

1. Es defineix la matriu de centrament com: $\mathbf{H} = \mathbf{H}_n = \mathbf{I}_n - n^{-1}\mathbf{J}_n$, $\mathbf{J}_n = \mathbf{1}'\mathbf{1}$.
- (a) Demostreu que el producte de la matriu de centrament i la matriu de disseny \mathbf{HX} dona com resultat la matriu de disseny centrada, és a dir, una matriu on a cada observació se li ha restat la mitjana de la seva columna.
- (b) Comproveu que l'equació de regressió obtinguda amb les dades centrades és equivalent a la que s'obté sense centrar-les. Com podem aprofitar aquesta propietat per simplificar els càlculs de regressió?
2. En un estudi per relacionar la longitud de la línia de la vida a la ma esquerra i la vida d'una persona, s'han obtingut dades de 50 persones amb els següents resultats (x = longitud de la línia en cm; y = edat al morir en anys):

$$\begin{aligned} \sum y &= 3333 & \sum y^2 &= 231933 & \sum xy &= 30549 \\ \sum x &= 459.9 & \sum x^2 &= 4308.57 & & \end{aligned}$$

- (a) Construïu una recta de regressió i analitzeu si existeix relació entre x i y . (b) Considereu si s'escau una previsió de l'edat a la que morirà una persona amb una línia de vida de 12 cm. (c) Tot suposant el model significatiu, a quina edat en promig preveieu que morin les persones amb una línia de la vida de 12 cm? A quina edat preveieu que mori el Sr. XX que té una línia de la vida de 12 cm? (d) Feu un interval de confiança al 90% per a les dues previsions anteriors i expliqueu quina diferència, matemàtica i conceptual, hi ha entre les dues.
3. El preu del lloguer fa uns anys en l'àrea metropolitana de Barcelona A en una mostra de 460 habitatges es relaciona amb la superfície en m^2 S , el número d'anys de lloguer per la mateixa persona N i l'edat de l'edifici en anys E , segons l'equació:

$$\ln A = 5.77 + 0.46 \ln S - 0.06 \ln N - 0.012E; \quad R^2 = 0.71, \hat{\sigma} = 0.48$$

(0.07) (0.03) (0.002)

- Els números entre parèntesis són els errors estàndard d'estimació dels coeficients. Es demana: (a) Indiqueu quins coeficients són significatius i interpreteu-los; (b) Construïu un interval de confiança del 90% per a l'efecte de la variable N ; (c) Calculeu la variació explicada i la no explicada; (d) Indiqueu el valor del contrast F per a que cap de les tres variables influeixi.
4. Un enginyer vol estudiar l'eficàcia de tres marques de benzina (A, B, C) a través dels km que pot recórrer un cotxe amb 1 litre d'aquella marca (podem suposar que només existeixen aquestes tres marques). Com sap que la mida del cotxe pot influir en el resultat separa els cotxes per a la prova en petits, mitjans o grans i a continuació assigna cada benzina a l'atzar a cadascuna de les mides, obtenint els resultats següents (el format és (*mida cotxe, marca benzina, km*): (*Petit, A, 12*), (*Petit, B, 16*), (*Petit, C, 15*), (*Mitjà, A, 8*), (*Mitjà, B, 14*), (*Mitjà, C, 15*), (*Gran, A, 7*), (*Gran, B, 12*), (*Gran, C, 10*).
- (a) Expliqueu el disseny experimental que ha fet servir l'enginyer. Escriviu el model lineal corresponent indicant quins són els factors, quins els seus nivells i quines suposicions fareu per analitzar-ho (b) Plantegeu la hipòtesi que vol contrastar l'enginyer i, a continuació, calculeu les sumes de quadrats i construïu la taula d'anàlisi de la variància necessària per contrastar-la (c) Digueu a quina conclusió us duen els resultats de l'apartat (b) i feu, si cal, les comparacions múltiples necessàries per completar l'estudi.

Feu només dos problemes entre l'1, el 2 i el 3. El 4 és obligatori.