MODELO LINEAL	Practica 3	8 de mayo de 2017

Ejercicio Sea el modelo lineal $E(Y) = \beta_0 + \beta_1 X_1 + ... + \beta_{p-1} X_{p-1} = X\beta$, donde X es una matriz de dimensión $n \times p$ y rango p e $Y \sim N(X\beta, \sigma^2 I_n)$. Definimos el coeficiente de correlación múltilpe como

$$R^2 = \frac{\sum_{i=1}^n (\widehat{y}_i - \overline{y})^2}{\sum_{i=1}^n (y_i - \overline{y})^2}.$$

Supongamos que se desea testear la hipótesis

$$H_o : \beta_1 = \beta_2 = \cdots = \beta_{p-1} = 0$$

a) Sea F el estadístico del test de F usado para testear H_o. Pruebe que

$$F = \frac{R^2}{(1 - R^2)} \frac{(n-p)}{(p-1)}$$

b) Halle una expresión para R^2 en términos de la representación canónica del modelo.

c) Muestre que bajo H_o, R^2 tiene distribución $Beta$.

d) Halle la esperanza de R^2 bajo H_o.

Ayuda: recuerde que si $X \sim \Gamma(\alpha, \lambda)$ es independiente de $Y \sim \Gamma(\theta, \lambda)$, entonces $\frac{X}{X+Y} \sim Beta(\alpha, \theta)$ y que si $U \sim Beta(\alpha, \theta)$, $E(U) = \frac{\alpha}{\alpha + \theta}$.

Ejercicio Continuación ejercicio 20 P2 (Cemento) para ser usado en el **ej 1 de la práctica 3** entre otros.

e) Si tuviera que elegir 1 sola variable ¿cuál elegiría? Considerar el modelo que tiene esa única variable regresora. Estimar los parámetros

f) Calcule, para cada punto del diseño, el intervalo de confianza de nivel 0.95 para la respuesta.

g) Calcule, para cada punto del diseño, el intervalo de predicción de nivel 0.95 para la respuesta.

h) Realice un scatterplot en el que se representen los pares de puntos (x,y), la recta de mínimos cuadrados y los límites de los intervalos obtenidos en f) y g) para cada punto del diseño de nivel individual 0.95.

i) Calcule, para cada punto del diseño, el intervalo de confianza para la respuesta de manera que el nivel global de los 14 intervalos obtenidos sea 0.95. Idem con predicción.

j) Se hace una nueva observación con X3=70%, independiente de las observaciones anteriores. Es, decir se mide el calor generado al fraguar una nueva muestra de cemento en la que 70% del peso corresponde a la componente x3. Hallar un intervalo de confianza para la esperanza de esta nueva observación. Dar un predictor para el calor generado al fraguar esta muestra. Hallar un intervalo de predicción para esta nueva observación.

k) Grafique un scatterplot en el que se representen los pares de puntos (x,y), la recta de mínimos cuadrados y las curvas que se obtendrían si se unieran los límites superiores por un lado y los inferiores por otro, de los intervalos de confianza computados en i) de nivel global 0.95. Superponga la banda de confianza de nivel total 0.95 para la recta ajustada.

Ejercicio Terminar **el ítem c)** de la clase del 25 de marzo para ser usado en el **ej 3 y 5 de la práctica 3** por ejemplo.

Supongamos que queremos comparar 2 rectas de regresión dadas por

$$Y = \alpha_i + \beta_i x + \epsilon \quad i = 1, 2,$$

siendo $E(\epsilon) = 0$ y $Var(\epsilon) = \sigma^2$. Para ello tomamos n pares (x_{1j}, y_{1j}) correspondientes a la recta 1 y n pares (x_{2j}, y_{2j}) correspondientes a la recta 2, de manera que

$$Y_{ij} = \alpha_i + \beta_i x_{ij} + \epsilon_{ij}$$

donde los $\epsilon_{ij} \sim N(0, \sigma^2)$.

a) Encuentre una expresión matricial adecuada para plantear este problema. ¿Qué observa de particular en la estructura de la matriz de diseño?

b) Encuentre los estimadores de mínimos cuadrados de los parámetros. ¿Cómo influye la estructura de la matriz de diseño en la determinación de estos estimadores?

c) Se desea testear que las 2 rectas son paralelas. Exprese la hipótesis nula y alternativa para este problema y deduzca un test de nivel α para decidir entre H_0 y H_1.

Ejercicio Analizaremos diferentes ej de la práctica 3

Ej 2 P3: Incorporación de variables dummy

Ej 8 P3: diseño, lm, aov, intervalos

Ej 10 P3: Elipsoide de confianza