ES. ~~[illegible]~~ REGR. MULTIPLA INVENTATO - RISOLUZIONE

→ I dati a disposizione sono:

α	x_1	x_2	x_3	y
1	9	10	2	8
1	12	16	5	6
1	9	8	4	14
1	8	6	7	12
1	6	12	7	15
1	4	14	8	17

$\overline{x_1} = 8$; $\overline{x_2} = 11$; $\overline{x_3} = 5.5$; $\overline{y} = 12$; $n=6$; $K=3$

$$(X' \cdot X)^{-1} = \begin{bmatrix} 8.586 & -0.524 & -0.0817 & -0.605 \\ -0.524 & 0.0465 & -0.00512 & 0.0379 \\ -0.0817 & -0.00512 & 0.0153 & -0.0082 \\ -0.605 & 0.0379 & -0.0082 & 0.0712 \end{bmatrix}$$

→ Le domande sono:

1) Calcolare l'equazione della retta di previsione
2) Applicare il ~~[illegible]~~ t-test per il coefficiente angolare di X_2
3) Sapendo che $SQREG_c = 74.703$, applicare il test F per il modello completo e prendere una decisione per $\alpha = 0.05$
4) Confrontare il modello completo con quello ridotto con solo X_1, sapendo che $SQTOT = 90$ e che $SQREG_r = 71.158$. Applicare il test F e prendere una decisione per $\alpha = 0.05$.

SVOLGIMENTO:

1) $\hat{y} = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 = BX + \varepsilon$ → $\boxed{B = (X' \cdot X)^{-1} \cdot (X' \cdot Y)}$

$$\Rightarrow (X' \cdot Y) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 9 & 12 & 9 & 8 & 6 & 4 \\ 10 & 16 & 8 & 6 & 12 & 14 \\ 2 & 5 & 4 & 7 & 7 & 8 \end{bmatrix} \cdot \begin{bmatrix} 8 \\ 6 \\ 14 \\ 12 \\ 15 \\ 17 \end{bmatrix} = \begin{bmatrix} 72 \\ 524 \\ 778 \\ 427 \end{bmatrix}$$

$$\Rightarrow (X' \cdot X)^{-1} \cdot (X' \cdot Y) = \begin{bmatrix} 21.672 \\ -1.173 \\ -0.202 \\ 0.351 \end{bmatrix} (= B)$$

→ Quindi l'equazione della retta è: $\hat{y} = 21.672 - 1.173\, x_1 - 0.202\, x_2 + 0.351\, x_3$

2) $\boxed{t = \dfrac{\beta_i - 0}{\sqrt{Se^2 \cdot (X' \cdot X)^{-1}_{ii}}}}$ dove $\boxed{Se^2 = \dfrac{e' \cdot e}{n-k-1}}$ e $\boxed{e' \cdot e} = \boxed{SQERR_c} = \boxed{\sum (y - \hat{y})^2 = (y-\hat{y})' \cdot (y - \hat{y})}$

→ La matrice degli $\hat{y}$ si può trovare o calcolando con la retta trovata al punto 1 il valore atteso di y per ogni X o ~~[illegible]~~ ~~[illegible]~~ facendo un prodotto fra matrici: $\hat{y} = X \cdot B$.

(Utilizzo il secondo metodo).

$$\Rightarrow X = \begin{bmatrix} 1 & 9 & 10 & 2 \\ 1 & 12 & 16 & 5 \\ 1 & 9 & 8 & 4 \\ 1 & 8 & 6 & 7 \\ 1 & 6 & 12 & 7 \\ 1 & 4 & 14 & 8 \end{bmatrix} \qquad B = \begin{bmatrix} 21.672 \\ -1.173 \\ -0.202 \\ 0.351 \end{bmatrix} \Rightarrow \underbrace{X \cdot B}_{=\hat{y}} = \begin{bmatrix} 9.800 \\ 6.124 \\ 10.905 \\ 13.534 \\ 14.671 \\ 16.965 \end{bmatrix}$$

$$\Rightarrow (y-\hat{y})' \cdot (y-\hat{y}) = \left(\begin{bmatrix} 8 \\ 6 \\ 14 \\ 12 \\ 15 \\ 17 \end{bmatrix} - \begin{bmatrix} 9.800 \\ 6.124 \\ 10.905 \\ 13.534 \\ 14.671 \\ 16.965 \end{bmatrix} \right)' \cdot \left(\begin{bmatrix} 8 \\ 6 \\ 14 \\ 12 \\ 15 \\ 17 \end{bmatrix} - \begin{bmatrix} 9.800 \\ 6.124 \\ 10.905 \\ 13.534 \\ 14.671 \\ 16.965 \end{bmatrix} \right) = \left(\begin{bmatrix} -1.800 \\ -0.124 \\ 3.095 \\ -1.534 \\ 0.329 \\ 0.035 \end{bmatrix} \right)' \cdot \left(\begin{bmatrix} -1.800 \\ -0.124 \\ 3.095 \\ -1.534 \\ 0.329 \\ 0.035 \end{bmatrix} \right) =$$

$$= \begin{bmatrix} -1.800 & -0.124 & 3.095 & -1.534 & 0.329 & 0.035 \end{bmatrix} \cdot \begin{bmatrix} -1.800 \\ -0.124 \\ 3.095 \\ -1.534 \\ 0.329 \\ 0.035 \end{bmatrix} = 15.297$$

$$\downarrow$$

$$e' \cdot e = 15.297$$

$$\Rightarrow S_e^2 = \frac{15.297}{6-3-1} = 7.649$$

$$\rightarrow t = \frac{-0.202}{\sqrt{7.649 \cdot (X' \cdot X)^{-1}_{33}}} \rightarrow \boxed{(X' \cdot X)^{-1}_{33}}$$ è la cella corrispondente all'incrocio fra terza riga e terza colonna. (della matrice fornita nei dati).

$$\Rightarrow t = \frac{-0.202}{\sqrt{7.649 \cdot 0.0153}} = -0.59$$

3) $$\boxed{F_{(k;\, n-k-1)} = \frac{SQ_{REGc}}{SQ_{ERRc}} \cdot \frac{n-k-1}{k}}$$ → ricordo che $SQ_{ERRc} = e' \cdot e = 15.297$

$$\Rightarrow F_{(3;2)} = \frac{74.703}{15.297} \cdot \frac{2}{3} = 3.256$$

→ Dalla tavola D a pag. 529 del libro trovo che per $\alpha = 0.05$, per $(3;2)$ gdl, la statistica F critica ha un valore di 19.16. $3{,}256 < 19.16$ quindi non rifiuto H_0

4) $$\boxed{F_{(q;\, n-k-1)} = \frac{R^2_c - R^2_r}{1 - R^2_c} \cdot \frac{n-k-1}{q}}$$ dove $\boxed{R^2_c = \frac{SQ_{REGc}}{SQ_{TOT}}}$ e $\boxed{R^2_r = \frac{SQ_{REGr}}{SQ_{TOT}}}$; $q = 2$

↓ perché il modello completo ha due variabili in più.

$$\rightarrow R^2_c = \frac{74.703}{90} = 0.83 \; ; \; R^2_r = \frac{71.158}{90} = 0.79$$

$$\Rightarrow F_{(2;2)} = \frac{0.83 - 0.79}{1 - 0.83} \cdot \frac{2}{2} = 0.232$$

→ Dalla tavola D a pag. 529 del libro trovo che per $\alpha = 0.05$, per $(2,2)$ gdl, la statistica F critica ha un valore di 19.00. $0{,}232 < 19.00$ quindi non rifiuto H_0.