

1. Enuncieu el Teorema de Gauss-Markov. Expliqueu breument les condicions en que es verifica i les seves conseqüències per a la Teoria del Models Lineals.

2. Considereu el model lineal:

$$y_1 = \beta_1 + \beta_2 + \epsilon_1$$

$$y_2 = \beta_1 + \beta_3 + \epsilon_2$$

$$y_3 = \beta_1 + \beta_2 + \epsilon_3$$

- (a) Proveu que $\lambda_1\beta_1 + \lambda_2\beta_2 + \lambda_3\beta_3$ és estimable si i només si $\lambda_1 = \lambda_2 + \lambda_3$.
- (b) Quina és la covariància entre els estimadors lineals MQ de les funcions paramètriques $\beta_1 - \beta_2$ i $\beta_1 - \beta_3$, si aquestes són estimables?
- (c) Quina és la dimensió de l'espai paramètric? Quines son les conseqüències pràctiques de la teva resposta?
3. La variable y representa, en milers, el nombre d'ases (*Equus asinus*) en Espanya i la x el tant per cent del pressupost de l'Estat dedicat a educació en diferents anys.

any	1920	1925	1930	1935	1940	1945	1950
y	1006	1162	1479	805	795	747	732
x	5.5	4.8	7.8	8.2	8.6	9.7	9.6
any	1955	1960	1965	1970	1975	1980	1985
y	683	686	493	476	386	368	
x	8.9	11.4	10,6	12.7	11.5	11.4	13

Es demana:

- (a) Construïu una recta de regressió per estimar el nombre d'ases com funció del pressupost i interpreteu els resultats. Podeu fer dues interpretacions. Una on no tingueu en compte el significat de les variables i una altra on sí el tingueu.
- (b) Calculeu intervals de confiança al 95% per als paràmetres.
- (c) Calculeu la variància residual, el coeficient de correlació i el coeficient de determinació. Expliqueu breument el significat de cadascun.
- (d) L'any 1985 el ministeri d'educació va rebre un 13% del pressupost de l'estat. Quin és el nombre d'ases previst per l'any 1985? Construïu un interval de confiança al 90% per a la predicció obtinguda. Hi ha algun problema addicional per fer aquesta previsió?
4. Un experimentador estudia com afecta l'exposició a una substància determinada (ricí) a l'expressió dels gens en el ratolí. Per a això realitza un experiment on prova dues menes de ricí, A i B, i un grup control (sense ricí) sobre 6 microxips de DNA (que a efectes pràctics son com "parcelles", on es mesura el grau d'expressió gènica). Les dades *resultants* (no sabem res de com s'ha dissenyat l'experiment) son les següents:

	Ricí A	Ricí B	Control
xip 1	110	105	95
xip 2	90	95	83
xip 3	105	79	82
xip 4	85	85	73
xip 5	83	80	74
xip 6	93	102	82

Análisis de varianza de un factor

ANÁLISIS DE VARIANZA

<i>Origen de las variaciones</i>	<i>Suma de cuadrados</i>	<i>g.d.l.</i>	<i>Promedio de los cuadrados</i>	<i>F</i>	<i>Probabilidad</i>	<i>Valor crítico para F</i>
Entre grupos	521.4444	2	260.7222	2.49680783	0.11587757	3.68231667
Dentro de los grupos	1566.333	15	104.4222			
Total	2087.778	17				

Análisis de varianza de dos factores con una sola muestra por grupo

ANÁLISIS DE VARIANZA

<i>Origen de las variaciones</i>	<i>Suma de cuadrados</i>	<i>g.d.l.</i>	<i>Promedio de los cuadrados</i>	<i>F</i>	<i>Probabilidad</i>	<i>Valor crítico para F</i>
Filas	1162.444	5	232.4889	5.7562586	0.00931008	3.32583738
Columnas	521.4444	2	260.7222	6.45529574	0.01584231	4.10281586
Error	403.8889	10	40.38889			
Total	2087.778	17				

Figura 1: Resultats obtinguts amb Excel

L'estadístic que col·labora amb l'experimentador està de vacances, però per adelantar feina aquest decideix anar fent els càlculs. Quan agafa l'Excel es troba que no sap ben bé quina opció agafar i es decideix per les dues que li "sonen" més. Els resultats apareixen a la figura 1.

- Quina diferència hi ha entre les dues resolucions? Quin és el disseny experimental i el model lineal corresponent a cadascuna?
- Quan creus que té sentit fer servir la primera resolució? I la segona? Quina et sembla més adient en aquest problema?
- A quines conclusions arribes fent servir la primera resolució? I la segona? Compara les conclusions i explica les diferències que observis entre elles.
- Quines suposicions has hagut de fer per respondre les qüestions anteriors? Com les verificaries?