

1. Un transportista fa diversos trajectes entre tres poblacions A , B i C . En quatre dies consecutius ha fet els recorreguts que reflexa la taula següent:

| trajecte | km |
|---|------|
| $A \rightarrow B \rightarrow A \rightarrow C$ | 533 |
| $C \rightarrow A \rightarrow C \rightarrow B$ | 583 |
| $B \rightarrow C \rightarrow A \rightarrow C \rightarrow A \rightarrow B \rightarrow A$ | 1111 |
| $A \rightarrow B \rightarrow A \rightarrow C \rightarrow A \rightarrow B \rightarrow A$ | 1069 |

on el quilometratge és, per diverses causes, aproximat.

- (a) Proposeu un model lineal, amb la matriu de disseny i les hipòtesis que calguin, per tal d'estimar les distàncies quilomètriques entre les tres poblacions. Amb les dades proporcionades, és possible estimar les distàncies entre les tres poblacions? Quines són les distàncies o funcions paramètriques estimables (fpe) en aquest model?
- (b) Podeu estimar el quilometratge del trajecte $M_{BC} \Rightarrow B \Rightarrow A \Rightarrow C \Rightarrow M_{AC}$, on M_{IJ} és el punt mig entre dues poblacions? És una bona estimació? Quin és l'error d'aquesta estimació?
- (c) El transportista discuteix amb un amic que afirma que el doble de la distància entre A i B és equivalent a la distància del trajecte $A \Rightarrow C \Rightarrow B$. Podem aclarir en termes estadístics la seva discussió?
2. Es planteja un experiment per a avaluar l'efecte de diferents quantitats de fertilitzant per a un cultiu determinat amb els següents resultats:

| Quantitat de fertilitzant | Collita (kg) | | |
|---------------------------|--------------|------|------|
| 2 | 11.6 | 10.5 | 9.3 |
| 3 | 13.7 | 13.7 | 13.5 |
| 4 | 15.4 | 14.3 | 14.3 |
| 5 | 15.8 | 14.4 | 13.4 |

- (a) Si tractem el fertilitzant com un factor amb 4 nivells el model és:

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad i = 1, 2, 3, 4; \quad j = 1, 2, 3$$

La seva estimació proporciona la següent taula ANOVA:

| Font de variació | Suma de quadrats | Graus de llibertat | MQ | F | $Pr(F)$ |
|------------------|------------------|--------------------|--------|--------|---------|
| Fertilitzant | 34.576 | 3 | 11.525 | 14.437 | 0.0014 |
| Error | 6.387 | 8 | 0.798 | | |

- i. Si $\mathbf{y} = (11.6, 10.5, 9.3, 13.7, \dots, 14.4, 13.4)$, escriviu la matriu de disseny \mathbf{X} del model. Quin és el seu rang?
- ii. Expliqueu el significat de la taula ANOVA i del número 0.798.
- (b) Considereu ara que les columnes de la taula de dades es refereixen a 3 varietats diferents: la columna 1 correspon a una estàndard i les altres dues són noves. La suma de quadrats de les varietats és 4.56. Indiqueu:

- i. Quin és ara el model? Hi ha diferències significatives entre els fertilitzants i entre les varietats?
 - ii. Quin fertilitzant, tractat com un factor a 4 nivells, dona resultats significativament millors que els altres?
3. Si amb les dades del problema anterior considerem que la quantitat de fertilitzant és una variable quantitativa amb tres respostes per a cada valor (no hi ha varietats), es demana:
- (a) Reformuleu el model. Escriviu la matriu de disseny d'aquest model.
 - (b) Estimeu els principals paràmetres d'aquest model i estudieu la significació de la regressió.
 - (c) Calculeu el coeficient de determinació i comenteu el seu valor.
 - (d) Feu una previsió amb interval de confiança per a la collita quan la quantitat de fertilitzant és 6. Alguna observació?
 - (e) Amb aquestes dades ens podem plantejar el contrast sobre la linealitat de la regressió. Com es fa? (No cal fer els càlculs)

4. En funció de les variables

MaxAT, MinAT: Màxima i mínima temperatura de l'aire.

MaxST, MinST: Màxima i mínima temperatura del sòl.

MAT, MST: Temperatures mitjanes de l'aire i del sòl.

MaxH, MinH: Màxima i mínima humitats relatives.

MH: Humitat mitjana durant el dia.

V: Força del vent.

es disposa de la següent informació sobre l'evaporació:

| Font | SS | df | MS | F |
|-----------|---------|----|--------|-------|
| Regressió | 8159.83 | 10 | 815.98 | 19.27 |
| Error | 1482.27 | 35 | 42.35 | |
| Total | 9642.11 | 45 | | |

| Variable | $\hat{\beta}$ | t | VIF |
|----------|---------------|-------|-------|
| MaxAT | 0.5011 | 0.88 | 8.83 |
| MinAT | 0.3041 | 0.39 | 8.89 |
| MAT | 0.0930 | 0.42 | 21.21 |
| MaxST | 2.2320 | 2.22 | 39.29 |
| MinST | 0.2050 | 0.19 | 14.08 |
| MST | -0.7400 | -2.12 | 52.36 |
| MaxH | 1.1100 | 0.98 | 1.98 |
| MinH | 0.7510 | 1.54 | 25.38 |
| MH | -0.5560 | -3.44 | 24.12 |
| V | 0.0089 | 0.97 | 1.98 |

Es demana:

- (a) Calculeu R^2 i verifiqueu la significació de la regressió amb $\alpha = 0.01$.
- (b) Calculeu R_j^2 per a cada variable explicativa
- (c) Hi ha variables implicades en una possible multicolinealitat?
- (d) Què penseu dels coeficients negatius? Discutiu els resultats obtinguts.