



Università di Cagliari
Dipartimento di Matematica e Informatica



Piano Lauree Scientifiche 2019 - Matematica

Seminari Formazione Insegnanti

Modelli statistico-probabilistici per v.a. discrete e continue

Accenni (*pochi e ... sparsi*) all'inferenza statistica

Walter Racugno

Cagliari, 6 dicembre 2019

Impostazione di questa “conversazione”

- Non vuole essere una lezione “organica” sulle distribuzioni di probabilità e sulla *Normale* in particolare, (per questo si rimanda alle poche pagine di un qualunque libro di testo di ogni ordine di scuola: secondaria, università).
- Mi rivolgerò in particolare agli insegnanti ma con lo sguardo ai loro studenti. Lo scopo infatti è di fornire alcuni spunti didattici per l’attività in classe: semplici riflessioni e stimoli su come porgere concetti e nozioni per arrivare dagli esempi alla formalizzazione.

Prerequisiti

- Gli studenti dovrebbero conoscere le nozioni elementari e intuitive sulla probabilità, la rappresentazione di dati statistici attraverso semplici grafici, e possibilmente avere le basi per la costruzione del grafico di una funzione.

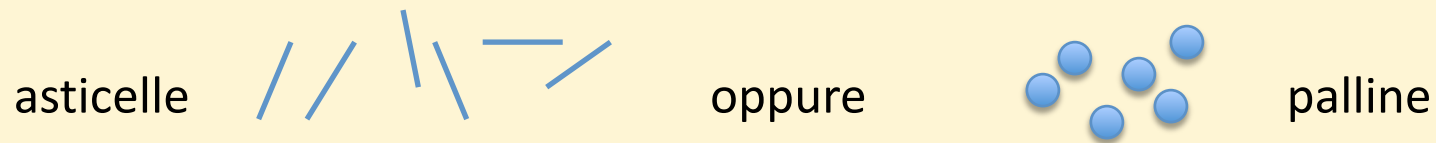
Indice dei temi

- Richiami sparsi sulla nozione di probabilità
masse, frequenze, probabilità: analogie formali
 - nel discreto
 - nel continuo
- Distribuzioni di probabilità di v.a. discrete:
 - Binomiale; Geometrica; di Poisson
- Carl Friedrich Gauss (1777 – 1855)
- La curva degli errori accidentali: funzione di densità di probabilità di una v.a. normale (gaussiana)
- La normale standard
 - Uso delle tavole per il calcolo delle probabilità
- Esempi applicativi
- Cenni di inferenza statistica e alcune proprietà teoriche

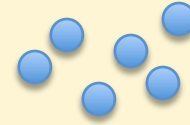
Premessa

(con qualche digressione)

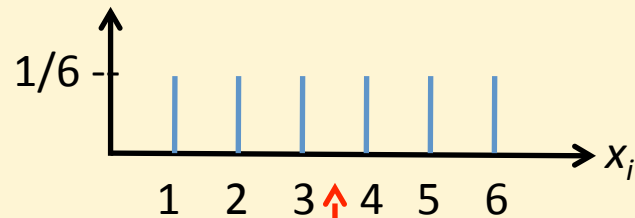
- Immaginiamo di avere 6 oggetti (palline o asticelle di metallo) di uguale massa tra loro e tali che la somma di tutti i pesi sia uguale a 1, in unità di misura qualsiasi (*kg* , *hg* , *g*);



- il peso di ciascun oggetto è = $1/6$, che è anche, ovviamente, il peso medio



- volendo rappresentare tali masse su un asse, indicandole con $x_1=1, x_2=2, x_3=3, x_4=4, x_5=5, x_6=6$



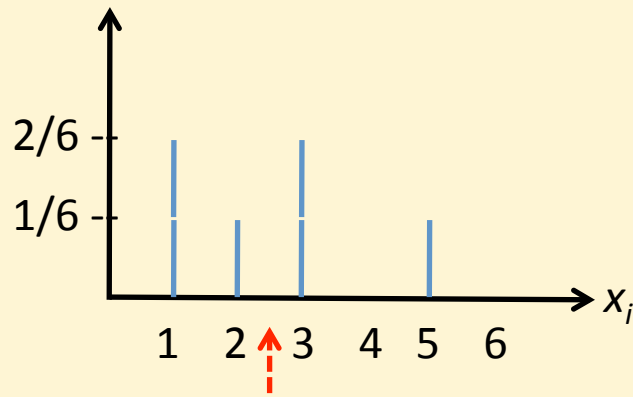
o anche



- il “centro di massa” (**baricentro**) della distribuzione dei pesi è rappresentato sull’asse dal punto di equilibrio \bar{x}

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = \frac{1}{6} (1 + 2 + 3 + 4 + 5 + 6) = 3.5$$

... se le masse non avessero lo stesso peso ...



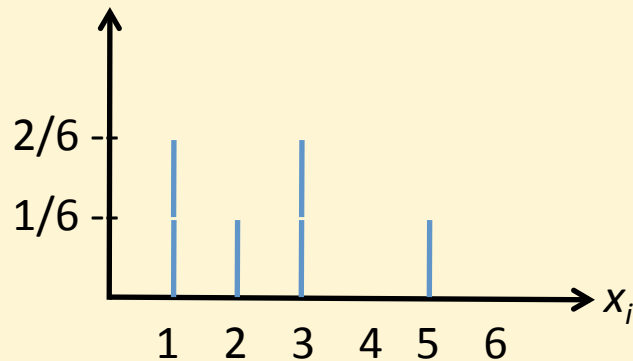
m_i masse (*pesi*)

$$\bar{x} = \sum_{i=1}^k x_i \times m_i = 1 \times \frac{2}{6} + 2 \times \frac{1}{6} + 3 \times \frac{2}{6} + 5 \times \frac{1}{6} = 2.5$$

Osservazione

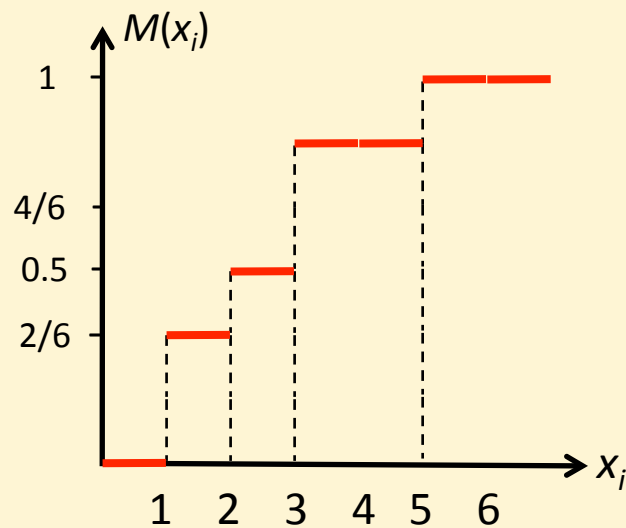
- i pesi si “addensano” maggiormente nella parte sinistra dell’asse x_i
- è interessante studiare come avviene tale accumulo, anche attraverso una rappresentazione grafica

Funzione cumulativa delle masse (*pesi*): $M(x_i)$



m_i masse (*pesi*)

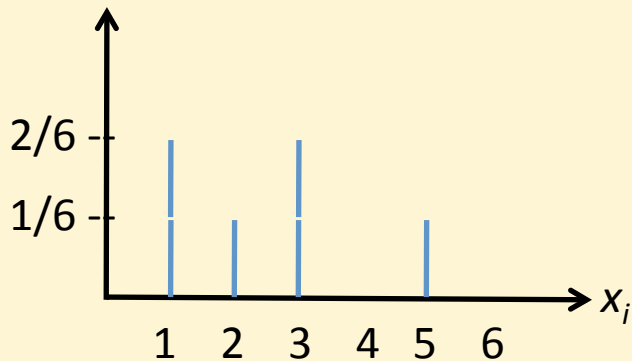
$$\bar{x} = \sum_{i=1}^k x_i \times m_i = 1 \times \frac{2}{6} + 2 \times \frac{1}{6} + 3 \times \frac{2}{6} + 5 \times \frac{1}{6} = 2.5$$



$$M(x_i) = \sum_{j=1}^i m_j$$

x_i	$M(x_i)$
1	$m_1 = 0.33$
2	$m_1 + m_2 = 0.5$
3	$m_1 + m_2 + m_3 = 0.83$
4	$m_1 + \dots + m_4 = 0.83$
5	$m_1 + \dots + m_5 = 1$

Quartili: Q_1 (25%), Q_2 (50%)=Mediana, Q_3 (75%)

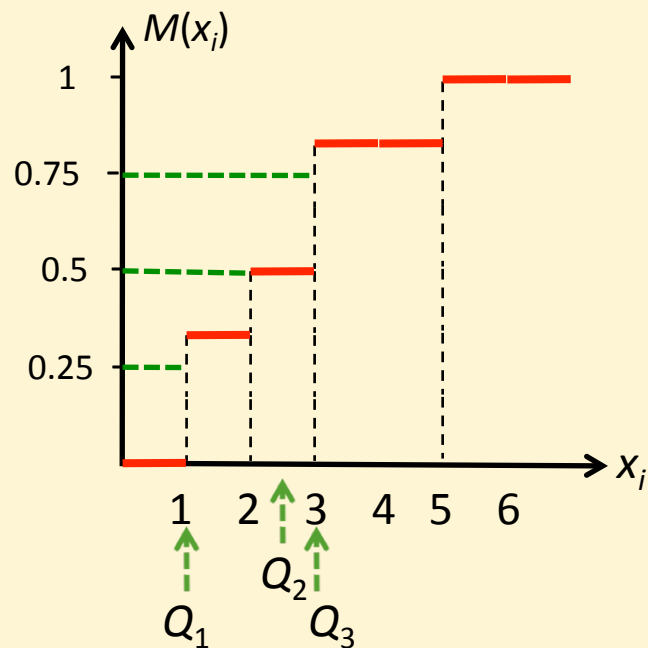


la metà delle masse è concentrata in x_1 e x_2

→ Mediana = $Q_2 = 2.5$

il 25% nel punto x_1 → primo quartile $Q_1 = 1$

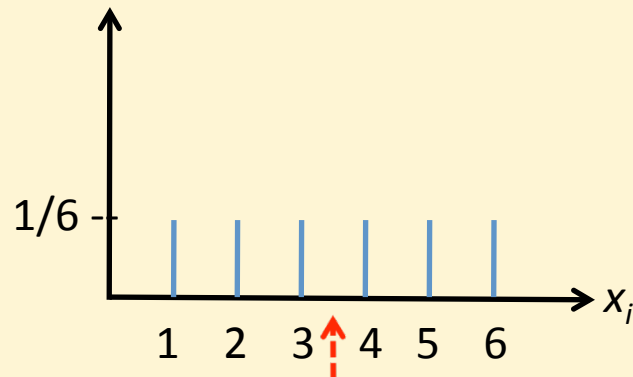
il 75% in x_1, x_2 e x_3 → terzo quartile $Q_3 = 3$



$$M(x_i) = \sum_{j=1}^i m_j$$

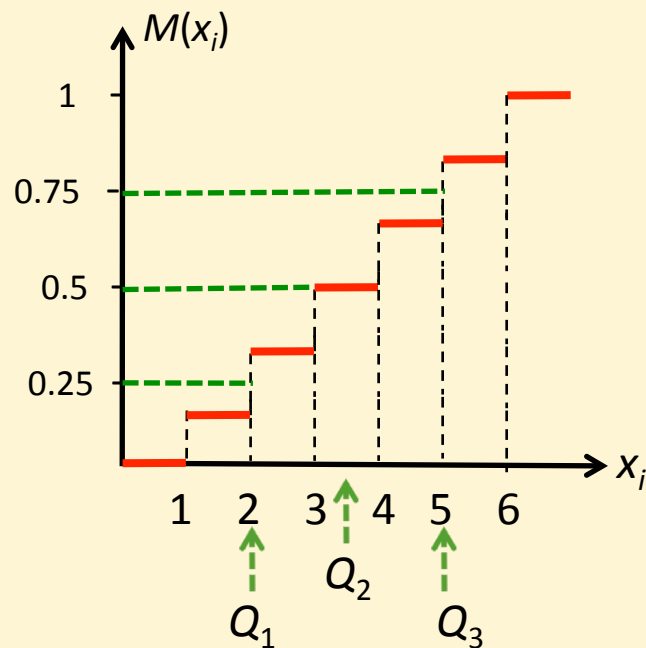
x_i	$M(x_i)$	
1	$m_1 = 0.33$	←--- Q_1
2	$m_1 + m_2 = 0.5$	←--- Q_2
3	$m_1 + m_2 + m_3 = 0.83$	←--- Q_3
4	$m_1 + \dots + m_4 = 0.83$	
5	$m_1 + \dots + m_5 = 1$	

Quartili: Q_1 (25%), Q_2 (50%)=Mediana, Q_3 (75%)
pesi tutti uguali tra loro



m_i masse (pesi)

$$\bar{x} = \sum_{i=1}^6 x_i \times m_i = 3.5$$

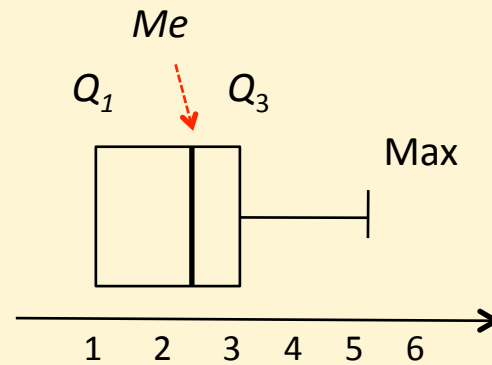
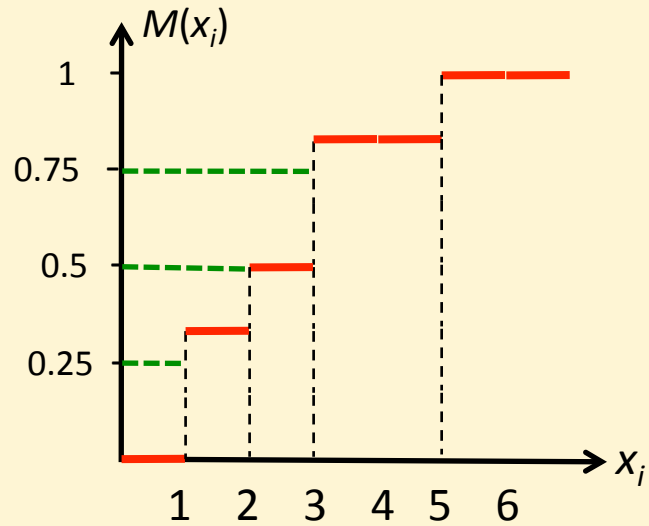


$$M(x_i) = \sum_{j=1}^i m_j$$

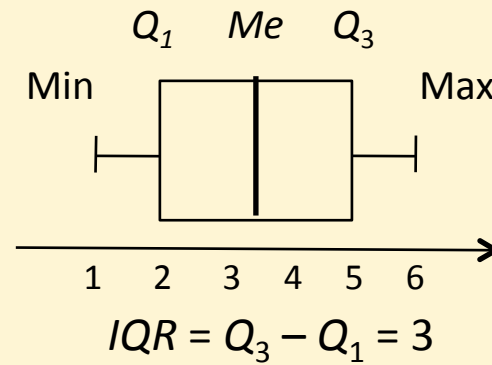
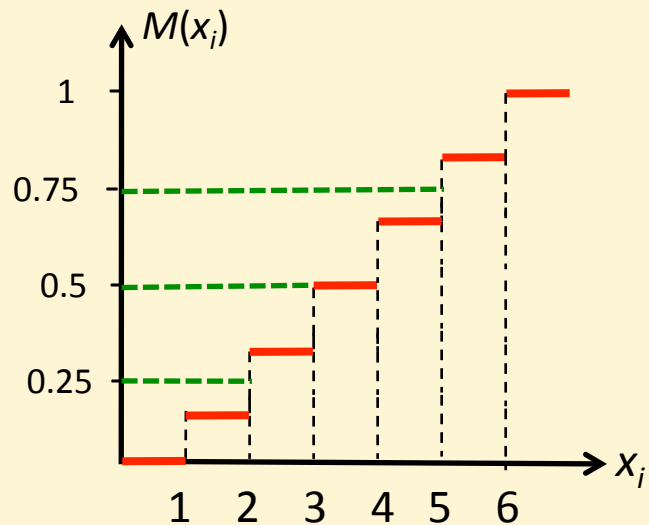
x_i	$M(x_i)$
1	$m_1 = 0.17$
2	$m_1 + m_2 = 0.33$ ← Q_1
3	$m_1 + m_2 + m_3 = 0.5$ ← Q_2
4	$m_1 + \dots + m_4 = 0.67$
5	$m_1 + \dots + m_5 = 0.83$ ← Q_3
6	$m_1 + \dots + m_6 = 1$

Una sintesi

il Box-plot (diagramma a scatola e baffi)



Distanza interquartile $IQR = Q_3 - Q_1 = 2$

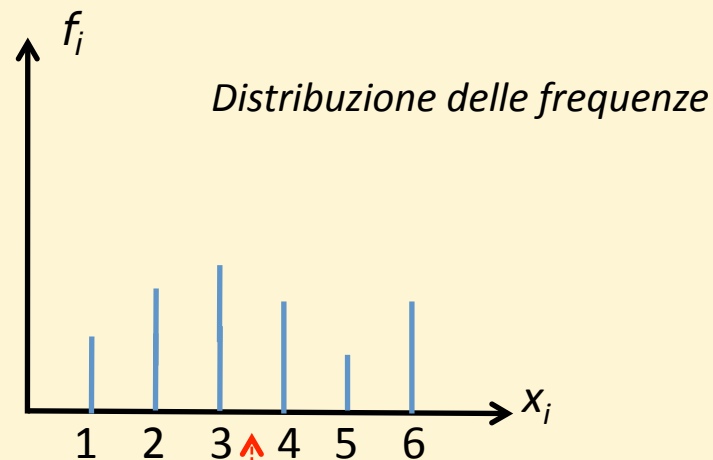


$IQR = Q_3 - Q_1 = 3$

Lancio di un dado cubico equilibrato (reale)

- Supponiamo di effettuare $N=60$ lanci di un dado, ottenendo ad es. i risultati sintetizzati in tabella, dove x_i , con $i=1, 2, \dots, 6$, sono i punteggi realizzati, n_i le corrispondenti frequenze assolute e $f_i = n_i/N$ le frequenze relative

x_i	n_i	$f_i = \frac{n_i}{N}$
1	8	0.133
2	12	0.20
3	14	0.233
4	10	0.167
5	6	0.10
6	10	0.167
	60	1



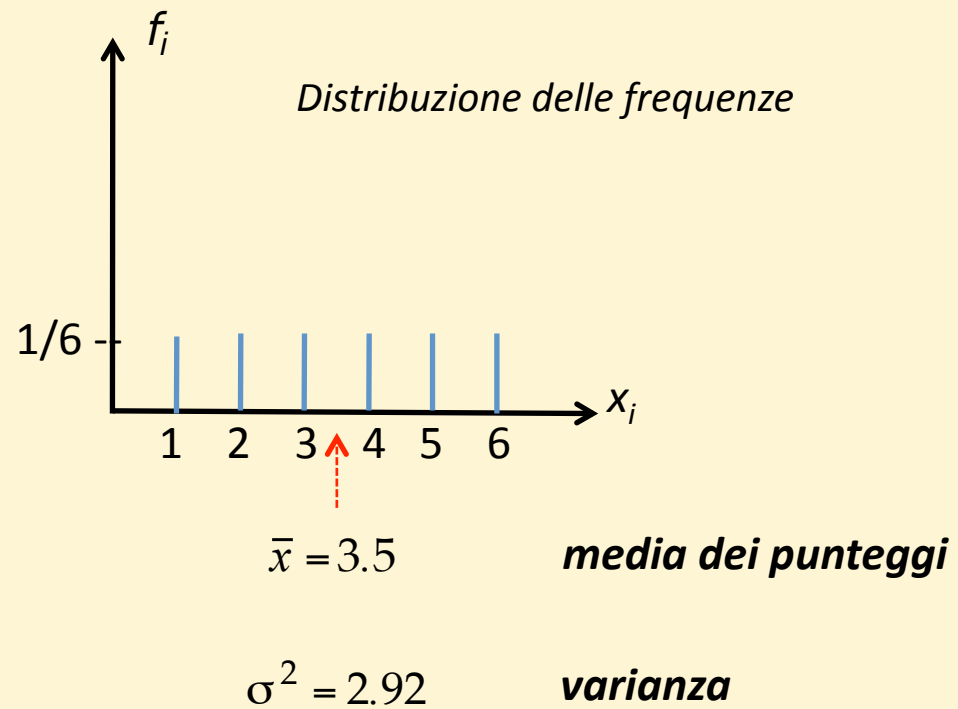
$$\bar{x} = \sum_{i=1}^6 x_i \times f_i = 3.4 \quad \text{media dei punteggi}$$

$$\sigma^2 = \sum_{i=1}^6 (x_i - \bar{x})^2 f_i = \sum_{i=1}^6 x_i^2 f_i - \bar{x}^2 = 2.64 \quad \text{varianza}$$

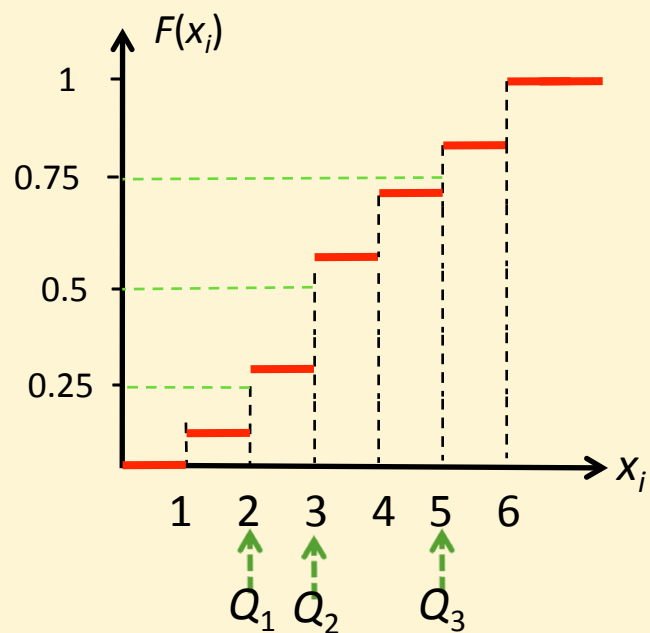
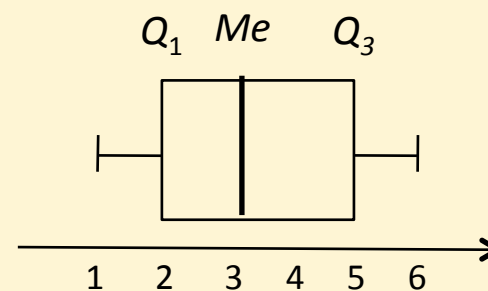
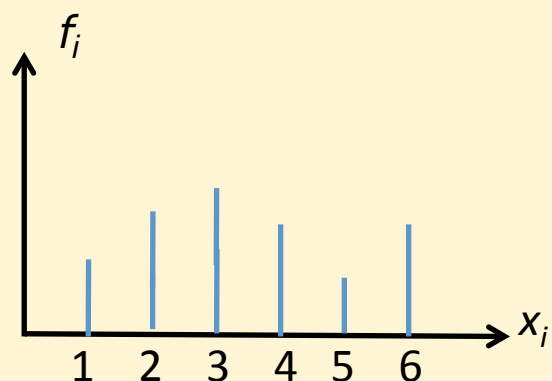
Lancio di un dado cubico equilibrato (ideale)

- Immaginiamo che negli $N=60$ lanci di un dado, ognuno dei punteggi possibili (*determinazioni della variabile statistica*) si verifichi lo stesso numero di volte. Si avrebbe la seguente sintesi dei dati:

x_i	n_i	$f_i = \frac{n_i}{N}$
1	10	0.167
2	10	0.167
3	10	0.167
4	10	0.167
5	10	0.167
6	10	0.167



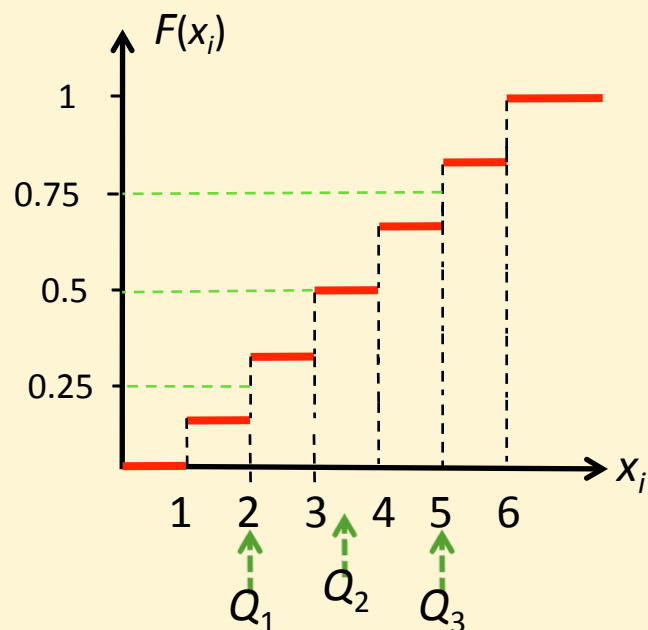
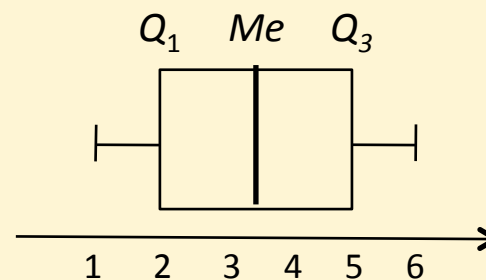
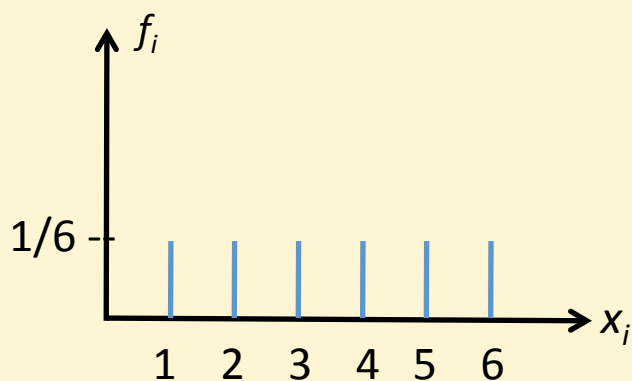
Funzione cumulativa delle frequenze: $F(x_i)$ (esempio di lancio "reale")



$$F(x_i) = \sum_{j=1}^i f_j$$

x_i	$F(x_i)$	
1	$f_1 = 0.133$	
2	$f_1 + f_2 = 0.333$	←--- Q_1
3	$f_1 + f_2 + f_3 = 0.566$	←--- Q_2
4	$f_1 + \dots + f_4 = 0.733$	
5	$f_1 + \dots + f_5 = 0.833$	←--- Q_3
6	$f_1 + \dots + f_5 + f_6 = 1$	

Funzione cumulativa delle frequenze: $F(x_i)$ (ideale)

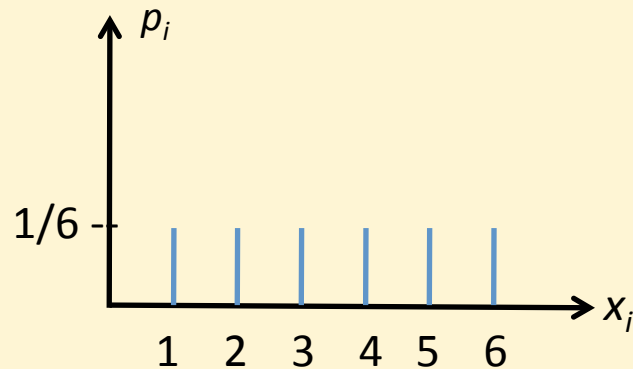


$$F(x_i) = \sum_{j=1}^i f_j$$

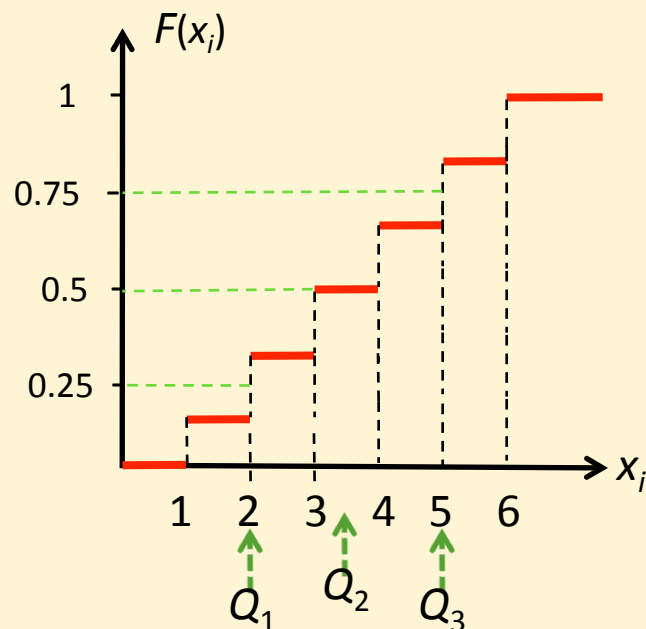
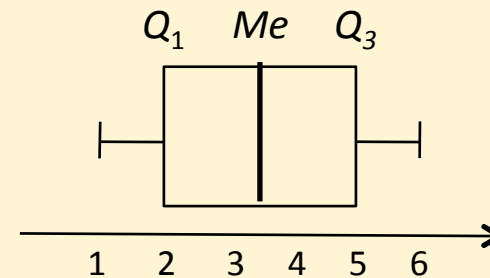
x_i	$F(x_i)$
1	$f_1 = 0.167$
2	$f_1 + f_2 = 0.334$ ← Q_1
3	$f_1 + f_2 + f_3 = 0.50$ ← Q_2
4	$f_1 + \dots + f_4 = 0.668$
5	$f_1 + \dots + f_5 = 0.835$ ← Q_3
6	$f_1 + \dots + f_5 + f_6 = 1$

Lancio di un dado cubico equilibrato

(prima del lancio ogni faccia x_i ha **probabilità = 1/6** di mostrarsi)



p_i probabilità



$$F(x_i) = \sum_{j=1}^i p_j$$

x_i	$F(x_i)$
1	$p_1 = 0.167$
2	$p_1 + p_2 = 0.334$ ← Q_1
3	$p_1 + p_2 + p_3 = 0.50$ ← Q_2
4	$p_1 + \dots + p_4 = 0.668$
5	$p_1 + \dots + p_5 = 0.835$ ← Q_3
6	$p_1 + \dots + p_5 + p_6 = 1$

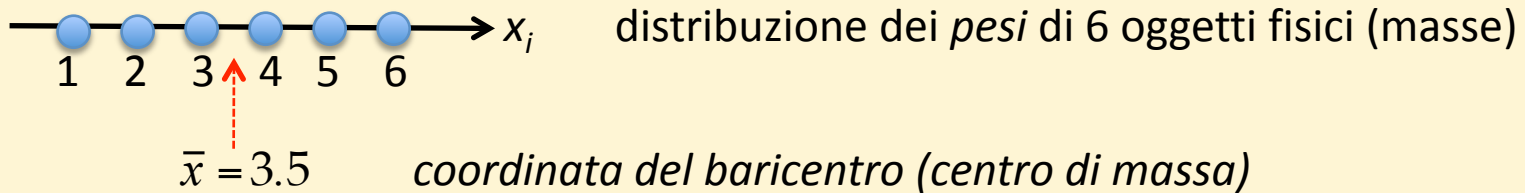
Un poco di terminologia

- $p_i = P(X = x_i)$ rappresenta la *probabilità dell'evento* x_i (del valore x_i di X)
 - per $i = 1, 2, 3, 4, 5, 6$ si ha la *distribuzione della probabilità*
 - in generale $P(X \leq x)$ indica la probabilità dell'intervallo $(-\infty, x]$
 - la funzione $F(x) : R \rightarrow [0, 1]$ $F(x) = P(X \leq x)$ che fa corrispondere ai valori x della *variabile aleatoria* X le probabilità cumulate $P(X \leq x)$ è detta *funzione di ripartizione*
 - nel caso specifico del dado (*variabile casuale discreta*): $F(x_i) = \sum_{j=1}^i p_j$
 - la media della variabile X : $\mu = E(X) = \sum_{i=1}^6 x_i \times p_i$ rappresenta il *valore atteso* (o *speranza matematica*)
- se i punteggi delle facce del dado indicassero la somma in Euro che un giocatore incasserebbe in un futuro lancio, egli dovrebbe fare il seguente calcolo: $1\text{€} \times 1/6 + 2\text{€} \times 1/6 + 3\text{€} \times 1/6 + 4\text{€} \times 1/6 + 5\text{€} \times 1/6 + 6\text{€} \times 1/6 = 3.5\text{€}$

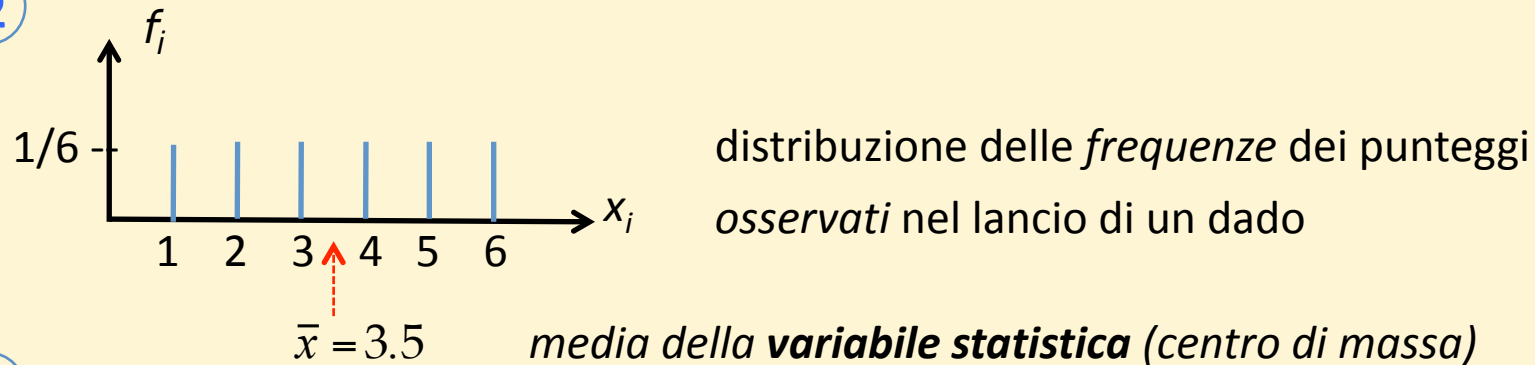
$$\mu = E(X) = 3.5\text{€} \text{ incasso atteso (o sperato)}$$

Alcune osservazioni di sintesi sui tre casi considerati

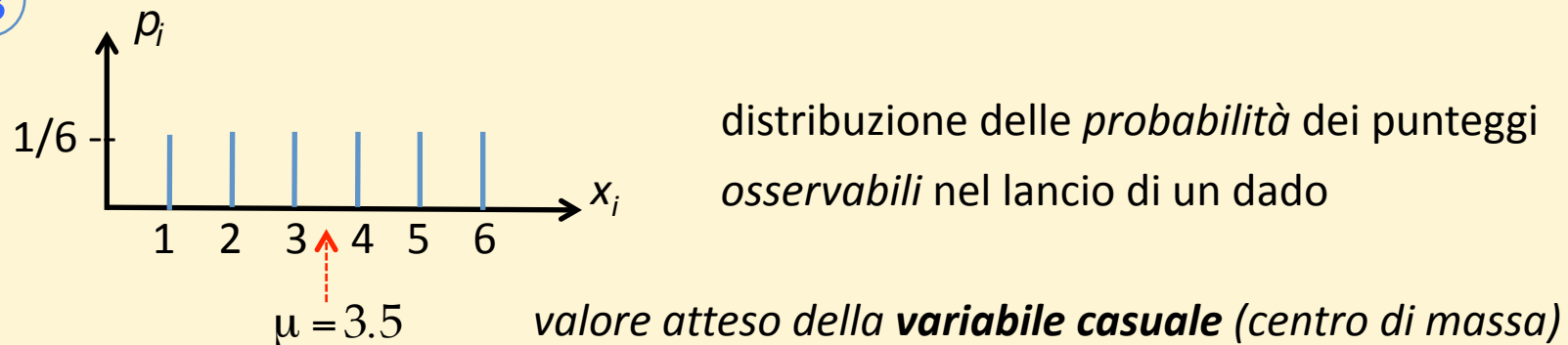
1



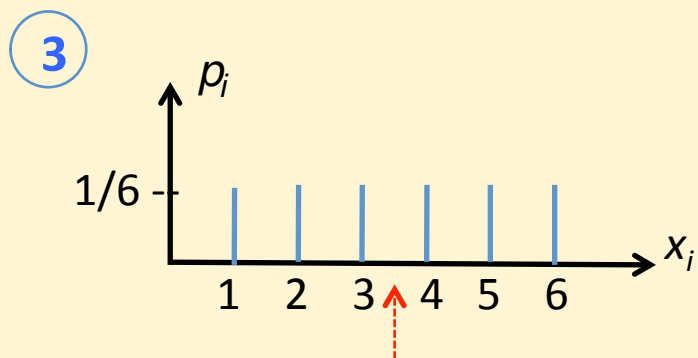
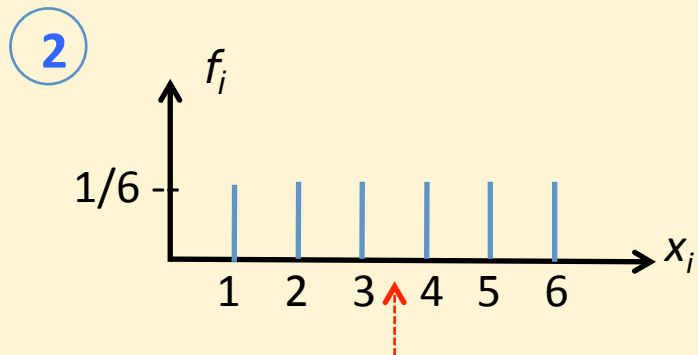
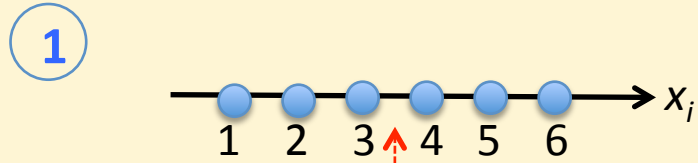
2



3



Alcune osservazioni di sintesi sui tre casi considerati



① ② ③

casi concettualmente differenti

ma

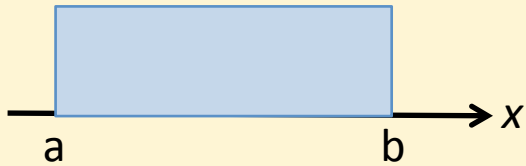
formalmente analoghi



analogia con la *fisica delle masse*

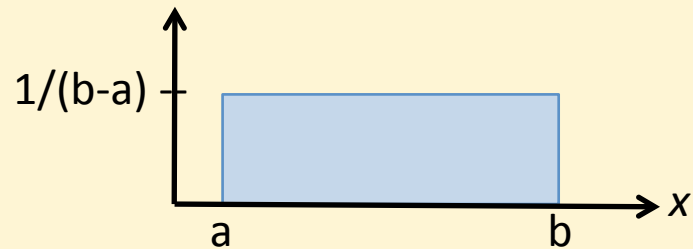
Passaggio al continuo

- supponiamo di avere una lamina (o barretta di metallo) di lunghezza $L = (b - a)$ e peso uguale a 1, in unità di misura qualsiasi (kg , hg , g)



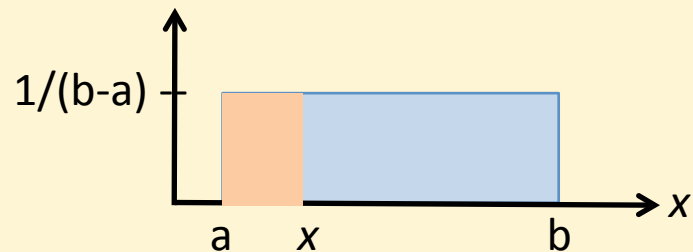
distribuzione della massa

- assumendo lo spessore costante, la sua altezza deve essere $h = 1/(b - a)$



$$\text{Area} = L \times h = 1$$

- peso di una *massa* di lunghezza $(x - a) \rightarrow m_x = (x - a) \frac{1}{b - a}$

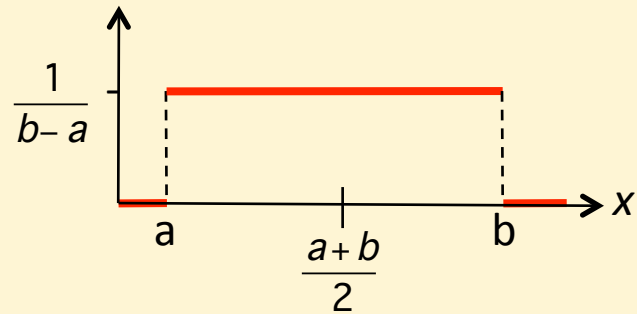


l'altezza rappresenta la **densità di massa**

$$\frac{1}{b - a} = \frac{m_x}{x - a}$$

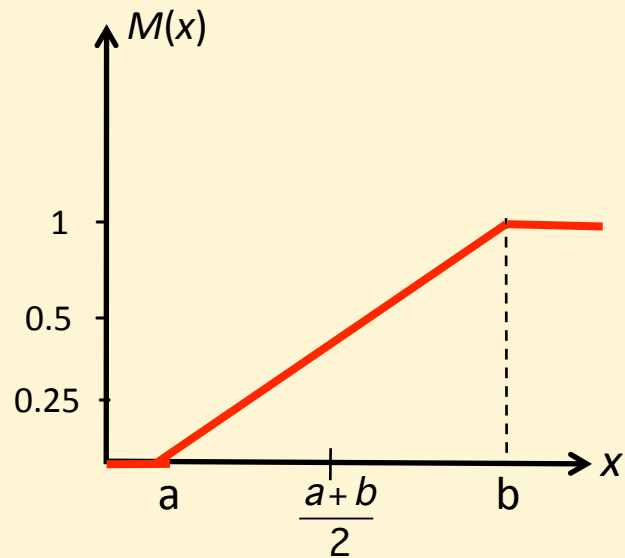
Nota: *ogni singolo punto x ha peso nullo.*

Funzione cumulativa delle masse (*pesi*): $M(x)$



m_x masse (*pesi*)

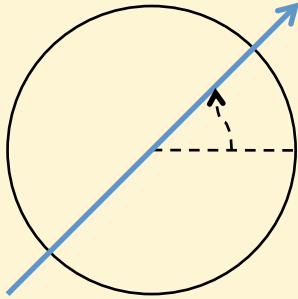
$$\bar{x} = \frac{a+b}{2} \quad \text{baricentro}$$



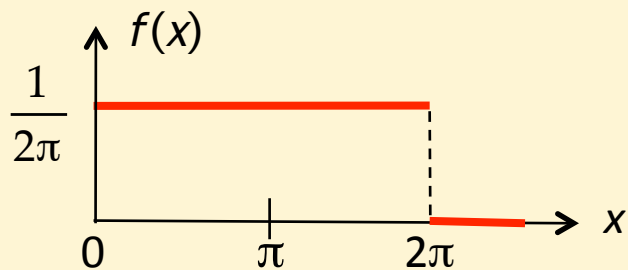
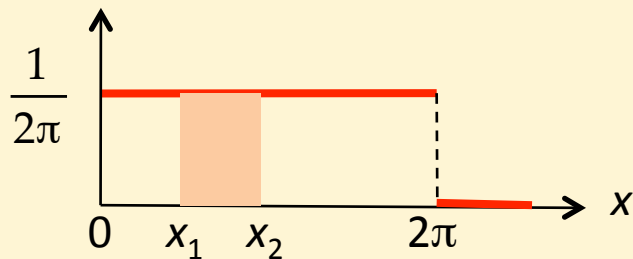
$$M(x) = \int_{-\infty}^x \frac{1}{b-a} dt$$

$$M(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$$

La ruota della fortuna



- La freccia ruota intorno al punto fisso centrale e si arresta per attrito.
- Il giocatore dà un piccolo colpo alla freccia. L'ampiezza dell'angolo tra la freccia e il raggio fisso determina l'entità della vincita.
- Se il raggio del cerchio è = 1, l'angolo di arresto è una **variabile aleatoria X continua** che varia **uniformemente** tra 0 e 2π



probabilità che la freccia si fermi tra x_1 e x_2

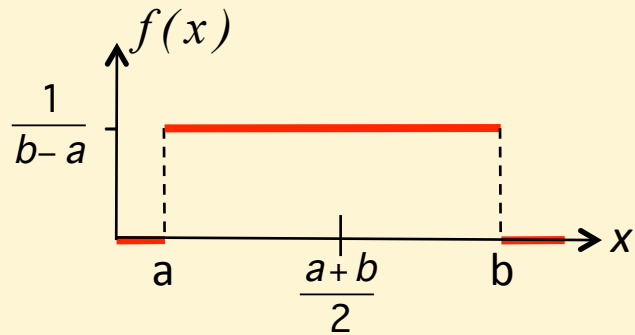
$$P(x_1 < X < x_2) = P(x_1 \leq X \leq x_2) = \text{area} = (x_2 - x_1) \frac{1}{2\pi}$$

$f(x)$ funzione di densità della probabilità

Nota: $P(X = x) = 0 \rightarrow$ ogni punto isolato ha probabilità nulla (base = 0)

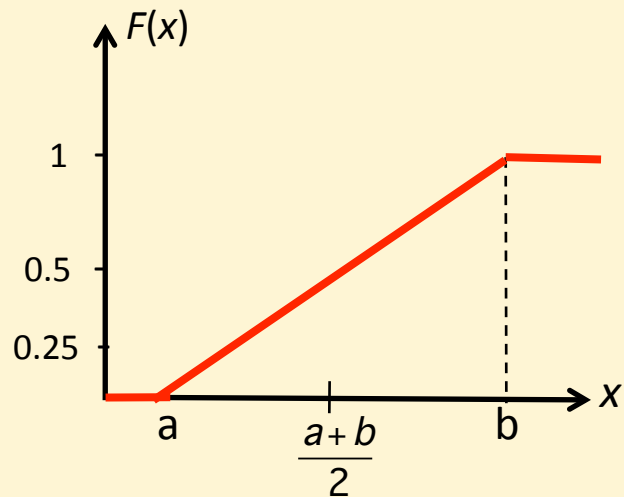
In generale

distribuzione di probabilità della v.a. **Uniforme continua** $X \sim \mathcal{U}(a,b)$



$f(x)$ funzione di densità

$$\mu = \frac{a+b}{2} \quad \text{valore atteso (media)}$$



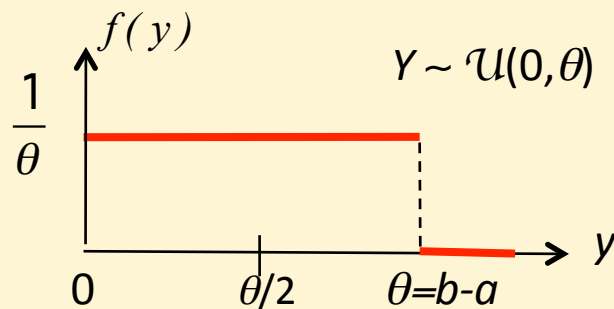
$$F(x) = \int_{-\infty}^x \frac{1}{b-a} dt \quad \text{funzione di ripartizione}$$

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & x \geq b \end{cases}$$

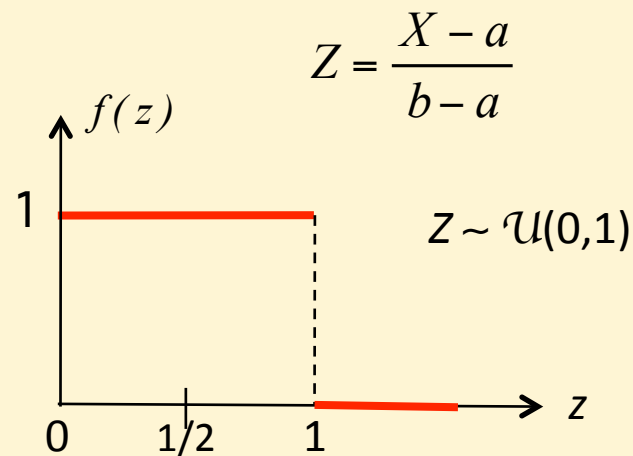
Osservazioni sui *parametri* del modello Uniforme: $\mathcal{U}(a,b)$

- $X \sim \mathcal{U}(a,b)$ è un modello (funzione) che dipende dai due **parametri** a e b
→ la distribuzione di probabilità è perfettamente nota se a e b sono noti
- il numero dei parametri può essere ridotto attraverso opportune *riparametrizzazioni*, o trasformazioni, della variabile senza alterare la distribuzione di probabilità

traslazione: $Y = X - a$



cambiamento di scala (*standardizzazione*)



Modello *Binomiale*

Urna con *composizione nota* \longrightarrow Probabilità di *Bianca* = $P(B) = p$ nota

Ci domandiamo:

- se estraiamo un certo numero di palline (es. n), qual è la probabilità che ve ne siano $0, 1, 2, \dots, n$ *Bianche*?
(più semplicemente x_i *Bianche* su n estratte , con $x_i = 0, 1, 2, \dots, n$)

Un poco di terminologia

- p proporzione di *Bianche* nell'urna (probabilità di estrarre *Bianca* in una estrazione).
- Se l'estrazione viene ripetuta dopo avere reimbussolato ogni volta la pallina estratta, la composizione dell'urna non cambia: $p =$ costante.
- Il “numero di palline *Bianche* su n estratte”, tutte con reimbussolamento, è una variabile X che può assumere i valori x_i tutti incerti prima del verificarsi dell'evento: X è una **variabile aleatoria** (casuale), ogni valore x_i ha una sua probabilità p_i di verificarsi.

Modello Binomiale

Urna che contiene il 30% di palline *Bianche* e il restante 70% *Nere*

- Probabilità di *Bianca* in **una estrazione**: $p = 0.30$

In altri termini:

- Quali sono le probabilità p_0 di osservare 0 *Bianche* e p_1 di 1 *Bianca* in **una estrazione**?

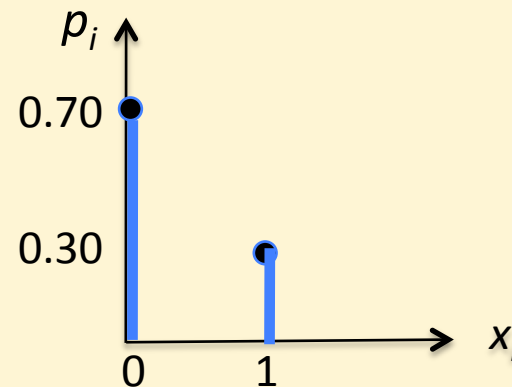
ovvero

- Quali sono le probabilità di $x_0 = 0$ e di $x_1 = 1$?

o ancora

- Quali sono le probabilità $P(X=0)$ e $P(X=1)$?

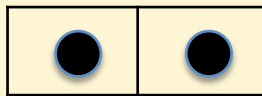
x_i	p_i
0	0.70
1	0.30



Modello Binomiale

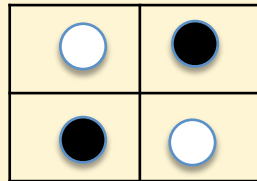
Urna che contiene il 30% di palline *Bianche* e il restante 70% *Nere*

- **Due estrazioni** (sempre con reimbussolamento)
- Quali sono le probabilità p_0 di osservare 0 *Bianche*, p_1 1 *Bianca* e p_2 2 *Bianche* in **due estrazioni** ?



$$x_0 = 0$$

$$p_0 = 0.7 \times 0.7$$



$$x_1 = 1$$

$$p_1 = (0.3 \times 0.7) + (0.7 \times 0.3)$$

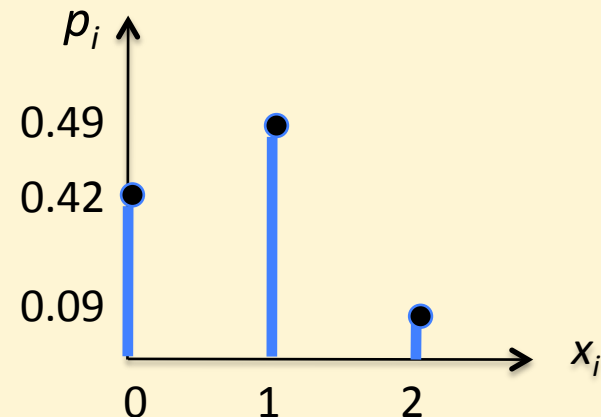


$$x_2 = 2$$

$$p_2 = 0.3 \times 0.3$$

x_i	p_i
0	0.49
1	0.42
2	0.09

1



Modello Binomiale

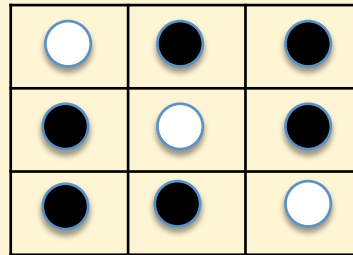
Urna che contiene il 30% di palline *Bianche* e il restante 70% *Nere*

- **Tre estrazioni** (sempre con reimbussolamento)
- Quali sono le probabilità p_0 di osservare 0 *Bianche*, p_1 1 *Bianca*, p_2 2 *Bianche* e p_3 3 *Bianche* in **tre estrazioni** ?



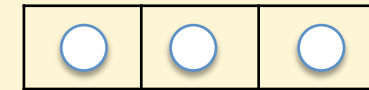
$$x_0 = 0$$

$$p_0 = 0.7 \times 0.7 \times 0.7 = (0.7)^3$$



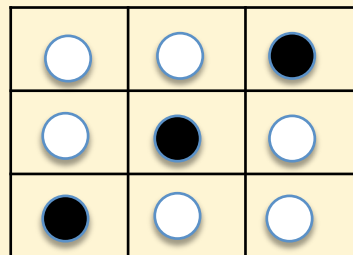
$$x_1 = 1$$

$$p_1 = (0.3 \times 0.7 \times 0.7) \times 3$$



$$x_3 = 3$$

$$p_3 = 0.3 \times 0.3 \times 0.3 = (0.3)^3$$



$$x_2 = 2$$

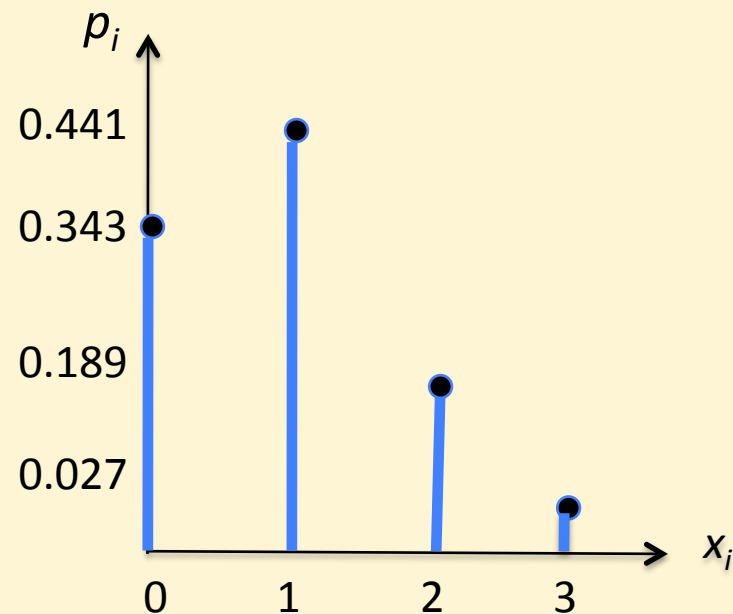
$$p_2 = (0.3 \times 0.3 \times 0.7) \times 3$$

Modello Binomiale

Urna che contiene il 30% di palline *Bianche* e il restante 70% *Nere*

- **Tre estrazioni** (sempre con reimbussolamento)
- Quali sono le probabilità p_0 di osservare 0 *Bianche*, p_1 di 1 *Bianca*, p_2 di 2 *Bianche* e p_3 di 3 *Bianche* in **tre estrazioni** ?

x_i	p_i
0	0.343
1	0.441
2	0.189
3	0.027



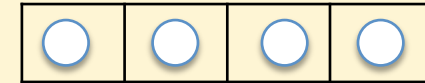
Modello Binomiale

Urna che contiene il 30% di palline *Bianche* e il restante 70% *Nere*

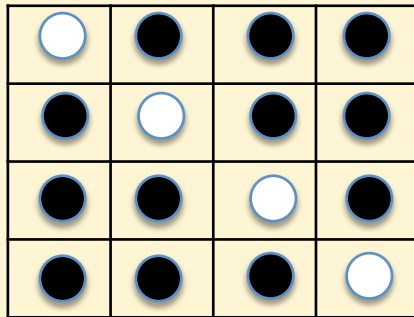
- **Quattro estrazioni** (sempre con reimbussolamento)



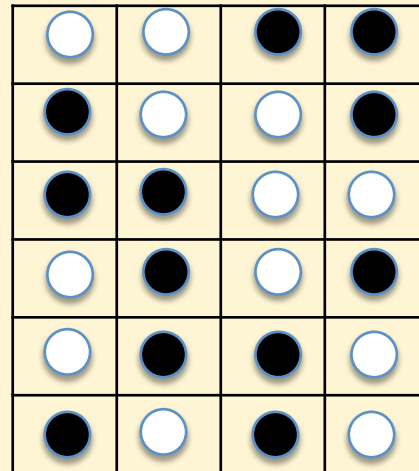
$$x_0 = 0$$



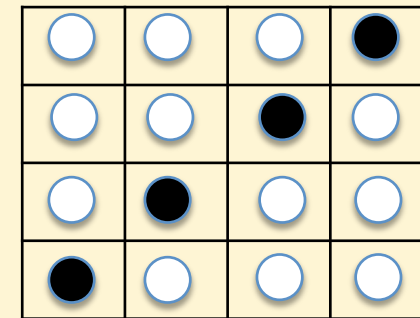
$$x_4 = 4$$



$$x_1 = 1$$



$$x_2 = 2$$



$$x_3 = 3$$

... ma ...

... ma è possibile andare avanti così ...?

Serve una regola!

Modello Binomiale

Urna con *composizione nota* $\begin{cases} \longrightarrow \text{Probabilità di Bianca} = P(B) = \mathbf{p} \text{ nota} \\ \longrightarrow \text{Probabilità di Nera} = \mathbf{1 - p} \end{cases}$

- *Evento* : estrazione di n palline (con reimbussolamento)
- *Variabile X* : numero di *Bianche* tra le n
- *Esiti (valori) possibili* : $x_i = 0, 1, 2, \dots, n$

Per $n = 2$, $x_i = 0, 1, 2$

- $X = 0 \longrightarrow \text{Nera, Nera} \longrightarrow P(X = 0) = (1-p) \times (1-p) = (1-p)^2$

- $X = 1 \begin{cases} \longrightarrow B, N \longrightarrow P(B, N) = p \times (1-p) \\ \longrightarrow N, B \longrightarrow P(N, B) = p \times (1-p) \end{cases} \longrightarrow = 2 p (1-p)$

- $X = 2 \longrightarrow B, B \longrightarrow P(B, B) = p \times p = p^2$

Modello Binomiale

Urna con *composizione nota* \longrightarrow Probabilità di *Bianca* = $P(B) = p$ nota
 \searrow Probabilità di *Nera* = $1 - p$

Per $n = 3$, $x_i = 0, 1, 2, 3$

• $X = 0 \longrightarrow N, N, N \longrightarrow P(N, N, N) \longrightarrow P(X = 0) = (1-p)^3$

• $X = 1$ $\begin{cases} \nearrow B, N, N \longrightarrow P(B, N, N) = p (1-p)^2 \\ \longrightarrow N, B, N \longrightarrow P(N, B, N) = p (1-p)^2 \\ \searrow N, N, B \longrightarrow P(N, N, B) = p (1-p)^2 \end{cases} \longrightarrow P(X = 1) = 3 p (1-p)^2$

• $X = 2$ $\begin{cases} \nearrow B, B, N \longrightarrow P(B, B, N) = p^2 (1-p) \\ \longrightarrow B, N, B \longrightarrow P(B, N, B) = p^2 (1-p) \\ \searrow N, B, B \longrightarrow P(N, B, B) = p^2 (1-p) \end{cases} \longrightarrow P(X = 2) = 3 p^2 (1-p)$

• $X = 3 \longrightarrow B, B, B \longrightarrow P(B, B, B) \longrightarrow P(X = 3) = p^3$

Modello Binomiale ... sintetizzando ...

$$P(X = x) = \text{Coefficiente } p^x (1-p)^{n-x}$$

- Nel caso $n = 2$
 - per $x = 0 \longrightarrow$ Coefficiente = 1
 - $x = 1 \longrightarrow$ Coefficiente = 2
 - $x = 2 \longrightarrow$ Coefficiente = 1
- Nel caso $n = 3$
 - per $x = 0 \longrightarrow$ Coefficiente = 1
 - $x = 1 \longrightarrow$ Coefficiente = 3
 - $x = 2 \longrightarrow$ Coefficiente = 3
 - $x = 3 \longrightarrow$ Coefficiente = 1
- Nel caso $n = 4$
 - per $x = 0 \longrightarrow$ Coefficiente = 1
 - $x = 1 \longrightarrow$ Coefficiente = 4
 - $x = 2 \longrightarrow$ Coefficiente = 6
 - $x = 3 \longrightarrow$ Coefficiente = 4
 - $x = 4 \longrightarrow$ Coefficiente = 1

Modello *Binomiale* (2)

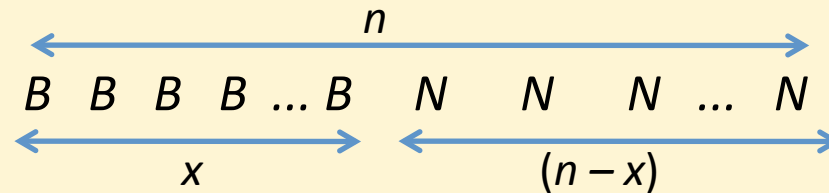
Il triangolo di Tartaglia
Due esercizi esemplificativi

Modello *Binomiale*

Urna con *composizione nota* $\begin{cases} \longrightarrow \text{Probabilità di } \textit{Bianca} = P(B) = \mathbf{p} \text{ nota} \\ \searrow \text{Probabilità di } \textit{Nera} = \mathbf{1 - p} \end{cases}$

- *Evento* : estrazione di n palline (con reimbussolamento)
- *Variabile X* : numero di *Bianche* tra le n

Immaginiamo di ottenere la seguente sequenza di x *Bianche* e $(n - x)$ *Nere*



Modello *Binomiale*

Urna con *composizione nota* $\begin{cases} \longrightarrow \text{Probabilità di } \textit{Bianca} = P(B) = \mathbf{p} \text{ nota} \\ \searrow \text{Probabilità di } \textit{Nera} = \mathbf{1 - p} \end{cases}$

- *Evento* : estrazione di n palline (con reimbussolamento)
- *Variabile X* : numero di *Bianche* tra le n

Immaginiamo di ottenere la seguente sequenza di x *Bianche* e $(n - x)$ *Nere*

Con probabilità

	$\overleftrightarrow{\hspace{10em} n \hspace{10em}}$										
	<i>B</i>	<i>B</i>	<i>B</i>	<i>B</i>	...	<i>B</i>	<i>N</i>	<i>N</i>	<i>N</i>	...	<i>N</i>
	$\overleftrightarrow{\hspace{6em}}$						$\overleftrightarrow{\hspace{5em}}$				
	p	p	p	p	...	p	$(1-p)$	$(1-p)$	$(1-p)$...	$(1-p)$

Modello Binomiale

Urna con *composizione nota* $\begin{cases} \longrightarrow \text{Probabilità di Bianca} = P(B) = p \text{ nota} \\ \searrow \text{Probabilità di Nera} = 1 - p \end{cases}$

- *Evento* : estrazione di n palline (con reimbussolamento)
- *Variabile X* : numero di *Bianche* tra le n

Immaginiamo di ottenere la seguente sequenza di x *Bianche* e $(n - x)$ *Nere*

Con probabilità

$$\begin{array}{ccccccc} & \xleftarrow{n} & & & & & \xrightarrow{} \\ B & B & B & B & \dots & B & N & N & N & \dots & N \\ p & p & p & p & \dots & p & (1-p) & (1-p) & (1-p) & \dots & (1-p) \\ & \xleftarrow{p^x} & & & & & \xrightarrow{(1-p)^{n-x}} & & & & \end{array}$$

Ogni altra sequenza con x *Bianche* e $(n - x)$ *Nere* ha la stessa probabilità $p^x(1 - p)^{n - x}$ di verificarsi: si tratta quindi di eventi equiprobabili la cui probabilità totale è data dalla somma delle singole probabilità, tutte uguali tra loro.

Quante sono queste sequenze equivalenti ?

Quante sono le sequenze equivalenti

- Tante quante le *combinazioni* che si possono formare con le x *Bianche* nella sequenza delle n palline estratte (*combinazioni* di n oggetti a x a x)
- Questo numero è rappresentato con il simbolo $\binom{n}{x}$ “ n sopra x ”

$$\binom{n}{x} = \frac{n!}{x! (n-x)!}$$

- $n! = n \times (n-1) \times (n-2) \times \dots \times 3 \times 2 \times 1$ cioè n per tutti i suoi precedenti

Esempio: $\binom{5}{3} = \frac{5!}{3! (5-3)!} = 10$

Distribuzione (modello) *Binomiale* $X \sim \text{Bin}(n, p)$

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \quad \text{con } x = 0, 1, 2, \dots, n; 0 \leq p \leq 1$$

Osservazioni

(una coincidenza davvero sorprendente ?!)

Chiamiamo $\binom{n}{x}$ *coefficiente binomiale*

- per $n = 2$ i coefficienti sono
1 2 1
- per $n = 3$
1 3 3 1
- per $n = 4$
1 4 6 4 1
- per $n = 5$
1 5 10 10 5 1
- per $n = 6$
1 6 15 20 15 6 1
-
... ..

sono gli stessi coefficienti della potenza n -esima di un *binomio*!

Osservazioni (potenze di un binomio)

$\binom{n}{x}$ *coefficiente binomiale*

- $(a + b)^2 = aa + ab + ba + bb$
 $\begin{array}{ccccccc} & \longleftrightarrow & & \longleftrightarrow & & \longleftrightarrow & \\ & 1 & & 2 & & 1 & \end{array}$

- $(a + b)^3 = aaa + aab + aba + baa + abb + aba + bba + bbb$
 $\begin{array}{ccccccccccc} & \longleftrightarrow & & \longleftrightarrow & & \longleftrightarrow & & \longleftrightarrow & & \longleftrightarrow & \\ & 1 & & 3 & & 3 & & 1 & & & \end{array}$

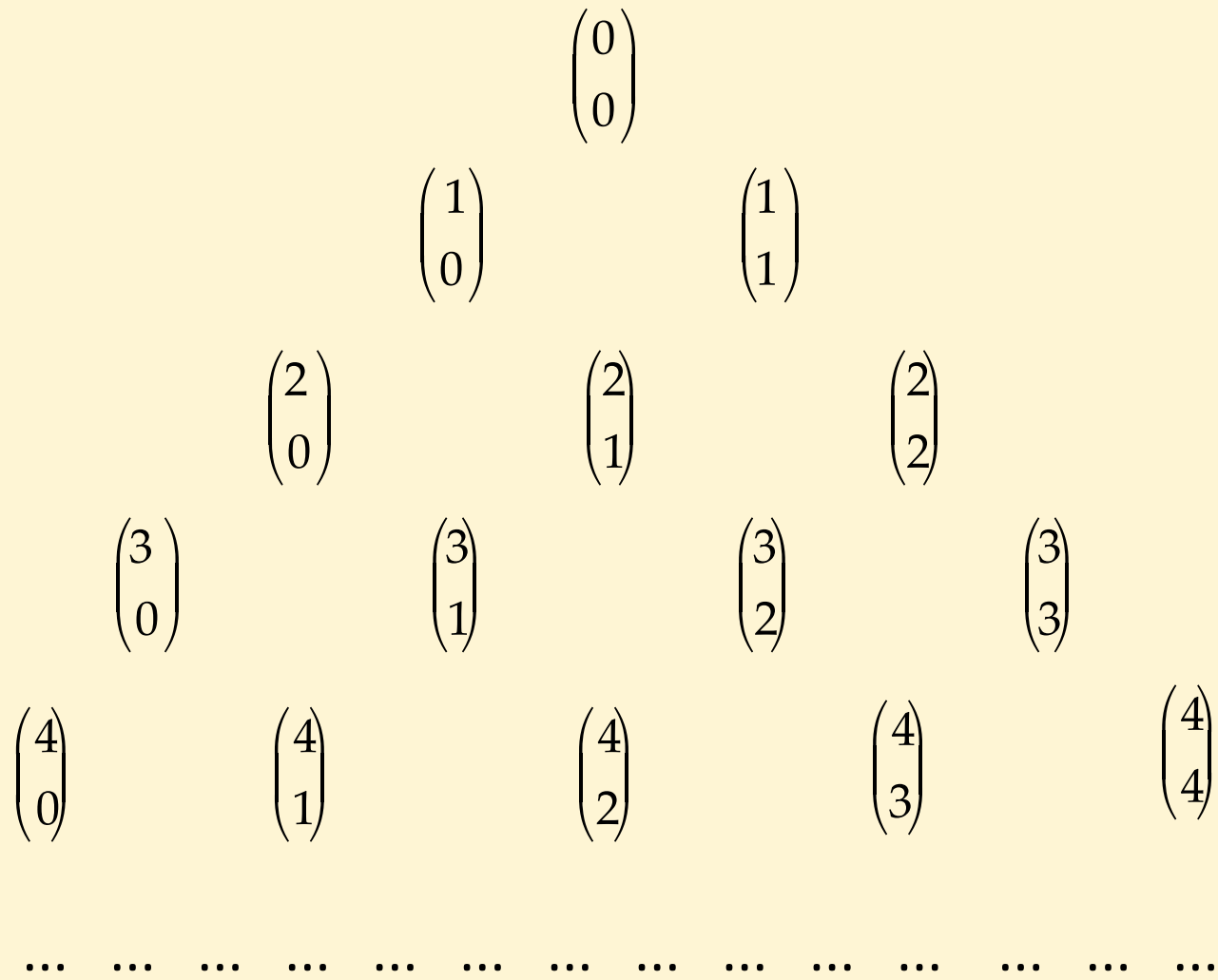
... ..

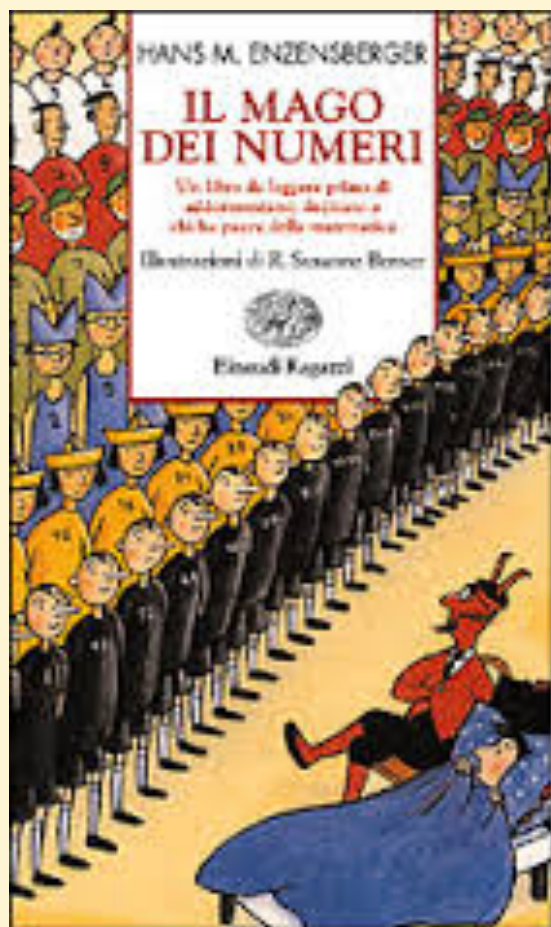
Per completezza

- $(a + b)^0 = 1$

- $(a + b)^1 = a + b$
 $\begin{array}{cc} \longleftrightarrow & \longleftrightarrow \\ 1 & 1 \end{array}$

Il triangolo di Tartaglia
(un automatismo: geniale!)





Torniamo alla realtà

Esercizio su un caso reale

Un'antica sala conferenze che deve ospitare una tavola rotonda è attrezzata con 8 microfoni, 3 dei quali sono mal funzionanti.

I 5 relatori invitati si recano, ciascuno **indipendentemente** dagli altri, a provare i microfoni prima dell'inizio. Ogni relatore accende uno dei microfoni scelto a caso e ne verifica il funzionamento.

- a) Con quale probabilità i 5 relatori potrebbero non accorgersi che vi sono alcuni microfoni fuori uso?
- b) Con quale probabilità tutti i relatori potrebbero convincersi che non vi sono microfoni funzionanti?

Torniamo alla realtà (*soluzione esercizio*)

Urna contenente 8 “palline” di cui 3 *Bianche* (difettosi), il 37.5%.

Estrazioni con reimbussolamento. X = numero microfoni difettosi su $n = 5$ “estratti”

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} \quad ; \quad x = 0, 1, 2, 3, 4, 5$$

a) Nessuno dei microfoni “estratti” è difettoso:

$$P(X = 0) = \binom{5}{0} 0.375^0 \times 0.625^5 = 0.095$$

b) Tutti i microfoni “estratti” sono difettosi:

$$P(X = 5) = \binom{5}{5} 0.375^5 \times 0.625^0 = 0.007$$

Osservazioni (*soluzione esercizio*)

In via più diretta:

a) Probabilità di microfono “non difettoso”: $(1 - p) = 0.625$

→ 5 non difettosi tra i 5 estratti: $(1 - p)^5$

b) Probabilità di 5 difettosi tra i 5 estratti: p^5

Naturalmente ci si può domandare qual è la probabilità di trovarne 0 , 1 , 2 ...



Altro esercizio

Esercizio 2

Un venditore di piccoli elettrodomestici che vende porta-a-porta, in base alla propria esperienza, valuta pari a 0.4 la probabilità che bussando alla porta di un cliente, concluda positivamente la vendita.

Nel corso della mattina deve visitare 7 possibili clienti scelti a caso da un elenco numerosissimo.

- a) *Calcolare la probabilità che il venditore concluda al massimo una vendita.*
- b) *Calcolare la probabilità che le vendite realizzate siano almeno 5.*

Soluzione esercizio 2

Urna contenente “palline” (clienti) di cui il 40% *Bianche* (clienti che acquistano).

NOTA: i clienti sono “estratti” a caso da un elenco molto numeroso, quindi estrarne un piccolo numero non cambia sensibilmente la composizione dell’urna, anche se essi non vengono “reimbussolati” a ogni estrazione.

Si possono dunque considerare le estrazioni con probabilità di *Bianca* $p = 0.4$ costante.

X = numero di clienti che acquistano su $n = 7$ “estratti”

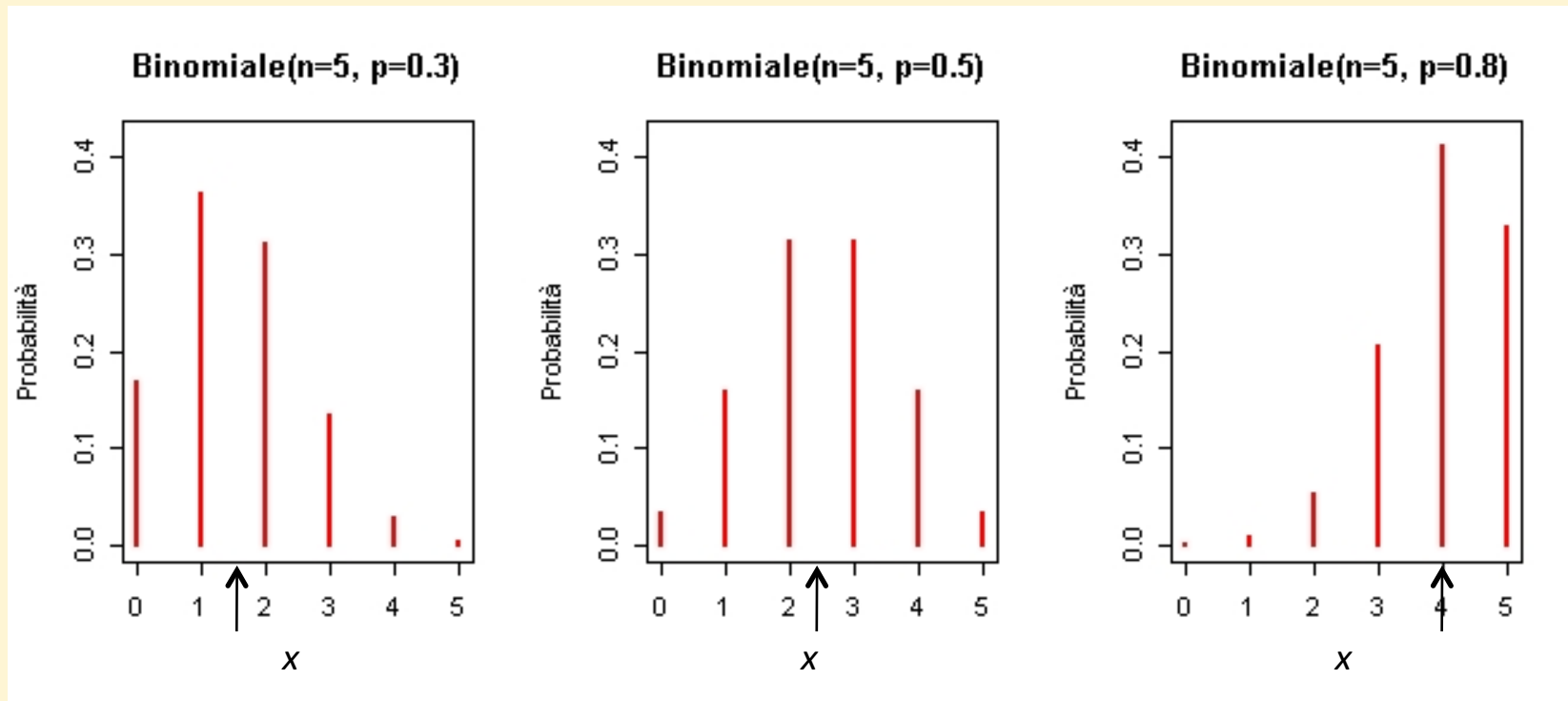
a) Una vendita al massimo: $P(X \leq 1) = p_0 + p_1$

$$P(X = 0) + P(X = 1) = \binom{7}{0} 0.4^0 \times 0.6^7 + \binom{7}{1} 0.4 \times 0.6^6 = 0.159$$

b) Almeno 5 vendite: $P(X \geq 5) = p_5 + p_6 + p_7$

$$P(X = 5) + P(X = 6) + P(X = 7) = 0.096$$

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x} ; \mu_x = np ; \sigma_x^2 = np(1 - p) \quad (*)$$

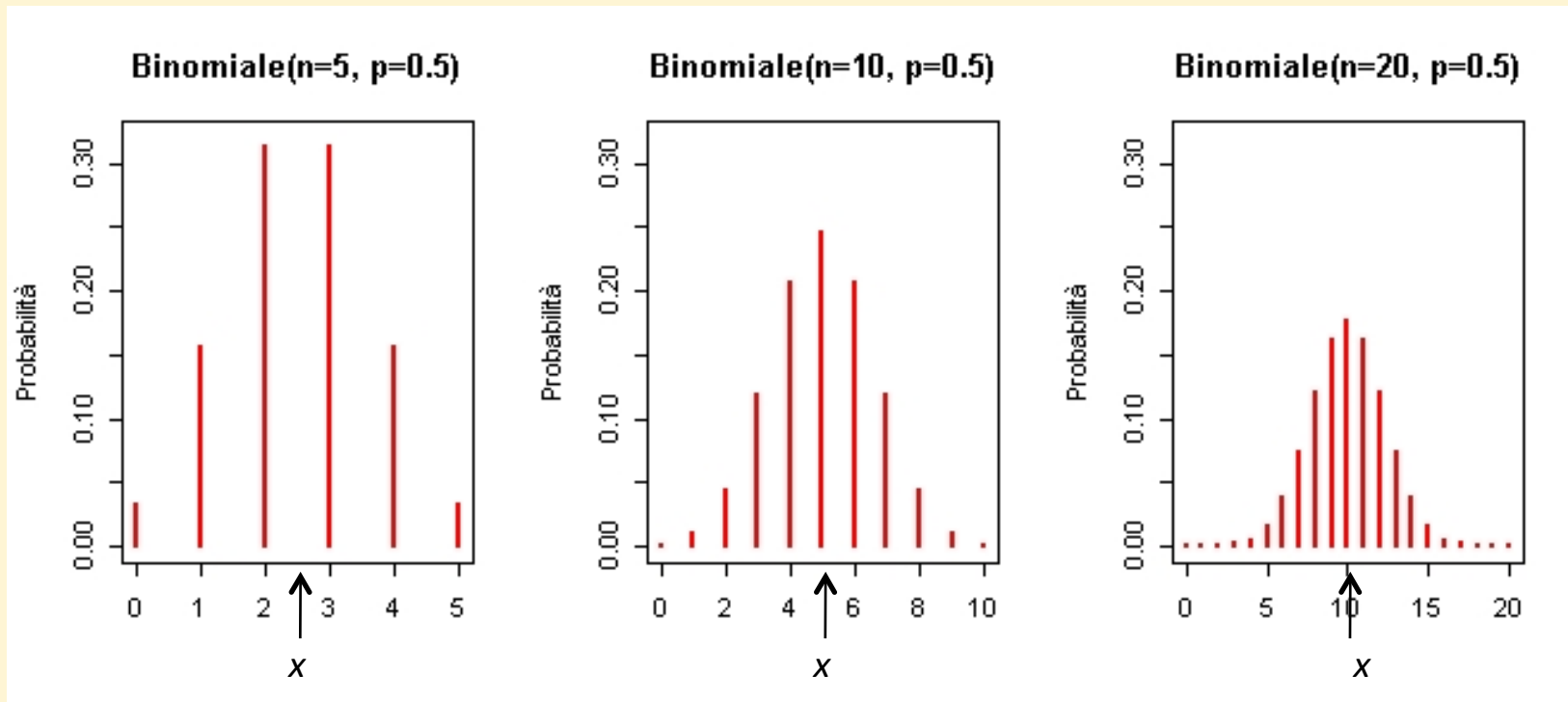


$$\mu_x = 1.5 ; \sigma_x^2 = 1.05$$

$$\mu_x = 2.5 ; \sigma_x^2 = 1.25$$

$$\mu_x = 4 ; \sigma_x^2 = 0.80$$

(*) vedi appendice



$$\mu_x = 2.5 ; \sigma_x^2 = 1.25$$

$$\mu_x = 5 ; \sigma_x^2 = 2.5$$

$$\mu_x = 10 ; \sigma_x^2 = 0.5$$

Appendice

(media e varianza della distribuzione - modello - Binomiale)

- Caso particolare ($n=1$): $p^x (1-p)^{1-x}$ $x = 0,1 \rightarrow$ modello Bernoulliano
o
indicatore di evento

$$\text{media: } \mu = \sum_{x=0}^1 x \cdot p^x (1-p)^{1-x} = 0 \cdot p^0 (1-p) + 1 \cdot p(1-p)^0 = p$$

$$\text{varianza: } \sigma^2 = \sum_{x=0}^1 (x - \mu_x)^2 \cdot p^x (1-p)^{1-x} = p(1-p)$$

- Osservando che la *Binomiale* è la somma di n *Bernoulliane* tra loro indipendenti segue che la media e la varianza della *Binomiale* sono rispettivamente:

$$\mu_x = np \ ; \ \sigma_x^2 = np(1-p)$$

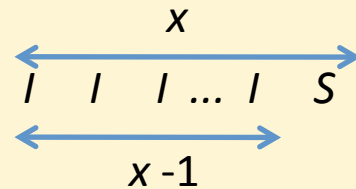
Distribuzione *Geometrica*

Urna con composizione assegnata; estrazioni con reimbussolamento.

Quante estrazioni devo effettuare per avere 1 *Bianca*?

Es. moneta: quanti lanci per 1 *Testa*? Quante prove per 1 *Successo (S)*?

$X = \#$ prove (ripetute) per ottenere 1 S ; $P(S) = p$, $P(\bar{S} = I) = 1 - p$

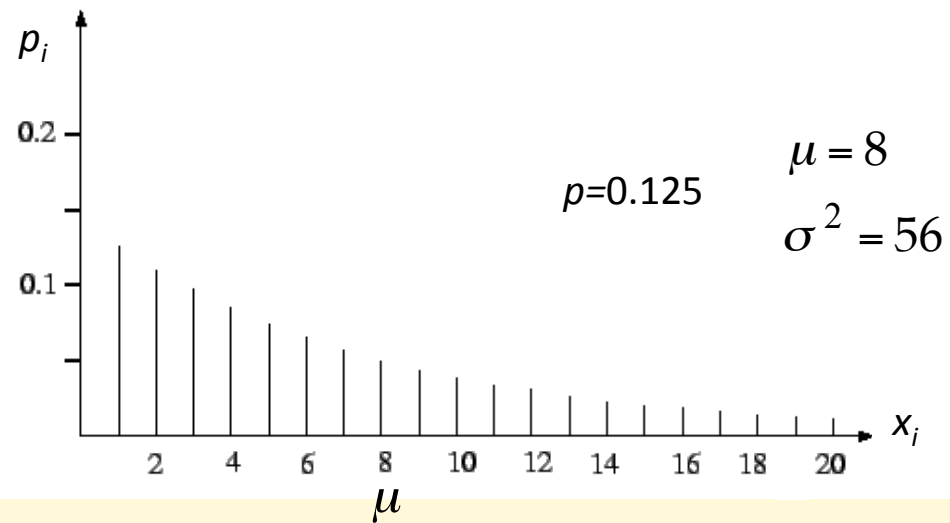
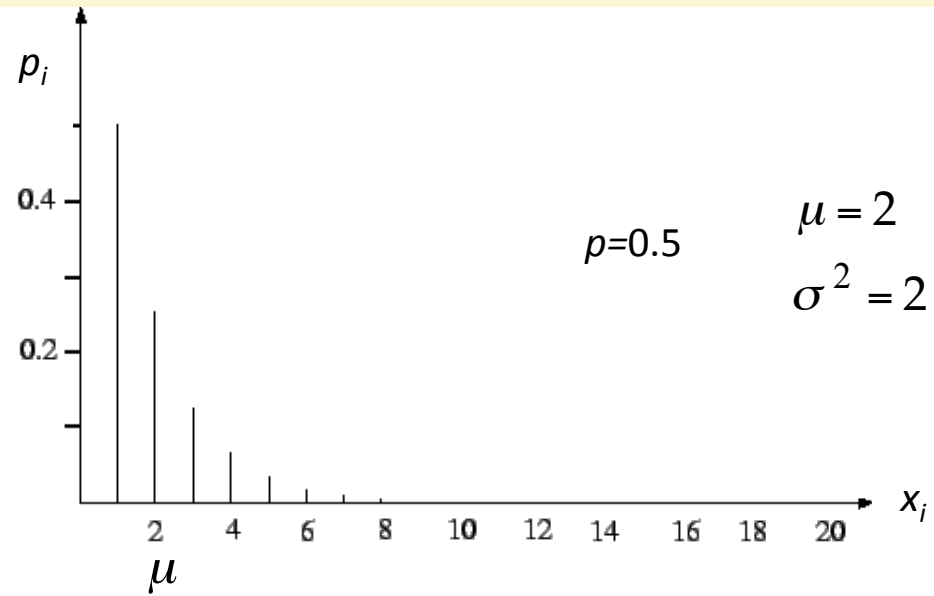


$$P(X = x) = \underbrace{(1-p) \cdot (1-p) \cdot (1-p) \cdots (1-p)}_{x-1} \cdot p = p \cdot (1-p)^{x-1}$$

$$X \sim Ge(p) ; x = 1, 2, \dots$$

$$E(X) = \mu = \frac{1}{p} ; \sigma^2 = \frac{1-p}{p^2}$$

Distribuzione Geometrica



Esempio di applicazione

Uno studente vuole trovare una ragazza come amica. Ha deciso di dare appuntamento singolarmente a ogni ragazza che conosce. Le fanciulle non si conoscono tra di loro, e l'esito di ogni incontro non influenza l'esito del successivo. Si assume dunque che gli esiti di questi incontri siano tra loro indipendenti.

Lo studente (ottimista) valuta pari a 0.2 (20%) la probabilità che un appuntamento vada bene.

Qual è la probabilità che siano necessari 5 tentativi perché un incontro “riesca”?

$X = \#$ incontri per ottenere 1 S ; $P(S) = 0.2$, $P(I) = 0.8$; $X \sim Ge(p)$; $x = 1, 2, \dots$

$$P(X = 5) = 0.2 \cdot 0.8^4 = 0.08$$

... che vada bene al secondo incontro?

$$P(X = 2) = 0.2 \cdot 0.8 = 0.16$$

Formuliamo la domanda in altro modo (diversa *v.a.*)

Uno studente vuole trovare una ragazza come amica. Ha deciso di dare 5 appuntamenti singolarmente a delle ragazze. Le fanciulle non si conoscono tra di loro, e l'esito di ogni incontro non influenza l'esito del successivo. Si assume dunque che gli esiti di questi incontri siano tra loro indipendenti.

Lo studente (ottimista) valuta pari a 0.2 (20%) la probabilità che un appuntamento vada bene.

Qual è la probabilità che 1 tentativo "riesca"

$X = \#$ Successi su 5 incontri ; $P(S) = 0.2$, $P(I) = 0.8$; $X \sim Bi(5, 0.2)$; $x = 0, 1, \dots, 5$

$$P(X = 1) = \binom{5}{1} 0.2 \cdot 0.8^4 = 0.41$$

... che vadano bene 2 incontri?

$$P(X = 2) = \binom{5}{2} 0.2^2 \cdot 0.8^3 \cong 0.20$$

... altro modo ancora (diversa v.a.)

Uno studente vuole trovare una ragazza come amica. Ha deciso di dare 5 appuntamenti singolarmente a delle ragazze. Le fanciulle non si conoscono tra di loro, e l'esito di ogni incontro non influenza l'esito del successivo. Si assume dunque che gli esiti di questi incontri siano tra loro indipendenti.

Lo studente (ottimista) valuta pari a 0.2 (20%) la probabilità che un appuntamento vada bene.

Qual è la probabilità che almeno un tentativo riesca?

$$X \geq 1 ; P(S) = 0.2 , P(I) = 0.8 ; \quad X \sim Bi(5, 0.2) ; x = 0, 1, \dots, 5$$

$$\begin{aligned} P(X \geq 1) &= \sum_{x=1}^5 \binom{5}{x} 0.2^x \cdot 0.8^{5-x} = \\ &= \binom{5}{1} 0.2^1 \cdot 0.8^{5-1} + \binom{5}{2} 0.2^2 \cdot 0.8^{5-2} + \dots + \binom{5}{5} 0.2^5 \cdot 0.8^0 = 0.67 \end{aligned}$$

$$\text{altra via} \quad P(X \geq 1) = 1 - P(X = 0) = 1 - 0.8^5 = 0.67$$

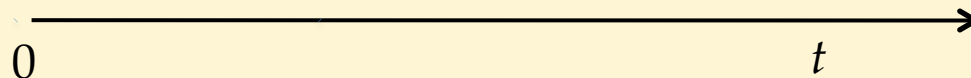
Modello di *Poisson* $P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda} \quad x = 0, 1, 2, \dots ; \lambda \in R^+$

Genesi:

- un evento si può ripetere un numero x di volte in un prefissato intervallo di tempo $[0, t]$, ad es. arrivi casuali a uno sportello, fulmini durante temporale, chiamate a un centralino telefonico, contatti su un sito internet, ...

assunzioni

- $i)$ eventi indipendenti tra loro
- $ii)$ la prob. di un evento nell'intervallo Δt è proporzionale a Δt : $\lambda \Delta t + o(\Delta t)$
- $iii)$ la prob. di più eventi in Δt è infinitesima : $o(\Delta t)$



dividiamo l'intervallo $[0, t]$ in n intervallini di ampiezza Δt
tali che valgano le i, ii, iii

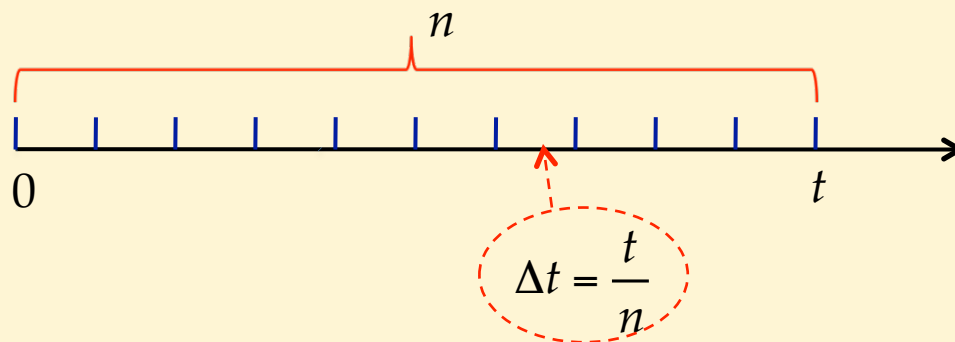
Modello di *Poisson* $P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}$ $x = 0, 1, 2, \dots$; $\lambda \in \mathbb{R}^+$

Genesis:

- un evento si può ripetere un numero x di volte in un prefissato intervallo di tempo $[0, t]$, ad es. arrivi casuali a uno sportello, fulmini durante temporale, chiamate a un centralino telefonico, contatti su un sito internet, ...

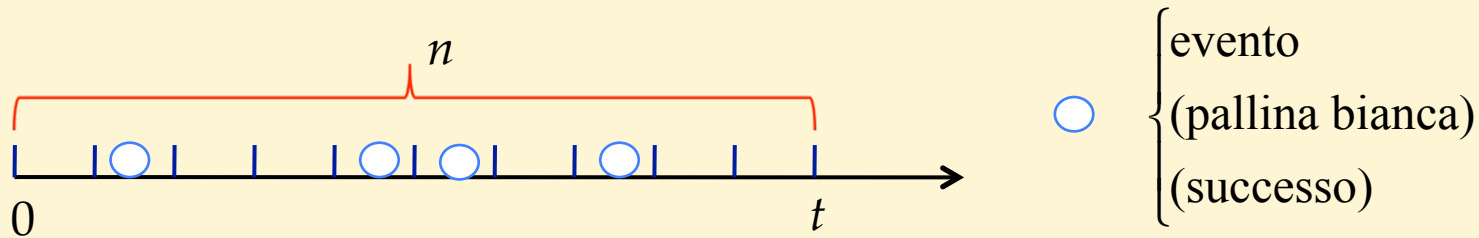
assunzioni

- i)* eventi indipendenti tra loro
- ii)* la prob. di un evento nell'intervallino Δt è proporzionale a Δt : $\lambda \Delta t + o(\Delta t)$
- iii)* la prob. di più eventi in Δt è infinitesima : $o(\Delta t)$



La prob. che si verifichi un evento in Δt (intervallino "pieno") è pari a $\lambda \frac{t}{n}$ (*ii*)

Modello di *Poisson* $P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}$ $x = 0, 1, 2, \dots$; $\lambda \in \mathbb{R}^+$



Urna con composizione assegnata, n estrazioni indipendenti: $X = \#$ di bianche su n

$$X \sim Bi(n, p = \lambda \frac{t}{n})$$

$$P(X = x) = \binom{n}{x} \left(\lambda \frac{t}{n}\right)^x \left(1 - \lambda \frac{t}{n}\right)^{n-x} = \text{poniamo } t = 1, \text{ per semplicità}$$

$$= \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-x} =$$

$$= \frac{n(n-1)(n-2) \cdots (n-x+1)}{n^x} \cdot \frac{\lambda^x}{x!} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-x}$$

passando al lim $n \rightarrow \infty$

1 $e^{-\lambda}$ 1

Modello di *Poisson* $P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}$ $x = 0, 1, 2, \dots$; $\lambda \in \mathbb{R}^+$

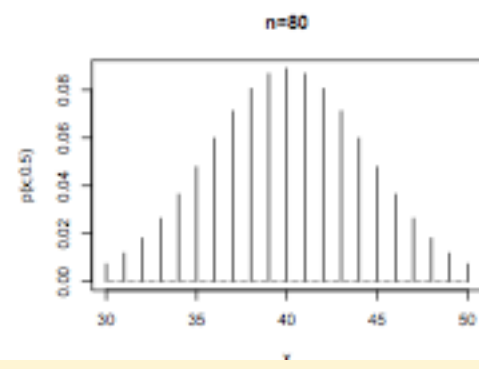
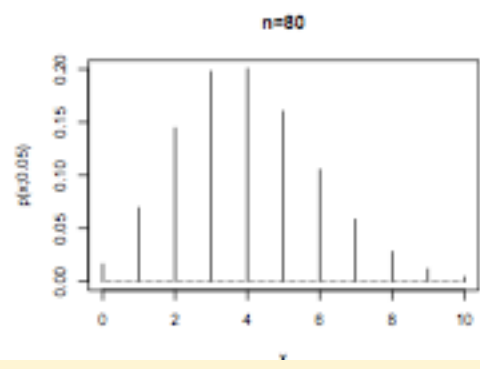
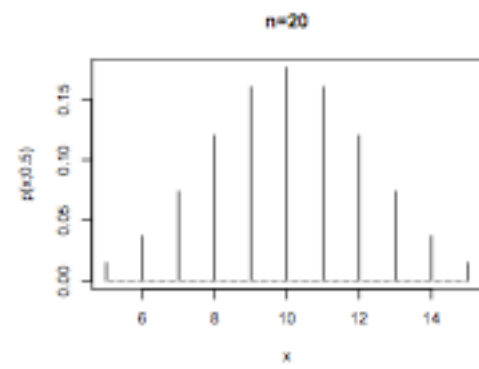
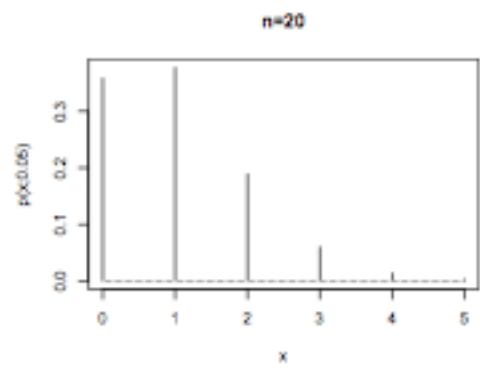
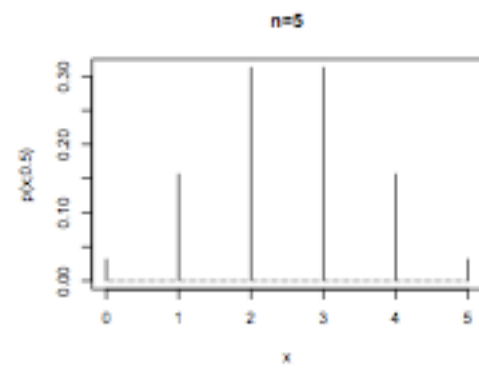
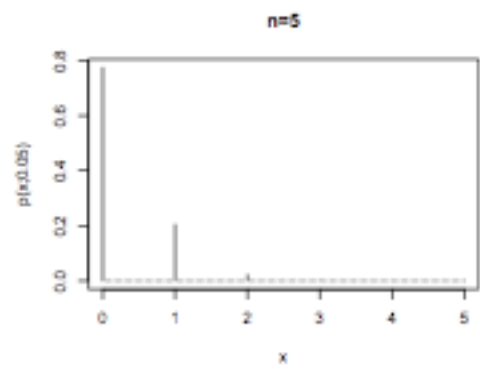
In definitiva, all'*infittirsi* degli intervallini $\Delta t = \frac{1}{n}$, mantenendo λ costante

$$P(X = x) = \binom{n}{x} \left(\lambda \frac{1}{n}\right)^x \left(1 - \lambda \frac{1}{n}\right)^{n-x} \rightarrow \frac{\lambda^x}{x!} e^{-\lambda}$$

$$\mu = \lambda$$

$$\sigma^2 = \lambda$$

proprietà esclusiva della distribuzione di *Poisson*



2 Esercizi (*Poisson*)

- 1) Un commerciante di elettrodomestici sa che la domanda settimanale di apparecchi televisivi Sogny segue una distribuzione di *Poisson* con media pari a 1. Avendo in magazzino una scorta di 3 televisori Sogny, qual è la probabilità che in una settimana il commerciante non riesca a soddisfare la richiesta di quei televisori?

$X = \#$ di televisori richiesti in una settimana $\rightarrow X \sim Po(\lambda = 1)$

$$P(X = x) = \frac{1^x}{x!} e^{-1}$$

$$P(X > 3) = 1 - P(X \leq 3) = 1 - (p_0 + p_1 + p_2 + p_3) = 0.019$$

$$p_0 = \frac{1}{e} = 0.368$$

$$p_1 = \frac{1}{e} = 0.368$$

$$p_2 = \frac{1}{2e} = 0.184$$

$$p_3 = \frac{1}{6e} = 0.061$$

$$\rightarrow p_0 + p_1 + p_2 + p_3 = 0.981$$

2) Il numero delle spedizioni giornaliere di un'azienda è una *v.a.* del tipo di *Poisson* con media pari a 0.7.

a) Si calcoli la probabilità che in una certa giornata debba effettuare esattamente 3 spedizioni.

b) Sapendo che ogni spedizione costa 4.50 €, si determini la somma che l'azienda dovrà prevedere in bilancio per le spedizioni settimanali.

a) $X = \#$ di spedizioni giornaliere $\rightarrow X \sim Po(\lambda = 0.7)$

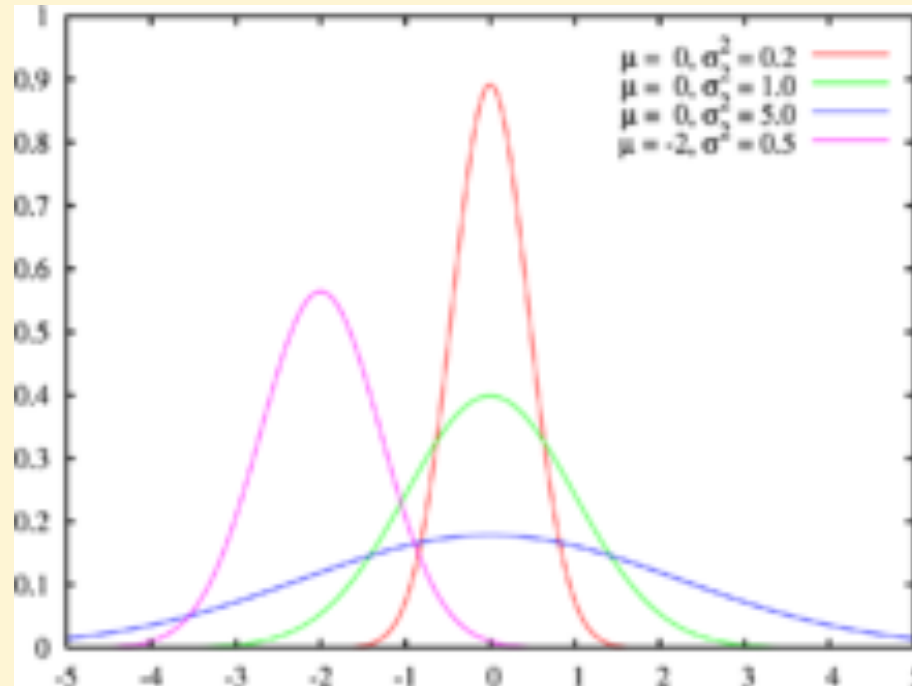
$$P(X = x) = \frac{0.7^x}{x!} e^{-0.7}$$

$$P(X = 3) = \frac{0.7^3}{3!} e^{-0.7} = 0.028$$

b) numero medio di spedizioni in una settimana (5 gg) $= 5 \times 0.7 = 3.5$

Spesa settimanale attesa $= 4.50 \times 3.5 = 15.75$

La curva degli errori di Gauss (*Gaussiana* o *Normale*)



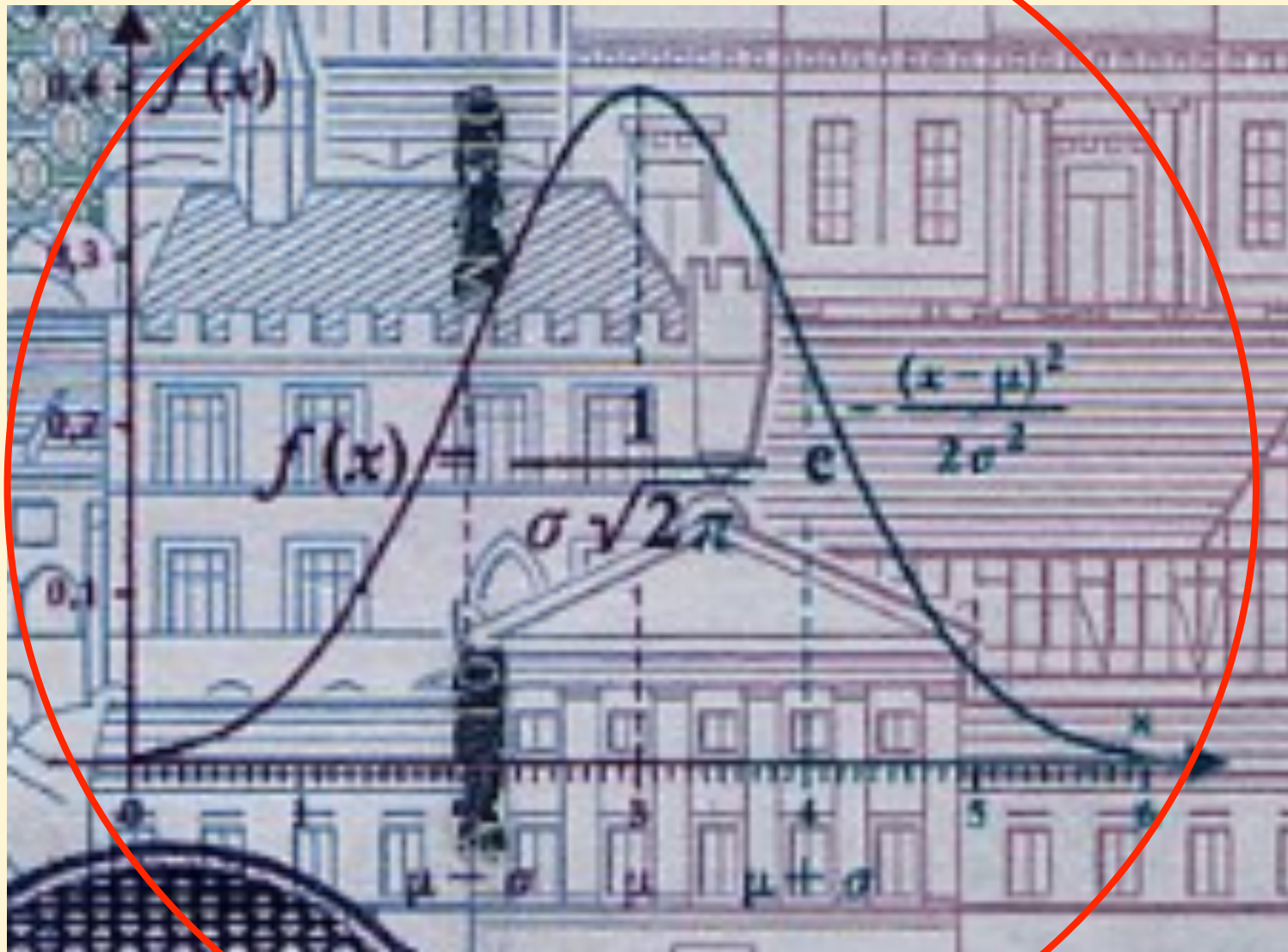
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \begin{array}{l} -\infty < x < \infty \\ -\infty < \mu < \infty \\ \sigma > 0 \end{array}$$

GN4480100S8

Deutsche Bundesbank
Wolfgang Krauß
Frankfurt am Main
1. September 1999



sino al 1° gennaio 2002



La curva degli errori accidentali
(*Curva di Gauss; Curva di Gauss-Laplace*)
Distribuzione di probabilità *Normale* (o *Gaussiana*)

- La prima formulazione è attribuibile a Abraham de Moivre (1667 Francia – 1754 Inghilterra) che la costruì nel 1733 ma i suoi scritti andarono persi sino al loro ritrovamento nel 1924 da parte di Karl Pearson (1857 – 1936 Londra) che gliene restituì il merito e la denominò estensivamente con il termine “*normale*” già coniato da altri (Charles Sanders Peirce, Wilhelm Lexis, Francis Galton intorno al 1875)
- Nel 1783 Pierre Simon de Laplace (1749 – 1827 Francia) la utilizzò per descrivere la distribuzione degli errori accidentali di misura.
- Nel 1809 Carl Friedrich Gauss (1777 – 1855 Germania) la utilizzò per lo studio di dati astronomici e ne approfondì e divulgò le proprietà

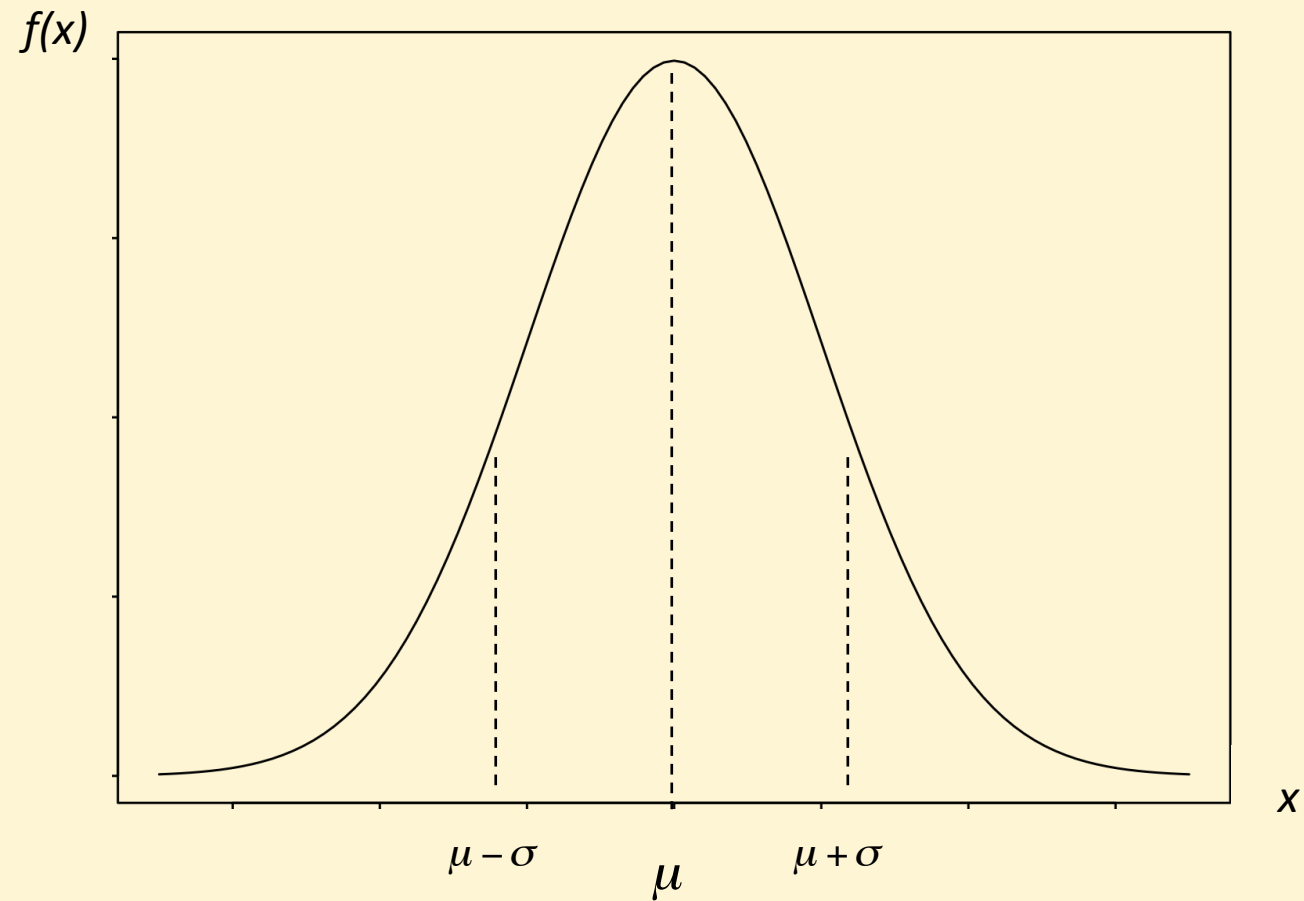
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty \ ; \ -\infty < \mu < \infty \ ; \ \sigma > 0$$

La funzione $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Tralasciandone la genesi e la natura probabilistica e riguardando la $f(x)$ come una funzione reale di variabile reale, possiamo studiarla e costruirne il grafico

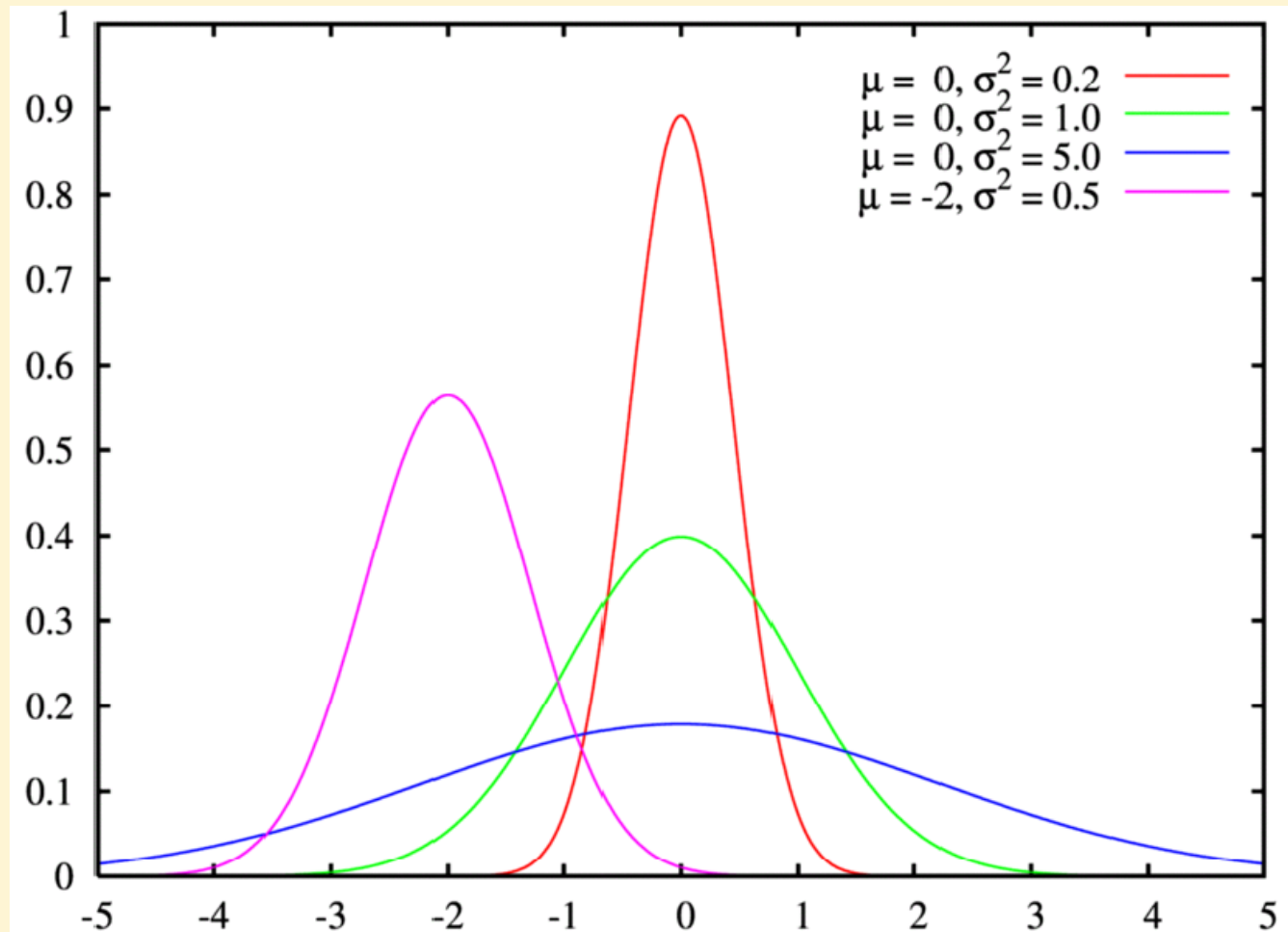
- *campo di definizione* : $-\infty < x < \infty$
- *positiva* : $f(x) \geq 0$, $\forall x$
- *simmetrica rispetto all'asse* $x = \mu$: $f[(x - \mu)] = f[-(x - \mu)]$
- *asintoto orizzontale* $y = 0$: $\lim_{x \rightarrow \pm\infty} f(x) = 0$
- *punto di massimo in* $x = \mu$: $\max_x f(x) = f(\mu)$
- *flessi ascendente e discendente, rispettivamente, in* $x = \mu \mp \sigma$

La funzione $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$



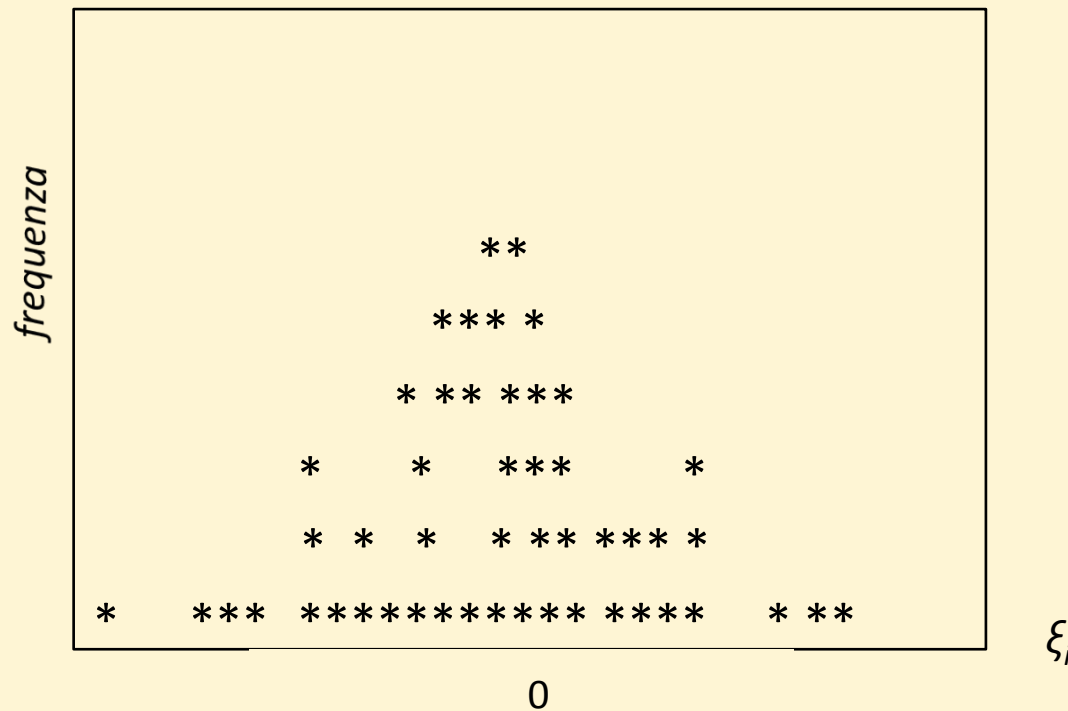
La funzione $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

dipendenza dai parametri μ e σ

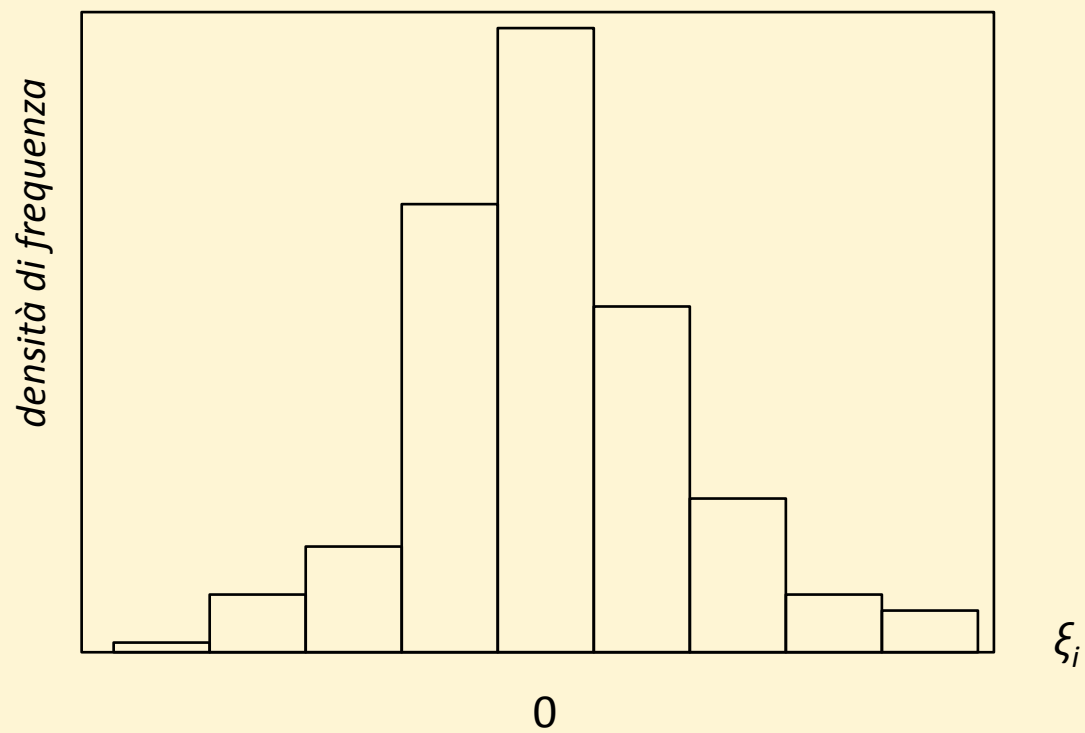


verso la funzione di densità di probabilità $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

n misure *ripetute* su una grandezza di valore incognito μ (*lunghezza, peso, tempo, ecc.*), denotando i valori osservati con $x_i = \mu + \xi_i$ ($i = 1, 2, \dots, n$) con ξ_i *errori accidentali* di misura



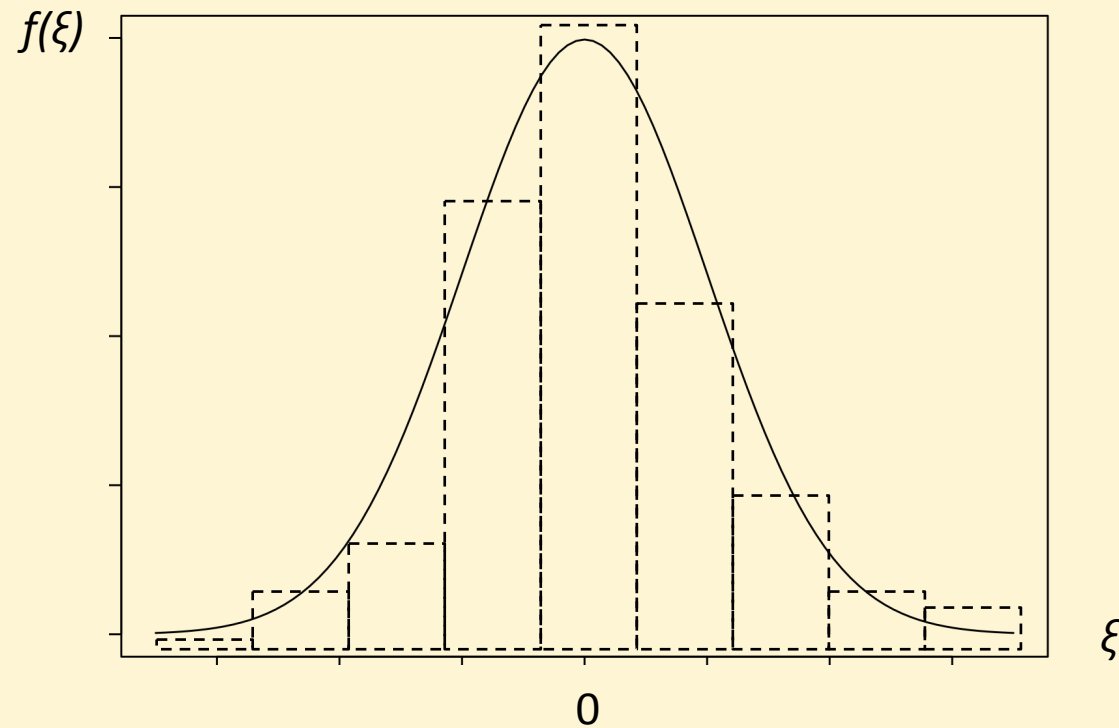
all'aumentare del numero n delle misure è opportuno raggruppare i dati in classi e costruire l'istogramma delle frequenze



Infittendo gli intervalli l'istogramma tende ad "assumere la forma" di una *curva gaussiana* con media (valore atteso) $\mu = 0$ e deviazione standard σ .

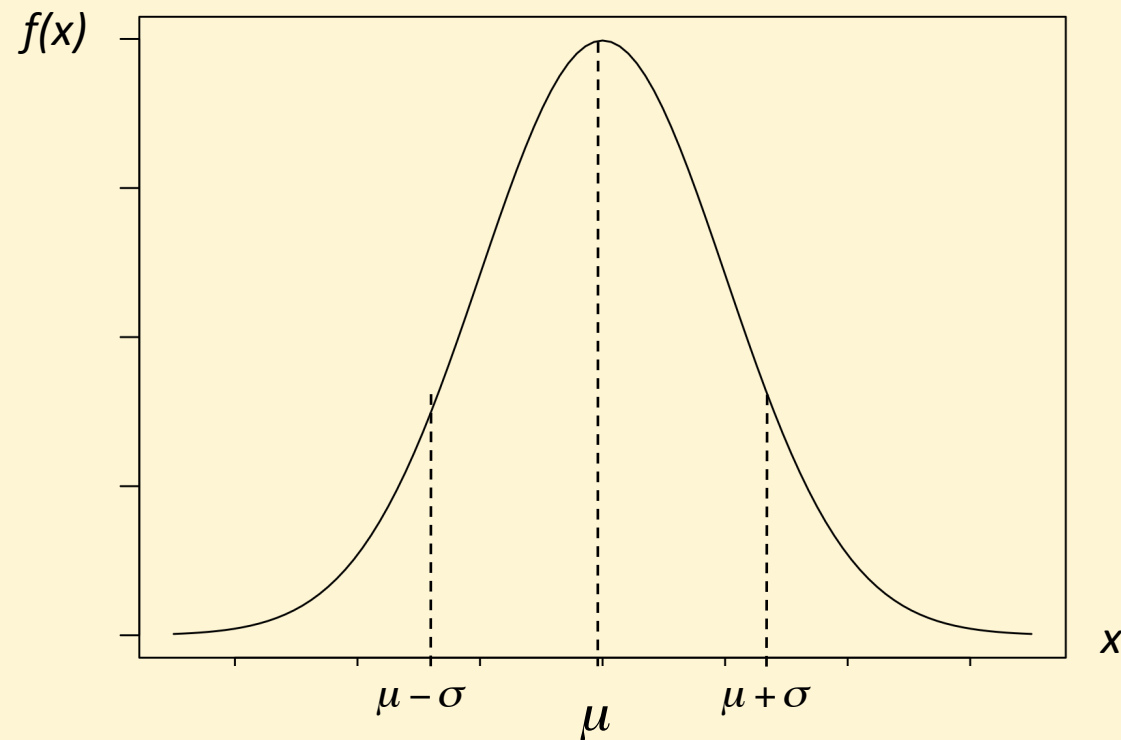
Nota: gli errori ξ sono *puramente accidentali*: se fosse $\mu \neq 0$ ci sarebbe una componente d'errore sistematica. σ rappresenta inversamente la precisione della misura.

$$f(\xi) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{\xi^2}{2\sigma^2}}$$



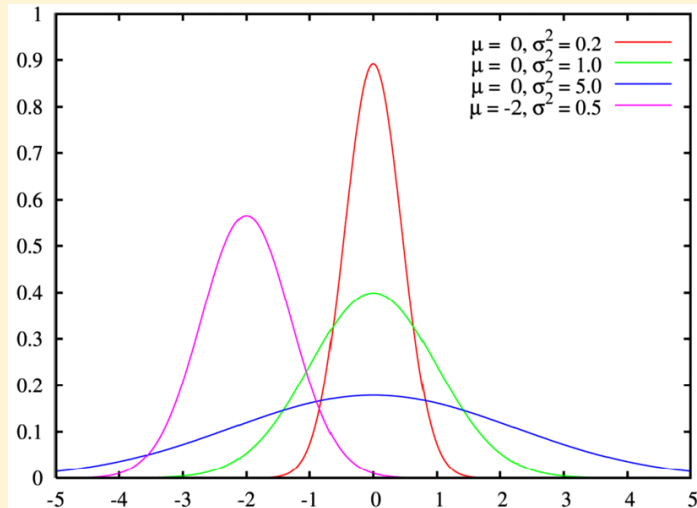
Dal modello probabilistico degli errori $f(\xi)$ si ottiene di conseguenza la distribuzione di probabilità dei valori di misura $x_i = \mu + \xi_i \rightarrow (\xi = x - \mu)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{con media } \mu \text{ e s.d.} = \sigma$$

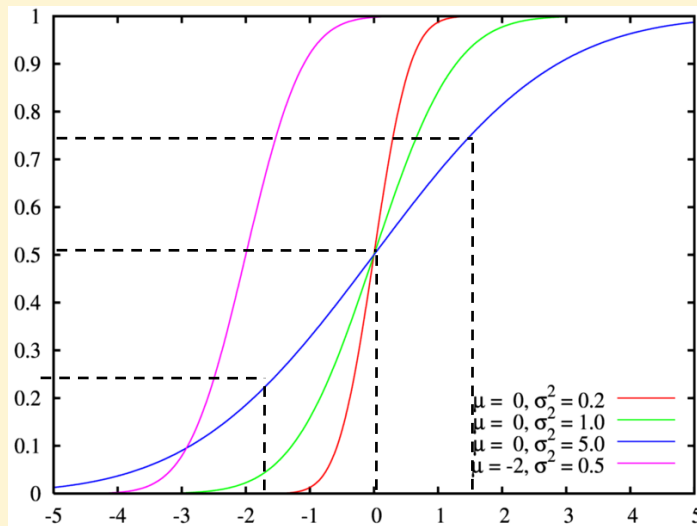


$X \sim \mathcal{N}(\mu, \sigma^2)$
v.a. *Normale*
(o *gaussiana*)

Densità e funzione di ripartizione di $X \sim \mathcal{N}(\mu, \sigma^2)$



densità di probabilità



funzioni di ripartizione (o cumulative)

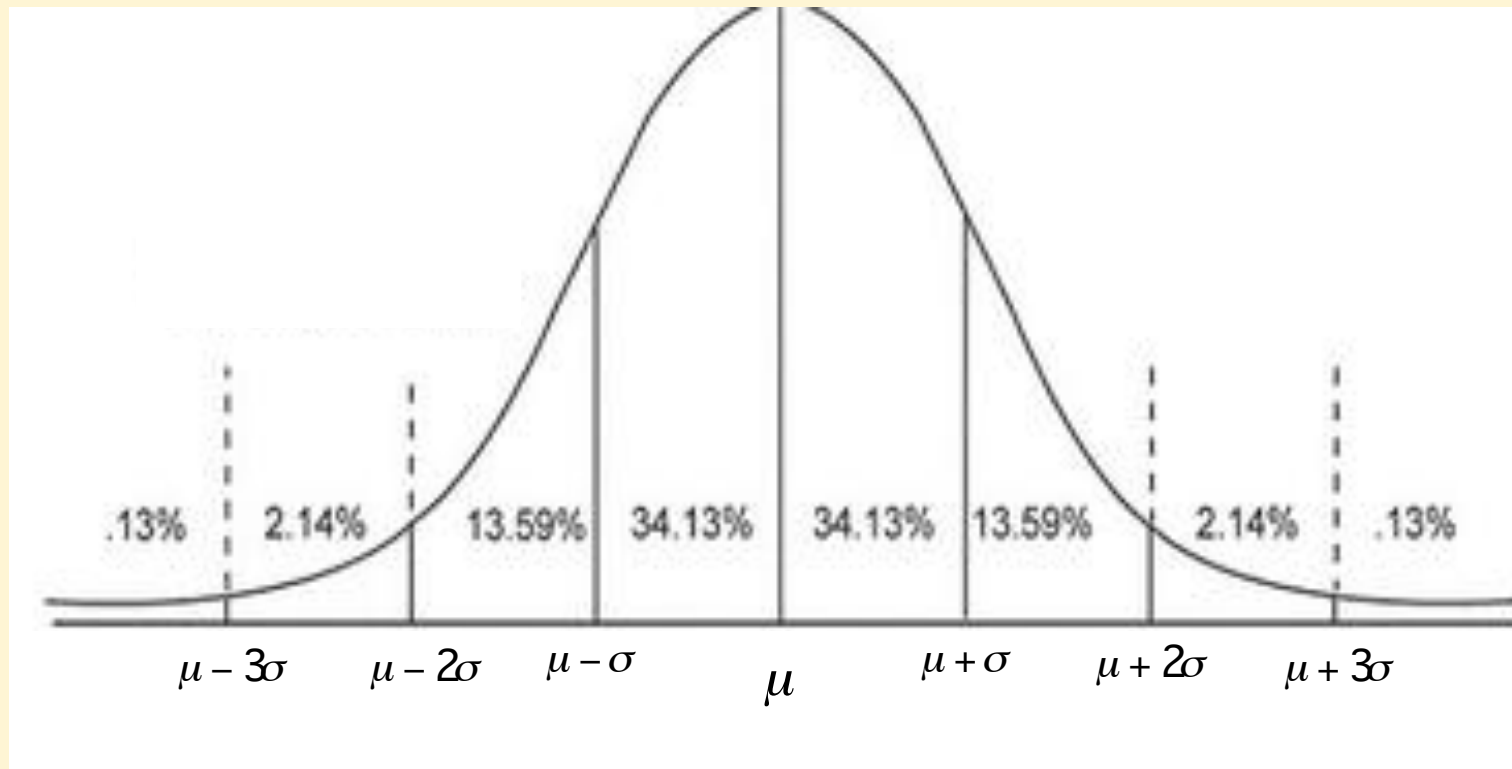
$$Q_1, Q_2=Me, Q_3$$

quartili per la curva blu

(analogamente per le altre)

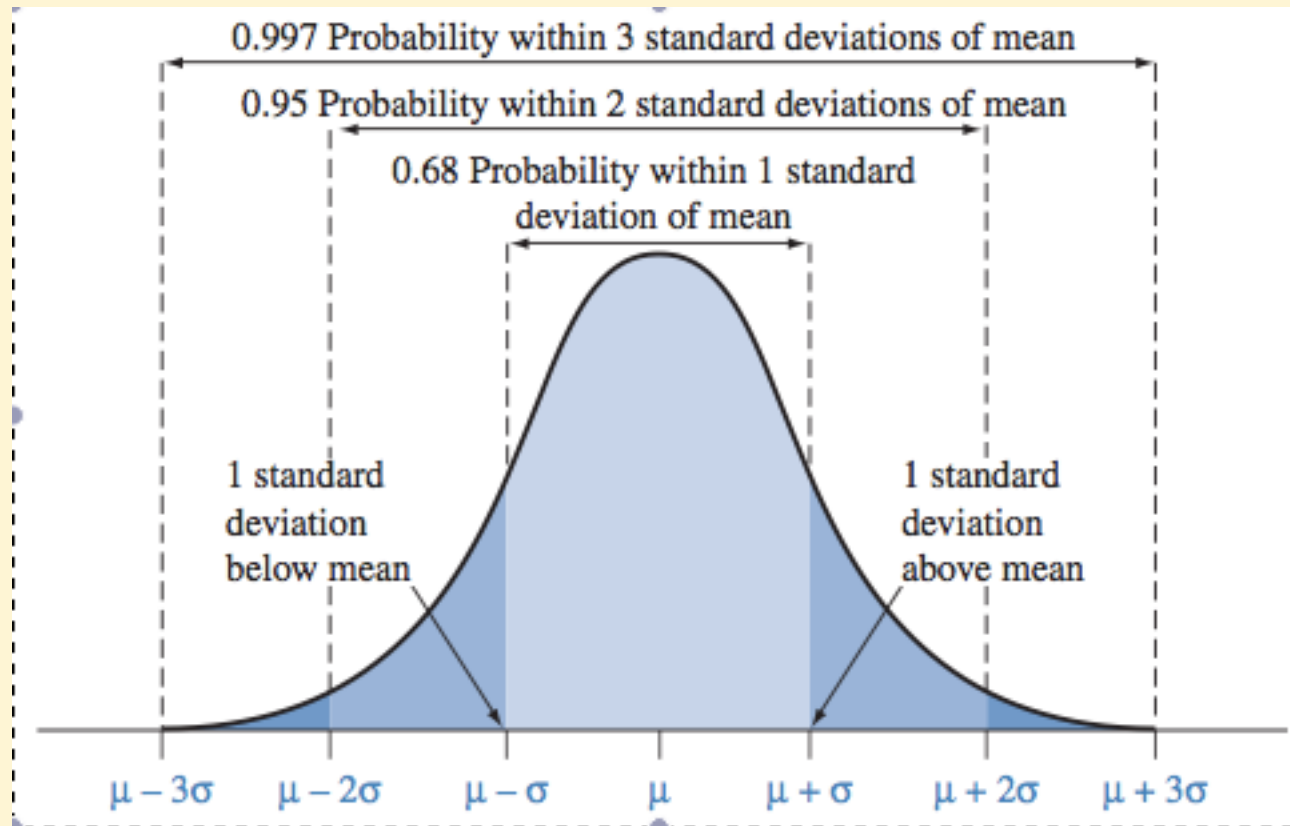
Q_1 Me Q_3

Aree sotto la curva (= probabilità)



Osservazione. I valori di x compresi in un intorno della media di 3σ raccolgono il 99.72% della probabilità: questo spiega perché, anche se per definizione x varia tra $-\infty$ e $+\infty$, il modello può rappresentare grandezze positive come lunghezze, pesi, tempi, ecc.

Aree sotto la curva (= probabilità)
(in sintesi)



Una osservazione che cade a una distanza $>$ di 3σ dalla media può essere considerata come un “*valore anomalo*” rispetto al modello adottato

Lo z-score

- Lo **z-score** per un'osservazione è il numero di deviazioni standard dalla media in cui essa cade. Uno z positivo indica che l'osservazione è alla destra della media, se invece è negativo essa cade alla sua sinistra

$$x - \mu = z \sigma \rightarrow z = \frac{x - \mu}{\sigma}$$

- Oltre a consentire di interpretare con immediatezza come si concentrano i dati e individuare i valori estremi, z converte i dati in una scala di misura adimensionale, permettendo il confronto tra dati provenienti da grandezze diverse.

Ad esempio:

- se l'altezza media di una classe di studenti è pari a 165 cm con s.d. = 10 cm, possiamo aspettarci che sia poco probabile che uno studente abbia un'altezza > 195 o < 135 cm
- analogamente: se in una classe primaria l'altezza media degli alunni è di 120 cm con s.d. = 6 (quindi più omogenei rispetto al caso precedente), sarà poco probabile trovare un alunno di altezza > 138 cm o < 102 cm.

Un'interessante proprietà (generale) sulla relazione tra σ e la distribuzione di probabilità di una v.a. \rightarrow *la diseguaglianza di Cebycev*

- Sia X una generica v.a. (continua o discreta) con media μ e s.d. σ , indichiamo con k un numero reale > 0 , vale la seguente relazione:

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

o, equivalentemente (considerando l'evento complementare a 1):

$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

esplicitando il valore assoluto:

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

es:

$$k = 1 \rightarrow P(\mu - \sigma < X < \mu + \sigma) \geq 0 \quad \text{ovvio!}$$

$$k = 2 \rightarrow P(\mu - 2\sigma < X < \mu + 2\sigma) \geq 1 - 1/4 = 0.75$$

$$k = 3 \rightarrow P(\mu - 3\sigma < X < \mu + 3\sigma) \geq 1 - 1/9 = 0.89$$

si ritrova, in *forma più debole*, quanto abbiamo visto nel caso di $X \sim \mathcal{N}(\mu, \sigma)$

Dalla *normale* $X \sim \mathcal{N}(\mu, \sigma^2)$ alla *Normale standard* $Z \sim \mathcal{N}(0, 1)$

- Per il calcolo delle aree sottese alla curva normale (probabilità di un certo intervallo $x_1 < X < x_2$) è necessario calcolare l'integrale

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

che però non è risolvibile analiticamente (in via elementare).

- E' necessario ricorrere a metodi numerici e tabulare i valori delle aree, in altri termini tabellare la funzione di ripartizione.
- Poiché è insensato pensare di poter fare questo per tutti i possibili valori dei parametri μ e σ è opportuna una riparametrizzazione, standardizzando la X :

$$Z = \frac{X - \mu}{\sigma} \rightarrow f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad (\mu_z = 0, \sigma_z^2 = 1)$$

Dalla *normale* $X \sim \mathcal{N}(\mu, \sigma^2)$ alla *Normale standard* $Z \sim \mathcal{N}(0, 1)$

Una volta calcolata numericamente la funzione di ripartizione della variabile Z

$$\int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \quad (\text{o, grazie alla simmetria, } \int_0^z \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz)$$

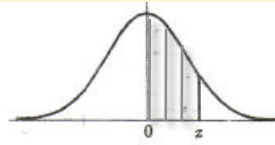
è possibile ricavare l'area sottesa a qualunque curva $\mathcal{N}(\mu, \sigma^2)$

$$P(x_1 < X < x_2) = P\left(\frac{x_1 - \mu}{\sigma} < Z < \frac{x_2 - \mu}{\sigma}\right) = \int_{z_1}^{z_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

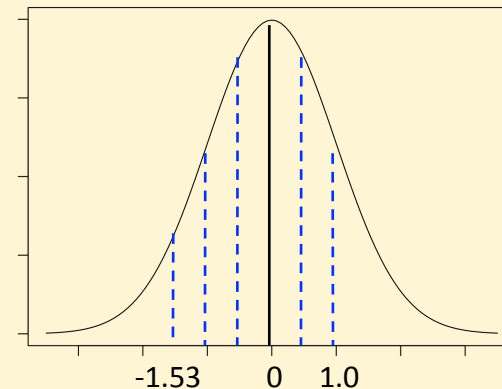
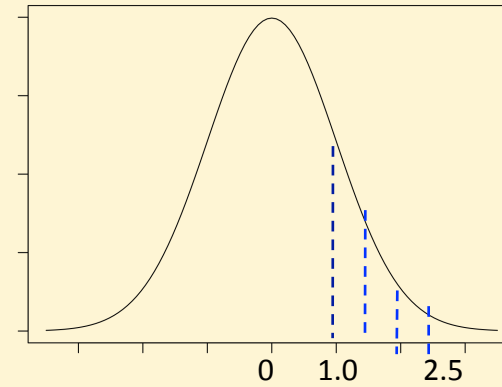
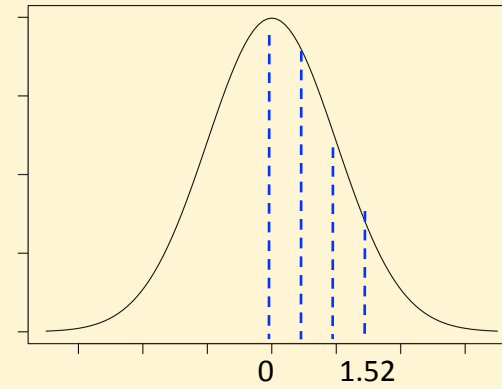
e, inversamente, risalire ai valori originari della X dalla funzione di ripartizione della Z.

Uso delle tavole della distribuzione *Normale standard*

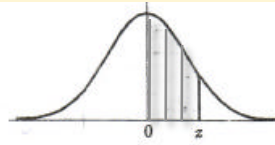
AREE
sotto la
CURVA NORMALE
STANDARDIZZATA
da 0 a z



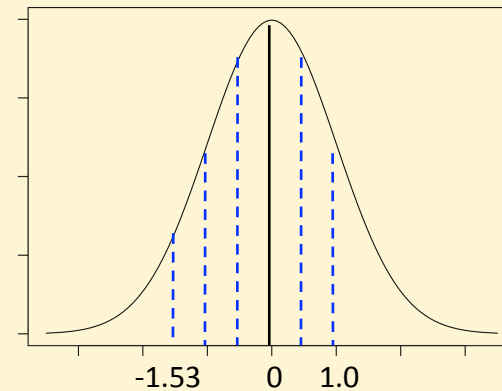
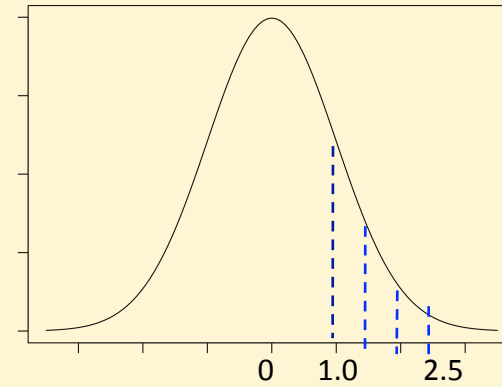
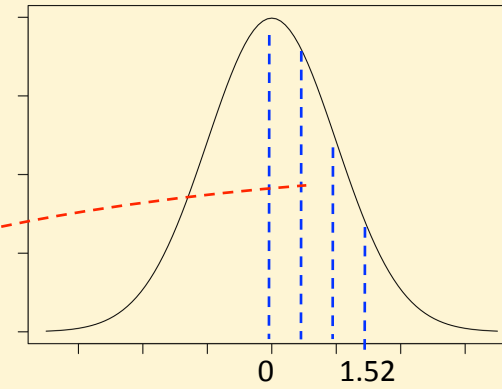
z	0	1	2	3	4	5	6	7	8	9
0.0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0.1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0754
0.2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0.3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0.4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0.5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0.6	0,2258	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0.7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0.8	0,2881	0,2910	0,2939	0,2967	0,2996	0,3023	0,3051	0,3078	0,3106	0,3133
0.9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1.0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1.1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1.2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1.3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1.4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1.5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1.6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1.7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1.8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1.9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2.0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2.1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2.2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2.3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2.4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2.5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2.6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2.7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2.8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2.9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3.0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3.1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3.2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3.3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3.4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3.5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3.6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000



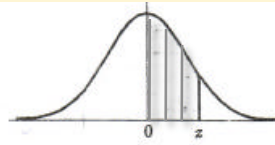
AREE
sotto la
CURVA NORMALE
STANDARDIZZATA
da 0 a z



z	0	1	2	3	4	5	6	7	8	9
0.0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0.1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0754
0.2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0.3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0.4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0.5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0.6	0,2258	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0.7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0.8	0,2881	0,2910	0,2939	0,2967	0,2996	0,3023	0,3051	0,3078	0,3106	0,3133
0.9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1.0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1.1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1.2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1.3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1.4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1.5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1.6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1.7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1.8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1.9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2.0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2.1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2.2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2.3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2.4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2.5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2.6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2.7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2.8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2.9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3.0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3.1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3.2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3.3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3.4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3.5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3.6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

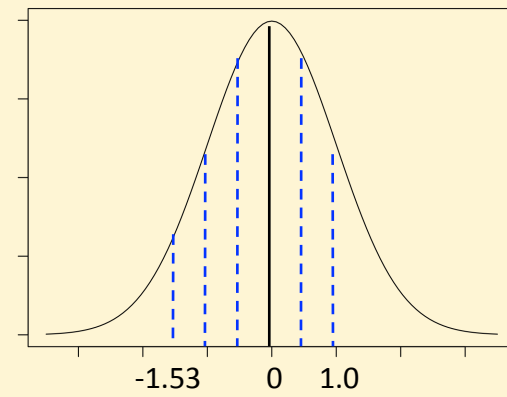
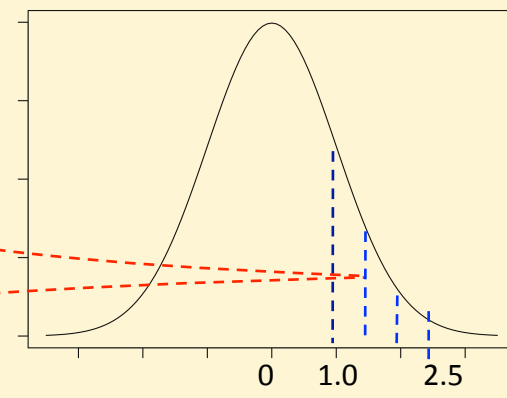
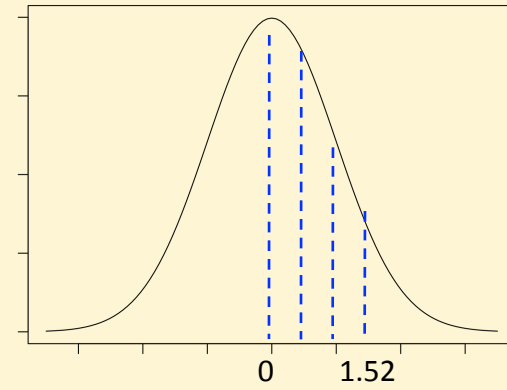


AREE
sotto la
CURVA NORMALE
STANDARDIZZATA
da 0 a z

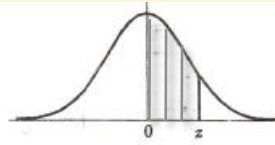


z	0	1	2	3	4	5	6	7	8	9
0.0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0.1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0754
0.2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0.3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0.4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0.5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0.6	0,2258	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0.7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0.8	0,2881	0,2910	0,2939	0,2967	0,2996	0,3023	0,3051	0,3078	0,3106	0,3133
0.9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1.0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1.1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1.2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1.3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1.4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1.5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1.6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1.7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1.8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1.9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2.0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2.1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2.2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2.3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2.4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2.5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2.6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2.7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2.8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2.9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3.0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3.1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3.2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3.3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3.4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3.5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3.6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

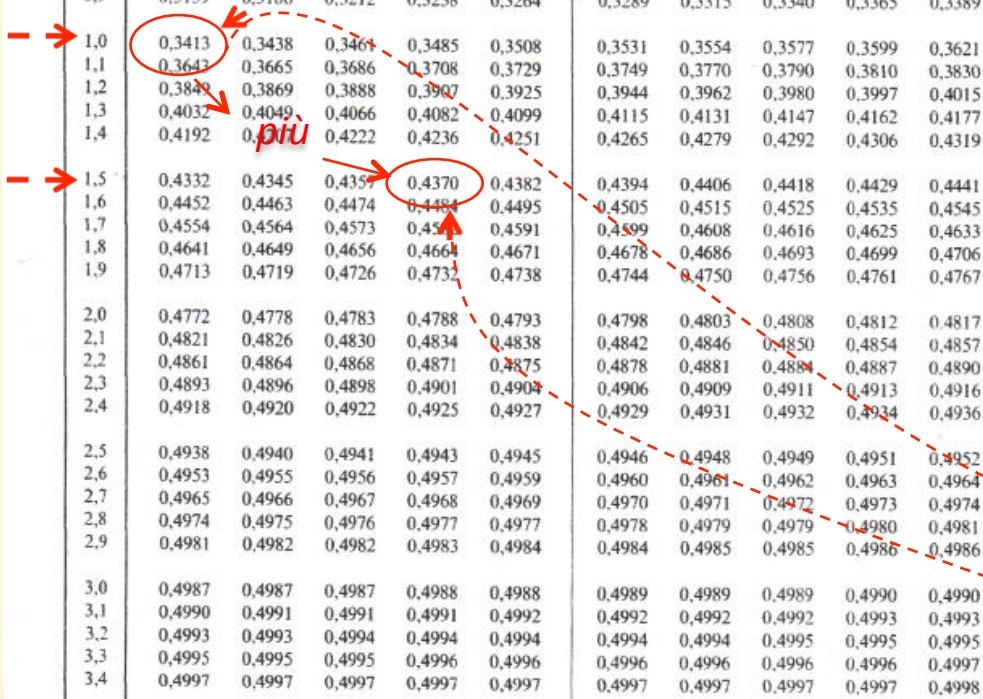
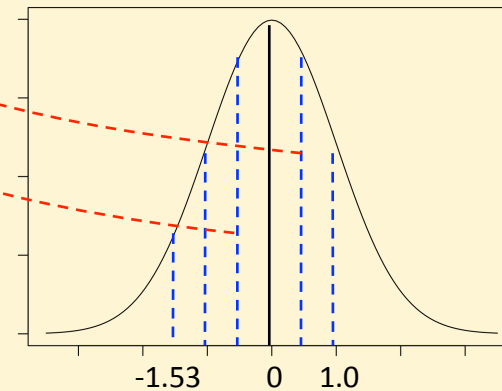
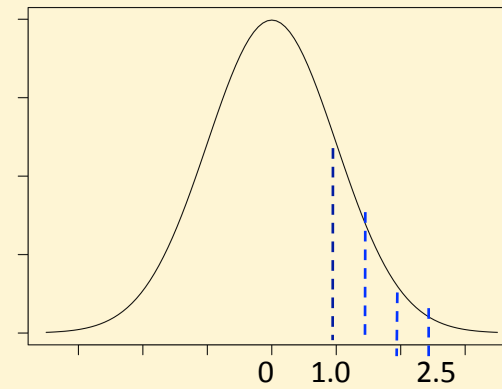
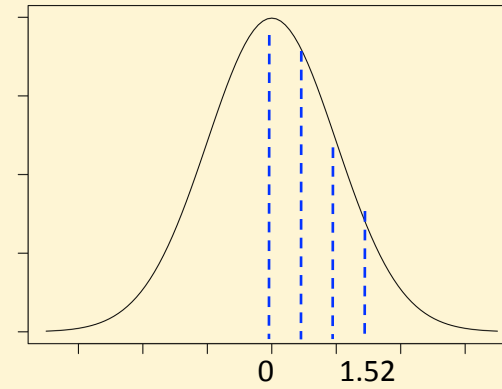
meno



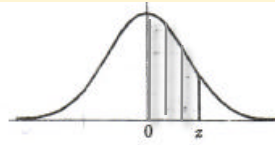
AREE
sotto la
CURVA NORMALE
STANDARDIZZATA
da 0 a z



z	0	1	2	3	4	5	6	7	8	9
0.0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0.1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0754
0.2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0.3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0.4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0.5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0.6	0,2258	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0.7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0.8	0,2881	0,2910	0,2939	0,2967	0,2996	0,3023	0,3051	0,3078	0,3106	0,3133
0.9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1.0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1.1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1.2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1.3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1.4	0,4192	0,4209	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1.5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1.6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1.7	0,4554	0,4564	0,4573	0,4581	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1.8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1.9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2.0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2.1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2.2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2.3	0,4893	0,4896	0,4898	0,4901	0,4903	0,4906	0,4909	0,4911	0,4913	0,4916
2.4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2.5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2.6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2.7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2.8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2.9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3.0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3.1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3.2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3.3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3.4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3.5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3.6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000



AREE
sotto la
CURVA NORMALE
STANDARDIZZATA
da 0 a z



z	0	1	2	3	4	5	6	7	8	9
0.0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0.1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0754
0.2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0.3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0.4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0.5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0.6	0,2258	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0.7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0.8	0,2881	0,2910	0,2939	0,2967	0,2996	0,3023	0,3051	0,3078	0,3106	0,3133
0.9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1.0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1.1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1.2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1.3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1.4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1.5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1.6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1.7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1.8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1.9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2.0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2.1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2.2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2.3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2.4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2.5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2.6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2.7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2.8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2.9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3.0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3.1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3.2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3.3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3.4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3.5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3.6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3.9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

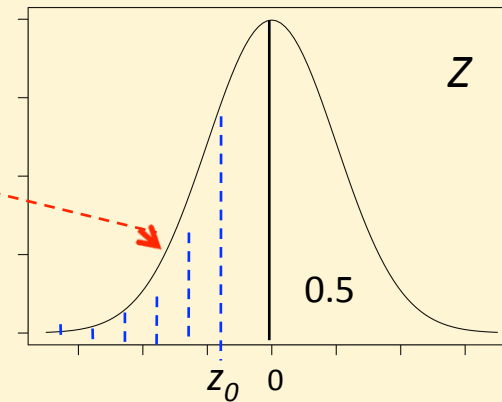
$X \sim \mathcal{N}(\mu, \sigma^2)$ con $\mu = 3.2, \sigma = 0.8$

per quale valore x_0 si ha

$$P(X < x_0) = 0.308 ?$$

$$P(X < x_0) = P\left(Z < \frac{x_0 - \mu}{\sigma} = z_0\right) = 0.308$$

$$0.5 - 0.308 = 0.192$$



$$z_0 = -0.505 = \frac{x_0 - 3.2}{0.8} \Rightarrow x_0 \cong 2.8$$

Esempi applicativi dell'uso della distribuzione *Normale*

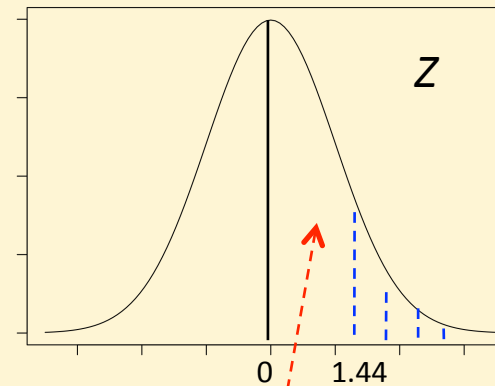
- E' noto che il punteggio – in unità convenzionali – che gli studenti riportano nella prova d'ingresso in una certa Università si distribuisce secondo una legge normale con media 420 punti e *s.d.* pari a 90 punti.
 - a. Quale probabilità ha uno studente di ottenere un punteggio > 550 ?
(in altri termini: *qual è la percentuale di studenti che ottengono > 550 ?*)
 - b. Se l'Università ammette soltanto l'80% dei candidati, quale dovrà essere il punteggio di soglia affinché uno studente sia ammesso?
 - c. Quale è il punteggio minimo atteso dal 10% degli studenti più *bravi*?
Quale il punteggio massimo del 10% degli studenti *meno bravi*?

Esempi applicativi dell'uso della distribuzione *Normale*

- a. Quale probabilità ha uno studente di ottenere un punteggio > 550 ?
(in altri termini: *qual è la percentuale di studenti che ottengono > 550 ?*)

$$X \sim \mathcal{N}(\mu=420, \sigma^2=8100)$$

$$\begin{aligned} P(X > 550) &= P\left(Z > \frac{550 - 420}{90}\right) = \\ &= P(z > 1.44) = 0.075 \Rightarrow 7.5\% \end{aligned}$$



$$0.5 - 0.4251 \approx 0.075$$

Esempi applicativi dell'uso della distribuzione *Normale*

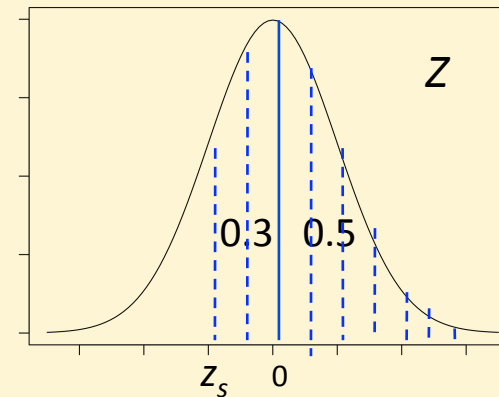
- b. Se l'Università ammette soltanto l'80% dei candidati, quale dovrà essere il punteggio di soglia affinché uno studente sia ammesso?

$$X \sim \mathcal{N}(\mu=420, \sigma^2=8100)$$

$$P(X > x_s) = P(Z > \frac{x_s - 420}{90} = z_s) = 0.80$$

$$\frac{x_s - 420}{90} = -0.845$$

$$x_s = 343.95 \approx 344$$



Esempi applicativi dell'uso della distribuzione *Normale*

- c. Quale è il punteggio minimo atteso dal 10% degli studenti più *bravi*?
Quale il punteggio massimo del 10% degli studenti *meno bravi*?

$$X \sim \mathcal{N}(\mu=420, \sigma^2=8100)$$

$$P(X > x_{mM}) = P(Z > \frac{x_{mM} - 420}{90} = z_*) = 0.10$$

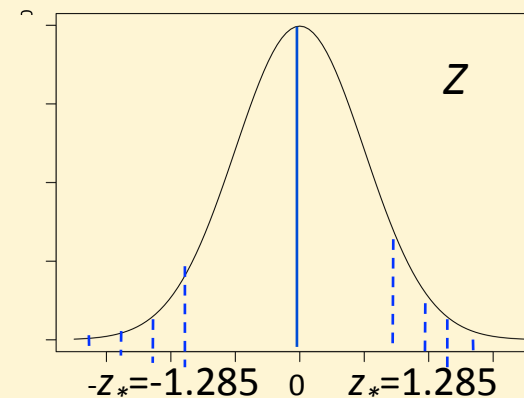
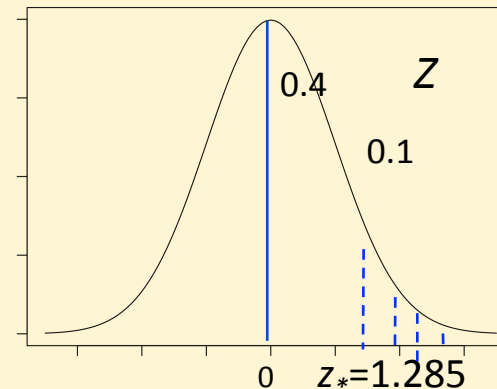
$$\frac{x_{mM} - 420}{90} = 1.285 \Rightarrow x_{mM} = 535.65$$

... per simmetria

$$P(X < x_{Mm}) = P(Z < \frac{x_{Mm} - 420}{90} = -z_*) = 0.10$$

$$\frac{x_{Mm} - 420}{90} = -1.285 \Rightarrow x_{Mm} = 304.35$$

punteggio tra 304.35 e 535.65 con prob.= 0.8



Esempi applicativi dell'uso della distribuzione *Normale*

- Una fabbrica di detersivi vende il prodotto in confezioni il cui peso lordo è una *v.a.* distribuita normalmente con media 600 gr e *s.d.* = 8 gr.
Il peso del contenitore è anch'esso distribuito normalmente con media e *s.d.* di 50 e 1.8 gr, rispettivamente.

Con quale probabilità un generico acquirente porterà a casa una quantità netta di detersivo maggiore di 560 gr?

peso lordo $L \sim \mathcal{N}(\mu_L=600, \sigma_L^2=64)$

tara $T \sim \mathcal{N}(\mu_T=50, \sigma_T^2=3.24)$

peso netto $X = (L - T)$

qual è la distribuzione di probabilità della v.a. $X = \text{peso netto}$?

Proprietà *riproduttiva* della *Normale*

- Date due (o più) v.a. tra loro *indipendenti*, distribuite normalmente

$$X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2) ; X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$$

la somma algebrica è una v.a. $X \sim \mathcal{N}(\mu_1 \pm \mu_2, \sigma_1^2 + \sigma_2^2)$

Nota: $\mu_1 \pm \mu_2, \sigma_1^2 + \sigma_2^2$

Soluzione del problema: *con quale probabilità un generico acquirente porterà a casa una quantità netta di detersivo maggiore di 560 gr?*

peso lordo $L \sim \mathcal{N}(\mu_L=600, \sigma_L^2=64)$; tara $T \sim \mathcal{N}(\mu_T=50, \sigma_T^2=3.24)$

peso netto $X = (L - T) \sim \mathcal{N}(\mu_L - \mu_T=550, \sigma_L^2 + \sigma_T^2=67.24)$

$$P(X > 560) = P\left(Z > \frac{560 - 550}{8.2}\right) = P(Z > 1.22) = 0.111 \Rightarrow \cong 11\%$$

$$(0.5 - 0.3888 = 0.1112)$$

Cenno (*intuitivo*) all'inferenza mediante la proprietà riproduttiva

- Osservazioni casuali di un campione di numerosità n di una *v.a. normale*

$X \sim \mathcal{N}(\mu, \sigma^2)$ di cui non è nota la media μ

$$X_1, X_2, \dots, X_n$$

- ogni osservazione X_i , ($i=1,2, \dots, n$), è una *v.a.* identica alla variabile *madre*

$$X_i \sim \mathcal{N}(\mu, \sigma^2) \text{ per ogni } i$$

- la media (*campionaria*) delle X_i è $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

- per la proprietà riproduttiva della normale $\bar{X} \sim \mathcal{N}(\mu, \sigma^2/n)$

Note:

- con n cresce l'informazione e diminuisce l'incertezza su μ : $\sigma^2/n \rightarrow 0$
- la distribuzione di \bar{X} tende a concentrarsi intorno a μ per $n \gg$;

Una generalizzazione: il Teorema del Limite Centrale (versione semplificata)

- X_1, X_2, \dots, X_n campione di numerosità n di una v.a. **qualsunque** (non necessariamente normale) con media μ e s.d. σ

Teorema del Limite Centrale

- per $n \gg$ la distribuzione della media campionaria $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ é approssimata da una distribuzione normale $\mathcal{N}(\mu, \sigma^2/n)$

Note:

- il risultato è *indipendente* dalla forma della distribuzione della v.a. *madre*
- la distribuzione della media \bar{X} diventa nota ($\mathcal{N}(\mu, \sigma^2/n)$) sebbene non si conosca la forma della v.a. madre, che ha originato i campioni x_i ($i= 1, 2, \dots$)

Particolari conseguenze del Teorema del Limite Centrale

Consideriamo una v.a. *Binomiale* $X \sim \mathcal{B}in(n, p)$: essa è interpretabile come somma di n v.a. *Bernoulliane* $X_i \sim \mathcal{B}er(1, p)$ tra loro indipendenti.

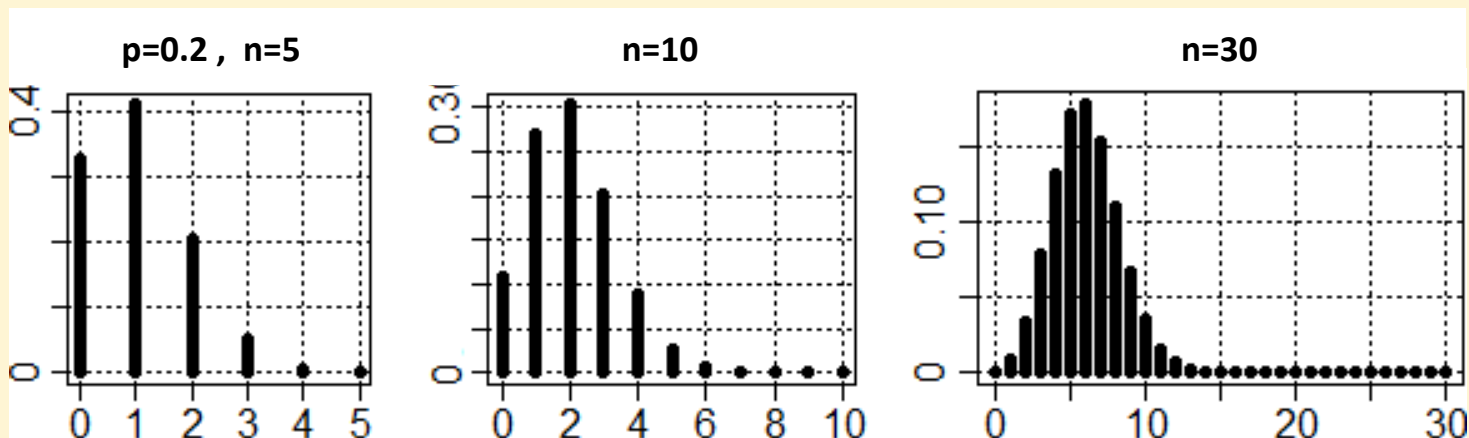
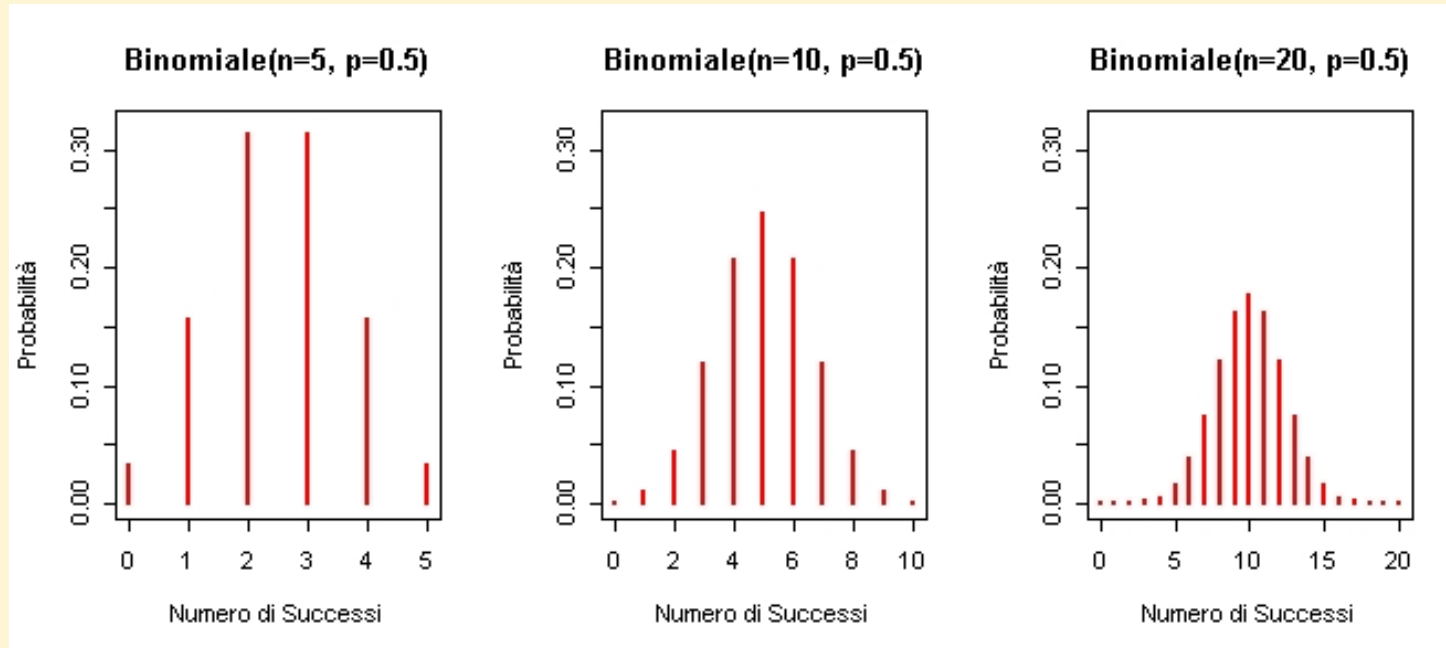
- Grazie al Teorema del Limite Centrale la *Binomiale* può essere approssimata da una normale $\mathcal{N}(np, np(1-p))$, dove p e $p(1-p)$ indicano, rispettivamente, media e varianza della *Binomiale*.

Analogamente per la distribuzione di *Poisson* $X \sim \mathcal{P}o(\lambda)$

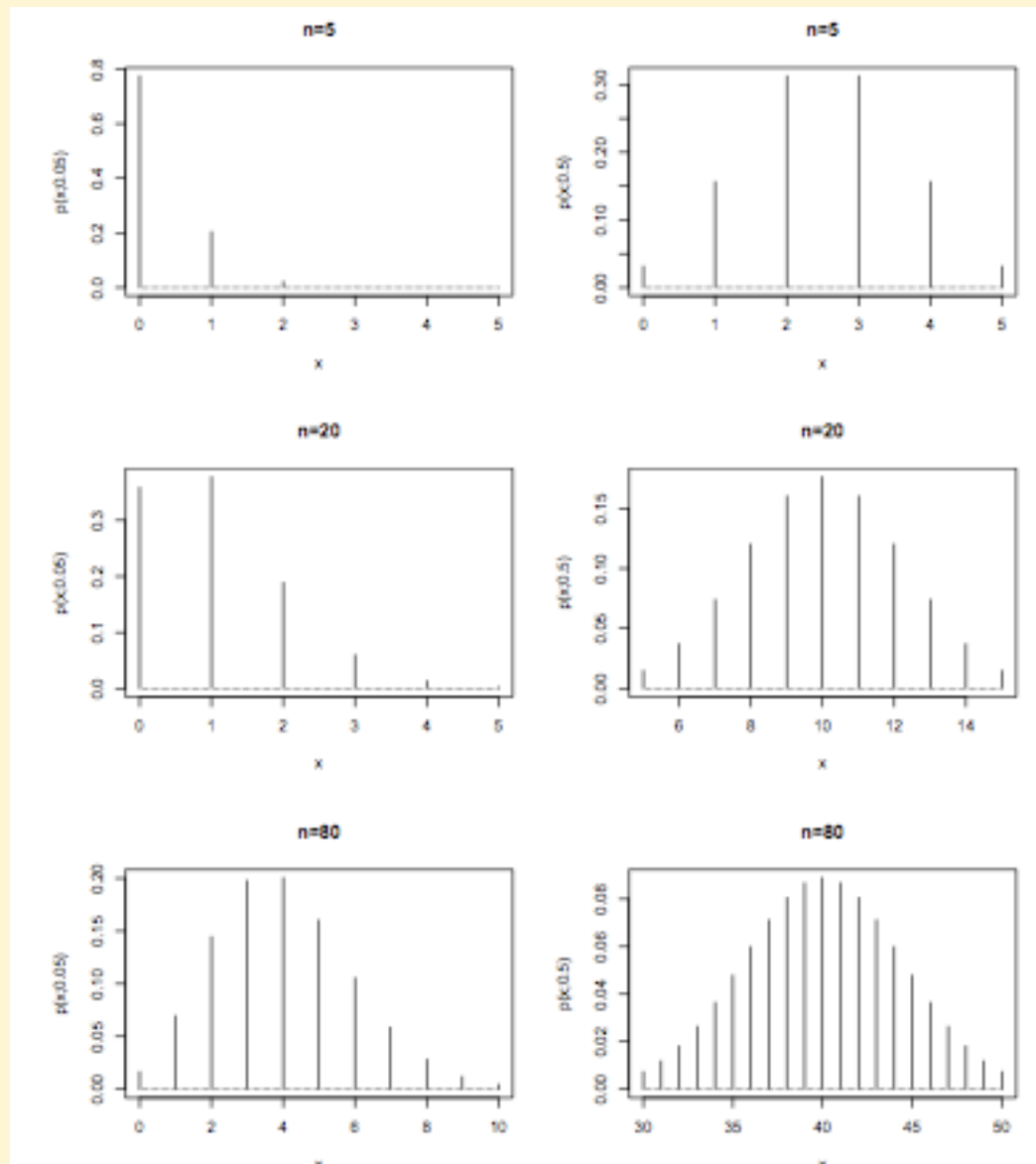
- Sapendo che la *Binomiale* tende ad approssimarsi alla *Poisson* per $n \gg$ e $p \ll$ con $\lambda = np$, ricorrendo a un semplice teorema sui limiti (teorema del confronto o “*dei carabinieri*”) si può affermare che la *Poisson* può essere approssimata da una *Normale* $\mathcal{N}(\lambda, \lambda)$, dove λ indica sia la media sia la varianza

$$\mathcal{B}in(n, p) \rightarrow \mathcal{P}o(\lambda) \rightarrow \mathcal{N}(\lambda, \lambda)$$

Convergenza della *Binomiale* alla *Normale* (puramente illustrativa)



Convergenza della *Poisson* alla *Normale* (puramente illustrativa)



Stima di parametri

perché “*parametri*”

- X v.a. che caratterizza una popolazione (oggetto dello studio)
- X è modellizzata attraverso un modello statistico-probabilistico
- Il modello dipende da uno o più parametri: genericamente θ

Es.

$$X \sim Bi(n, p) ; X \sim Ge(p) ; X \sim Po(\lambda) ; X \sim N(\mu, \sigma^2)$$

λ e μ rappresentano la media della Po e della N , rispettivamente;
 p è legata alle medie di Bi e Ge

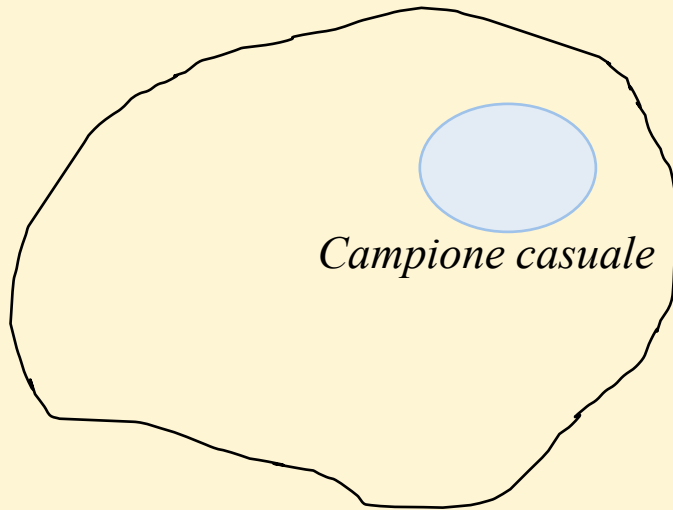
- Ottenere una stima della media di una popolazione è dunque molto importante in molti casi (ovviamente anche la stima degli altri parametri)

ci limitiamo qui ad accennare al processo inferenziale che conduce alla stima della media di una popolazione

Campionamento casuale: ogni elemento della *Popolazione* ha la stessa probabilità d'essere osservato ("campionato")

Popolazione N elementi

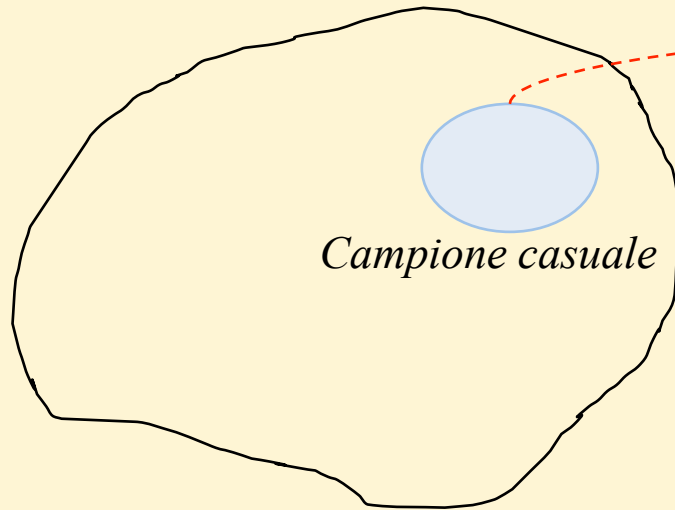
v.a. X , media $\mu = E(X)$



Campionamento casuale: ogni elemento della *Popolazione* ha la stessa probabilità d'essere osservato ("campionato")

Popolazione N elementi

v.a. X , media $\mu = E(X)$



Campione 1

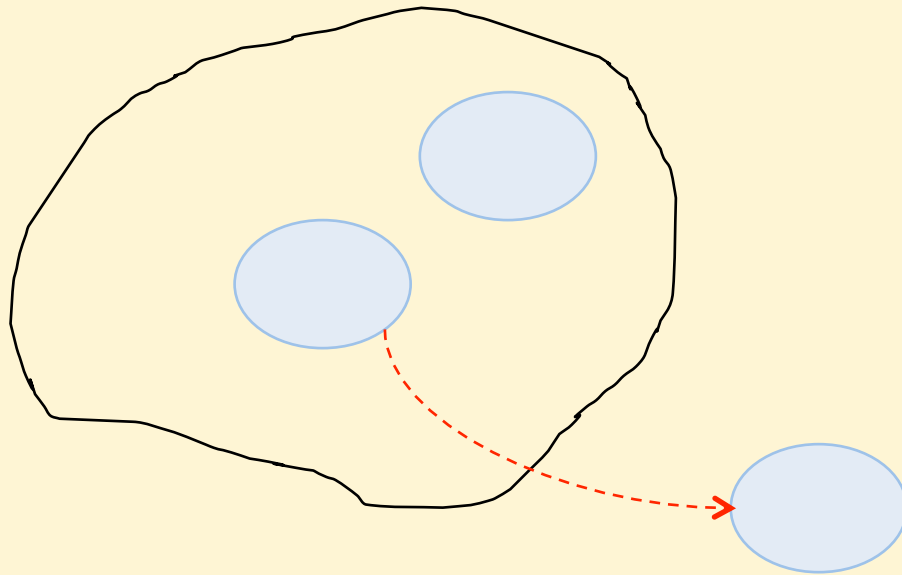
(n osservazioni x_i con $n \ll N$)

media del campione $\bar{x}_1 = \frac{1}{n} \sum_{i=1}^n x_i$

Campionamento casuale: ogni elemento della *Popolazione* ha la stessa probabilità d'essere osservato ("campionato")

Popolazione N elementi

v.a. X , media $\mu = E(X)$



Campione 1

(n osservazioni x_i con $n \ll N$)

media del campione $\bar{x}_1 = \frac{1}{n} \sum_{i=1}^n x_i$

Campione 2, $\bar{x}_2 = \frac{1}{n} \sum_{i=1}^n x_i$

\vdots

Campione k , $\bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_i$

La media campionaria è dunque una *v.a.* \bar{X} con $E(\bar{X}) = \mu$ (*correttezza*)

Osservazioni sulla stima della media

- perché proprio la media campionaria?

Problema: stima dell'altezza media degli studenti di questa classe

a) mediana campionaria ; b) media campionaria ; c) semi-range

la media gode di proprietà *ottimali* (*correttezza* e *consistenza* tra le altre)

- perché proprio un valore *puntuale*?

Problema 1: stima dell'altezza media degli studenti di questa classe

Problema 2: stima della mia altezza

Problema 3: stima del numero di passeggeri su un bus

Problema 4: stima del costo di una pizzata

... ..

Problema k : stima del tempo di attesa per una visita medica

Stima per intervalli (intervallo di confidenza)

Scopo

costruire dai dati campionari un intervallo $[T_1, T_2]$ entro cui cada il parametro incognito con prefissata probabilità $1-\alpha$

$$P(T_1 \leq \theta \leq T_2) = 1 - \alpha$$

dove $T_1(x_1, \dots, x_n)$ e $T_2(x_1, \dots, x_n)$ sono gli estremi dell'intervallo stimati dai dati e $1-\alpha$ è il *livello di confidenza*

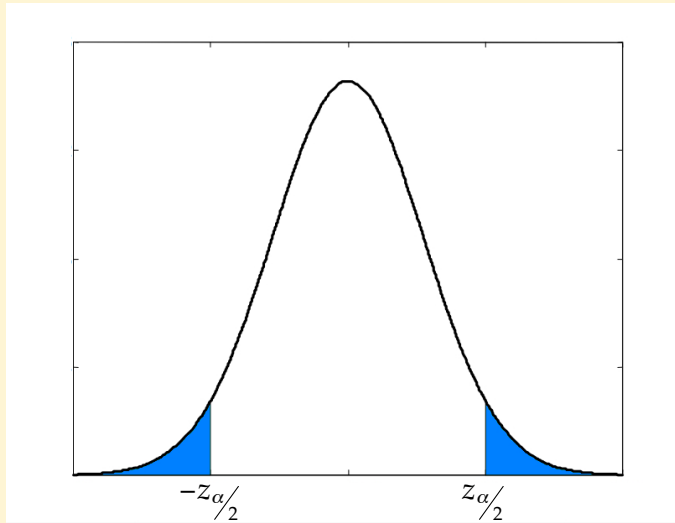
Nota: θ è un numero reale (incognito); T_1, T_2 variabili aleatorie

Intervallo di confidenza per la media di una *Normale* $N(\mu, \sigma^2)$ con σ^2 nota

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \quad \text{standardizzando} \quad \rightarrow \quad Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad ; \quad Z \sim N(0,1)$$

$$P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha$$

Intervallo di confidenza per la media di una *Normale* (σ^2 nota)



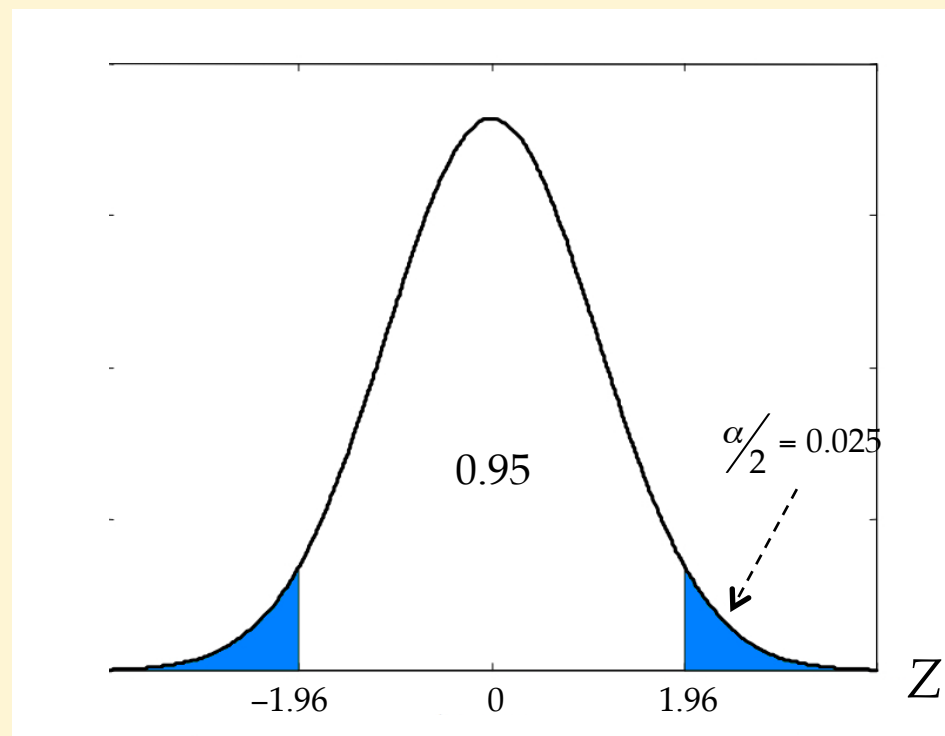
$$P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha$$

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

$$P(-z_{\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \leq z_{\alpha/2}) = 1 - \alpha$$

$$P(\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

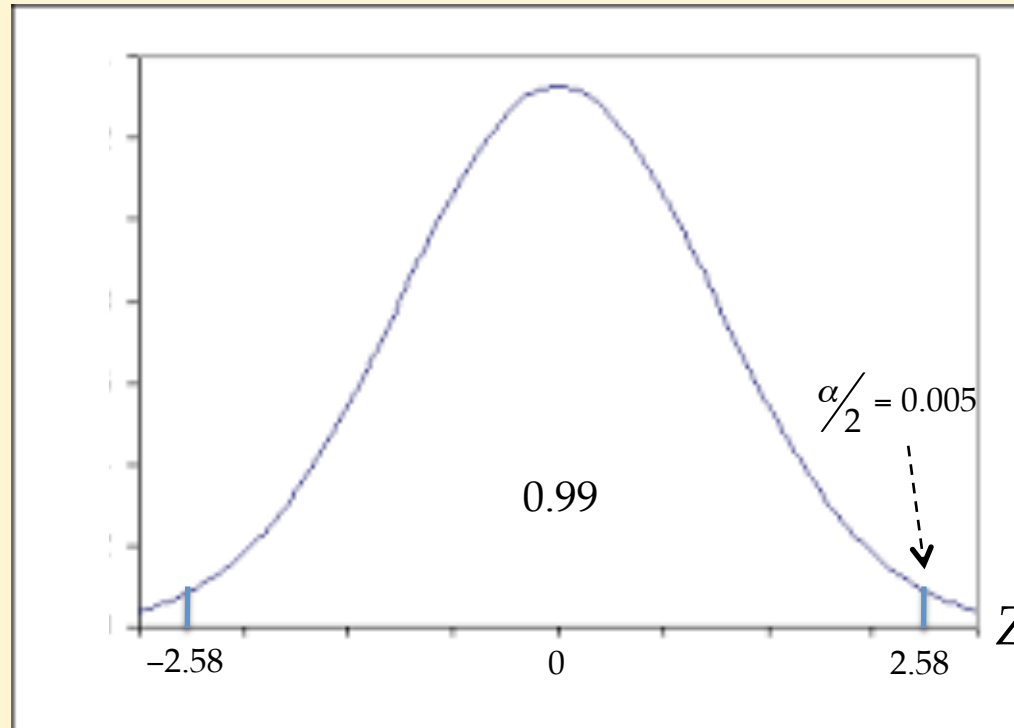
Esempio per $1-\alpha=0.95$



$$\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}$$

con livello di confidenza del 95%

Esempio per $1-\alpha=0.99$



$$\bar{X} - 2.58 \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + 2.58 \frac{\sigma}{\sqrt{n}}$$

con livello di confidenza del 99%

Ampiezza dell'intervallo

L'ampiezza $A = (T_2 - T_1)$ dell'intervallo $[T_1, T_2]$ dipende, oltre che da α

$$P(T_1 \leq \theta \leq T_2) = 1 - \alpha$$

anche dall'espressione degli estremi e dai dati campionari.

Nel caso in studio $P(\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$

$$A = 2z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

quindi, a parità di α , l'intervallo aumenta con σ e diminuisce al crescere di n .

Osservazioni

- σ caratterizza il modello $N(\mu, \sigma^2)$ mentre n dipende dal campione:
→ un maggiore contenuto informativo produce una stima più "precisa"
- gli estremi dell'intervallo dipendono dalle osservazioni campionarie (\bar{X})
- quale n per avere una determinata A ? → $n = \left(2z_{\alpha/2} \frac{\sigma}{A}\right)^2$

ampiezza dell'intervallo

Esempi

$$\sigma^2 = 9 ; 1 - \alpha = 0.95 \quad (z_{0.025} = 1.96); \quad \text{se } \bar{x} = 4.7$$

$$\left[\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}} , \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}} \right]$$

per $n = 10 \rightarrow [2.84 , 6.56]$
per $n = 100 \rightarrow [4.11 , 5.29]$
per $n = 1000 \rightarrow [4.51 , 4.88]$

$$\sigma^2 = 9 ; 1 - \alpha = 0.99 \quad (z_{0.005} = 2.58); \quad \text{se } \bar{x} = 4.7$$

$$\left[\bar{x} - 2.58 \frac{\sigma}{\sqrt{n}} , \bar{x} + 2.58 \frac{\sigma}{\sqrt{n}} \right]$$

per $n = 10 \rightarrow [2.25 , 7.14]$
per $n = 100 \rightarrow [3.93 , 5.47]$
per $n = 1000 \rightarrow [4.46 , 4.94]$

Qualche esercizio

1. Il dirigente di un ufficio pubblico per consentire ai dipendenti di entrare in ufficio con qualche minuto di ritardo (o di anticipo!) rispetto all'orario stabilito e proporre un orario flessibile, deve stimare il ritardo medio dei dipendenti. A questo scopo rileva i ritardi di un campione di essi scelti a caso.

Volendo che la stima abbia uno scarto di $\pm 2 \text{ min.}$, si stabilisca la numerosità campionaria necessaria a garantire un livello di confidenza del 95% (si assuma che i ritardi siano distribuiti normalmente con s.d. = 9 min.)

$$P\left(\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right) = 0.95$$

$$A = \frac{2z_{\alpha/2}\sigma}{\sqrt{n}} = 4 \quad ; \quad z_{\alpha/2} = 1.96 \quad ; \quad \sigma = 9$$

$$n \geq \left(\frac{2 \cdot 1.96 \cdot 9}{4}\right)^2 = 77.8 \quad \Rightarrow \quad n = 78$$

Se ci si "accontentasse" di uno scarto di $\pm 5 \text{ min.}$, ($A=10$), $n \geq 12.4 \Rightarrow n = 13$

Qualche esercizio

2. Una fabbrica metalmeccanica costruisce cuscinetti a sfera. La produzione avviene su due linee di lavorazione distinte. Da ciascuna linea è rilevato indipendentemente un campione casuale indipendente di numerosità $n_1=12$ e $n_2=18$, con medie dei diametri $\bar{x}_1 = 8.75$, $\bar{x}_2 = 8.63$, rispettivamente, (in unità convenzionali). Sapendo che la varianza comune dei diametri prodotti nelle due linee è $\sigma^2 = 3.2$, si costruisca l'intervallo di confidenza per la differenza delle medie dei diametri prodotti dalle due linee.

$$\bar{X}_1 \sim N\left(\mu_1, \frac{\sigma^2}{n_1}\right) ; \bar{X}_2 \sim N\left(\mu_2, \frac{\sigma^2}{n_2}\right) \rightarrow (\bar{X}_1 - \bar{X}_2) \sim N\left(\mu_1 - \mu_2, \sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right)$$

$$(\bar{X}_1 - \bar{X}_2) - z_{\alpha/2} \sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leq \mu_1 - \mu_2 \leq (\bar{X}_1 - \bar{X}_2) + z_{\alpha/2} \sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

con $1 - \alpha = 0.95$ ($z_{0.025} = 1.96$) e $\sigma = \sqrt{3.2}$

$$0.12 - 1.96 \cdot 1.79 \sqrt{\frac{1}{12} + \frac{1}{18}} \leq \mu_1 - \mu_2 \leq 0.12 + 1.96 \cdot 1.79 \sqrt{\frac{1}{12} + \frac{1}{18}}$$

$$-1.19 \leq \mu_1 - \mu_2 \leq 1.42$$

Il dramma del professore
(per spiegare la significatività statistica)

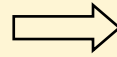
Dramma aperto in VI atti
e un epilogo

Protagonisti: *lo studente; il professore*

Atto I (il contesto)	<ul style="list-style-type: none"> Lo studente deve sostenere un esame Può essere preparato o non-preparato 	
Atto II (l'azione)	Il professore deve compiere l'azione: { promuovere bocciare	
Atto III (la trama si sviluppa)	<ul style="list-style-type: none"> Lo studente e il professore si incontrano Il prof non sa se lo studente è preparato o no (forse neppure lo studente lo sa!) Non potendo esplorare "tutta" la preparazione dello studente, il prof ha la possibilità di fare alcune domande (ad es. 5) 	
Atto IV (il dilemma)	<p>Il professore pensa: quante risposte esatte dovrà darmi lo studente per convincermi d'essere preparato?</p> <p>(Nota: professore "buono"=2 risposte su 5; "severo" =3 su 5; "cattivo"=4 su 5; "terribile"=5 su 5).</p>	
Atto V (il fatto)	Lo studente risponde a 3 domande → il prof lo promuove	Lo studente risponde a meno di 3 domande → il prof lo boccia
Atto VI (il dramma del prof)	Ha promosso un non-preparato	Ha bocciato un preparato
Epilogo	Il professore si rivolge al suo statistico di fiducia !	

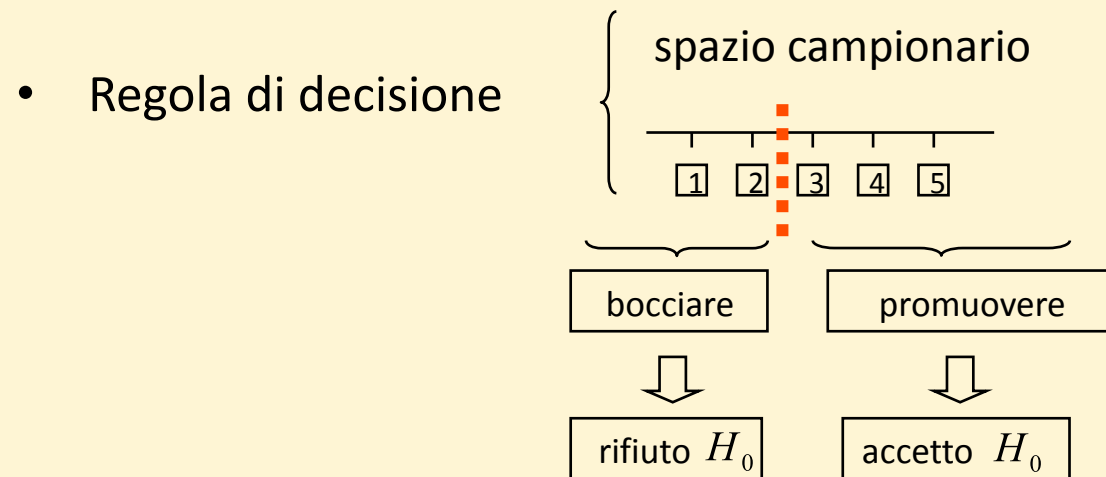
(per spiegare la significatività statistica)

- Studente { preparato
non-preparato



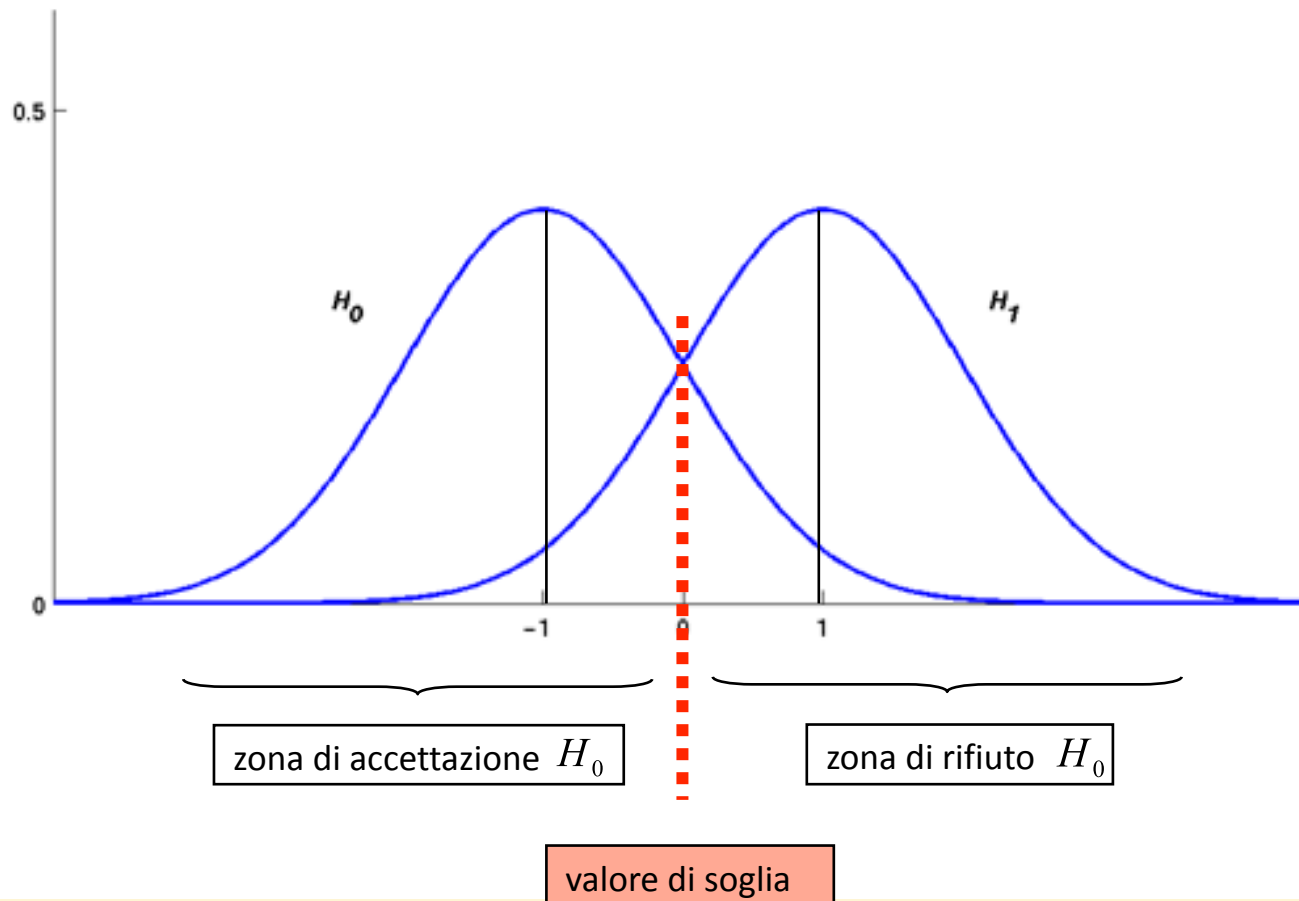
Stati di natura $\Omega : \{H_0, H_1\}$
 $H_0 : \text{preparato}, H_1 : \text{non-preparato}$

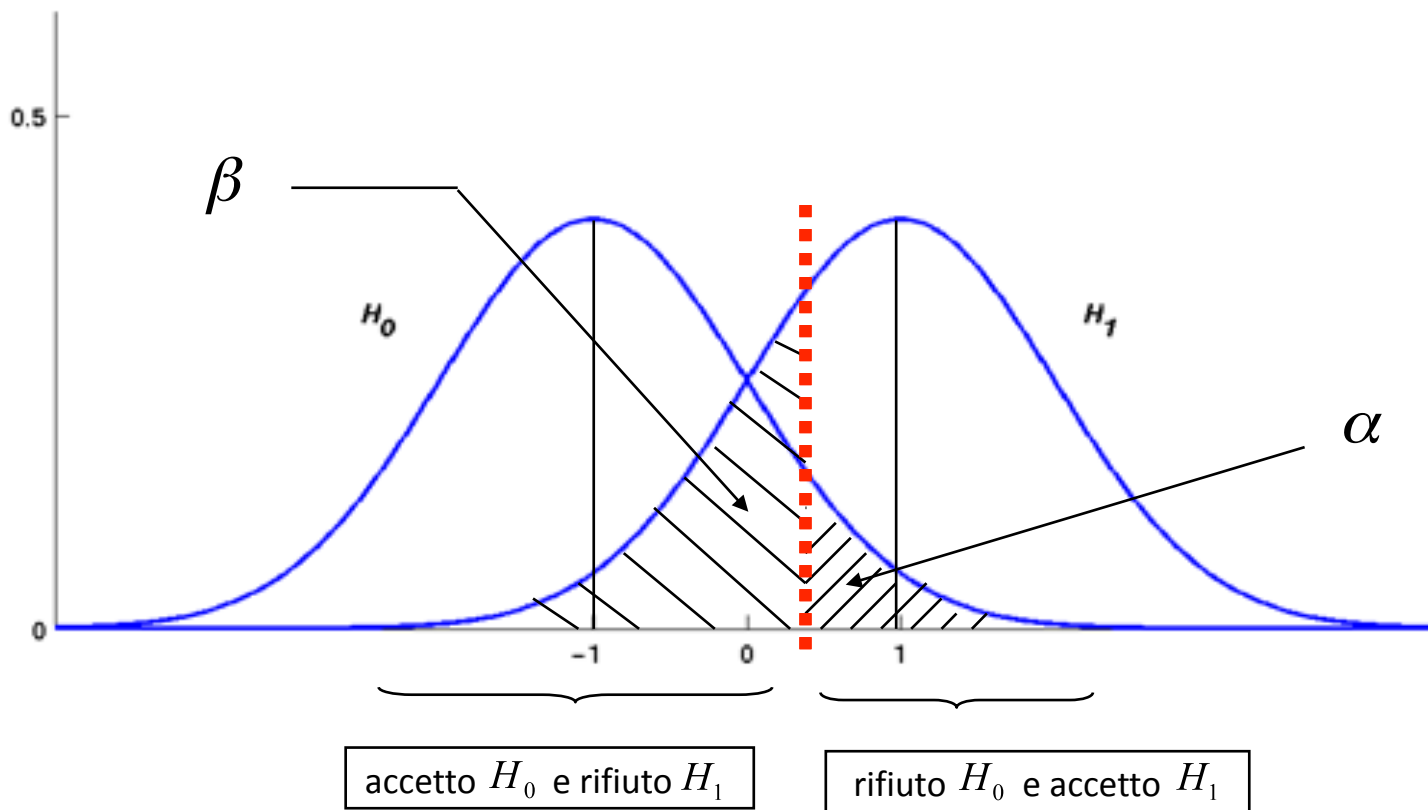
- Azioni { promuovere
bocciare



... in sintesi

	H_0	H_1
	preparato	non-preparato
promuovo (b)	$P(pr H_0)$	$\beta = P(pr H_1)$
boccio (b)	$\alpha = P(b H_0)$	$P(b H_1)$





$$\alpha = P(\text{rifiutare } H_0 \text{ quando è vera}) = P(\text{rifiuto} | H_0)$$

Nella teoria della verifica (test) d'ipotesi di *Neyman-Pearson-Wald*

- α livello di significatività del test: livello d'errore con cui siamo "disposti" a rifiutare l'ipotesi H_0 .
- α è usualmente molto piccolo: valori standard 0.05; 0.01; anche 0.001.

Jerzy Neyman, (1894 – 1981)

Sir Ronald Aylmer Fisher, (1890 – 1962)

Egon Sharpe Pearson, (1895 – 1980)

Abraham Wald, (1902 – 1950)

racugno@unica.it