Día	Mes	Año	Hora	Institución
08	04	22		Universidad del Rosario

Alumno	Código	Materia
Laura Sofia Rincón Sierra		Análisis Estadístico de Datos

Curso	Bimestre	Semestre	Salón	Hoja No ___ de ___	CALIFICACIÓN

Profesor: Juan Camilo Yepes

Parcial 2

① Tres grupos de un equipo de fútbol.

Método grupo 1 $\longrightarrow$ Sesiones largas de carrera de resistencia

Método grupo 2 $\longrightarrow$ Series cortas de alta intensidad

Método gupo 3 $\longrightarrow$ Trabajo de fuerza en el gimnasio.

Tiempos en un test de rendimiento en carrera de 3 Km

Método I	Método II	Método III
15	14	13
16	13	12
14	15	11
15	16	14
17	14	11

A un nivel de confianza del 95 %. ¿Puede considerarse que los 3 métodos producen resultados equivalentes?

- Tenemos 1 equipo dividido en 3 grupos.
  - * 3 poblaciones
  - * 3 Tratamientos

★ Realizamos prueba ANOVA (Prueba de hipótesis) ✓

- $H_0: \mu_1 = \mu_2 = \mu_3$

  $H_0: \tau_1 = \tau_2 = \tau_3$

- Estadístico de prueba

$$F = \frac{\frac{SS_{treatment}}{g-1}}{\frac{SS_{res}}{n-g}}$$

- n observaciones $\longrightarrow$ 5
- g poblaciones $\longrightarrow$ 3

$$\star\ SS_{treatment} = \sum_{l=1}^{g} n_l (\bar{x}_l - \bar{x})^2$$

$$= n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2 + n_3(\bar{x}_3 - \bar{x})^2$$

donde: $\bar{x}_1 = 15.4$, $\bar{x}_2 = 14.4$, $\bar{x}_3 = 12.2$, $\bar{x} = 14$, $n_l = 5$ $(l=1,2,3)$

$$SS_{treatment} = 5(15.4-14)^2 + 5(14.4-14)^2 + 5(12.2-14)^2$$

$$= 5(1.96) + 5(0.16) + 5(3.24)$$

$$= 9.8 + 0.8 + 16.2$$

$$\boxed{= 26.8}$$ ✓

$$\star \ SS_{res} = \sum_{l=1}^{g} \sum_{j=1}^{n_l} (x_{lj} - \bar{x}_l)^2$$

$$SS_{res} = (15-15.4)^2 + (16-15.4)^2 + (14-15.4)^2 + (15-15.4)^2 + (17-15.4)^2 + (14-14.4)^2 + (13-14.4)^2 + (15-14.4)^2 + (16-14.4)^2 + (14-14.4)^2 + (13-12.2)^2 + (12-12.2)^2 + (11-12.2)^2 + (14-12.2)^2 + (11-12.2)^2$$

$$= 0.16 + 0.36 + 1.96 + 0.16 + 2.56 + 0.16 + 1.96 + 0.36 + 2.56 + 0.16 + 0.64 + 0.04 + 1.44 + 3.24 + 1.44$$

$$= 5.2 + 5.2 + 6.8$$

$$\boxed{= 17.2} \ \checkmark$$

$$\rightarrow F = \frac{\frac{26.8}{2}}{\frac{17.2}{\cancel{3}\ 12}} = \frac{13.4}{5.73} \boxed{= 2.338}$$

- $F \approx F_{g-1,n-g}$ ↗ $n-g = 15-3 = 12 \quad (n = \sum n_l)$

  $F \approx F_{2,3} \longrightarrow F \approx \cancel{9.552}\ 3.88 \longrightarrow$ Con $\alpha = 0.05$

15/20

- Región de Rechazo

  Rechazamos Ho si $F > F_{2,3}$

★ Como $2.338 < 9.552$ con $\alpha = 0.05$ concluimos que los 3 metodos son equivalentes. Aceptamos Ho con 95% confianza

② Estimar los gastos en alimentación de una familia en base a los ingresos mensuales y número de miembros de la familia. Para ello, se recoge una muestra aleatoria simple de 15 familias cuyos resultados se recogen en la siguiente tabla

12/20

Gasto alimentación	Ingresos	tamaño familia
0.43	2.10	3
0.31	1.10	4
0.52	1.80	6
0.29	1	5
0.35	2.40	2

a) Estime el modelo de regresión para poder predecir los gastos en alimentación.

- $Y = Z\beta$

$$\begin{bmatrix} 0.43 \\ 0.31 \\ 0.52 \\ 0.29 \\ 0.35 \end{bmatrix} = \begin{bmatrix} 1 & 2.10 & 3 \\ 1 & 1.10 & 4 \\ 1 & 1.80 & 6 \\ 1 & 1 & 5 \\ 1 & 2.40 & 2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} \ \checkmark$$

★ Necesitamos estimar β

$$\longrightarrow \beta = (Z'Z)^{-1} Z'Y$$

$$\mathbb{B} = \left(\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2.1 & 1.1 & 1.8 & 1 & 2.4 \\ 3 & 4 & 6 & 5 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2.1 & 3 \\ 1 & 1.1 & 4 \\ 1 & 1.8 & 6 \\ 1 & 1 & 5 \\ 1 & 2.4 & 2 \end{bmatrix} \right)^{-1} \mathbb{Z}' \mathbb{Y}$$

$$= \left(\begin{bmatrix} 5 & 8.4 & 20 \\ 8.4 & 15.62 & 31.3 \\ 20 & 31.3 & 90 \end{bmatrix} \right)^{-1} \left(\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 2.1 & 1.1 & 1.8 & 1 & 2.4 \\ 3 & 4 & 6 & 5 & 2 \end{bmatrix} \begin{bmatrix} 0.43 \\ 0.31 \\ 0.52 \\ 0.29 \\ 0.35 \end{bmatrix} \right)$$

Inversa

$$\star \quad \left[\begin{array}{ccc|ccc} 5 & 8.4 & 20 & 1 & 0 & 0 \\ 8.4 & 15.62 & 31.3 & 0 & 1 & 0 \\ 20 & 31.3 & 90 & 0 & 0 & 1 \end{array}\right] = A$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} A = \left[\begin{array}{ccc|ccc} 5 & 8.4 & 20 & 1 & 0 & 0 \\ 8.4 & 15.62 & 31.3 & 0 & 1 & 0 \\ 0 & -2.3 & 10 & -4 & 0 & 1 \end{array}\right] = A$$

$$\begin{bmatrix} 1 & 0 & 0 \\ -1.68 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A = \left[\begin{array}{ccc|ccc} 5 & 8.4 & 20 & 1 & 0 & 0 \\ 0 & 1.508 & -2.3 & -1.68 & 1 & 0 \\ 0 & -2.3 & 10 & -4 & 0 & 1 \end{array}\right] = A$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1.52 & 1 \end{bmatrix} A = \left[\begin{array}{ccc|ccc} 5 & 8.4 & 20 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1.68 & 1 & 0 \\ 0 & 0 & 6.504 & -6.55 & 1.52 & 1 \end{array}\right] = A$$

$$\begin{bmatrix} 1 & -8.4 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A = \left[\begin{array}{ccc|ccc} 5 & 0 & 20 & 15.112 & -8.4 & 0 \\ 0 & 1 & 0 & -1.68 & 1 & 0 \\ 0 & 0 & 6.504 & -6.55 & 1.52 & 1 \end{array}\right] = A$$

$$\begin{bmatrix} 1 & 0 & -3.07 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} A = \left[\begin{array}{ccc|ccc} 5 & 0 & 0 & 35.22 & -13.06 & -3.07 \\ 0 & 1 & 0 & -1.68 & 1 & 0 \\ 0 & 0 & 6.504 & -6.55 & 1.52 & 1 \end{array}\right] = A$$

$$\begin{bmatrix} 1/5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/6.504 \end{bmatrix} A = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 7.044 & -2.612 & -0.614 \\ 0 & 1 & 0 & -1.68 & 1 & 0 \\ 0 & 0 & 1 & -1.007 & 0.233 & 0.154 \end{array}\right]$$

$$\mathbb{B} \approx \begin{bmatrix} 7.044 & -2.612 & -0.614 \\ -1.68 & 1 & 0 \\ -1.007 & 0.233 & 0.154 \end{bmatrix} \begin{bmatrix} 1.9 \\ 3.31 \\ 7.8 \end{bmatrix} \approx \boxed{\begin{bmatrix} -0.05132 \\ 0.118 \\ 0.059 \end{bmatrix} = \mathbb{B}}$$

$$\mathbb{Y} = \mathbb{Z}\,\mathbb{B}$$

b) Calcule R^2 del modelo.

c) ¿Qué quiere decir β_0?

$\star$ Que la intersección del modelo es negativa.

d) ¿Que dice β_1?

* Que los gastos aumentan 0.118 por cada unidad que aumenta ingresos y ~~la~~ otra variable permanece constante!

Observaciones