Esercizio 1

In una località balneare sono stati raccolti 50 campioni di acqua per valutare la presenza di un certo inquinante . La tabella riporta i livelli di inquinante rilevati:

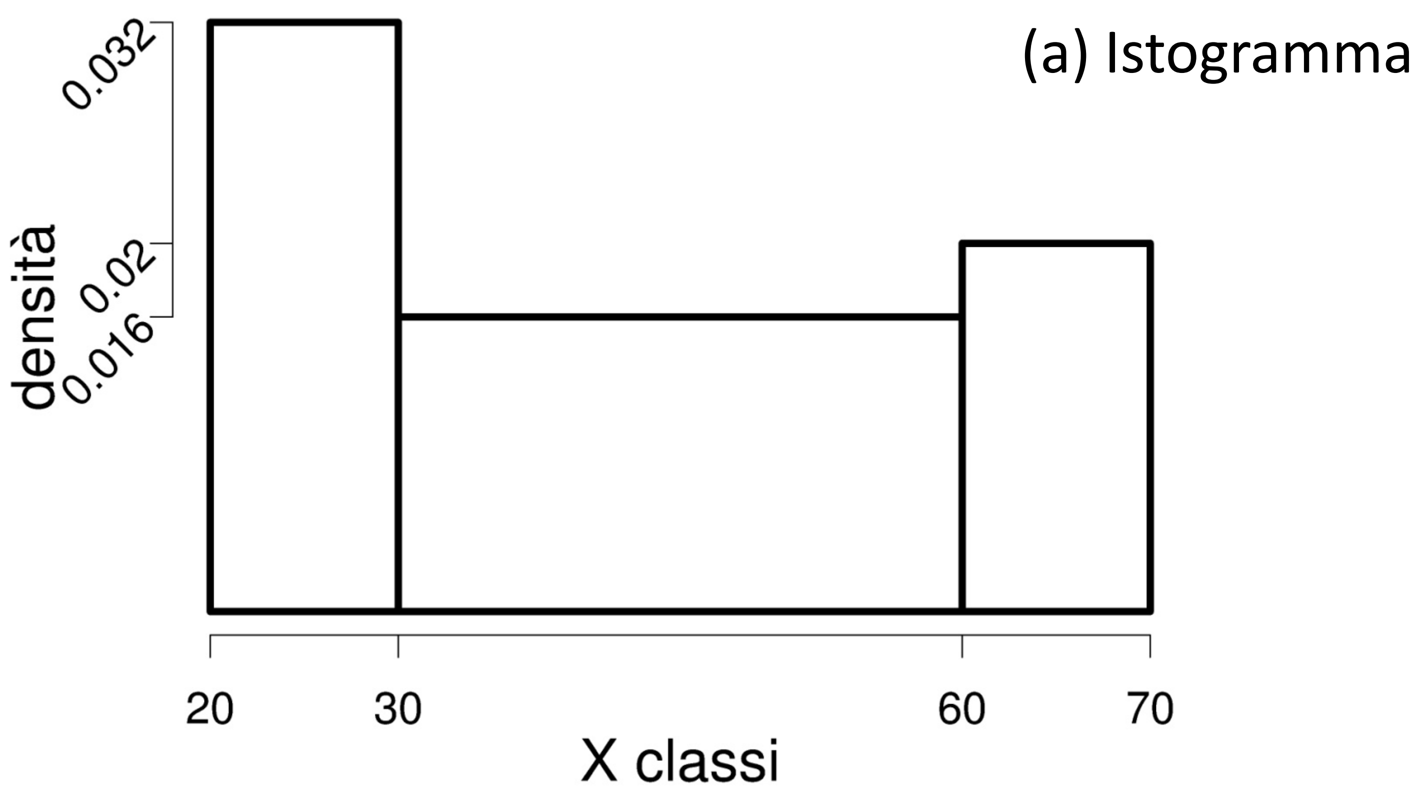
classi x_i	frequenze n_i
20 --| 30	16
30 --| 60	24
60 --| 70	10

(a) Fornire una rappresentazione grafica opportuna

(b) Disegnare il box-plot

(c) Calcolare moda e coefficiente di variazione

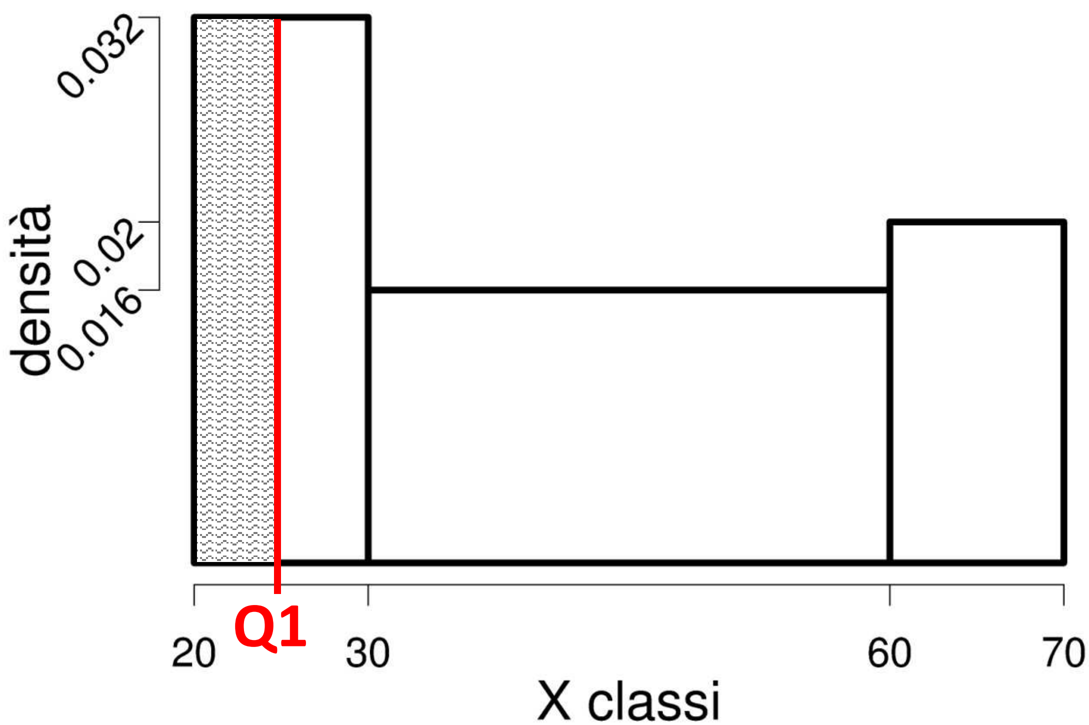
x_i classi	n_i freq. assolute	f_i freq. relative	a_i ampiezza classi	$d_i = f_i / a_i$ densità
20 |-- 30	16	16/50 = 0.32	30-20 = 10	0.32/10 = 0.032
30 |-- 60	24	24/50 = 0.48	60-30 = 30	0.48/30 = 0.016
60 |-- 70	10	10/50 = 0.20	70-60 = 10	0.20/10 = 0.020
	50	**1**		



x_i	n_i	f_i	a_i	$d_i = f_i / a_i$
20 |-- 30	16	16/50 = 0.32	30-20 = 10	0.32/10 = 0.032
30 |-- 60	24	24/50 = 0.48	60-30 = 30	0.48/30 = 0.016
60 |-- 70	10	10/50 = 0.20	70-60 = 10	0.20/10 = 0.020
	50	**1**		

Primo quartile $\mathbf{F(X < Q1) = 0.25}$

La prima colonna dell'istogramma ha area **0.32** e Q1 stacca un'area di **0.25**. Quindi Q1 sta nella prima classe.



$$(Q1 - 20) \times 0.032 = 0.25$$

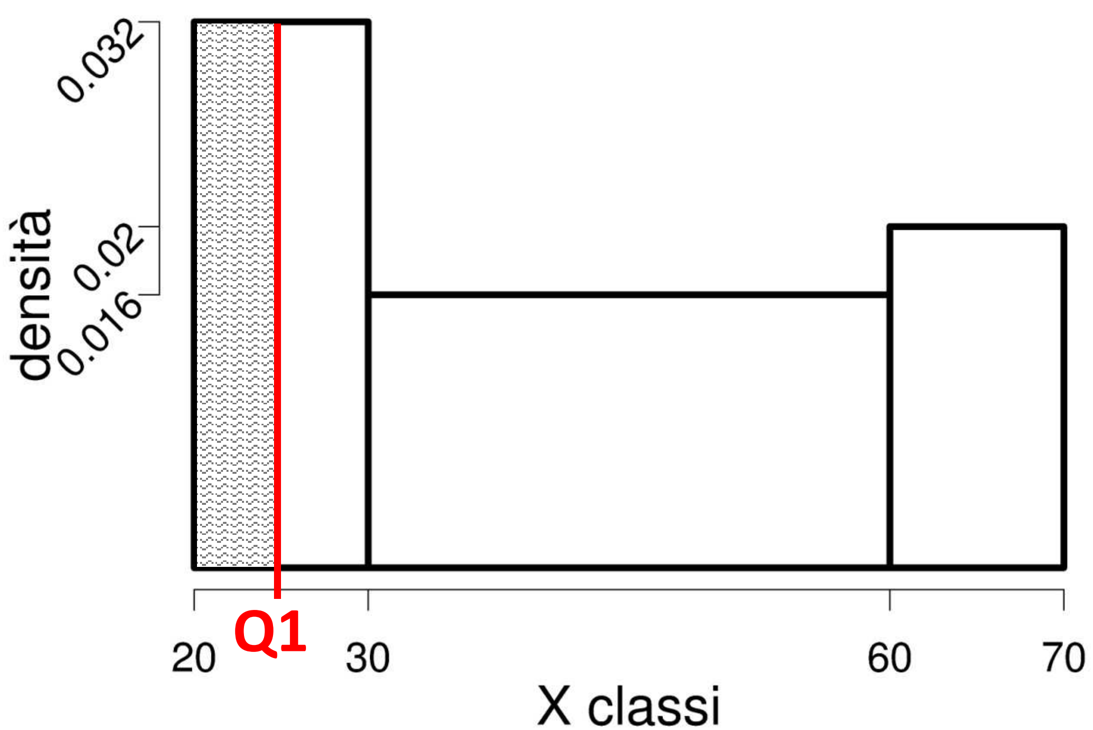
$$(Q1 - 20) = \frac{0.25}{0.032}$$

$$Q1 = \frac{0.25}{0.032} + 20 = 27.8$$

Oppure :

Primo quartile

x_i	n_i	f_i	a_i	N_i freq. assoluta cumulata	$l_i = n_i / a_i$
20 |-- 30	16	0.32	10	16	16/10 = **1.6**
30 |-- 60	24	0.48	30	40	24/10 = 0.8
60 |-- 70	10	0.20	10	50	10/10 = 1.0
	50	**1**			



$$\frac{n}{4} = \frac{50}{4} = \mathbf{12.5}$$

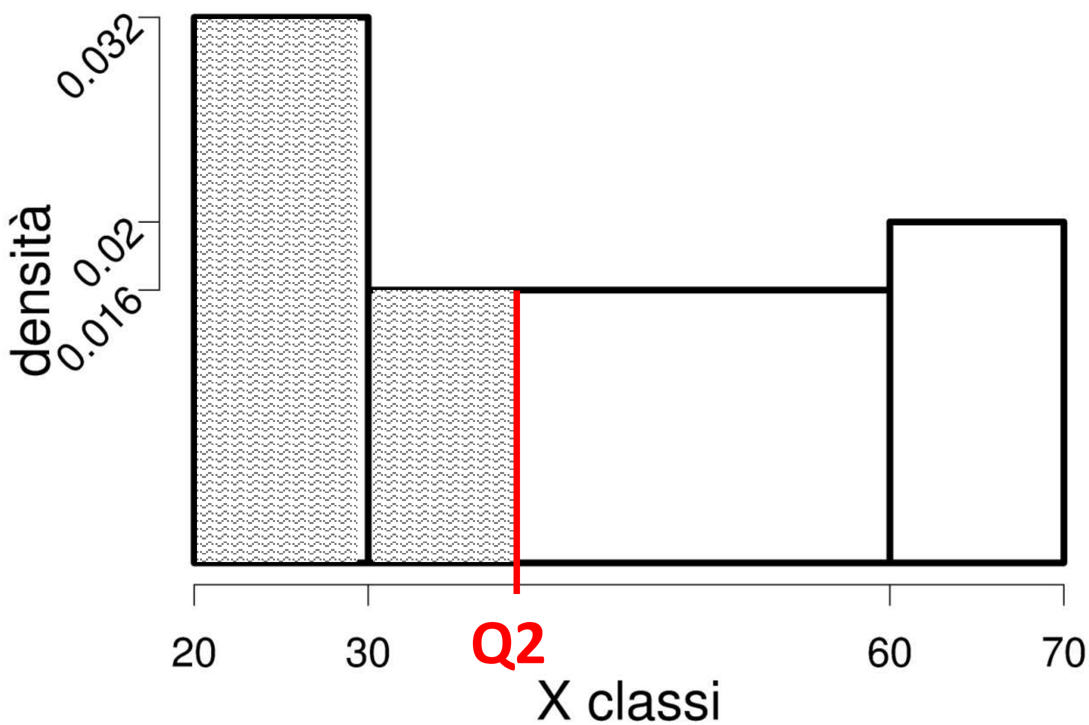
$$Q1 = x_i^{(inf)} + \frac{\frac{n}{4} - N_{i-1}}{l_i} =$$

$$= \mathbf{20} + \frac{\mathbf{12.5 - 0}}{\mathbf{1.6}} = 27.8$$

x_i	n_i	f_i	a_i	$d_i = f_i / a_i$
20 |-- 30	16	16/50 = 0.32	30-20 = 10	0.32/10 = 0.032
30 |-- 60	24	24/50 = 0.48	60-30 = 30	0.48/30 = 0.016
60 |-- 70	10	10/50 = 0.20	70-60 = 10	0.20/10 = 0.020
	50	**1**		

Mediana F$(\mathbf{X < Q2}) = \mathbf{0.50}$

0.50 sta tra 0.32 e 0.32+0.48=0.80, quindi Q2 sta nella **seconda classe**

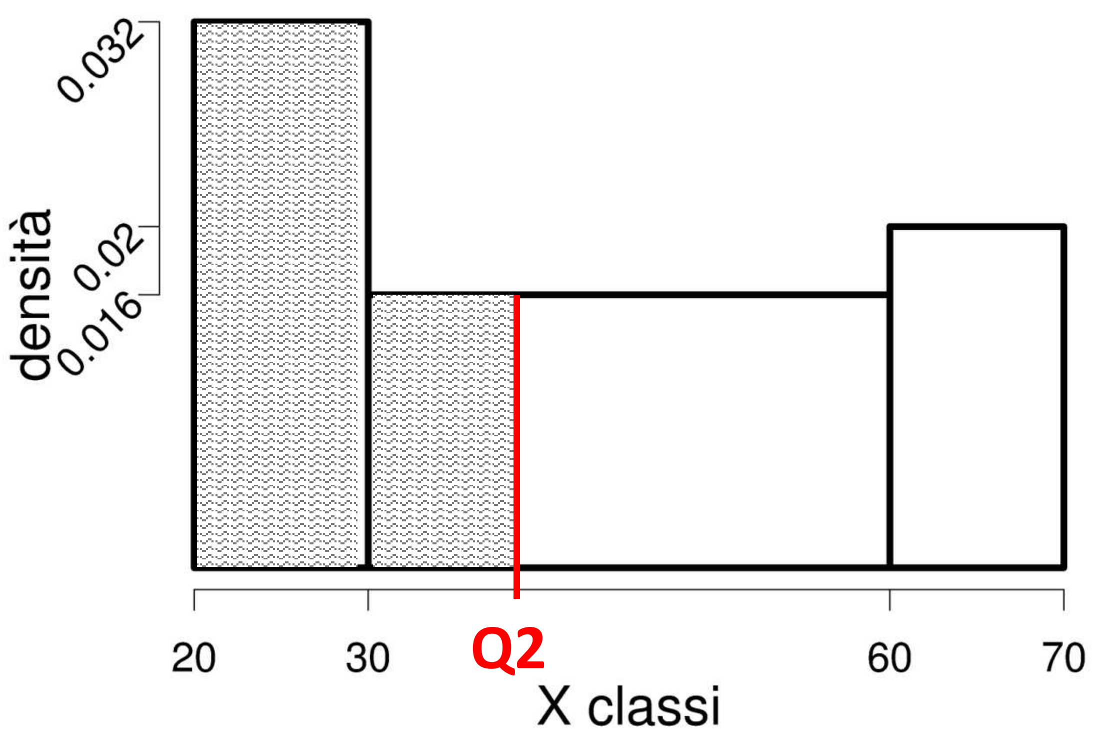


$$0.32 + (Q2 - 30) \times 0.016 = 0.50$$

$$Q2 = \frac{0.50 - 0.32}{0.016} + 30 = 41.25$$

Oppure : **Mediana**

x_i	n_i	f_i	a_i	N_i **freq. assoluta cumulata**	l_i **densità di freq. assol.**
20 |-- 30	16	0.32	10	**16**	16/10 = 1.6
30 |-- 60	24	0.48	30	40	24/10 = **0.8**
60 |-- 70	10	0.20	10	50	10/10 = 1.0
	50	**1**			



$$\frac{n}{2} = \frac{50}{2} = \mathbf{25}$$

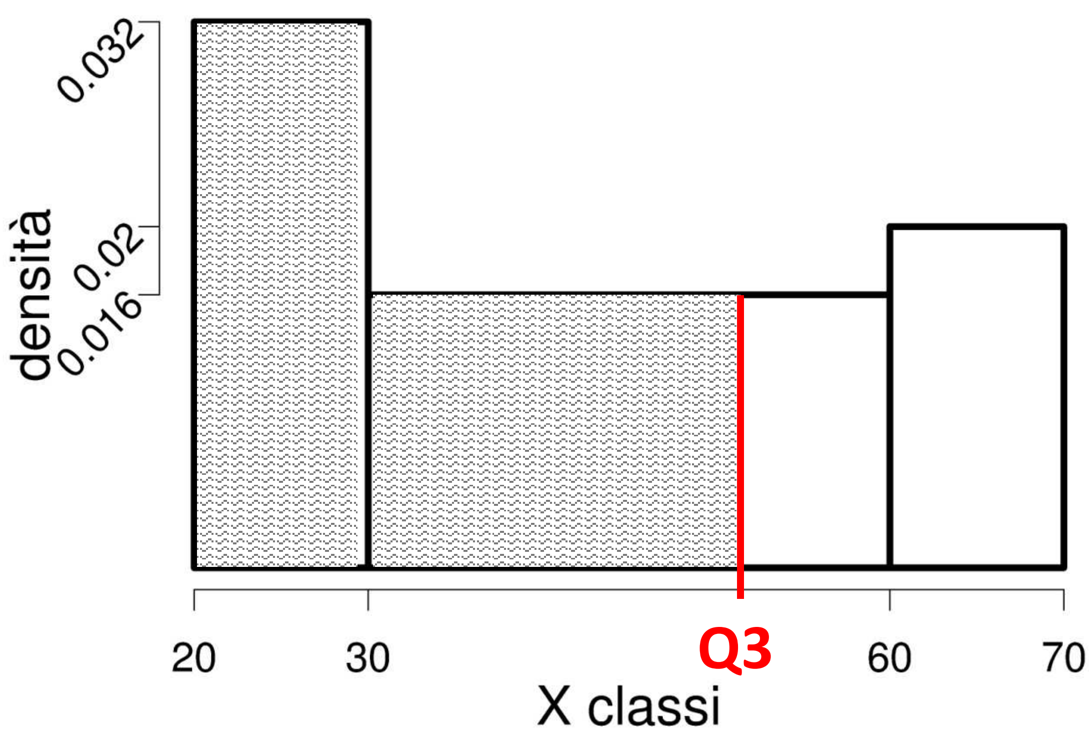
$$Q2 = x_i^{(inf)} + \frac{\frac{n}{2} - N_{i-1}}{l_i} =$$

$$= \mathbf{30} + \frac{\mathbf{25 - 16}}{\mathbf{0.8}} = 41.25$$

x_i	n_i	f_i	a_i	$d_i = f_i / a_i$
20 |-- 30	16	16/50 = 0.32	30-20 = 10	0.32/10 = 0.032
30 |-- 60	24	24/50 = 0.48	60-30 = 30	0.48/30 = 0.016
60 |-- 70	10	10/50 = 0.20	70-60 = 10	0.20/10 = 0.020
	50	**1**		

Terzo quartile $\mathbf{F(X < Q3) = 0.75}$

0.75 sta tra 0.32 e 0.32+0.48=0.80, quindi Q3 sta nella **seconda classe**

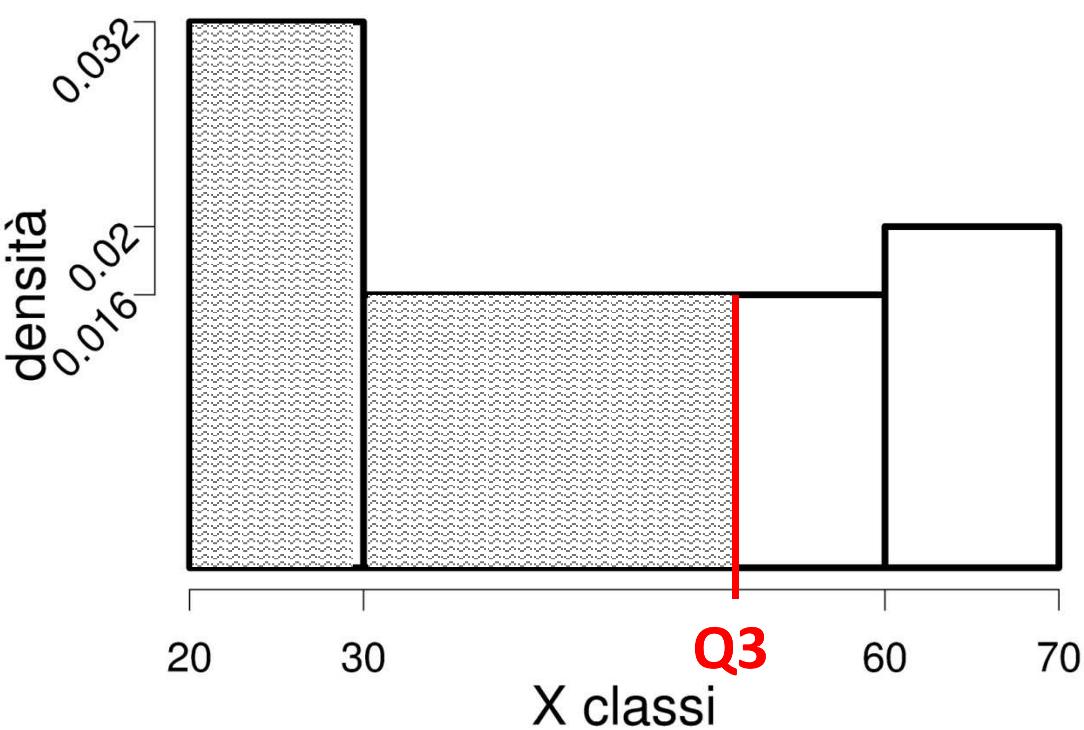


$$0.32 + (Q3 - 30) \times 0.016 = 0.75$$

$$Q3 = \frac{0.75 - 0.32}{0.016} + 30 = 56.9$$

Oppure : **Terzo quartile**

x_i	n_i	f_i	a_i	N_i freq. assoluta cumulata	l_i densità di freq. assol.
20 |-- 30	16	0.32	10	**16**	16/10 = 1.6
30 |-- 60	24	0.48	30	40	24/10 = **0.8**
60 |-- 70	10	0.20	10	50	10/10 = 1.0
	50	**1**			



$$\frac{3}{4}n = \frac{3 \times 50}{4} = \mathbf{37.5}$$

$$Q3 = x_i^{(inf)} + \frac{\frac{3}{4}n - N_{i-1}}{l_i} =$$

$$= \mathbf{30} + \frac{\mathbf{37.5 - 16}}{\mathbf{0.8}} = 56.9$$

Baffi

Abbiamo calcolato:

Q1 = 27.8 mediana = 41.25 Q3 = 56.9

Serve il range interquartile:

1.5 x IQR = 1.5 x (Q3-Q1) = 1.5 x (56.9 – 27.8) = 43.65

Baffo superiore = Q3 + 1.5 x IQR = 56.9 + 43.65 = 97.55 $\longrightarrow$ = 70

Baffo inferiore = Q1 - 1.5 x IQR = 27.8 - 43.65 = -15.85 $\longrightarrow$ = **20**

I dati osservati variano da 20 a 70.

I valori calcolati dei baffi sono fuori dall'intervallo 20-70 dei dati,

quindi i baffi sono posti uguali a 20 e 70.

In tal caso non ci sono outliers.

x_i classi	n_i	f_i Frequenze relative	a_i	d_i	m_i punto medio classi
20 |-- 30	16	0.32	10	0.032	(20+30)/2 = 25
30 |-- 60	24	0.48	30	0.016	(30+60)/2 = 45
60 |-- 70	10	0.20	10	0.020	(60+70)/2 = 65
	50	**1**			

Media

$$\bar{x} = media = \sum_{\mathrm{i}} m_i \, \mathrm{f_i} =$$

$$= 25 \times 0.32 + 45 \times 0.48 + 65 \times 0.20 = 42.6$$

oppure

$$\bar{x} = media = \frac{\sum_i m_i n_i}{n}$$

$$= (25 \times 16 + 45 \times 24 + 65 \times 10)/50 = 42.6$$

Baffo superiore = 70

Q3 = 56.9

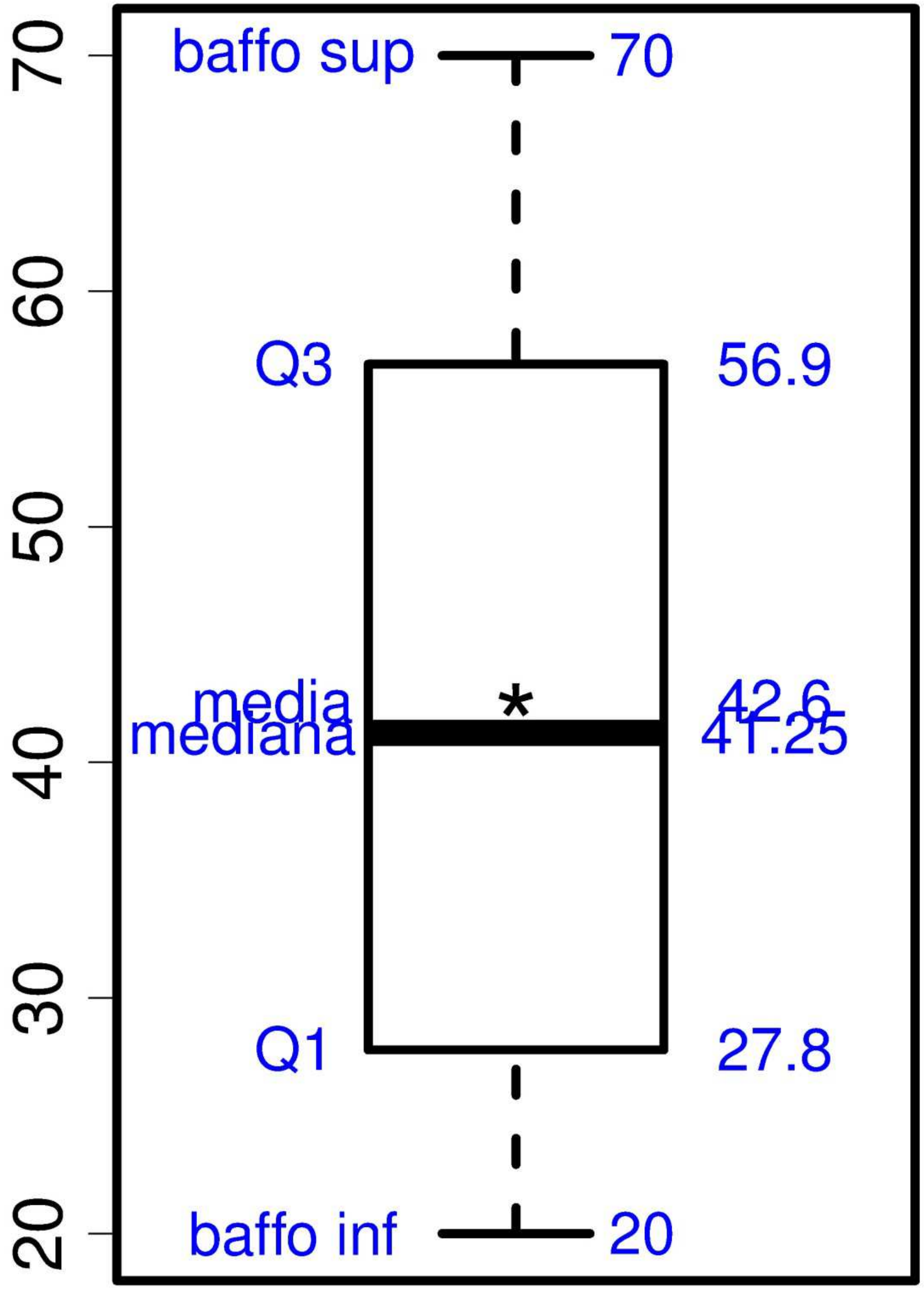
Mediana = 41.25

Q1 = 27.8

Baffo inferiore = 20

media = 42.6

Non ci sono outliers



x_i	n_i	f_i	a_i	d_i	m_i
20 |-- 30	16	0.32	10	0.032	25
30 |-- 60	24	0.48	30	0.016	45
60 |-- 70	10	0.20	10	0.020	65
	50	**1**			

(c) Moda

La classe modale è la classe a cui è associata la massima densità e la moda è il punto medio della classe modale.

La **massima densità** è 0.032, quindi

la **classe modale** è 20|--30 da cui

$$\mathbf{moda = 25}$$

x_i	n_i	f_i	a_i	d_i	m_i
20 |-- 30	16	0.32	10	0.032	25
30 |-- 60	24	0.48	30	0.016	45
60 |-- 70	10	0.20	10	0.020	65
	50	**1**			

(c) Coefficiente di Variazione

$$\mathrm{CV} = \frac{\sigma_x}{|\bar{x}|} = \frac{14.2}{42.6} = 0.33$$

$\bar{x} = \text{media} = 42.6$

$$\sigma_x{}^2 = \text{varianza} = \left(\sum_i m_i{}^2 \, f_i \right) - \bar{x}^2 =$$

$$= (25^2 \times 0.32 + 45^2 \times 0.48 + 65^2 \times 0.2) - (42.6)^2 =$$

$$= 202.24$$

$\sigma_x = \text{scarto quadratico medio} = \sqrt{202.24} = 14.2$

x_i	n_i	f_i	a_i	d_i	m_i
20 |-- 30	16	0.32	10	0.032	25
30 |-- 60	24	0.48	30	0.016	45
60 |-- 70	10	0.20	10	0.020	65
	50	**1**			

La varianza si può calcolare anche così:

$$\sigma_x{}^2 = \text{varianza} = \sum_i (m_i - \bar{x})^2\, \mathrm{f_i} =$$

$$= (25 - 42.6)^2 \times 0.32 + (45 - 42.6)^2 \times 0.48 + (65 - 42.6)^2 \times 0.20 =$$

$$= 202.24$$

oppure

$$\sigma_x{}^2 = \text{varianza} = \left[\sum_i (m_i - \bar{x})^2\, \mathrm{n_i}\right] / n =$$

$$= [(25 - 42.6)^2 \times 16 + (45 - 42.6)^2 \times 24 + (65 - 42.6)^2 \times 10]/50 =$$

$$= 202.24$$

QUIZ 1

Sia X una variabile con media $\mu_x = -3$ e varianza ${\sigma_x}^2 = 2$

Se $Y = -X - 9$, allora:

(a) ${\sigma_y}^2 = -7$

(b) $\mu_y = -12$

$\boxed{\text{(c) } \mu_y = -6}$

(d) ${\sigma_y}^2 = -7$

$$Y = -X - 9 = (-1) \cdot X - 9$$

$$\boldsymbol{\mu_y} = (-1) \cdot \mu_x - 9 = (-1) \cdot (-3) - 9 = 3 - 9 = \mathbf{-6}$$

$$\boldsymbol{{\sigma_y}^2} = (-1)^2 \cdot {\sigma_x}^2 = (-1)^2 \cdot 2 = 1 \cdot 2 = \mathbf{2}$$

Esercizio 2

Siano A e B eventi **indipendenti** con P(A)=0.2 e P(B)=0.1.

Calcolare $P(A \cap B)$, $P(A \cup B)$ e $P(A|B)$.

Eventi indipendenti → $P(A \cap B) = P(A) \times P(B) = 0.2 \times 0.1 = 0.02$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.2 + 0.1 - 0.02 = 0.28$$

$$P(A|B) = P(A) = 0.2$$

Esercizio 3

Siano A e B eventi **incompatibili** con P(A)=0.2 e P(B)=0.1.

Calcolare $P(A \cap B)$, $P(A \cup B)$ e $P(A|B)$.

Eventi incompatibili → $A \cap B = \emptyset$ → $\mathbf{P}(A \cap B) = 0$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.2 + 0.1 - 0 = 0.3$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0}{0.1} = 0$$

QUIZ 2

Siano A e B due eventi con $P(A) = 0.4$ e $P(B) = 0.2$

Quale affermazione è vera?

(a) $P(A \cup B) = 2$

(b) $P(A \cap B) = 0$

(c) $\boxed{P(A \cup B) \leq 0.6}$

(d) $P(A \cap B) = 0.08$

Si sceglie (c) per esclusione:

(a) Falso. La probabilità è sempre un valore tra 0 e 1

(b) Falso. Sarebbe vero se A e B fossero incompatibili, ma questo non è specificato nel testo. Quindi in generale (b) è falso.

(d) Falso. Sarebbe vero se A e B fossero indipendenti, ma questo non è specificato nel testo. Quindi in generale (d) è falso.

QUIZ 2

Siano A e B due eventi con $P(A) = 0.4$ e $P(B) = 0.2$

Quale affermazione è vera?

(a) $P(A \cup B) = 2$

(c) $\boxed{P(A \cup B) \leq 0.6}$

(b) $P(A \cap B) = 0$

(d) $P(A \cap B) = 0.08$

Con le informazioni disponibili, cosa si può dire di $\boldsymbol{P(A \cup B)}$ e $\boldsymbol{P(A \cap B)}$?

Sappiamo che $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, quindi

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

$$P(A \cup B) \leq 0.4 + 0.2 = 0.6$$

Sappiamo solo che $P(A \cap B)$ è inferiore sia a $P(A)$ sia a $P(B)$, quindi

$$P(A \cap B) \leq 0.2$$

Esercizio 4

Un'urna contiene 50 palline, di cui 40 bianche e 10 nere.

- Qual è la probabilità di estrarre una pallina bianca?
- E la probabilità di estrarne una nera?

B = estrazione pallina bianca $P(B) = 40/50 = 0.8$

N = estrazione pallina nera $P(N) = 1\text{-}P(B) = 1\text{-}0.8 = 0.2$

- Si estrae una pallina e, dopo averla rimessa nell'urna, se ne estrae una seconda . Qual è la probabilità che siano entrambe nere?

Estrazioni con reimmissione (indipendenti):

stesse probabilità ad ogni estrazione

Estrazione n.1 $P(N1) = 0.2$

Estrazione n.2 $P(N2) = 0.2$

$\mathbf{P(N1 \cap N2)} = P(N1) \times P(N2) = 0.2 \times 0.2 = \mathbf{0.04}$

- Si estraggono 10 palline con reimmissione. Calcolare la probabilità di estrarne esattamente 7 bianche.

X = numero di palle bianche su 10 estrazioni con reimmissione

X ha distribuzione Binomiale(n = 10, p = 0.8)

$$P(X = i) = \binom{n}{i} p^i (1 - p)^{n-i} \qquad \mathrm{i} = 0,1,2, \dots, \mathrm{n}$$

Quindi si tratta di calcolare:

$$P(X = 7) = \binom{10}{7} 0.8^7 (1 - 0.8)^{10-7}$$

$$\boldsymbol{P(X=7)} = \binom{10}{7} 0.8^7 (1-0.8)^{10-7} =$$

$$= \frac{10!}{7!\,(10-7)!} \cdot 0.8^7 \cdot 0.2^3 =$$

$$= \frac{1\times2\times3\times4\times5\times6\times7\times8\times9\times10}{(1\times2\times3\times4\times5\times6\times7)\times(1\times2\times3)} \cdot 0.8^7 \cdot 0.2^3 =$$

$$= \frac{\overset{\textbf{1}}{\cancel{1\times2\times3\times4\times5\times6\times7}}\times\overset{\textbf{4x3x10}}{\cancel{8\times9\times10}}}{(\underset{\textbf{1}}{\cancel{1\times2\times3\times4\times5\times6\times7}})\times(\underset{\textbf{1}}{\cancel{1\times2\times3}})} \times 0.210 \times 0.008 =$$

$$= \boldsymbol{0.202}$$

La probabilità di pescare esattamente 7 palline bianche su 10 estrazioni con reimmissione è 0.202

- Calcolare la probabilità approssimata che, estraendo con reinserimento 35 palline, al massimo 6 siano nere.

X = numero di palle nere su 35 estrazioni con reimmissione

X ha distribuzione Binomiale(n = 35, p = 0.2)

Quindi si tratterebbe di calcolare:

$$P(X \leq 6) = P(X = 0) + P(X = 1) + P(X = 2) + \dots + P(X = 6) =$$

$$= \sum_{i=0}^{6} \binom{35}{i} 0.2^{i} (1 - 0.2)^{35-i}$$

È laborioso!

Si può ricorrere all'approssimazione della distribuzione Binomiale con la distribuzione Normale?

$$\boldsymbol{X \sim Binomiale(n, p) \quad \rightarrow \quad Z = \frac{X - np}{\sqrt{np(1-p)}} \sim Normale(0, 1)}$$

qualora $\mathbf{n \geq 30}$, $\mathbf{np \geq 5}$, $\mathbf{n(1-p) \geq 5}$.

Si può perché per n=35 e p=0.2 risulta

$$\boldsymbol{n = 35} \quad \geq 30$$

$$\boldsymbol{np} = 35 \times 0.2 = \boldsymbol{7} \quad \geq 5,$$

$$\boldsymbol{n(1-p)} = 35 \times (1-0.2) = \boldsymbol{28} \quad \geq 5$$

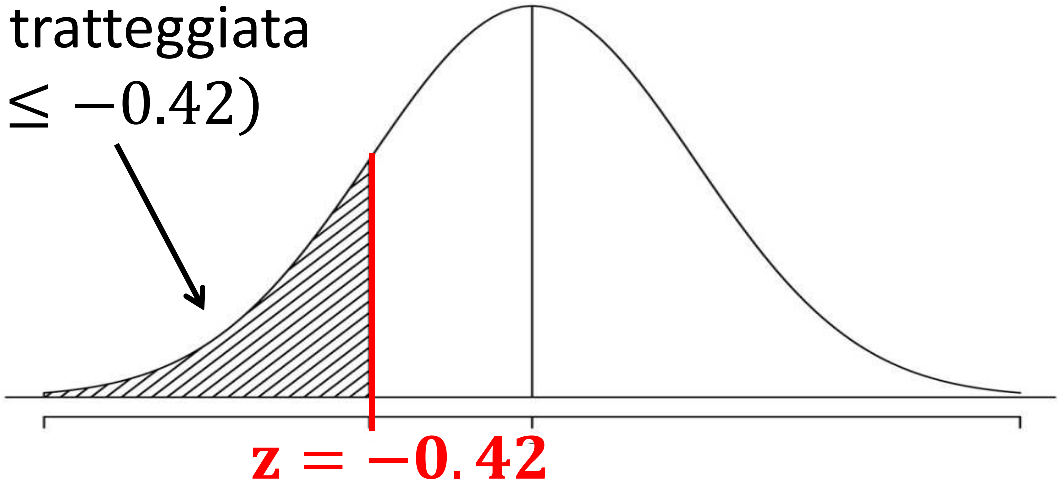
Quindi si applica questa approssimazione:

$$\boldsymbol{X \sim Binomiale(n,p)} \quad \rightarrow \quad \boldsymbol{Z} = \frac{\boldsymbol{X - np}}{\sqrt{\boldsymbol{np(1-p)}}} \sim \boldsymbol{Normale(0,1)}$$

$$\boldsymbol{np} = \boldsymbol{7}$$

$$\sqrt{\boldsymbol{np(1-p)}} = \sqrt{35 \times 0.2 \times 0.8} = \sqrt{5.6} = \boldsymbol{2.37}$$

$$\boldsymbol{P(X \leq 6)} = P\left(Z \leq \frac{6-7}{2.37}\right) = P(Z \leq -0.42) = \boldsymbol{0.34}$$



$P(Z \leq -0.42) = \mathbf{0.34}$

z	0.00	-0.01	-0.02	-0.03	-0.04
0.0	0.50000	0.49601	0.49202	0.48803	0.48405
-0.1	0.46017	0.45620	0.45224	0.44828	0.44433
-0.2	0.42074	0.41683	0.41294	0.40905	0.40517
-0.3	0.38209	0.37828	0.37448	0.37070	0.36693
-0.4	0.34458	0.34090	0.33724	0.33360	0.32997
-0.5	0.30854	0.30503	0.30153	0.29806	0.29460
-0.6	0.27425	0.27093	0.26763	0.26435	0.26109
-0.7	0.24196	0.23885	0.23576	0.23270	0.22965
-0.8	0.21186	0.20897	0.20611	0.20327	0.20045
-0.9	0.18406	0.18141	0.17879	0.17619	0.17361

Iacus “Statistica”, pag. 218

- Si estraggono 3 palline **distinte** dall'urna. Calcolare la probabilità che la prima sia nera e le altre bianche.

Palline **distinte** ⟶ Estrazioni **senza** reinserimento (dipendenti)

$$\boldsymbol{P(N1 \cap B2 \cap B3)} = P(N1) \times P(B2|N1) \times P(B3|N1 \cap B2) = \color{red}{?}$$

Estrazione n.1 — 50 palline = 40 bianche + 10 nere

P(N1) = 10/50 = **0.20**

Estrazione n.2 — 49 palline = 40 bianche + 9 nere

P(B2|N1) = 40/49 = **0.82**

Estrazione n.3 — 48 palline = 39 bianche + 9 nere

P(B3|N1 ∩B2) = 39/48 = **0.81**

$$\boldsymbol{P(N1 \cap B2 \cap B3)} = P(N1) \times P(B2|N1) \times P(B3|N1 \cap B2) =$$
$$= 0.18 \times 0.82 \times 0.81 = \mathbf{0.12}$$

QUIZ 3

Il carattere X associa ogni animale vertebrato ad una delle seguenti modalità: "pesci", "anfibi", "rettili", "uccelli", e "mammiferi".

Si tratta di un carattere :

(a) quantitativo discreto
(b) qualitativo/categorico nominale
(c) quantitativo continuo
(d) qualitativo ordinale

È possibile calcolare :

(a) la mediana
(b) il terzo quartile
(c) la moda
(d) la varianza

In tal caso, la rappresentazione grafica più opportuna è :

(a) il diagramma a bastoncini / aste / barre
(b) il box-plot
(c) il diagramma a torta
(d) l'istogramma

Esercizio 5

Il direttore del personale di una azienda esegue un'indagine sulla disponibilità dei dipendenti a turni di lavoro domenicali con i seguenti risultati:

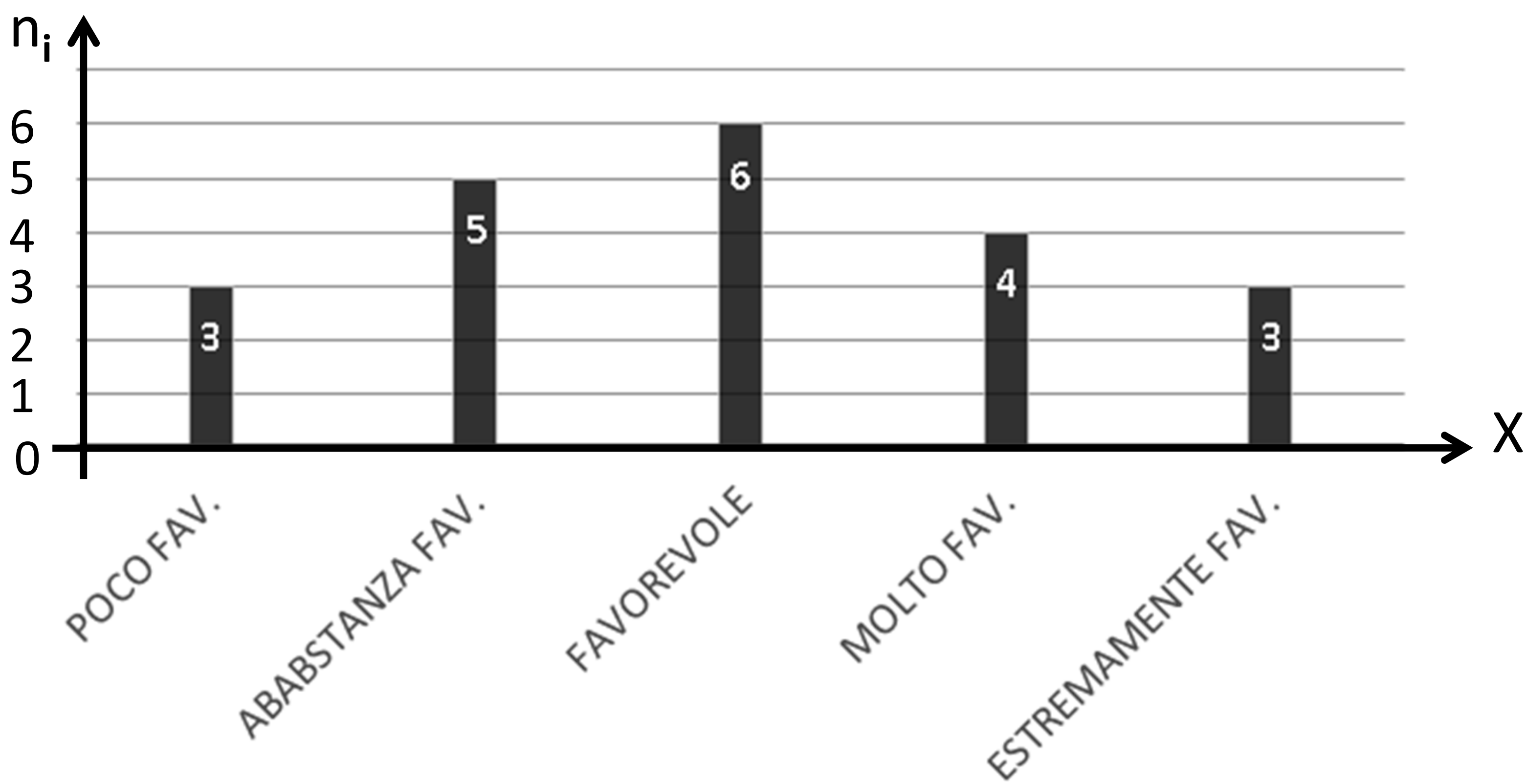
$X = x_i$ grado di disponibilità	n_i
Poco favorevole	3
Abbastanza favorevole	5
Favorevole	6
Molto favorevole	4
Estremamente favorevole	2
	20

(a) Di che tipo è il carattere X? Fornire una rappresentazione grafica opportuna.

$X=x_i$ grado di disponibilità	n_i
Poco favorevole	3
Abbastanza favorevole	5
Favorevole	6
Molto favorevole	4
Estremamente favorevole	2
	20

X è un **carattere qualitativo ordinale (o categorico ordinale)**

Si rappresenta con un **diagramma ad aste/bastoncini/barre**



(b) Calcolare moda, mediana, primo e terzo quartile.

$X=x_i$	n_i
Poco fav.	3
Abbastanza fav.	5
Favorevole	6
Molto fav.	4
Estremamente fav.	2
	20

La moda è la risposta (modalità) più frequente:

Moda = Favorevole

(b) Calcolare moda, mediana, primo e terzo quartile.

$X=x_i$	n_i	n_i cumulato
Poco fav.	3	3
Abbastanza fav.	5	8
Favorevole	6	14
Molto fav.	4	18
Estremamente fav.	2	20
	20	

La moda è la risposta (modalità) più frequente:

Moda = Favorevole

Nell'elenco ordinato delle risposte:

la mediana è la risposta in posizione $(n+1)/2 = (20+1)/2 = 10.5$

Mediana = Favorevole

La colonna delle frequenze assolute cumulate si interpreta così: nell'elenco ordinato dei dati,
le prime 3 persone hanno risposto "poco fav."
le persone dalla posizione 4 alla 8 hanno risposto "abbastanza fav."
le persone dalla posizione 9 alla 14 hanno risposto "favorevole"
etc…
Quindi la posizione cercata 10.5 (tra 10 e 11) corrisponde a "favorevole"

QUIZ 4

Se un fenomeno è qualitativo ordinale, è possibile calcolare:

(a) solo la moda

(b) la media

(c) moda, mediana, primo e terzo quartile

(d) la varianza

QUIZ 5

La mediana è:

(a) la modalità più frequente

(b) il secondo quartile

(c) la modalità con maggior densità

(d) maggiore della media

Esercizio 6   (Es. 1.10 pag.30)

E' noto che solo due virus A e B possono provocare una certa patologia P. I due virus non possono convivere nello stesso organismo. Il virus A e' presente nella popolazione nel 30% dei casi, il virus B nel 70% dei casi.

La probabilita' che si sviluppi la patologia P a seguito di infezione da virus A e' 2%, mentre per il virus B e' del 10%.

Con quale probabilita':

(a) preso a caso un individuo dalla popolazione degli infetti, l'individuo sviluppa la patologia P?

(b) sapendo che l'individuo infetto ha sviluppato la patologia P, risulta infettato dal virus A?

A = individuo con virus A $P(A) = 0.30$

B = individuo con virus B $P(B) = 0.70$

C = individuo senza virus A, B $P(C) = ?$

A, B, C sono eventi incompatibili tali che $A \cup B \cup C = \boldsymbol{\Omega}$:

$$P(A) + P(B) + P(C) = 1$$

$$0.30 + 0.70 + P(C) = 1 \longrightarrow P(C) = 0$$

Esercizio 6 (Es. 1.10 pag.30)

E' noto che solo due virus A e B possono provocare una certa patologia P.
I due virus non possono convivere nello stesso organismo.
Il virus A e' presente nella popolazione nel 30% dei casi, il virus B nel 70% dei casi.
La probabilita' che si sviluppi la patologia P a seguito di infezione da virus A e' 2%, mentre per il virus B e' del 10%.
Con quale probabilità :
(a) preso a caso un individuo dalla popolazione degli infetti, l'individuo sviluppa la patologia P?
(b) sapendo che l'individuo infetto ha sviluppato la patologia P, risulta infettato dal virus A?

A = individuo con virus A $\qquad P(A) = 0.30$

B = individuo con virus B $\qquad \mathrm{P}(B) = 0.70$

M = individuo malato $\qquad P(M|A) = 0.02$

S = individuo sano $\qquad P(M|B) = 0.10$

DOMANDA (a) $\longrightarrow$ $P(M) = ?$

A = individuo con virus A $P(A) = 0.30$

B = individuo con virus B $P(B) = 0.70$

M = individuo malato $P(M|A) = 0.02$

S = individuo sano $P(M|B) = 0.10$

DOMANDA (a) $\longrightarrow$ $P(M) = ?$

Si applica il **principio delle probabilità totali**

$$P(M) = P(M|A)P(A) + P(M|B)P(B) =$$
$$= 0.02 \times 0.30 + 0.10 \times 0.70 = 0.076$$

Preso a caso un individuo dalla popolazione degli infetti, l'individuo sviluppa la patologia P con probabilità 7.6%

Esercizio 6 (Es. 1.10 pag.30)

E' noto che solo due virus A e B possono provocare una certa patologia P.
I due virus non possono convivere nello stesso organismo.
Il virus A e' presente nella popolazione nel 30% dei casi, il virus B nel 70% dei casi.
La probabilita' che si sviluppi la patologia P a seguito di infezione da virus A e' 2%, mentre per il virus B e' del 10%.

Con quale probabilità,

(a) preso a caso un individuo dalla popolazione degli infetti, l'individuo sviluppa la patologia P?

(b) sapendo che l'individuo infetto ha sviluppato la patologia P, risulta infettato dal virus A?

A = individuo con virus A $P(A) = 0.30$

B = individuo con virus B $\mathrm{P}(B) = 0.70$

M = individuo malato $P(M|A) = 0.02$

S = individuo sano $P(M|B) = 0.10$ $P(M) = 0.076$

DOMANDA (b) $\longrightarrow$ $P(A|M) = ?$

A = individuo con virus A $P(A) = 0.30$

B = individuo con virus B $\mathrm{P}(B) = 0.70$

M = individuo malato $P(M|A) = 0.02$

S = individuo sano $P(M|B) = 0.10$ $P(M) = 0.076$

DOMANDA (b) $\longrightarrow$ $P(A|M) = ?$

Dal **teorema di Bayes**

$$P(A|M) = \frac{P(M|A)P(A)}{P(M)} = \frac{0.02 \times 0.30}{0.076} = 0.079$$

La probabilità che un individuo risulti infetto da virus A, sapendo che è malato, è pari al 7.9%

QUIZ 6

Una variabile aleatoria X ha distribuzione Binomiale con parametri n=7 e p=0.25

Il valore di $P(X = 5)$ è

(a) -0.11

(b) <u>0.01</u>

(c) 2.82

(d) 0.18

X ha distribuzione Binomiale(n = 7, p = 0.25)

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i} \qquad \mathrm{i} = 0,1,2, \dots, \mathrm{n}$$

Quindi si tratta di calcolare:

$$P(X = 5) = \binom{7}{5} 0.25^5 (1 - 0.25)^{7-5}$$

$$\boldsymbol{P(X=5)} = \binom{7}{5} 0.25^5 (1-0.25)^{7-5} =$$

$$= \frac{7!}{5!\,(7-5)!} \times 0.25^5 \times 0.75^2 =$$

$$= \frac{1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7}{(1 \times 2 \times 3 \times 4 \times 5) \times (1 \times 2)} \times 0.00098 \times 0.5625 =$$

$$= \frac{\overset{1}{\cancel{1 \times 2 \times 3 \times 4 \times 5}} \times \overset{3 \times 7}{\cancel{6 \times 7}}}{\underset{1}{\cancel{(1 \times 2 \times 3 \times 4 \times 5)}} \times \underset{1}{\cancel{(1 \times 2)}}} \times 0.00098 \times 0.5625 =$$

$$= 21 \times 0.00098 \times 0.5625 = \mathbf{0.01}$$

Esercizio 7

La roulette ha 18 numeri neri, 18 numeri rossi, il numero 0 che fa vincere il banco.

(a) In una partita il giocatore punta tutto sul rosso. Qual è la probabilità che il giocatore perda?

$$X = \text{esito di una puntata} = \begin{cases} 1 & se\ \ vince \\ 0 & se\ \ perde \end{cases}$$

X ha distribuzione Bernoulli (p) dove p è la probabilità di vincere.

$$p = P(X = 1) = \frac{18}{18 + 18 + 1} = 0.49$$

$$1 - p = P(X = 0) = 1 - 0.49 = 0.51$$

La probabilità che il giocatore perda è pari al 51%

Esercizio 7

La roulette ha 18 numeri neri, 18 numeri rossi, il numero 0 che fa vincere il banco.

(b) Si supponga che il giocatore punti 6 volte di seguito sul rosso. Con quale probabilità il giocatore perde tutte le puntate?

Siano $X_1, X_2, X_3, X_4, X_5, X_6$ gli esiti delle 6 puntate.

variabili **indipendenti** con ugual distribuzione **Bernoulli (p=0.49)**

$$P(X_1 = 0 \cap X_2 = 0 \cap X_3 = 0 \cap X_4 = 0 \cap X_5 = 0 \cap X_6 = 0) =$$

$$= P(0) \times P(0) \times P(0) \times P(0) \times P(0) \times P(0) =$$

$$= (1-p)^6 = 0.51^6 = 0.0176$$

La probabilità che il giocatore perda 6 volte su 6 è pari a 1.76%

Esercizio 7

La roulette ha 18 numeri neri, 18 numeri rossi, il numero 0 che fa vincere il banco.

(b) Si supponga che il giocatore punti 6 volte di seguito sul rosso. Con quale probabilità il giocatore perde tutte le puntate?

Altra possibile soluzione

Sia $Y = X_1 + X_2 + X_3 + X_4 + X_5 + X_6$ = numero di vincite su 6 giocate

Y ha distribuzione Binomiale(n=6, p=0.49)

$$P(Y = 0) = \binom{6}{0} 0.49^0 (1 - 0.49)^{6-0} = 0.0176$$

Si ricorda che $\mathbf{0! = 1}$ e $\mathbf{0.49^0 = 1}$

Esercizio 7

La roulette ha 18 numeri neri, 18 numeri rossi, il numero 0 che fa vincere il banco.

(b) Si supponga che il giocatore punti 6 volte di seguito sul rosso. Qual è la probabilità che il giocatore vinca una sola volta?

$Y = X_1 + X_2 + X_3 + X_4 + X_5 + X_6$ = numero di vincite su 6 giocate

Y ha distribuzione Binomiale(n=6, p=0.49)

$$P(Y = 1) = \binom{6}{1} 0.49^1 (1 - 0.49)^{6-1} = 0.10$$

Si ricorda che $\mathbf{1! = 1}$

La probabilità che il giocatore vinca una sola volta su 6 è pari a 10%

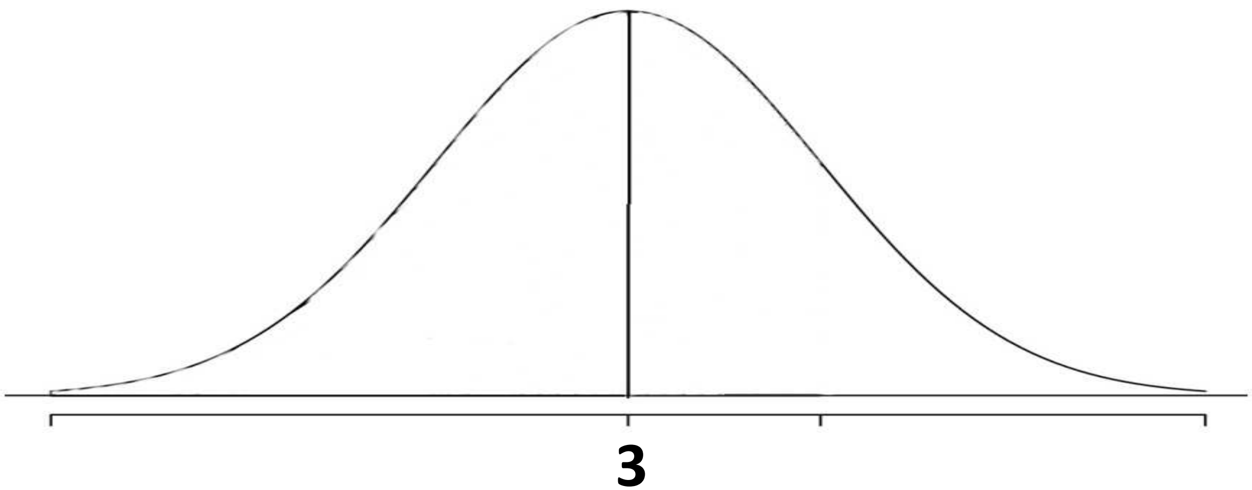
QUIZ 7

La variabile X ha distribuzione Normale con media 3 e varianza 5.
Il valore di $P(X \geq 5)$ è pari a :

(a) 0.6
(b) maggiore o uguale a 0.5
(c) -2
(d) minore di 0.5

Area = Probabilità

Area totale sottostante tutta la curva = Probabilità 1



QUIZ 7

La variabile X ha distribuzione Normale con media 3 e varianza 5.
Il valore di $P(X \geq 5)$ è pari a :

(a) 0.6

(b) maggiore o uguale a 0.5

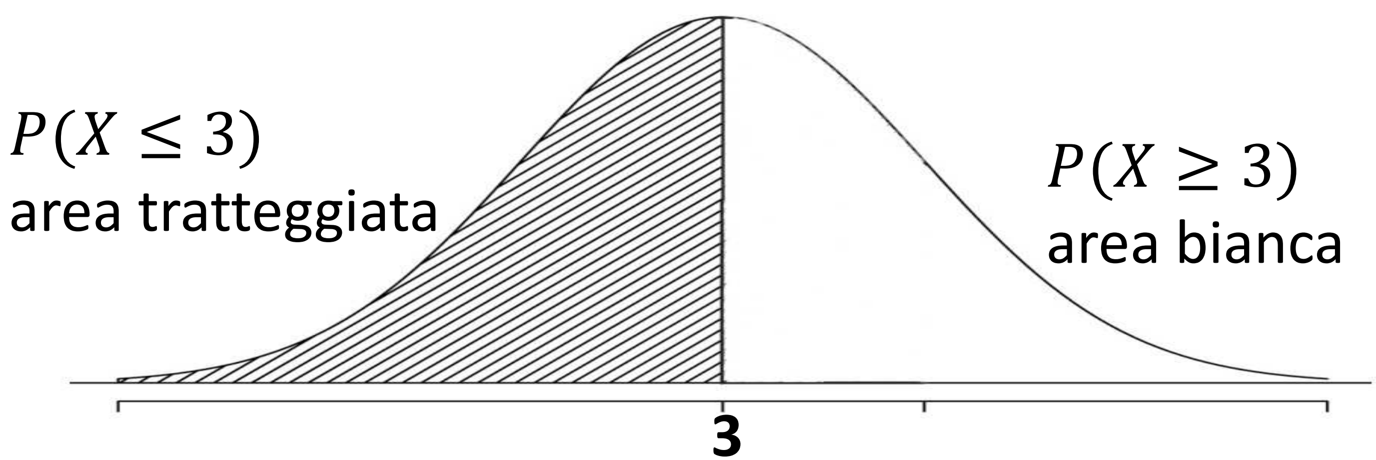
(c) -2

(d) minore di 0.5

Area = Probabilità

Area totale sottostante tutta la curva = Probabilità 1

Area sottostante la curva a destra di 3 = $P(X \geq 3)$ $\mathbf{= 0.5}$



QUIZ 7

La variabile X ha distribuzione Normale con media 3 e varianza 5.
Il valore di $P(X \geq 5)$ è pari a :

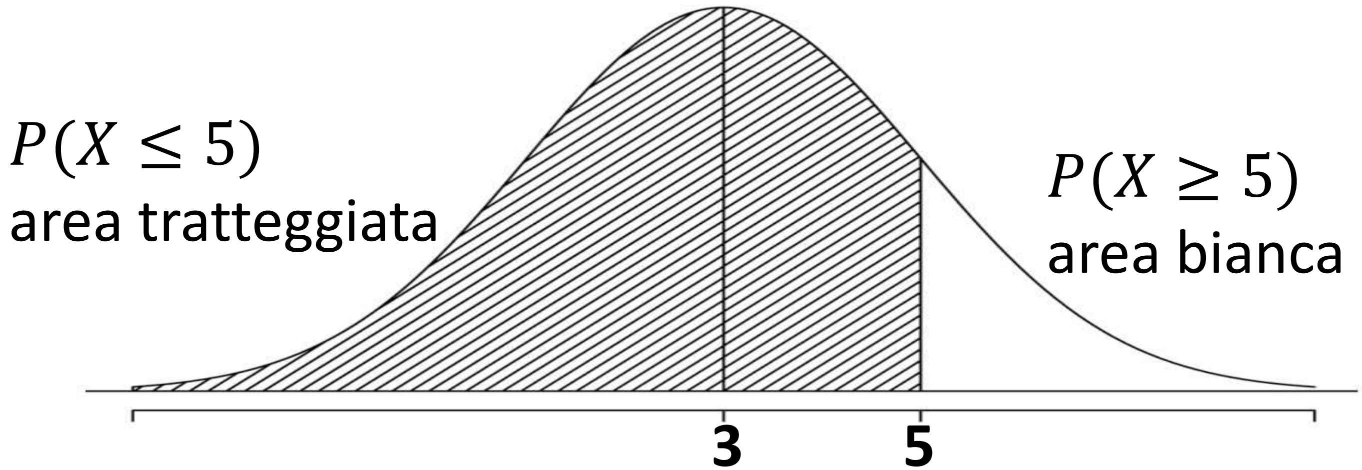
(a) 0.6
(b) maggiore o uguale a 0.5
(c) -2
(d) minore di 0.5

Area = Probabilità

Area totale sottostante tutta la curva = Probabilità 1

Area sottostante la curva a destra di 3 = $P(X \geq 3)$ $\mathbf{= 0.5}$

Area sottostante la curva a destra di 5 = $P(X \geq 5)$ $\mathbf{< 0.5}$



QUIZ 7

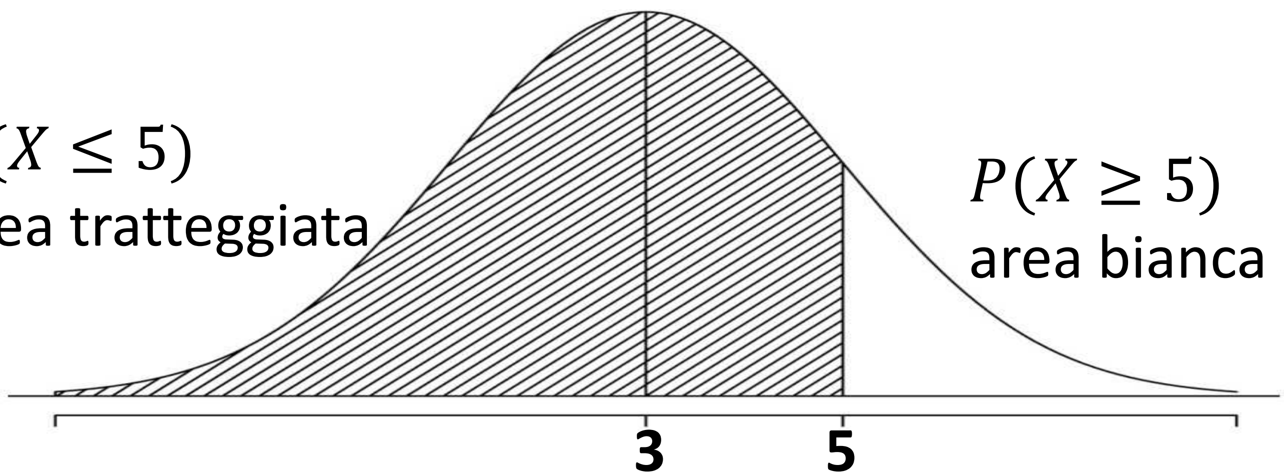
La variabile X ha distribuzione Normale con media 3 e varianza 5.
Il valore di $P(X \geq 5)$ è pari a :

(a) 0.6
(b) maggiore o uguale a 0.5
(c) -2
(d) <u>minore di 0.5</u>

Abbiamo capito che $P(X \geq 5) < \mathbf{0.5}$, quindi escludiamo (a) e (b)

Escludiamo anche (c) perché la probabilità è un numero tra 0 e 1

$P(X \leq 5)$
area tratteggiata



$P(X \geq 5)$
area bianca