to Statistics*.
- Douglas C. Montgomery-George C. Runger- Norma Faris Hubele: *Engineering Statistics*, ed. Wiley, New York.
- Ilia Negri: *Esercizi per il corso di Calcolo delle Probabilità e Statistica*, ed. Cusl Milano.

## Cosa è la Statistica?

*... quello che rende la Statistica unica è la capacità di quantificare l'incertezza;... essa mette gli statistici in grado di fare affermazioni categoriche, cioè in completa sicurezza, circa il loro grado di incertezza*

Alcune definizioni:

**Fenomeno statistico:** l'oggetto della nostra indagine

**Popolazione (statistica):** l'insieme degli *individui* portatori della caratteristica di interesse

**Campione:** sottogruppo della popolazione oggetto dello studio

**Modello:** oggetto teorico che descrive le caratteristiche salienti del fenomeno

Ambiti della statistica e loro interazioni:

**Statistica Descrittiva:** descrive il fenomeno sulla base dei risultati contenuti nel campione. Non si vuole estendere i risultati a tutta una popolazione. Strumenti: sintesi, indici, metodi grafici

**Statistica Inferenziale:** cerca di estendere i risultati ottenuti sul campione all'intera popolazione. Per poter estendere i risultati occorre che il campione sia scelto con criterio (casuale). Strumenti: stimatori, verifiche di ipotesi, intervalli di confidenza.

**Probabilità:** tecniche connesse al trattamento della casualità e alla descrizione dei modelli teorici.

Statistica Descrittiva – Probabilità



Statistica inferenziale

## Esempio 1

Dobbiamo verificare la resistenza del calcestruzzo da utilizzare per la costruzione di un edificio. La normativa prevede che sopra una certa metratura siano effettuati  $n$  provini sui quali valuteremo la resistenza.

**Fenomeno statistico:** resistenza del calcestruzzo

**Popolazione:** tutti i possibili campioni di calcestruzzo

**Campione:** gli  $n$  provini

- Si calcolano alcune quantità sui provini (media, varianza, percentili)
- Si sceglie un modello probabilistico in grado di descrivere il fenomeno (modello gaussiano)
- Si effettuano dei test per poter essere in grado di inferire sulla qualità globale del calcestruzzo



## Esempio 2

Abbiamo due macchine che effettuano la tessitura di un tessuto pregiato. Alla fine della giornata contiamo i difetti prodotti dalle due macchine e sono in numero diverso. Dobbiamo provvedere al controllo della macchina che produce più difetti?

**Fenomeno statistico:** numero dei difetti sulla stoffa prodotta in una giornata

**Popolazione:** le stoffe prodotte alla fine della giornata

**Campione:** per ogni macchina la rilevazione del numero di difetti alla fine di  $n$  giornate lavorative

- Si calcolano alcune quantità sulle stoffe prodotte dalle due macchine (media, varianza, percentili)
- Si sceglie un modello probabilistico in grado di descrivere il fenomeno (modello gaussiano)
- Si effettuano dei test per poter essere in grado di inferire sul funzionamento delle due macchine

# Probabilità

Studio delle leggi del caso.

Esempio: lancio di un dado

*Eventi*: entità per i quali è possibile calcolare le *probabilità*. Si indicano con le lettere  $E$ ,  $A$ , ...

$$E = \{\text{esce un numero pari}\}$$

L'insieme di tutti i risultati dell'esperimento:  $\Omega$   
detto *evento certo* o *spazio campionario*

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Ad ogni evento  $A \subset \Omega$  vogliamo assegnare una probabilità  $P(A)$ . Come?

- Sia  $A$  l'evento esce il numero 7
- Sia  $A$  l'evento non esce nessun numero
- Sia  $A$  l'evento esce un pari
- Sia  $A$  l'evento esce un numero tra -1 e 8
- Sia  $A$  l'evento esce almeno un numero

Alcune regole fondamentali:

- $P(A) = 0$      $A$  evento *quasi impossibile*
- $P(\emptyset) = 0$      $\emptyset$  evento *impossibile*
- $P(A) \geq 0$  per ogni  $A$
- $P(A) \leq 1$  per ogni  $A$
- $P(A) = 1$      $A$  evento *quasi certo*
- $P(\Omega) = 1$      $\Omega$  evento *certo*

Se due eventi  $A$  e  $B$  non possono verificarsi contemporaneamente diremo che sono *incompatibili*

$$A \cap B = \emptyset$$

### **eventi incompatibili**

$$A \cap B = \emptyset \Leftrightarrow P(A \cap B) = 0$$

Se  $A$  e  $B$  sono incompatibili è lecito assumere che

$$P(A \cup B) = P(A) + P(B)$$

Assiomi del calcolo delle probabilità (Kolmogorov 1933)

### **Assiomi della probabilità**

1.  $\forall A \subset \Omega, P(A) \geq 0$
2.  $P(\Omega) = 1$
3. Se  $A \cap B = \emptyset$  allora  $P(A \cup B) = P(A) + P(B)$

# Proprietà della probabilità

- $\bar{A}$  è l'evento *complementare* di  $A$

$$P(\bar{A}) = 1 - P(A)$$

- Se  $A \subset B$

$$P(A) \leq P(B)$$

- per ogni  $A \subset \Omega$

$$0 \leq P(A) \leq 1$$

- Se  $A$  e  $B$  sono eventi compatibili

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- Se  $A_1, A_2, \dots, A_n$  sono tali che  $A_i \cap A_j = \emptyset$ ,  $i \neq j$

$$P(A_1 \cup \dots \cup A_n) = P(A_1) + \dots + P(A_n)$$

ovvero

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$$

## Il caso degli eventi equiprobabili

In genere  $\Omega$  è descritto dagli *eventi elementari*

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\} \quad n \in \mathbb{N}$$

Se

$$P(\{\omega_1\}) = P(\{\omega_2\}) = \dots = P(\{\omega_n\})$$

allora la probabilità di  $A$ , composto da più eventi elementari, è data da

### eventi elementari equiprobabili

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\# \text{ casi favorevoli (all'evento)}}{\# \text{ casi possibili (dell'esperimento)}}$$

Esempio: lancio di due dadi

$$\Omega = \{(i, j) : 1 \leq i, j \leq 6\} \quad |\Omega| = 36$$

$$A = \{(i, j) : (i, j) = (6, 6)\} \quad |A| = 1$$

$$P(A) = \frac{1}{36}$$

# Eventi indipendenti

Due eventi sono indipendenti se il verificarsi dell'uno non influenza il verificarsi dell'altro e viceversa

Se la probabilità  $A$  varia sapendo che si è verificato  $B$  vuol dire che  $A$  “dipende” da  $B$

La probabilità di  $A$  in questo caso diventa  $P(A|B)$

## Probabilità condizionata

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \text{ con } P(B) > 0$$

Abbiamo quindi

## Indipendenza tra eventi

$A$  e  $B$  sono indipendenti se e solo se

$$P(A|B) = P(A) \text{ e } P(B|A) = P(B)$$

$A = \{2,3\}$ ,  $B =$  esce numero pari

$$P(A) = \frac{1}{3}, \quad P(A|B) = \frac{1}{3}, \quad P(B) = \frac{1}{2}, \quad P(A \cap B) = \frac{1}{6}$$

Poiché si ha che, se  $A$  e  $B$  sono indipendenti

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A) \Leftrightarrow P(A \cap B) = P(A) \cdot P(B)$$

allora un'altra definizione di indipendenza è la seguente:

**Indipendenza tra eventi (definizione alternativa)**

$A$  e  $B$  sono indipendenti se e solo se

$$P(A \cap B) = P(A) \cdot P(B)$$

Dalla definizione di probabilità condizionata segue

$$P(A \cap B) = P(A|B) \cdot P(B)$$

$$P(A \cap B) = P(B|A) \cdot P(A)$$



## Soluzione del problema di Chevalier De Mere

Siano

$E =$  “un 6 su 4 lanci” di un dado

$F =$  “un doppio 6 su 24 lanci di due dadi”

$A =$  “un 6 in un lancio”

$B =$  “un doppio 6 in un lancio di due dadi”

Abbiamo  $P(A) = 1/6$  e  $P(B) = 1/36$

Calcoliamo  $P(\bar{E})$  e  $P(\bar{F})$  al posto di  $P(E)$  e  $P(F)$ .

$$P(\bar{E}) = P(\bar{A}) \cdot P(\bar{A}) \cdots P(\bar{A}) \quad (\text{per 4 volte}) = \left(\frac{5}{6}\right)^4$$

$$P(E) = 1 - P(\bar{E}) = 1 - \left(\frac{5}{6}\right)^4 = 0.518$$

$$P(\bar{F}) = P(\bar{B}) \cdot P(\bar{B}) \cdots P(\bar{B}) \quad (\text{per 24 volte}) = \left(\frac{35}{36}\right)^{24}$$

$$P(F) = 1 - P(\bar{F}) = 1 - \left(\frac{35}{36}\right)^{24} = 0.491$$

Quindi  $P(F) < P(E)$