Statistica I modulo - CdS SIAFA *(2023/24)*
Esame 13 febbraio 2024 - Soluzioni

1. Si consideri per una specifica settimana la seguente distribuzione di frequenze del numero di interventi giornalieri (x_i) operati da 480 tecnici delle caldaie:

x_i	n_i
1	2
2	44
3	78
4	190
5	166
TOT	480

(a) Calcolare la media aritmetica del numero di interventi.

Per calcolare la media aritmetica si consideri la terza colonna ($x_i n_i$) della tabella seguente:

x_i	n_i	$x_i n_i$	N_i	p_i
1	2	2	2	0,42
2	44	88	46	9,17
3	78	234	124	16,25
4	190	760	314	39,58
5	166	830	480	34,58
TOT	480	1914		100,000

Dunque

$$\bar{x} = \frac{1914}{480} = 3,9875$$

(b) Individuare la mediana e la moda.

Per individuare la mediana si considerino le frequenze cumulate N_i nella quarta colonna della tabella sopra. Poiché n è pari, la mediana è la media aritemtica dei valori in posizione $n/2 = 240$ e $n/2 + 1 = 241$. Dalle frequenze cumulate, si osserva che il valore in posizione 240 è 4, il valore in posizione 241 è sempre 4, dunque la mediana è pari a 4.
La moda (il valore a cui è associata la massima frequenza) è pari a 4.

(c) Calcolare le frequenze percentuali e interpretarle.

Le frequenze percentuali (p_i) sono riportate nell'ultima colonna della tabella sopra. Interpretazione: lo 0,42% dei tecnici ha operato 1 intervento; il 9,17% dei tecnici ha operato 2 interventi; ecc.

2. Per 6 quartieri si registrano il reddito medio dei residenti (X) e il numero di guasti nella stagione invernale (Y):

Reddito medio (X)	Guasti (Y)
1240	75
1340	77
1420	92
1610	41
1970	12
1525	63

(a) Rappresentare il diagramma di dispersione.

Il diagramma di dispersione è riportato nella Figura 1.

(b) Identificare i coefficienti della regressione (con Y variabile dipendente) e rappresentare la retta di regressione.

Per individuare la retta di regressione si svolgono i calcoli sotto riportati dopo aver individuato $\bar{x} = 1517,5$ *e* $\bar{y} = 60$.

x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	Prodotto	$(x_i - \bar{x})^2$
1240	75	-277,5	15	-4162,5	77006,25
1340	77	-177,5	17	-3017,5	31506,25
1420	92	-97,5	32	-3120	9506,25
1610	41	92,5	-19	-1757,5	8556,25
1970	12	452,5	-48	-21720	204756,25
1525	63	7,5	3	22,5	56,25
TOT				-33755	331387,5

Dunque

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{6}[(x_i - \bar{x})(y_i - \bar{y})]}{\sum_{i=1}^{6}(x_i - \bar{x})^2} = \frac{-33755}{331387,5} = -0,10$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 60 - (-0,10) \cdot 1517,5 = 214,57$$

L'equazione della retta di regressione è dunque

$$\hat{Y} = 214,57 - 0,10X$$

La retta di regressione è riportata nella Figura 1

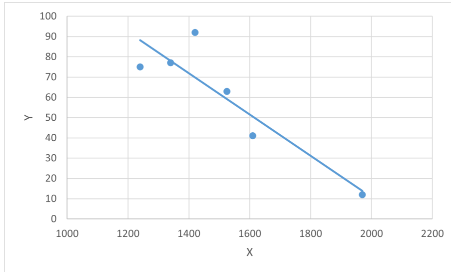

Figura 1: Diagramma di dispersione e retta di regressione.

(c) Le due variabili sono concordanti o discordanti? Qual è la variazione attesa del numero di guasti se il reddito medio (X) si riduce di 50?

Poiché $\hat{\beta}_1 < 0$*, le due variabili sono discordanti.*
Se il reddito medio (X) si riduce di 50, la variazione attesa del numero di guasti (Y) è data da $-50 \cdot \hat{\beta}_1 = 5$*, quindi se il reddito medio si riduce di 50, ci si aspetta un aumento pari a 5 del numero di guasti.*

(d) Prevedere il numero di guasti se $X = 1500$.

Con $X = 1500$*, il numero previsto di guasti è*

$$\hat{Y}(1500) = 214,57 - 0,10 \cdot 1500 = 64,57$$

3. La probabilità che una caldaia risulti in difetto di revisione è pari a 0.25. Considerando 5 controlli:

(a) qual è la probabilità che nessuna sia in difetto di revisione?

Si definisce la variabile casuale X che descrive il numero di caldaie in difetto di revisione su 5. X è una VC binomiale con parametri $\pi = 0.25$ e $n = 5$, quindi $X \sim Bin(0.25, 5)$.
La probabilità che nessuna caldaia sia in difetto ($X = 0$) è data da

$$P(X = 0) = \frac{5!}{0!(5-0)!}0.25^0(1-0.25)^{5-0} = 0,237$$

(b) qual è la probabilità che almeno 1 sia in difetto di revisione?

La probabilità che almeno 1 sia in difetto di revisione ($X \geq 1$) è data da

$$P(X \geq 1) = 1 - P(X = 0) = 1 - 0,237 = 0,763$$

(c) qual è la probabilità che meno di 2 siano in difetto di revisione?

La probabilità che meno di 2 siano in difetto di revisione ($X < 2$) è data da

$$P(X < 2) = P(X = 0)+P(X = 1) = 0,237+\frac{5!}{1!(5-1)!}0.25^1(1-0.25)^{5-1} = 0,237+0,396 = 0,633$$