



Models Lineals

PROVA PARCIAL A

9 de novembre de 2015

Grau d'Estadística

S'ha d'entregar un document PDF o HTML amb les respostes i un fitxer R ben ordenat per apartats i amb tots els comentaris (amb el símbol #) que considereu oportuns. En el nom dels fitxers no poseu accents, només símbols ASCII del 0 al 127 (sense ñ,ç,...).

Problema 1

Amb el model lineal normal que es deriva del següent disseny experimental amb 3 rèpliques per condició

$$\begin{aligned}\alpha - \beta - \gamma &\Rightarrow 0.95, 1.01, 0.98 \\ 2\alpha + \beta &\Rightarrow 15.03, 15.01, 14.99 \\ 3\alpha - \gamma &\Rightarrow 19.88, 19.97, 20.01 \\ \alpha + 2\beta + \gamma &\Rightarrow 14.02, 14.10, 13.98\end{aligned}$$

contesteu les següents qüestions:

- (a) Quina condició ha de verificar una funció paramètrica per a que sigui estimable en aquest model?
- (b) Indiqueu si les funcions paramètriques següents són estimables i calculeu l'estimador MQ quan sigui possible:

$$(i) 3\alpha + 3\beta + \gamma \qquad (ii) 3\alpha - 2\gamma$$

- (c) Calculeu l'estimació de la covariància entre els estimadors lineals òptims de $\alpha - \beta - \gamma$ i $3\alpha - \gamma$ i la variància de l'estimador lineal òptim de $3\alpha + 3\beta + \gamma$.
- (d) Feu el contrast de les hipòtesis $H_0^{(1)} : 3\alpha + 3\beta + \gamma = 29$ i $H_0^{(2)} : \alpha + 2\beta + \gamma = 14, \quad 2\alpha + \beta = 15$.

Problema 2

La base de dades **Robey** del paquet **car** presenta les dades de la taxa de reproducció (número mitjà de fills per dona) i del percentatge de dones que usen mètodes contraceptius en diversos països agrupats en grans zones geogràfiques (Àsia, Llatinoamèrica, Pròxim Orient i Àfrica).

Estem interessats en caracteritzar la tendència entre un augment de la contracepció i la disminució de la taxa reproductiva i verificar si les tendències són semblants en les diferents àrees geogràfiques.

- (a) Feu un resum numèric de les dues variables implicades en la regressió simple. Feu també un gràfic que relacioni les dues variables. Té sentit calcular la correlació entre ambdues variables? Quant val aquesta correlació?
- (b) Obteniu l'estimació dels paràmetres del model $(\beta_0, \beta_1, \sigma^2)$ i calculeu el coeficient de determinació. És significativa la regressió? Què significa això últim?
- (c) Investigueu la validesa del model lineal. Feu un test de normalitat. Sabeu fer el test¹ de Jarque-Bera?

Hi ha algun residu que podríem qualificar de atípic²?

¹El test de Jarque-Bera fa servir un estadístic que utilitza els coeficients d'asimetria i de kurtosi.

²Feu una definició del concepte *dada atípica* (outlier)

- (d) Doneu els intervals de confiança al 99% de β_0 , de β_1 i de σ^2 . Com interpretem β_0 , β_1 i σ^2 ? En quines unitats estan aquests coeficients?
- (e) Contrasteu la hipòtesi $H_0 : \beta_1 = -0.05$ amb dos tests diferents. Utilitzeu un nivell de significació $\alpha = 0.05$.
- (f) Quina seria la taxa de reproducció d'un país com Senegal si augmentés el percentatge de dones que usen mètodes contraceptius al 30%? Feu una predicció amb IC al 90%. Té sentit aquesta predicció?
- (g) Compareu les rectes de regressió dels països d'Àsia i Àfrica. Feu un test de paral·lelisme i un altre, si cal, de coincidència. Penseu que és necessari comprovar l'homocedasticitat de les variàncies dels residus de les dues rectes.