

Esame di Probabilità, distribuzioni e regressione multipla / Statistica II
Prova parziale di Regressione

02.02.06

COGNOME _____ **NOME** _____ **Matr.** _____

ATTENZIONE: Approssimare i calcoli alla quarta cifra decimale.

Un laboratorio di analisi chimica e fisica delle caratteristiche delle acque oligominerali di nove marche commerciali, dopo aver raccolto alcuni dati, decide di analizzarli.

In particolare vuole verificare l'ipotesi biologica che le acque che sgorgano da fonti ad altitudini maggiori (rilevata mediante la variabile: temperatura dell'acqua alla fonte) siano le più pure, cioè abbiano un contenuto di residuo fisso minore. Si ricorda che il residuo fisso è una quantità che si compone di diversi fattori, fra cui il sodio, ma in quantità diverse a seconda del tipo di acqua.

Le variabili rilevate sono le seguenti:

- X_1 : Contenuto di residuo fisso (mg/l);
- X_2 : Temperatura dell'acqua alla fonte (°C);
- X_3 : Contenuto di sodio (mg/l).

X_1	X_2	X_3
133	9.9	2
114	11.3	1.3
76.3	5.4	1.8
215.5	11	4.4
605	12.5	5
98	7	1.5
402	13	5.2
240	9.2	6.2
189	4.9	5.1
2072.8	84.2	32.5

$$\sum_{i=1}^9 x_{1i}^2 = 713500.94 \quad \sum_{i=1}^9 x_{2i}^2 = 858.76 \quad \sum_{i=1}^9 x_{3i}^2 = 147.03$$

$$\sum_{i=1}^9 x_{1i}x_{2i} = 21996.02 \quad \sum_{i=1}^9 x_{1i}x_{3i} = 9214.04 \quad \sum_{i=1}^9 x_{2i}x_{3i} = 315.24$$

Totali

- a) Si determinino i parametri delle due rette interpolanti che spiegano il contenuto di residuo fisso in funzione, rispettivamente, della temperatura e del contenuto di sodio. Se ne fornisca la relativa interpretazione;
- b) Si determinino i parametri del piano interpolante $\hat{X}_1 = b + \alpha_{12.3}X_2 + \alpha_{13.2}X_3$ e se ne fornisca la relativa interpretazione. Si commenti adeguatamente, inoltre, la differenza fra i coefficienti di regressione grezzi e parziali;
- c) Si valuti la bontà d'adattamento del piano determinato al punto b) e il grado di miglioramento rispetto alle rette del punto a) sia in termini di varianza spiegata che varianza residua, commentando i risultati e indicando quale modello scegliereste per spiegare il fenomeno.
- d) Si calcoli il coefficiente di correlazione parziale $r_{12.3}$, commentandolo adeguatamente ed evidenziandone le differenze rispetto i corrispettivi coefficienti di correlazione grezzi.
- e) Si consideri ora il carattere X_4 ="grado di Ph", per il quale si ha

$$r_{14} = -0.287 \quad r_{24} = -0.164 \quad r_{34} = -0.565$$

Si calcoli il miglioramento, in termini di varianza residua, che si ottiene passando dal piano al punto b) all'iperpiano che include anche X_4 fra le variabili esplicative.