Universidad de la República **Modelos Lineales**
Facultad de Ingeniería - IMERL **Primer semestre 2016**

Práctico 4

1. *a*) Se recuerda que el *criterio de información de Akaike* (AIC) se define:

$$AIC = 2d - 2\ln(L)$$

siendo d la cantidad de parámetros y L el máximo valor de la función de verosimilitud para el modelo estimado. Deducir que en el caso de la regresión lineal múltiple normal,

$$AIC = 2d + n \times \ln(SCR) + C$$

siendo C una constante.

   b) Buscar información sobre la función stepAIC de la librería MASS.

2. Leer el Apéndice 3.1 pág 102-103 del libro de Peña *Análisis de Datos Multivariantes* donde se prueba que si s^{jj} el elemento de la entrada jj de S^{-1}, entonces $s^{jj} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} e_i^2}$.

3. Simule el modelo siguiente:

```
set.seed(1)
x1=runif(100)
x2=0.5*x_1+rnorm(100)/10
y=2+2*x1+0.3*x2+rnorm(100)
```

   a) Ajuste un modelo lineal y halle los coeficientes de la regresión.

   b) ¿Puede rechazar la hipótesis nula $(H_0)\,\beta_1 = 0$? ¿$(H_0)\,\beta_2 = 0$?

   c) Ajuste un modelo lineal para predecir y solamente en función de x_1. ¿Puede rechazar la hipótesis nula $(H_0)\,\beta_1 = 0$?

   d) Ajuste un modelo lineal para predecir y solamente en función de x_2. ¿Puede rechazar la hipótesis nula $(H_0)\,\beta_1 = 0$?

   e) ¿Los resultados obtenidos en las preguntas anteriores son contradictorias?

   f) Supongamos que tenemos una observación adicional $(\mathsf{y}, \mathsf{x}_1, \mathsf{x}_2) = (6, 0{,}8, 0{,}1)$ Repita las preguntas anteriores con este nuevo conjunto de datos. ¿Qué efecto tiene la nueva observación sobre cada modelo? ¿Esta observación es un outlier? ¿Un punto influyente? Explique.

4. Una variable Y depende de otra x (variable control no aleatoria) que toma los valores $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4$ de acuerdo con el modelo lineal normal

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i$$

Encuentre la expresión del estadístico F para la hipótesis

$$(i)\,(H_0):\ \beta_2 = 0 \qquad (ii)\,(H_0):\ \beta_1 = \beta_2.$$

5. Dado el siguiente modelo lineal normal

$$\begin{cases} \beta_1 + \beta_2 = 6{,}6 \\ 2\beta_1 + \beta_2 = 7{,}8 \\ -\beta_1 + \beta_2 = 2{,}1 \\ 2\beta_1 - \beta_2 = 0{,}4 \end{cases}$$

estudie si se puede aceptar la hipótesis nula $(H_0): \ \beta_2 = 2\beta_1$

6. Considere el modelo lineal normal

$$Y = X\beta + \epsilon$$

Pruebe que para la hipótesis $(H_0): \ X\beta = 0$ se verifica

$$SCR_H - SCR = \widehat{\beta}' X' Y$$

y hallar el estadístico F correspondiente.

7. Hallar la descomposición SVD de la matriz $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$

8. Probar que en la regresión Ridge, si $X = UDV'$ entonces:

$$\widehat{\beta}^R = V \mathrm{diag}\left(\frac{d_j}{d_j^2 + \lambda}\right) U'Y \quad \text{y} \quad \widehat{Y}^R = \sum_{j=1}^{p} \left(u_j \frac{d_j^2}{d_j^2 + \lambda} u_j'\right) Y$$

9. Comparar técnicas clásicas de selección de variables y métodos de regularización sobre los datos *Hitters* de la biblioteca de R "ISLR" . Comparar en términos de predicción separando previamente los datos en muestra de entrenamiento y de evaluación (ver y seguir las secciones 6.5 y 6.6 del libro ISLR).

10. Ejercicio 8 página 262 del libro ISLR.

11. Se quiere predecir *per capita* la tasa de crimen a partir del Boston data set. Ajuste los métodos de regresión lineal vistos en esta primera parte del curso: mejor modelo, stepwise model, regresión ridge, regresión lasso. Presentar y discutir cada una de las técnicas. ¿Con cuál modelo se quedaría para tener la mejor predicción? Justificar.