STATISTICA

Esercizi

Esercizio 5

Un’urna contiene 400 palline nere e 100 palline rosse.
a) Si estraggono a caso 5 palline diverse: calcolare la probabilità che siano tutte nere.
b) Si estraggono a caso, con reimmissione, 15 palline: calcolare la probabilità che 8 siano rosse.
c) Si estraggono a caso, con reimmissione, 50 palline: calcolare la probabilità che meno di 8 siano rosse.

Esercizio 5

Un’urna contiene 400 palline nere e 100 palline rosse.

a) Si estraggono a caso **5 palline diverse**: calcolare la probabilità che siano **tutte nere**.

b) Si estraggono a caso, con reimmissione, 15 palline: calcolare la probabilità che 8 siano rosse.

c) Si estraggono a caso, con reimmissione, 50 palline: calcolare la probabilità che meno di 8 siano rosse.

a) $$\frac{400}{500} \times \frac{399}{499} \times \frac{398}{498} \times \frac{397}{497} \times \frac{396}{496} = 0.326$$

Esercizio 5

Un'urna contiene 400 palline nere e 100 palline rosse.
a) Si estraggono a caso 5 palline diverse: calcolare la probabilità che siano tutte nere.
b) Si estraggono a caso, **con reimmissione, 15 palline**: calcolare la probabilità che 8 siano rosse.
c) Si estraggono a caso, con reimmissione, 50 palline: calcolare la probabilità che meno di 8 siano rosse.

a) $\frac{400}{500} \times \frac{399}{499} \times \frac{398}{498} \times \frac{397}{497} \times \frac{396}{496} = 0.326$

b) $X =$ num. di palline rosse in un camp. cas. di 15 palline

$$X \sim Bin\left(15, \frac{\mathbf{100}}{\mathbf{500}}\right) \Rightarrow P(X = 8) = \binom{15}{8} 0.2^8 0.8^7 =$$

Esercizio 5

$$P(X=8)=\binom{15}{8}0.2^8 0.8^7=$$

$$=\frac{15\times 14\times 13\times 12\times 11\times 10\times 9\times 8}{8\times 7\times 6\times 5\times 4\times 3\times 2}\times 0.00000054$$

$$=\frac{15\times \overset{2}{\cancel{14}}\times 13\times \cancel{12}\times 11\times \cancel{10}\times \overset{\boxed{3}}{\cancel{9}}\times \cancel{8}}{\cancel{8}\times \cancel{7}\times \cancel{6}\times \cancel{5}\times \cancel{4}\times \cancel{3}\times \cancel{2}}\times 0.00000054$$

$$=\frac{15\times 13\times 11\times 3}{1}\times 0.00000054=0.003$$

Esercizio 5

Un'urna contiene 400 palline nere e 100 palline rosse.
a) Si estraggono a caso 5 palline diverse: calcolare la probabilità che siano tutte nere.
b) Si estraggono a caso, **con reimmissione, 15 palline**: calcolare la probabilità che 8 siano rosse.
c) Si estraggono a caso, con reimmissione, 50 palline: calcolare la probabilità che meno di 8 siano rosse.

a) $$\frac{400}{500} \times \frac{399}{499} \times \frac{398}{498} \times \frac{397}{497} \times \frac{396}{496} = 0.326$$

b) $X =$ num. di palline rosse in un camp. cas. di 15 palline

$$X \sim Bin\left(15, \mathbf{\frac{100}{500}}\right) \Rightarrow P(X = 8) = \binom{15}{8} 0.2^8 0.8^7 = 0.003$$

Esercizio 5

Un’urna contiene 400 palline nere e 100 palline rosse.
a) Si estraggono a caso 5 palline diverse: calcolare la probabilità che siano tutte nere.
b) Si estraggono a caso, con reimmissione, 15 palline: calcolare la probabilità che 8 siano rosse.
c) Si estraggono a caso, **con reimmissione, 50 palline**: calcolare la probabilità che meno di 8 siano rosse.

c) $X =$ num. di palline rosse in un camp. cas. di 50 palline

$$X \sim Bin\left(50, \frac{\mathbf{100}}{\mathbf{500}} = \mathbf{0.2}\right) \Rightarrow P(X < 8)$$

Esercizio 5

Un’urna contiene 400 palline nere e 100 palline rosse.
a) Si estraggono a caso 5 palline diverse: calcolare la probabilità che siano tutte nere.
b) Si estraggono a caso, con reimmissione, 15 palline: calcolare la probabilità che 8 siano rosse.
c) Si estraggono a caso, **con reimmissione, 50 palline**: calcolare la probabilità che meno di 8 siano rosse.

c) $X =$ num. di palline rosse in un camp. cas. di 50 palline

$$X \sim Bin\left(50, \frac{\mathbf{100}}{\mathbf{500}} = \mathbf{0.2}\right) \Rightarrow P(X < 8)$$

$$50 \times 0.2 = 10 > 5, 50 \times 0.8 = 40 > 5 \Rightarrow \textbf{GAUSSIANA}!$$

Esercizio 5

Un'urna contiene 400 palline nere e 100 palline rosse.
a) Si estraggono a caso 5 palline diverse: calcolare la probabilità che siano tutte nere.
b) Si estraggono a caso, con reimmissione, 15 palline: calcolare la probabilità che 8 siano rosse.
c) Si estraggono a caso, **con reimmissione, 50 palline**: calcolare la probabilità che meno di 8 siano rosse.

c) $X =$ num. di palline rosse in un camp. cas. di 50 palline

$$X \sim Bin\left(50, \frac{100}{500} = 0.2\right) \Rightarrow P(X < 8) \cong P\left(\boldsymbol{Z < \frac{8 - 50 \times 0.2}{\sqrt{50 \times 0.2 \times 0.8}}}\right)$$

$$P\left(Z < \frac{8 - 10}{\sqrt{8}}\right) = P(Z < -0.71) = \mathbf{0.23885}$$

Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5

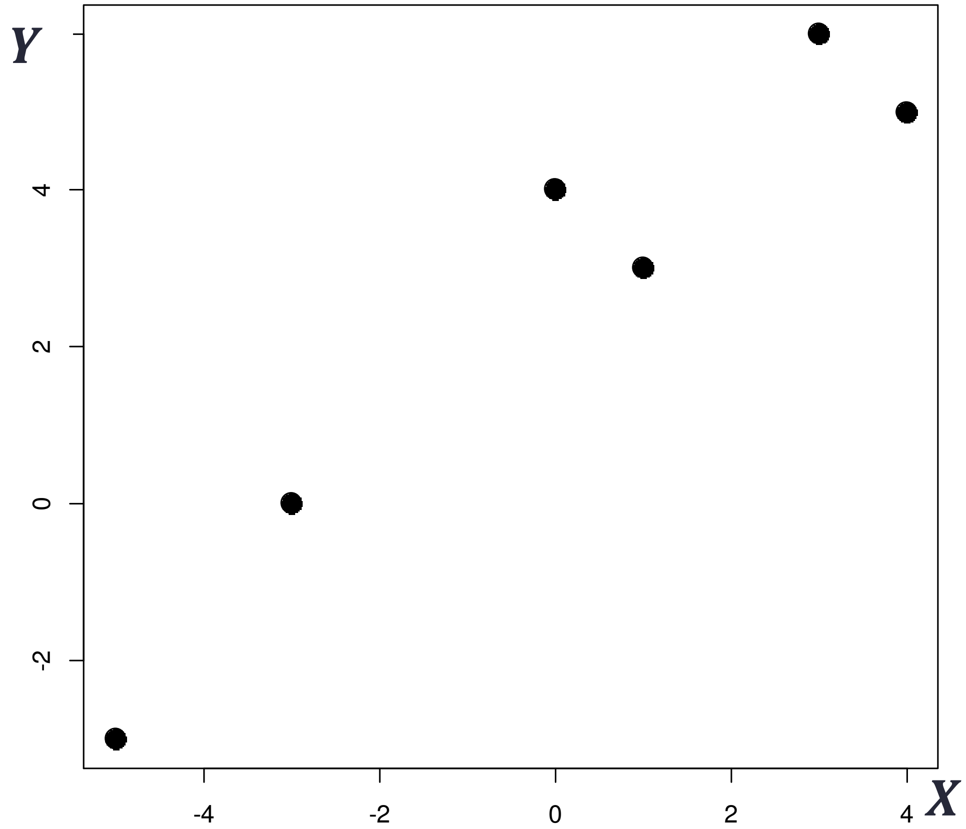
a) Rappresentare l'andamento di Y in funzione di X. Le due variabili sono correlate? Come e quanto?
b) Stimare i parametri della retta di regressione conseguente al punto a) e disegnarla sul grafico.
c) Quanto vale la bontà di adattamento del modello ai dati?
d) La regressione è statisticamente significativa al livello del 2%?
e) Se possibile, indicare l'intervallo di confidenza al 95% per la previsione del valore di Y in corrispondenza a $X = 2$ e $X = -10$.

Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5

a) Rappresentare l'andamento di Y in funzione di X. Le due variabili sono correlate? Come e quanto?



Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5

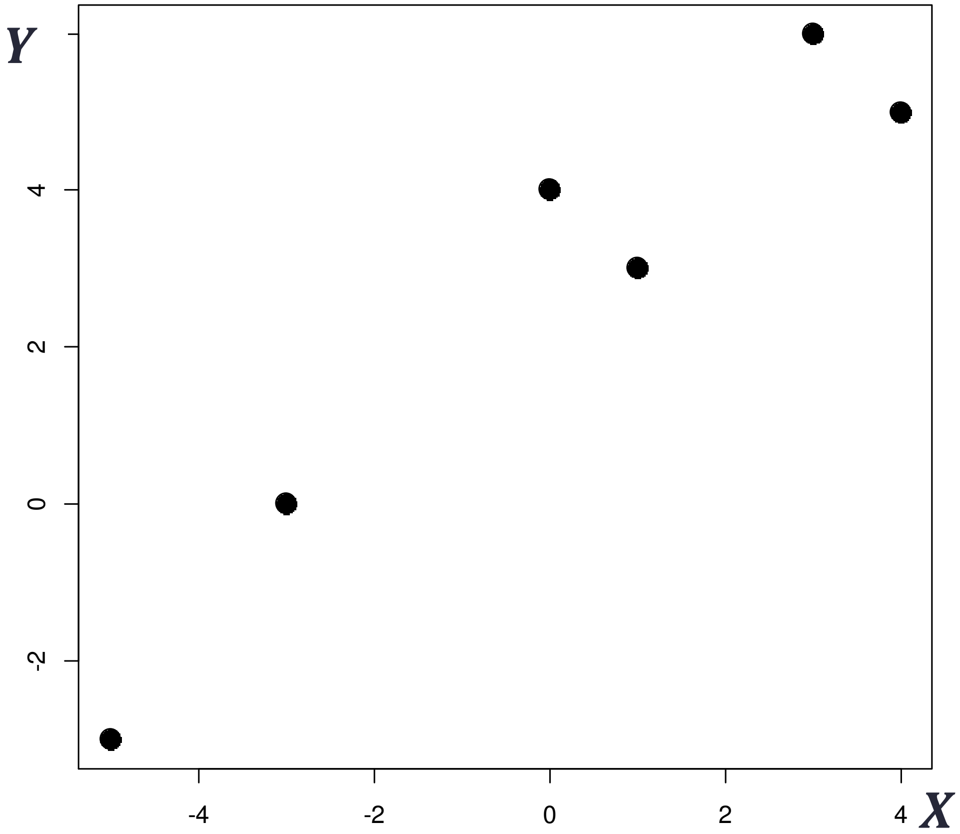
a) Le due variabili sono correlate? Come e quanto?

$$\bar{x} = 0, \qquad \bar{y} = 2.5$$

$$\sigma_x{}^2 = \frac{1}{6}\sum x_i{}^2 = 10 \ , \qquad \sigma_y{}^2 = 9.58$$

$$\sigma_{xy} = \frac{1}{6}\sum x_i y_i \ = 9.33$$

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{9.33}{\sqrt{10 \times 9.58}} = 0.95$$



Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

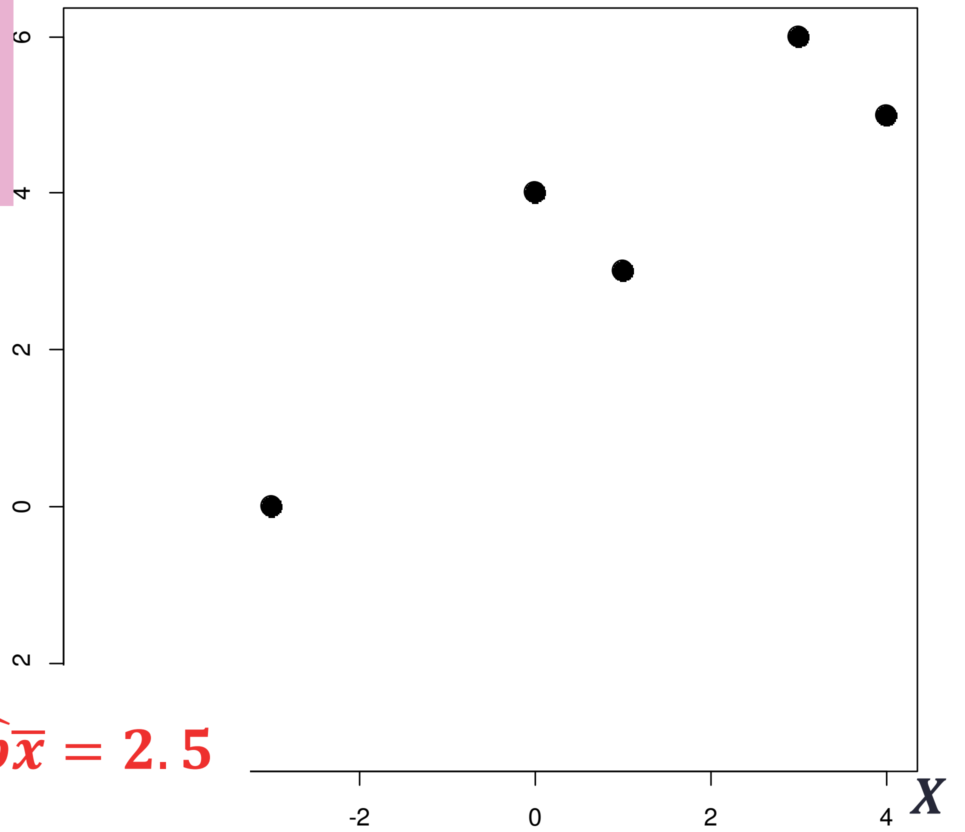
x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5

b) Stimare i parametri della retta di regressione conseguente al punto a) e disegnarla sul grafico.

$\bar{x} = 0, \qquad \bar{y} = 2.5$

$$\sigma_x{}^2 = \frac{1}{6}\sum x_i{}^2 = 10 \ , \qquad \sigma_y{}^2 = 9.58$$

$$\sigma_{xy} = \frac{1}{6}\sum x_i y_i \ = 9.33$$



$$\hat{b} = \frac{\sigma_{xy}}{\sigma_x{}^2} = \frac{9.33}{10} = 0.933, \qquad \hat{a} = \bar{y} - \hat{b}\bar{x} = 2.5$$

Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

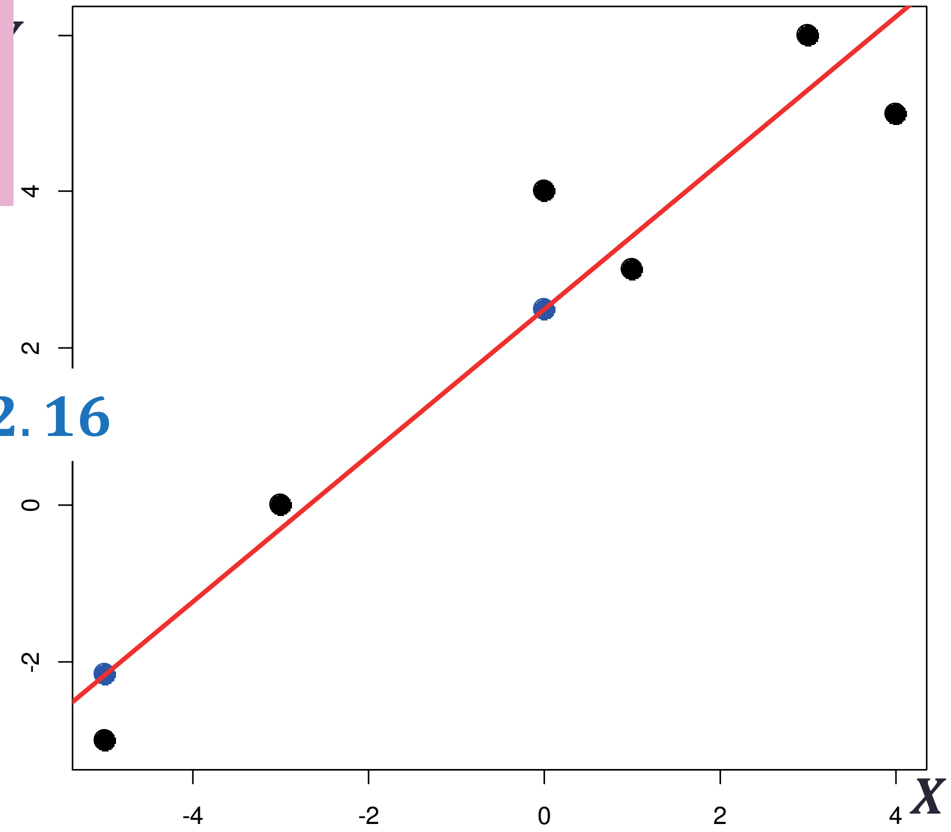
x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5

b) Stimare i parametri della retta di regressione conseguente al punto a) e disegnarla sul grafico.

$\bar{x} = 0,\qquad \bar{y} = 2.5$

$x = -5 \Rightarrow y = 2.5 + 0.933 \times (-5) = -2.16$

$\hat{b} = 0.933,\qquad \hat{a} = 2.5$

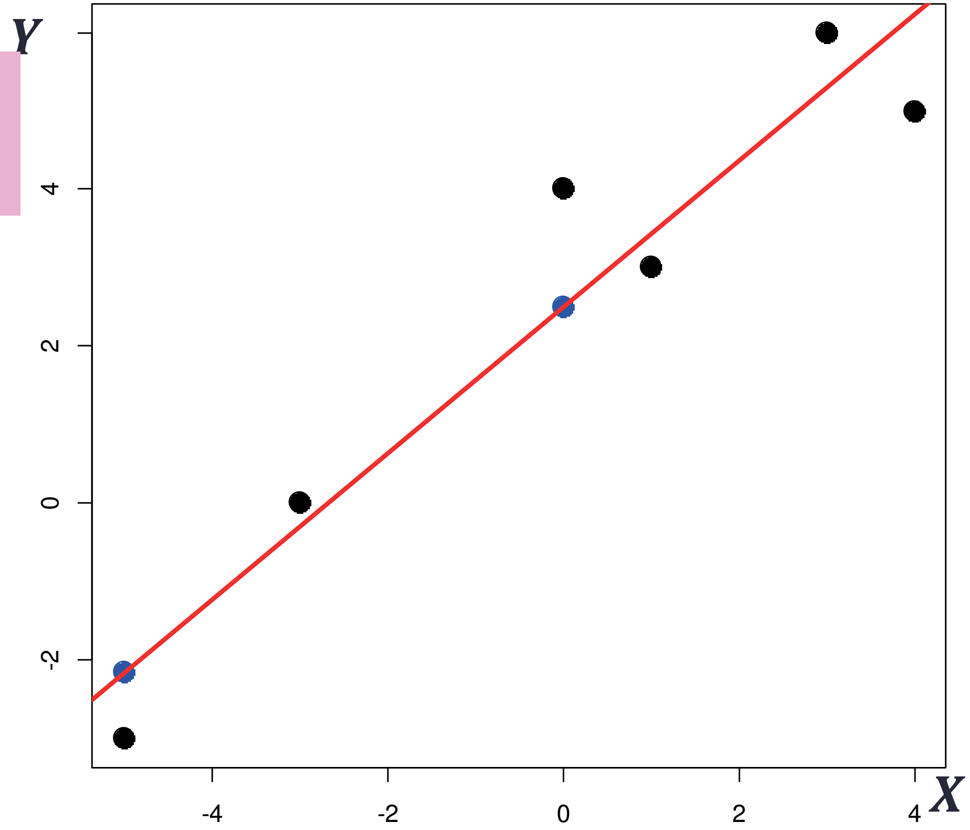


Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5

c) Quanto vale la bontà di adattamento del modello ai dati?



$$\hat{b} = 0.933, \qquad \hat{a} = 2.5$$

Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

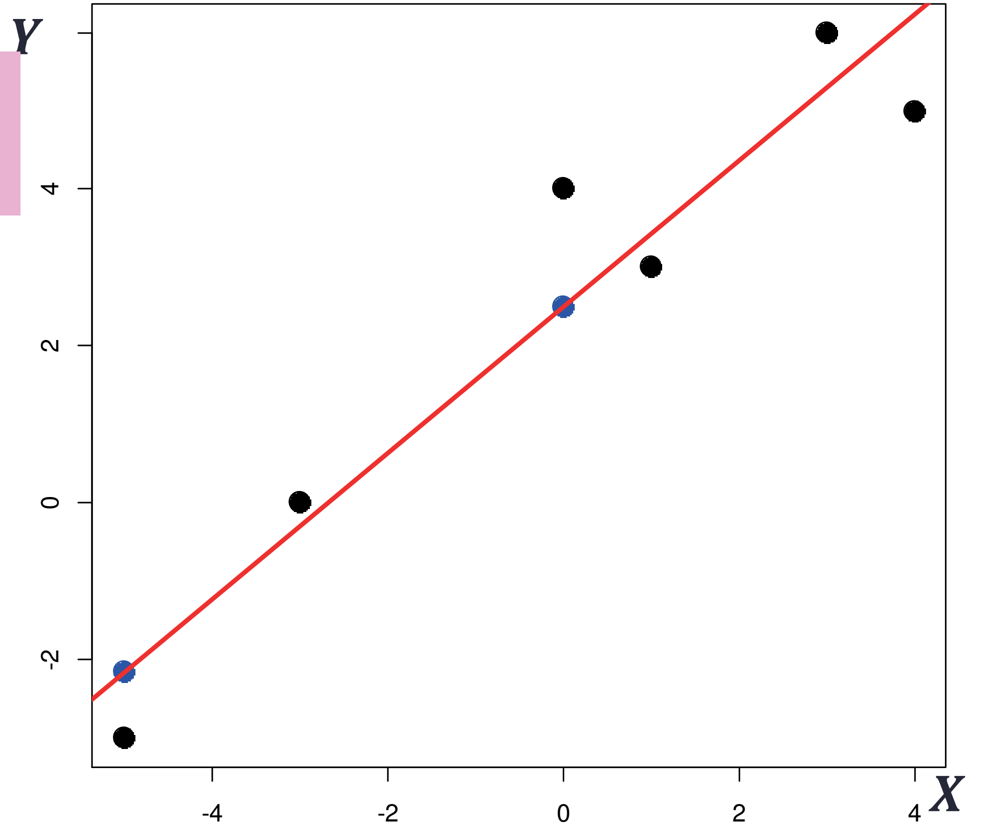
x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5

c) Quanto vale la bontà di adattamento del modello ai dati?

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{9.33}{\sqrt{10 \times 9.58}} = 0.95 \Rightarrow$$

$$R^2 = {\rho_{xy}}^2 = 0.95^2 = 0.90$$

$$\hat{b} = 0.933, \qquad \hat{a} = 2.5$$



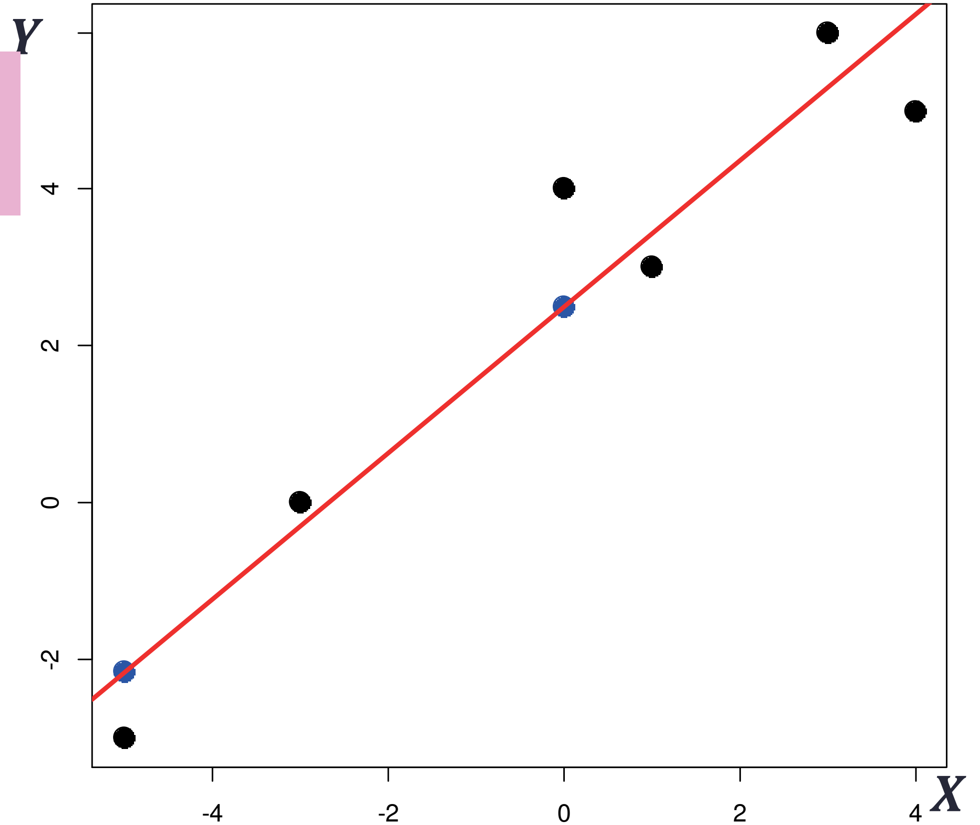
Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5

d) La regressione è statisticamente significativa al livello del 2%?

$$\hat{b} = 0.933, \qquad \hat{a} = 2.5$$



Esercizio 7

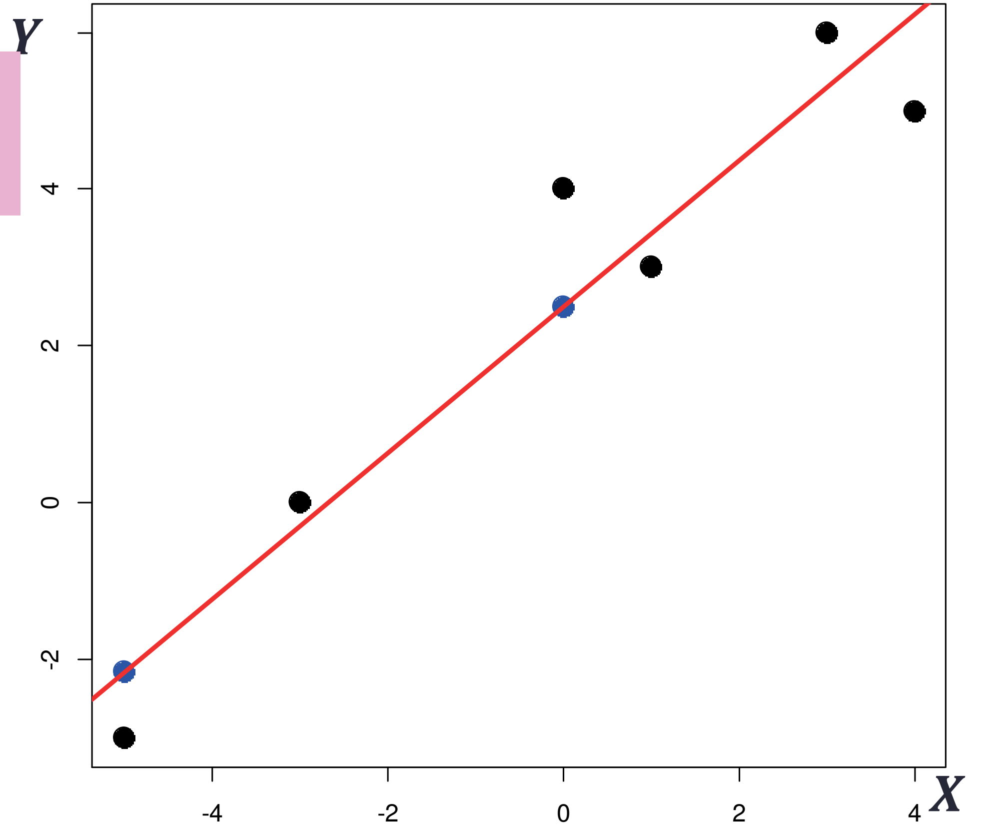
In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5

d) La regressione è statisticamente significativa al livello del 2%?

$$\hat{b} = 0.933, \qquad \hat{a} = 2.5 \qquad {\sigma_x}^2 = 10$$

$$\frac{|\hat{b}|}{\sqrt{\dfrac{s^2}{n{\sigma_x}^2}}}$$



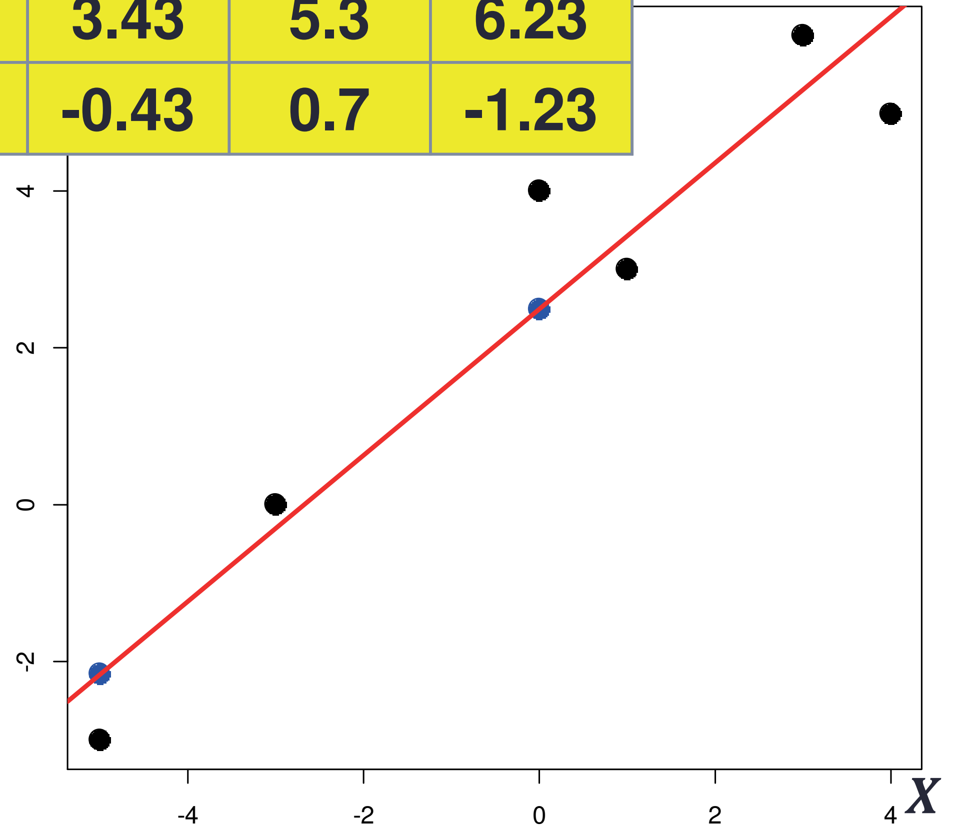
Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5
$\hat{y}_i$	-2.17	-0.3	2.5	3.43	5.3	6.23
e_i	-0.83	0.3	1.5	-0.43	0.7	-1.23

$\hat{b} = 0.933, \qquad \hat{a} = 2.5 \qquad {\sigma_x}^2 = 10$

$$\frac{|\hat{b}|}{\sqrt{\frac{s^2}{n{\sigma_x}^2}}}$$



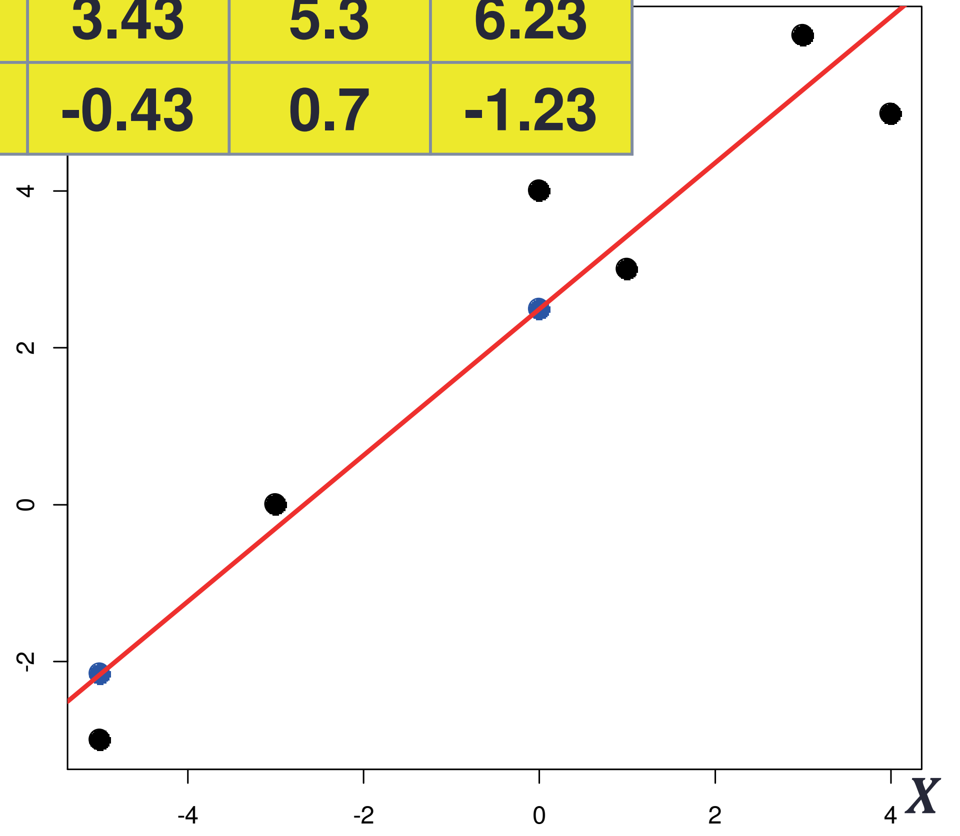
Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5
$\hat{y}_i$	-2.17	-0.3	2.5	3.43	5.3	6.23
e_i	-0.83	0.3	1.5	-0.43	0.7	-1.23

$$\hat{b} = 0.933, \qquad \hat{a} = 2.5 \qquad \sigma_x{}^2 = 10$$

$$s^2 = \frac{1}{n-2}\sum e_i{}^2 = \frac{1}{4} \times 5.22 = 1.31$$

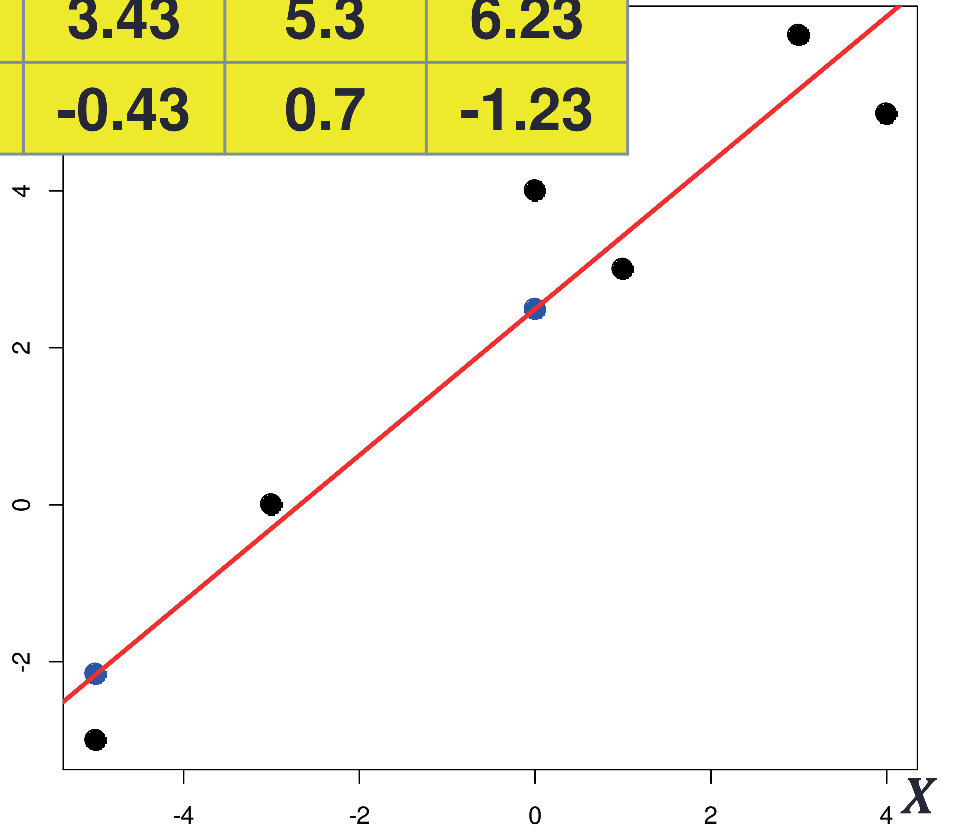


Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5
$\hat{y}_i$	-2.17	-0.3	2.5	3.43	5.3	6.23
				-0.43	0.7	-1.23

d) La regressione è statisticamente significativa al livello del 2%?



$$\hat{b} = 0.933, \qquad \hat{a} = 2.5 \qquad {\sigma_x}^2 = 10$$

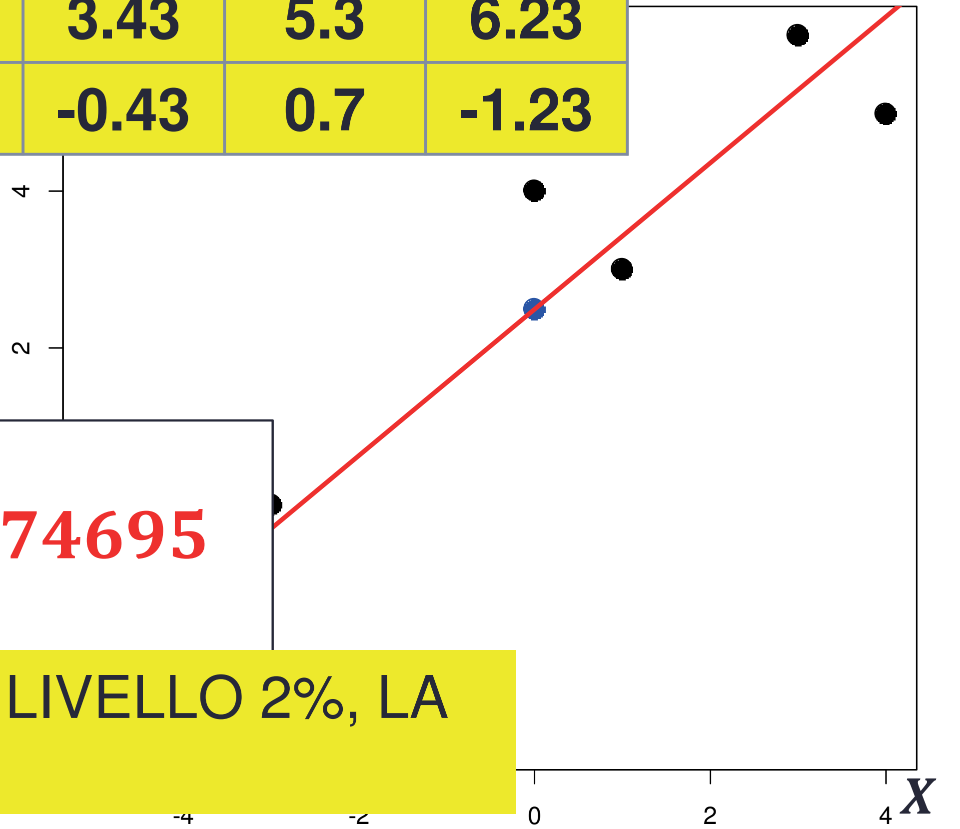
$$s^2 = 1.31$$

$$\frac{|\hat{b}|}{\sqrt{\frac{s^2}{n{\sigma_x}^2}}} = \frac{0.933}{\sqrt{\frac{1.31}{6 \times 10}}} = 6.31$$

Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5
$\hat{y}_i$	-2.17	-0.3	2.5	3.43	5.3	6.23
[illegible]	[illegible]	[illegible]	[illegible]	-0.43	0.7	-1.23



d) La regressione è statisticamente significativa al livello del 2%?

$\hat{b} = 0.933, \qquad \hat{a} = 2.5 \qquad \sigma_x{}^2 = 10$

$$\frac{|\hat{b}|}{\sqrt{\frac{s^2}{n\sigma_x{}^2}}} = 6.31 > t(4)_{1-\frac{0.02}{2}} = 3.74695$$

SI RIFIUTA L'IP. NULLA AL LIVELLO 2%, LA REG. **E'** SIGNIFICATIVA !

Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5

a) Rappresentare l'andamento di Y in funzione di X. Le due variabili sono correlate? Come e quanto?
b) Stimare i parametri della retta di regressione conseguente al punto a) e disegnarla sul grafico.
c) Quanto vale la bontà di adattamento del modello ai dati?
d) La regressione **è statisticamente significativa al livello del 2%! sì**
e) **Se possibile**, indicare l'intervallo di confidenza al 95% per la previsione del valore di Y in corrispondenza a $\boldsymbol{X = 2}$ ed a $\boldsymbol{X = -10}$.

non è nel *range* di X, quindi **non** possiamo fare previsione

Esercizio 7

In uno studio sul cambiamento climatico si sono effettuate alcune rilevazioni congiuntamente su due sostanze presenti in atmosfera, X e Y, ottenendo:

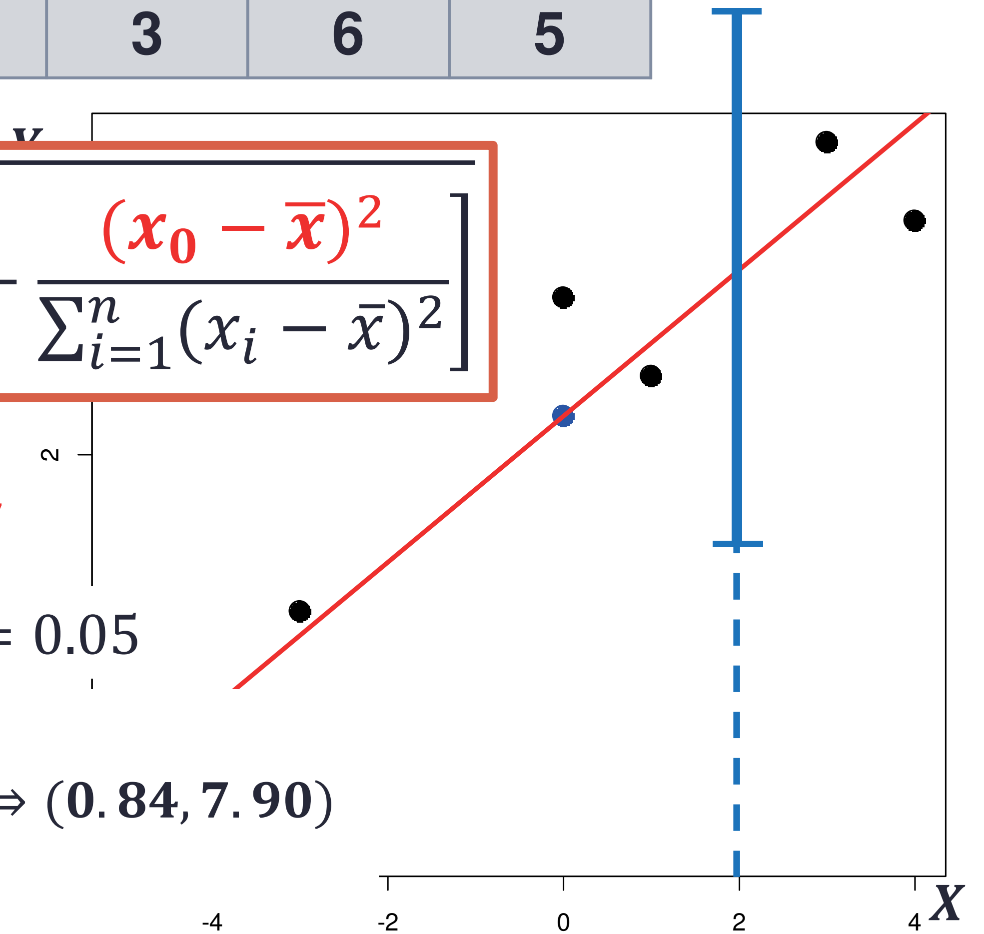
x_i	-5	-3	0	1	3	4
y_i	-3	0	4	3	6	5

$$\boxed{\hat{y}_0 \mp t(n-2)_{1-\frac{\alpha}{2}} \times \sqrt{s^2\left[1 + n^{-1} + \frac{(x_0 - \bar{x})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right]}}$$

$$x = 2 \Rightarrow y = 2.5 + 0.933 \times (2) = 4.37$$

$$\sigma_x{}^2 = 10, \quad s^2 = 1.31, \qquad \bar{x} = 0, \qquad \alpha = 0.05$$

$$4.37 \mp 2.77645 \times \sqrt{1.31\left[1 + \frac{1}{6} + \frac{(2)^2}{6 \times 10}\right]} \Rightarrow (0.84, 7.90)$$



Esercizio 10

In una città è stato rilevato il numero di guasti a mezzi pubblici di superficie in 10 giorni consecutivi

3, 0, 2, 1, 2, 0, 0, 1, 5, 1

a) Rappresentare con un opportuno grafico la distribuzione di frequenza dei dati
b) Calcolare moda, quartili, media e varianza dei dati
c) Ci sono *outlier* (valori anomali) nei dati ?

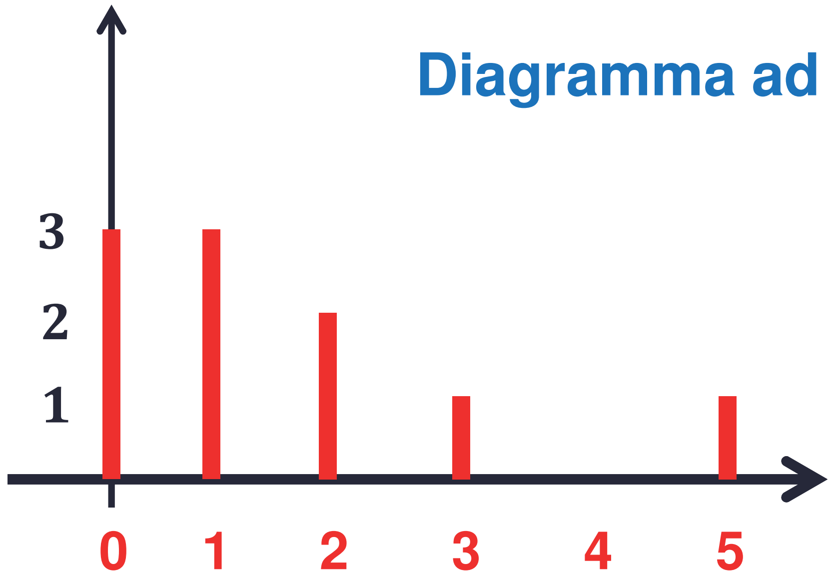
Esercizio 10

In una città è stato rilevato il numero di guasti a mezzi pubblici di superficie in 10 giorni consecutivi

0, 0, 0, 1, 1, 1, 2, 2, 3, 5

a) Rappresentare con **un opportuno grafico** la distribuzione di frequenza dei dati
b) Calcolare moda, quartili, media e varianza dei dati
c) Ci sono *outlier* (valori anomali) nei dati ?

Diagramma ad aste/bastoncini



Esercizio 10

In una città è stato rilevato il numero di guasti a mezzi pubblici di superficie in 10 giorni consecutivi

0, 0, 0, 1, 1, 1, 2, 2, 3, 5

a) Rappresentare con un opportuno grafico la distribuzione di frequenza dei dati
b) Calcolare **moda, quartili, media e varianza dei dati**

due mode: 0 e 1

x_i	n_i
0	3
1	3
2	2
3	1
5	1

$$\bar{x} = \frac{0 \times 3 + 1 \times 3 + 2 \times 2 + 3 + 5}{10} = 1.5$$

Esercizio 10

In una città è stato rilevato il numero di guasti a mezzi pubblici di superficie in 10 giorni consecutivi

0, 0, 0, 1, 1, 1, 2, 2, 3, 5

a) Rappresentare con un opportuno grafico la distribuzione di frequenza dei dati
b) Calcolare **moda, quartili, media e varianza dei dati**

x_i	n_i
0	3
1	3
2	2
3	1
5	1

due mode: 0 e 1

$$\bar{x} = 1.5$$

$$\sigma^2 = \frac{3(0-1.5)^2 + 3(1-1.5)^2 + 2(2-1.5)^2}{10} +$$

$$+ \frac{(3-1.5)^2 + (5-1.5)^2}{10} = 2.25$$

Esercizio 10

In una città è stato rilevato il numero di guasti a mezzi pubblici di superficie in 10 giorni consecutivi

0, 0, 0, 1, 1, 1, 2, 2, 3, 5

a) Rappresentare con un opportuno grafico la distribuzione di frequenza dei dati
b) Calcolare **moda, quartili, media e varianza dei dati**

x_i	n_i
0	3
1	3
2	2
3	1
5	1

Posizione dei quartili:

$$\frac{n+1}{4} = 2.75$$

$$\frac{n+1}{2} = 5.5$$

$$\frac{3(n+1)}{4} = 8.25$$

Esercizio 10

In una città è stato rilevato il numero di guasti a mezzi pubblici di superficie in 10 giorni consecutivi

0, 0, 0, 1, 1, 1, 2, 2, 3, 5

a) Rappresentare con un opportuno grafico la distribuzione di frequenza dei dati
b) Calcolare **moda, quartili, media e varianza dei dati**

x_i	n_i
0	3
1	3
2	2
3	1
5	1

Posizione dei quartili:

$$\frac{n+1}{4} = 2.75$$

$$\frac{n+1}{2} = 5.5$$

$$\frac{3(n+1)}{4} = 8.25$$

Valore dei quartili

$$\frac{\mathbf{0+0}}{\mathbf{2}} = \mathbf{0}$$

$$\frac{\mathbf{1+1}}{\mathbf{2}} = \mathbf{1}$$

$$\frac{\mathbf{2+3}}{\mathbf{2}} = \mathbf{2.5}$$

Esercizio 10

In una città è stato rilevato il numero di guasti a mezzi pubblici di superficie in 10 giorni consecutivi

0, 0, 0, 1, 1, 1, 2, 2, 3, 5

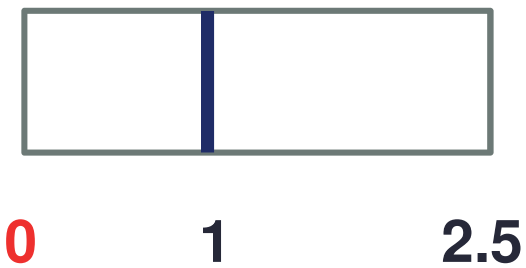
a) Rappresentare con un opportuno grafico la distribuzione di frequenza dei dati
b) Calcolare moda, quartili, media e varianza dei dati
c) Ci sono *outlier* (valori anomali) nei dati?

Valore dei quartili

$$\frac{0+0}{2} = 0$$

$$\frac{1+1}{2} = 1$$

$$\frac{2+3}{2} = 2.5$$



$$0 - 1.5 \times (2.5 - 0) < 0$$

Esercizio 10

In una città è stato rilevato il numero di guasti a mezzi pubblici di superficie in 10 giorni consecutivi

0, 0, 0, 1, 1, 1, 2, 2, 3, 5

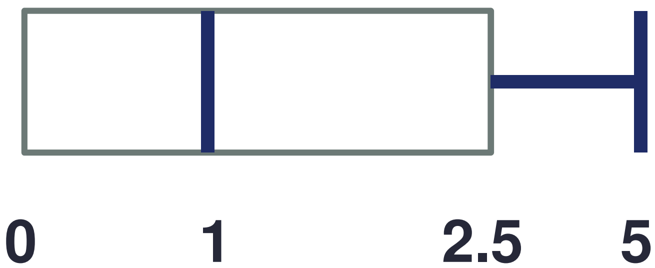
a) Rappresentare con un opportuno grafico la distribuzione di frequenza dei dati
b) Calcolare moda, quartili, media e varianza dei dati
c) Ci sono *outlier* (valori anomali) nei dati? NO

Valore dei quartili

$$\frac{0+0}{2}=0$$

$$\frac{1+1}{2}=1$$

$$\frac{2+3}{2}=2.5$$



$$2.5+1.5\times(2.5-0)=6.25>5$$

Quiz 4

Sia X una variabile gaussiana con media 2 e varianza 1. Se $Y = -2X + 1$ e le due variabili vengono osservate congiuntamente, allora

a) $\boldsymbol{\rho_{xy} = 1}$	**b)** $\boldsymbol{\rho_{xy} = -1}$
c) $\boldsymbol{\sigma_{xy} = 0}$	**d)** $\boldsymbol{R^2 = 0}$

Quiz 4

Sia X una variabile gaussiana con media 2 e varianza 1. Se $Y = -2X + 1$ e le due variabili vengono osservate congiuntamente, allora

a) $\boldsymbol{\rho_{xy} = 1}$	**b)** $\boldsymbol{\rho_{xy} = -1}$
c) $\boldsymbol{\sigma_{xy} = 0}$	**d)** $\boldsymbol{R^2 = 0}$

Quiz 5

In un test sulla media di ipotesi nulla $H_0 : \mu = 1$ contro l'alternativa $\mu < 1$ il p-valore vale 0.90. Allora:

a) Si rifiuta H_0 al livello del 5%	b) Non si può rifiutare H_0 al livello del 5%
c) $\alpha = 0.90$	d) Si rifiuta H_1 al livello del 90%

Quiz 5

In un test sulla media di ipotesi nulla $H_0 : \mu = 1$ contro l'alternativa $\mu < 1$ il p-valore vale 0.90. Allora:

a) Si rifiuta H_0 al livello del 5%	b) Non si può rifiutare H_0 al livello del 5%
c) $\alpha = 0.90$	d) Si rifiuta H_1 al livello del 90%

Domanda 11

Siano A e B due eventi incompatibili con la stessa probabilità, pari a 0.6. Allora

a) $P(A \cap B) = 0.36$	b) $P(A \cup B) = 1$
c) Due eventi del genere non esistono	d) $P(A \mid B) = 1$

Domanda 11

Siano A e B due eventi incompatibili con la stessa probabilità, pari a 0.6. Allora

a) $P(A \cap B) = 0.36$	b) $P(A \cup B) = 1$
c) Due eventi del genere non esistono	d) $P(A\|B) = 1$

$$\boldsymbol{P(A \cup B) = P(A) + P(B) = 0.6 + 0.6 = 1.2 > 1}\text{!!!}$$

Domanda 12

Se l'indice di connessione tra due variabili qualitative osservate congiuntamente vale 0, allora

a) $\rho = 0$	b) $\sigma_{xy} = 0$
c) $R^2 = 0$	d) $\chi^2 = 0$

Domanda 12

Se l'indice di connessione tra due variabili qualitative osservate congiuntamente vale 0, allora

a) $\rho = 0$	b) $\sigma_{xy} = 0$
c) $R^2 = 0$	d) $\chi^2 = 0$

Esercizio 9

In un campione casuale di 200 dipendenti di una grande azienda sono laureati in 70.

a) Stimare la percentuale di laureati nell'azienda.
b) Calcolare un intervallo di confidenza al livello del 99% per la percentuale stimata al punto a)
c) L'unione europea indica come azienda innovatrice una con almeno il 40% di laureati: sottoporre a verifica l'ipotesi che l'azienda in questione sia innovatrice, oppure no, al livello del 5%.

Esercizio 9

In un campione casuale di 200 dipendenti di una grande azienda sono laureati in 70.

a) Stimare la percentuale di laureati nell'azienda.
b) Calcolare un intervallo di confidenza al livello del 99% per la percentuale stimata al punto a)
c) L'unione europea indica come azienda innovatrice una con almeno il 40% di laureati: sottoporre a verifica l'ipotesi che l'azienda in questione sia innovatrice, oppure no, al livello del 5%.

$$\hat{p}_n = \frac{70}{200} = 0.35 = 35\%$$

Esercizio 9

In un campione casuale di 200 dipendenti di una grande azienda sono laureati in 70.

a) Stimare la percentuale di laureati nell'azienda.
b) Calcolare un intervallo di confidenza al livello del 99% per la percentuale stimata al punto a)
c) L'unione europea indica come azienda innovatrice una con almeno il 40% di laureati: sottoporre a verifica l'ipotesi che l'azienda in questione sia innovatrice, oppure no, al livello del 5%.

$$\hat{p}_n = 35\%, \qquad n\hat{p}_n = 70 > 5, n(1-\hat{p}_n) = 130 > 5, \alpha = 0.01$$

$$\hat{p}_n \mp z_{0.995}\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}} \Rightarrow 0.35 \mp \mathbf{2.575}\sqrt{\frac{0.35 \times 0.65}{200}}$$

Esercizio 9

In un campione casuale di 200 dipendenti di una grande azienda sono laureati in 70.

a) Stimare la percentuale di laureati nell'azienda.
b) Calcolare un intervallo di confidenza al livello del 99% per la percentuale stimata al punto a)
c) L'unione europea indica come azienda innovatrice una con almeno il 40% di laureati: sottoporre a verifica l'ipotesi che l'azienda in questione sia innovatrice, oppure no, al livello del 5%.

$$\hat{p}_n = 35\%, \qquad n\hat{p}_n = 70 > 5, n(1-\hat{p}_n) = 130 > 5, \alpha = 0.01$$

$$\hat{p}_n \mp z_{0.995}\sqrt{\frac{\hat{p}_n(1-\hat{p}_n)}{n}} \Rightarrow \mathbf{(0.26, 0.44)}$$

Esercizio 9

In un campione casuale di 200 dipendenti di una grande azienda sono laureati in 70.

a) Stimare la percentuale di laureati nell'azienda.
b) Calcolare un intervallo di confidenza al livello del 99% per la percentuale stimata al punto a)
c) L'unione europea indica come azienda innovatrice una con almeno il 40% di laureati: sottoporre a verifica l'ipotesi che l'azienda in questione sia innovatrice, oppure no, al livello del 5%.

$$\hat{p}_n = 35\%, \qquad n\hat{p}_n = 70 > 5, n(1-\hat{p}_n) = 130 > 5, \alpha = 0.05$$

$$\boldsymbol{H_0 : \quad p = 0.40 \quad vs \quad H_1 : \quad p < 0.40}$$

Esercizio 9

In un campione casuale di 200 dipendenti di una grande azienda sono laureati in 70.

a) Stimare la percentuale di laureati nell'azienda.
b) Calcolare un intervallo di confidenza al livello del 99% per la percentuale stimata al punto a)
c) L'unione europea indica come azienda innovatrice una con almeno il 40% di laureati: sottoporre a verifica l'ipotesi che l'azienda in questione sia innovatrice, oppure no, al livello del 5%.

$\widehat{\boldsymbol{p}}_{\boldsymbol{n}} = \mathbf{35}\%$, $\qquad n\hat{p}_n = 70 > 5, n(1-\hat{p}_n) = 130 > 5, \alpha = 0.05$

$$\boldsymbol{H_0 : \ p = 0.40 \ \ vs \ \ H_1 : \ p < 0.40}$$

$$\frac{\widehat{\boldsymbol{p}}_{\boldsymbol{n}} - \boldsymbol{p}_0}{\sqrt{\boldsymbol{p}_0(1-\boldsymbol{p}_0)/\boldsymbol{n})}} < -\boldsymbol{z}_{1-\boldsymbol{\alpha}} = -\mathbf{1.645}$$

Esercizio 9

In un campione casuale di 200 dipendenti di una grande azienda sono laureati in 70.

a) Stimare la percentuale di laureati nell'azienda.
b) Calcolare un intervallo di confidenza al livello del 99% per la percentuale stimata al punto a)
c) L'unione europea indica come azienda innovatrice una con almeno il 40% di laureati: sottoporre a verifica l'ipotesi che l'azienda in questione sia innovatrice, oppure no, al livello del 5%.

$\hat{\boldsymbol{p}}_{\boldsymbol{n}} = \mathbf{35\%},\qquad n\hat{p}_n = 70 > 5, n(1-\hat{p}_n) = 130 > 5, \alpha = 0.05$

$$\boldsymbol{H_0}: \quad \boldsymbol{p = 0.40} \quad \boldsymbol{vs} \quad \boldsymbol{H_1}: \quad \boldsymbol{p < 0.40}$$

$$\frac{\mathbf{0.35 - 0.40}}{\sqrt{\mathbf{0.40 \times 0.60/200}}} = \mathbf{-1.44 > -1.645}\ !$$

NON SI RIFIUTA!

Esercizio 9

In un campione casuale di 200 dipendenti di una grande azienda sono laureati in 70.

a) Stimare la percentuale di laureati nell'azienda.
b) Calcolare un intervallo di confidenza al livello del 99% per la percentuale stimata al punto a)
c) L'unione europea indica come azienda innovatrice una con almeno il 40% di laureati: sottoporre a verifica l'ipotesi che l'azienda in questione sia innovatrice, oppure no, al livello del 5%.

$$\hat{\boldsymbol{p}}_{\boldsymbol{n}} = \mathbf{35\%}, \qquad n\hat{p}_n = 70 > 5, n(1-\hat{p}_n) = 130 > 5, \alpha = 0.05$$

$$\boldsymbol{H_0} \ : \ \boldsymbol{p = 0.40} \ \ \boldsymbol{vs} \ \ \boldsymbol{H_1} : \ \boldsymbol{p < 0.40}$$

$$\frac{\mathbf{0.35 - 0.40}}{\sqrt{\mathbf{0.40 \times 0.60/200}}} = \mathbf{-1.44}$$

$\boldsymbol{p} -$ **valore?**

Esercizio 9

In un campione casuale di 200 dipendenti di una grande azienda sono laureati in 70.

a) Stimare la percentuale di laureati nell'azienda.
b) Calcolare un intervallo di confidenza al livello del 99% per la percentuale stimata al punto a)
c) L'unione europea indica come azienda innovatrice una con almeno il 40% di laureati: sottoporre a verifica l'ipotesi che l'azienda in questione sia innovatrice, oppure no, al livello del 5%.

$$\hat{\boldsymbol{p}}_{\boldsymbol{n}} = \mathbf{35\%}, \qquad n\hat{p}_n = 70 > 5, n(1 - \hat{p}_n) = 130 > 5, \alpha = 0.05$$

$$\boldsymbol{H_0} : \quad \boldsymbol{p = 0.40} \quad \boldsymbol{vs} \quad \boldsymbol{H_1} : \quad \boldsymbol{p < 0.40}$$

$$\frac{\mathbf{0.35 - 0.40}}{\sqrt{\mathbf{0.40 \times 0.60/200}}} = \mathbf{-1.44} \Rightarrow \approx \boldsymbol{P(Z < -1.44) = 0.07493}$$

Domanda 4

Per quale delle seguenti tabelle a doppia entrata l’associazione è massima?

(1)	A	B
S	100	100
N	100	100

(2)	A	B
S	100	0
N	0	100

(3)	A	B
S	100	100
N	0	100

(4)	A	B
S	1000	1000
N	1000	1000

Domanda 4

Per quale delle seguenti tabelle a doppia entrata l'associazione è massima?

(1)	A	B
S	100	100
N	100	100

(2)	A	B
S	100	0
N	0	100

(3)	A	B
S	100	100
N	0	100

(4)	A	B
S	1000	1000
N	1000	1000

Domanda 5

In un modello di regressione lineare $y_i = \hat{a} + \hat{b}x_i + \varepsilon_i$ la regressione è statisticamente significativa. Allora

a) $\hat{b} > 0$	b) $\hat{b} = 0$
c) p-valore < 0.05	d) $\rho = 0$

Domanda 5

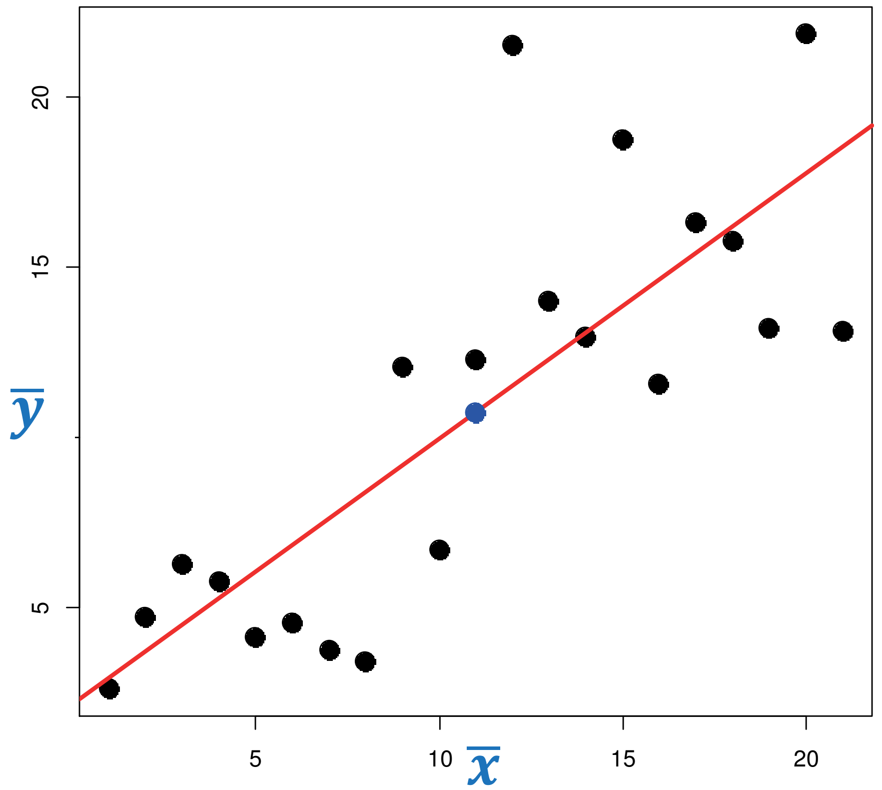
In un modello di regressione lineare $y_i = \hat{a} + \hat{b}x_i + \varepsilon_i$ la regressione è statisticamente significativa. Allora

a) $\hat{b} > 0$	b) $\hat{b} = 0$
c) p-valore < 0.05	d) $\rho = 0$

Domanda 6

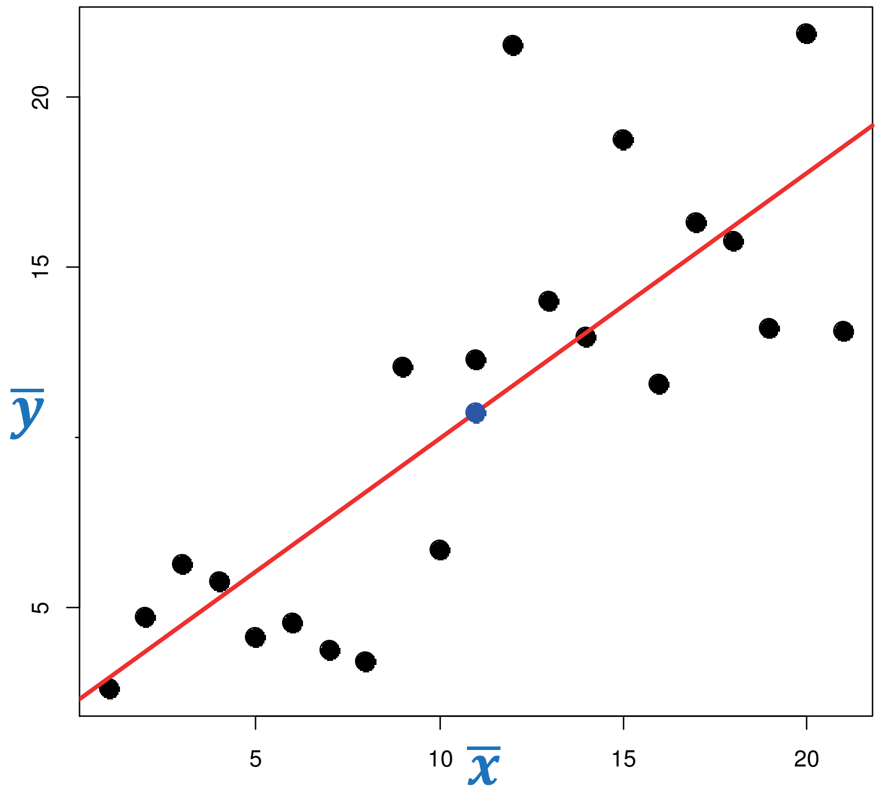
Con riferimento al grafico, relativo alla stima di un modello di regressione lineare, per quale dei seguenti valori di X la previsione è più attendibile?

a) $X = \bar{x}$	b) $X = 0$
c) $X = 20$	d) $X = 30$



Domanda 6

Con riferimento al grafico, relativo alla stima di un modello di regressione lineare, per quale dei seguenti valori di X la previsione è più attendibile?



a) $X = \bar{x}$	b) $X = 0$
c) $X = 20$	d) $X = 30$

$$\hat{y}_0 \mp t(n-2)_{1-\frac{\alpha}{2}} \times \sqrt{s^2\left[1 + n^{-1} + \frac{(\boldsymbol{x_0} - \overline{\boldsymbol{x}})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right]}$$