UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II

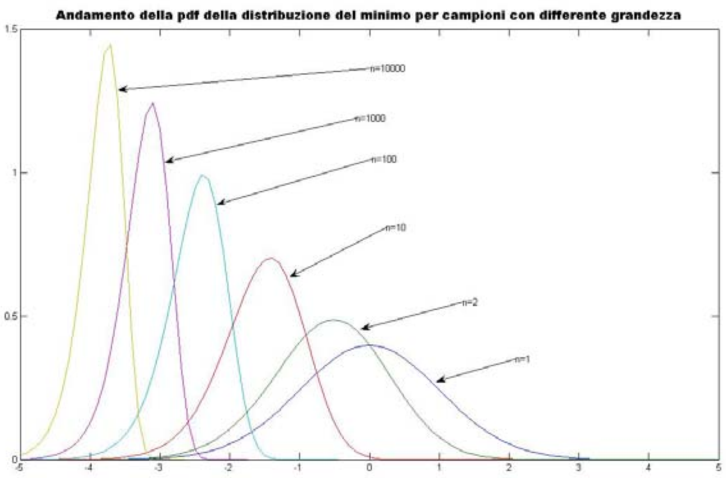
DIPARTIMENTO DI INGEGNERIA AEROSPAZIALE – D.I.A.S.

STATISTICA PER L'INNOVAZIONE

a.a. 2007/2008

DISTRIBUZIONE DEI VALORI ESTREMI

Prof. Antonio Lanzotti



A cura di: Ing. Andrea Colini

andrea.colini@unina.it

Statistica dei Valori Estremi

È applicata con successo per interpretare fenomeni di guasto provocati dai valori estremi (massimo o minimo) di determinati fattori fisici.

Ad esempio per i componenti meccanici soggetti al fenomeno di rottura per "fatica" il cui guasto è innescato dal difetto più grande; oppure i sistemi con struttura a "catena", la cui resistenza dipende ovviamente da quella del loro "anello" più debole.

Oltre che in ambito tecnologico, questi modelli trovano applicazione anche in contesti naturali per interpretare e/o controllare fenomeni quali i massimi annuali delle portate giornaliere dei fiumi (le piene) o i corrispondenti minimi (le secche), oppure le massime ampiezze dei sismi (le magnitudo) indotti da un certo fenomeno meccanico.

Applicazioni Ingegneristiche

In molte applicazioni ingegneristiche siamo interessati al Maggiore o al Minor valore di una variabile casuale:

- nella progettazione strutturale siamo interessati a conoscere la minima resistenze del materiale o i carichi massimi agenti sulla struttura;
- nei programmi di prevenzione delle inondazioni interessa conoscere il massimo flusso di un fiume o la massima altezza delle onde per correttamente dimensione altezza e resistenze di dighe e protezioni;
- la corrosione di materiale è associate alla formazione di piccoli crateri (micro-pits) di metallo corroso. La resistenza alla corrosione di un componente è governata dalla grandezza del massimo pit di corrosione;
- a progettazione di componenti di sicurezza (es. il braccio di sospensione di una autovettura) deve tenere conto dei carichi massimi cui il componente può essere sottoposto nella vita operativa del sistema;
- la progettazione degli edifici deve essere fatta tenendo conto dei carichi massimi dovuti ad azioni eoliche e sismiche.

Distribuzione del Massimo (1)

Consideriamo 2 variabili aleatorie i.i.d. X1, X2 da una distribuzione F(x)

Nel nostro caso rispettivamente:

- *X1 = risultato del lancio del primo dado*
- *X2 = risultato del lancio del secondo dado*

Punto di partenza per l'analisi dei valori estremi è lo studio di

$\mathbf{X_{max} = max\ (X_1, X_2)}$

La funzione distribuzione $\mathbf{F_x(x) = Pr\ \{X \leq x\}}$ sarà pari a:

$$F_{X_{\max}}(x) = \Pr\{X_{\max} \leq x\} = \Pr\{(X_1 \leq x) \cap (X_2 \leq x)\}$$

*<u>**Entrambi**</u> gli esiti devono essere minori o uguali a x

Distribuzione del Massimo (2)

$$F_{X_{\max}}(x) = \Pr\left\{X_{\max} \le x\right\} = \Pr\left\{\left(X_1 \le x\right) \cap \left(X_2 \le x\right)\right\}$$

Nell'ipotesi di indipendenza stocastica

$$\Pr\left\{\left(X_1 \le x\right) \cap \left(X_2 \le x\right)\right\} = \Pr\left\{X_1 \le x\right\} \cdot \Pr\left\{X_2 \le x\right\} = F_{X_1}(x) \cdot F_{X_2}(x)$$

Inoltre nell'ipotesi di identica distribuzione di X_1 e X_2

$$F_{X_{\max}}(x) = \left[F_X(x)\right]^2$$

Distribuzione del Minimo (1)

$$F_{X_{\min}}(x) = \Pr\{X_{\min} \leq x\} = \Pr\{(X_1 \leq x) \cup (X_2 \leq x)\}$$

<u>Almeno uno dei due</u> esiti deve essere minore o uguale a x

Distribuzione del Minimo (4)

$$F_{X_{\min}}(x) = \Pr\{X_{\min} \le x\} = \Pr\{(X_1 \le x) \cup (X_2 \le x)\} =$$
$$= 1 - \Pr\{(X_1 > x) \cap (X_2 > x)\}$$

Nell'ipotesi di indipendenza stocastica

$$1 - \Pr\{(X_1 > x) \cap (X_2 > x)\} = 1 - [\Pr\{X_1 > x\} \cdot \Pr\{X_2 > x\}] =$$
$$= 1 - [1 - F_{X_1}(x)] \cdot [1 - F_{X_2}(x)]$$

Inoltre nell'ipotesi di identica distribuzione di X_1 e X_2

$$F_{X_{\min}}(x) = 1 - [1 - F_X(x)]^2$$

Modello Classico dei Valori Estremi

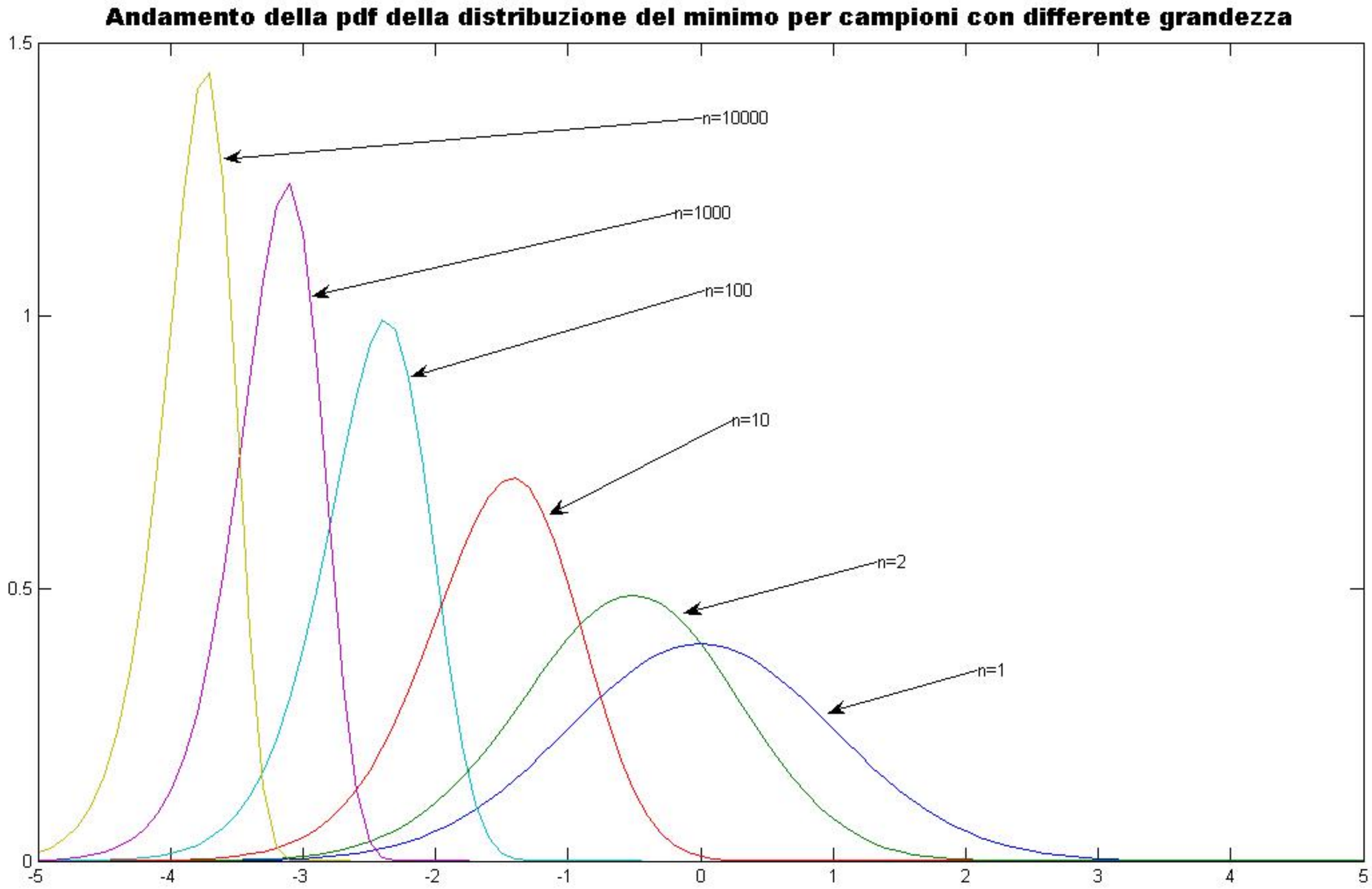
$$F_{X_{\max}}(x) = [F_X(x)]^2 \qquad F_{X_{\min}}(x) = 1 - [1 - F_X(x)]^2$$

Generalizzando X_{max} e X_{min} rappresentano, rispettivamente, il massimo e il minimo di un processo composto da n var aleatorie i.i.d.

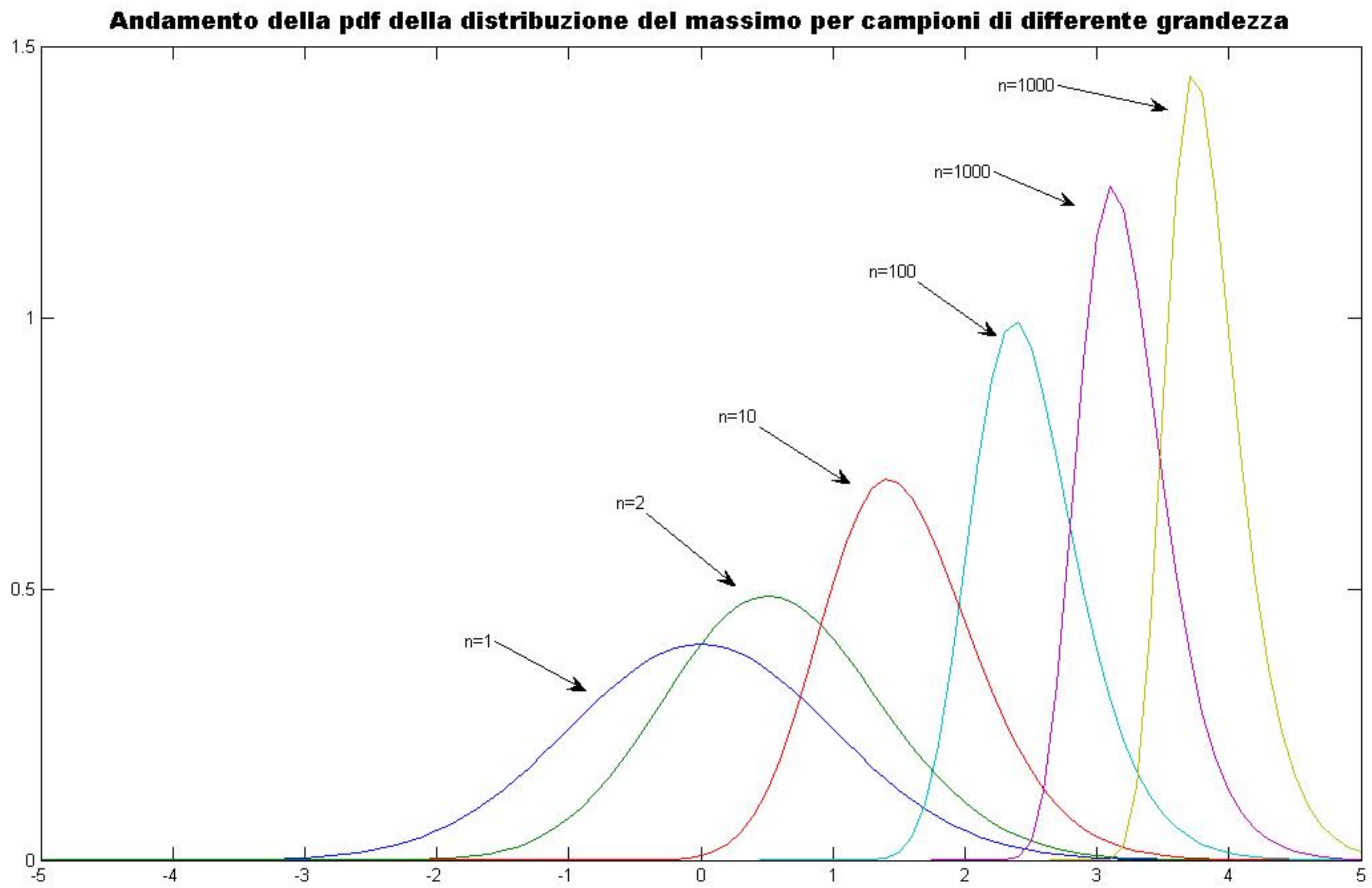
$$X_{max} = \max\{X_1, \ldots, X_n\}$$
$$X_{min} = \min\{X_1, \ldots, X_n\}$$

$$F_{X_{\max}}(x) = [F_X(x)]^n \qquad F_{X_{\min}}(x) = 1 - [1 - F_X(x)]^n$$

Variabili aleatorie Normali – pdf del Minimo



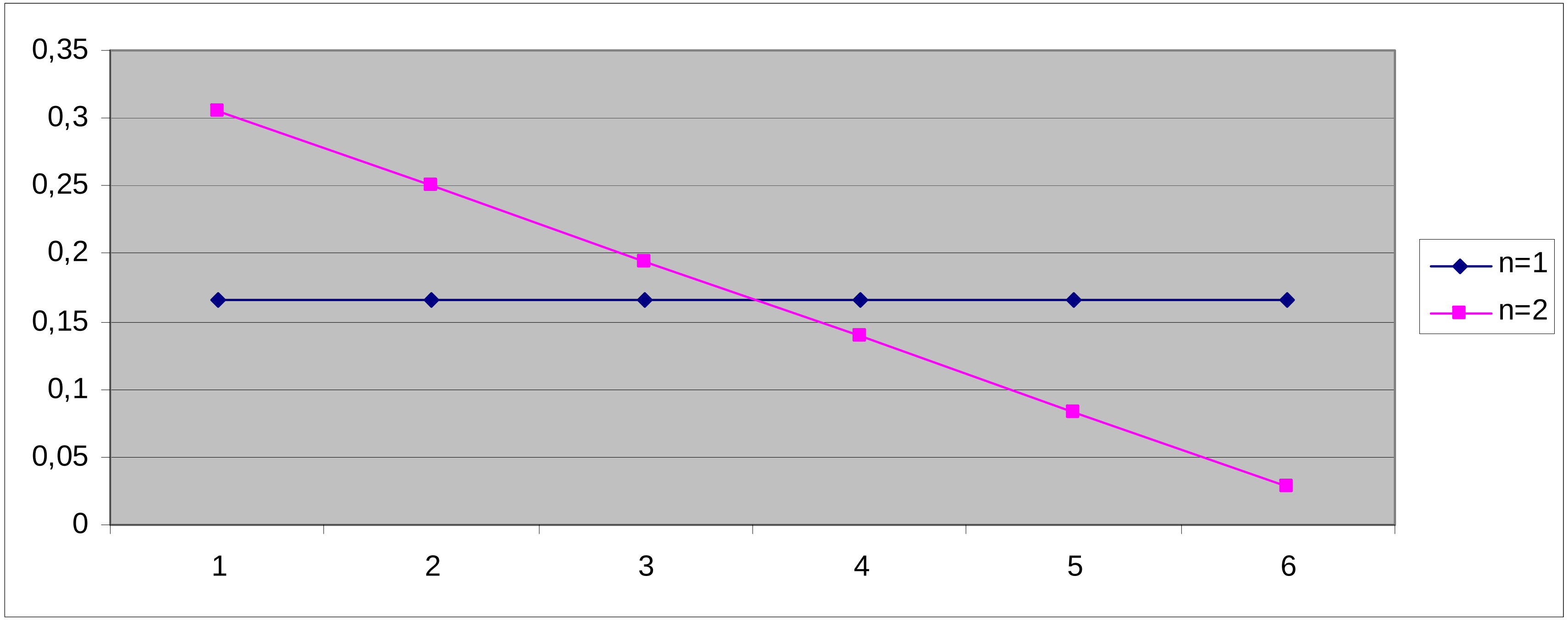
Variabili aleatorie Normali – pdf del Massimo



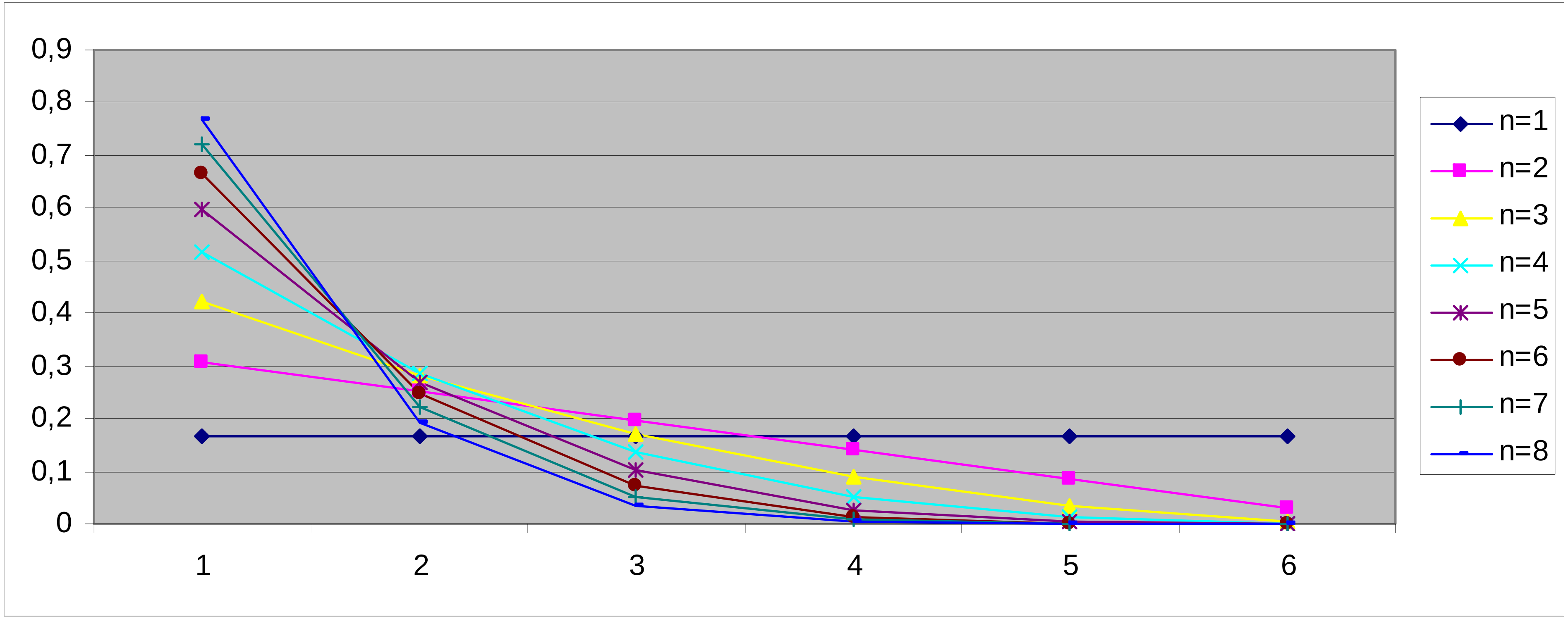
Esempio: Dadi regolari – pmf del Minimo

MIN	F(X)	$F_{Xmin}(X)=1-[1-F(X)]^2$ (2 dadi)	pmf min (2 dadi)	$F_{Xmin}(X)=1-[1-F(X)]^3$ (3 dadi)	pmf min (3 dadi)
1	1/6	0,31	0,31	0,42	0,42
2	2/6	0,56	0,56 – 0,31 = 0,25	0,70	0,28
3	3/6	0,75	0,19	0,87	0,17
4	4/6	0,89	0,14	0,96	0,09
5	5/6	0,97	0,08	0,99	0,03
6	6/6	1	0,02	1	0,004

Esempio: Dadi regolari – pmf del Minimo



Esempio: Dadi regolari – pmf del Minimo



Esempio sulla distribuzione del Minimo

100 lanci di 2 dadi a 6 facce:

lancio	Dado1	Dado 2	MIN
I	6	4	**4**
II	3	5	**3**
III	1	4	**1**
IV	3	5	**3**
V	4	3	**3**
VI	4	4	**4**
VII	4	1	**1**
VIII	2	5	**2**
IX	4	5	**4**
X	5	2	**2**

Osservazioni

Esito MIN	Freq	Freq Rel	F(X)	$[1-F(X)]^2$	$F_{Xmin}(X)=1-[1-F(X)]^2$	pmf Xmin
1	39	0,39	1/6 = 0,17	0,69	0,31	0,31
2	23	0,23	2/6 = 0,33	0,44	0,56	0,56 – 0,31 = 0,25
3	13	0,13	3/6 = 0,50	0,25	0,75	0,19
4	14	0,14	4/6 = 0,67	0,11	0,89	0,14
5	10	0,10	5/6 = 0,83	0,03	0,97	0,08
6	1	0,01	6/6 = 1	0	1	0,02

Esercizio sulla Distribuzione del Massimo

Altezze della popolazione

- Supponiamo di aver a disposizione un set di dati di dimensione *n* di misure (in cm) del carattere *altezza degli individui maschi* di una certa popolazione.
- L'esperienza ci sostiene nel ritenere che la distribuzione delle altezze segua un modello Normale.
- Avendo a disposizione stime attendibili della media e della varianza della popolazione Normale, calcolare la probabilità che l'individuo più alto del campione a disposizione sia più alto del valore soglia pari a 180 cm.

Sfruttiamo quanto appreso sulla distribuzione del massimo per valutare la probabilità richiesta

Osservazioni sperimentali

n	Altezza (cm)
1	177
2	183
3	170
4	175
5	173
6	176
7	181
8	167

Supponiamo quindi che le altezze si distribuiscano secondo una Normale di parametri:

$$\bar{x} = 175,25\,cm$$

$$s = 5,31\,cm$$

Soluzione

- Soglia altezza = x_s =180cm

Supponiamo quindi che le altezze si distribuiscano secondo una Normale di parametri:

$$\bar{x} = 175,25\,cm$$

$$s = 5,31\,cm$$

$F_X(x_s) = 0,8144$

$F_{Xmax}(x_s) = [F_X(x_s)]^8 = 0,1284$

$Pr\,\{X_{max} > x_s\} = 1 – Pr\,\{X_{max} \leq x_s\} = 1 – F_{Xmax}(x_s) = 1 – 0,1284 = 0,8716$

Nota

Raddoppiando i dati a disposizione ***(n = 16)*** e mantenendo gli stessi parametri della distribuzione, cosa ci aspettiamo sulla probabilità di $\{X_{max} > x_s\}$?

Nota

Raddoppiando i dati a disposizione ***(n = 16)*** e mantenendo gli stessi parametri della distribuzione, cosa ci aspettiamo sulla probabilità di $\{X_{max} > x_s\}$?

$$\Pr\{X_{\max} > 180cm\} = 1 - [F_X(x)]^{16} = 1 - [0{,}8716]^{16} = 0{,}9626$$

Il risultato non ci sorprende poiché aumentando gli individui esaminati tenderà ad essere asintoticamente pari all'unità (evento certo) la probabilità di trovare l'individuo più alto del gruppo più alto di 180 cm.

Riferimenti sul Libro

Pasquale Erto
"Probabilità e statistica per le scienze e l'ingegneria"
***McGraw Hill* – seconda edizione**

- **Capitolo 10**
  - **Test chi-quadrato pag. 224**
  - **Test di *Kolmogorov-Smirnov* pag. 236**
- **Capitolo 9**
  - **Grafici di probabilità pag. 185**
  - **Metodo dei momenti pag. 179**
- **Capitolo 5**
  - **Distribuzione del max pag. 97**

Mandare gli eventuali file excel all'indirizzo:

giovanna.matrone@unina.it
antonio.lepore@unina.it
andrea.colini@unina.it

Orario di ricevimento:

Giovedì ore 16:30-18:30
P.le Tecchio X piano
Stanza Dottorandi DIAS