

**CO3321-CO3322. Segundo Examen Parcial, 40%.**

1. La variable aleatoria  $X$  tiene distribución  $\text{Bin}(n, p)$ , siendo  $n$  conocida. Se quiere probar la hipótesis nula  $H_0: p = 1/2$  contra la alternativa  $H_a: p = 3/4$ . Verifique que la región de rechazo para la correspondiente prueba de Neyman-Pearson es de la forma

$$\text{RR} = \{X > c\}$$

para alguna constante  $c$ . Sabiendo que  $n = 20$ , halle (aproximadamente)  $c$  para error tipo I  $\alpha = .05$ . ¿Cual es la potencia de la prueba para la alternativa considerada?

(10 pts)

2. La siguiente tabla muestra la cilindrada  $x$  (en litros) y el rendimiento de gasolina  $y$  (en kms/litro) de nueve modelos de automóvil compacto.

$i$	1	2	3	4	5	6	7	8	9
$x$	0.97	0.85	0.98	1.05	1.20	1.51	1.40	1.34	1.46
$y$	10.1	12.2	10.9	10.1	10.1	9.2	9.6	9.6	8.8

Para estos datos se tiene  $\bar{x} = 1.195$ ,  $\bar{y} = 10.07$ ,  $S_{xx} = \sum(x_i - \bar{x})^2 = 0.4694$ ,  $\sum x_i^2 = 13.33$ ,  $S_{yy} = \sum(y_i - \bar{y})^2 = 8.04$  y  $S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y}) = -1.699$ .

Ajuste una recta  $y = \beta_0 + \beta_1 x$  a estos datos. Se piensa que, para este tipo de auto, la eficiencia disminuye con la cilindrada. Haga una prueba de hipótesis, a nivel 5% para  $H_0: \beta_1 = 0$  contra  $\beta_1 < 0$  e interprete el resultado. Indique, aproximadamente, el p-valor de su estadístico de prueba.

(10 pts)

3. Consideremos el conjunto de datos en la tabla siguiente

$y$	26	24	175.0	160.0	163.0	55.0	62.0	100.0	26.0	30.0	70.0	71.0
$x_1$	1	1	1.5	1.5	1.5	0.5	1.5	0.5	1.0	0.5	1.0	0.5
$x_2$	1	1	4.0	4.0	4.0	2.0	2.0	3.0	1.5	1.5	2.5	2.5

Para estos datos se han ajustado los dos modelos:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_2^2 + \beta_5 x_1 x_2 + \epsilon \text{ (modelo completo) y}$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon \text{ (modelo reducido).}$$

Sean  $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots)^t$  el vector de estimadores,  $\mathbf{X}$  la matriz de diseño (incluyendo la columna de unos) y  $\mathbf{Y}$  el vector de respuestas. Entre los cálculos realizados se tiene lo siguiente:

$$\sum y_i^2 = 112432;$$

$$\text{Para el modelo completo: } \hat{\beta}_0 = 24.4, \hat{\beta}_1 = -38, \hat{\beta}_2 = 0.72, \hat{\beta}_3 = 35.0, \hat{\beta}_4 = 11.1, \hat{\beta}_5 = -10.0, \hat{\beta}^t \mathbf{X}^t \mathbf{Y} = 112213;$$

$$\text{Para el modelo reducido: } \hat{\beta}_0 = -49.6, \hat{\beta}_1 = 18.4, \hat{\beta}_2 = 46.1 \text{ y } \hat{\beta}^t \mathbf{X}^t \mathbf{Y} = 111622.$$

se ha ajustado el modelo  $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \epsilon$ , obteniéndose el siguiente vector de estimadores:

$$\vec{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3)^t = (1.078, 1.17, -0.60, -0.712)^t$$

Sean  $\vec{Y}$  el vector de respuestas y  $\mathbf{X}$  la matriz de diseño (incluyendo la columna de unos). Se tiene que

$$\vec{Y}^t \vec{Y} = 41.15 \quad \text{y} \quad \vec{\beta}^t \mathbf{X}^t \vec{Y} = 38.6$$

Además, la matriz  $(\mathbf{X}^t \mathbf{X})^{-1}$  es una matriz diagonal, con diagonal (0.0833, 0.5, 0.0927, 0.0282). Con esta información,

- (i) Realice una prueba de hipótesis a nivel 5% para decidir si hay evidencia de que  $\beta_1$  es distinto de cero. Indique, aproximadamente, el  $p$ -valor de su estadístico de prueba.
- (ii) Se van a producir nuevos valores de  $y$  para  $x^* = (x_1^*, x_2^*, x_3^*)^t = (0, 0.5, 1)^t$ . Proporcione un intervalo del 95% de confianza para estas nuevas observaciones. (10 pts)

4. La tabla siguiente presenta la resistencia a la tensión del cemento producido con cuatro métodos diferentes de mezclado, haciendo cuatro pruebas para cada método de mezcla. Escriba la tabla ANOVA para estos datos y realice la prueba de hipótesis para  $H_0$ : Los efectos de los métodos de mezclado sobre la resistencia a la tensión del cemento son todos iguales. ¿Que se concluye de esta tabla? Gráfique los residuos del modelo contra  $\hat{y}_{ij}$ . ¿Podemos suponer que se cumplen las hipótesis estadísticas correspondientes a ANOVA?

Método de Mezcla	Resistencia a la tensión (lb/pulg <sup>2</sup> )			
1	3129	3000	2865	2890
2	3200	3300	2975	3150
3	2800	2900	2985	3050
4	2600	2700	2600	2765

Para aligerar sus cálculos, puede usar que  $\bar{y}_{1.} = 2971$ ,  $\bar{y}_{2.} = 3156.25$ ,  $\bar{y}_{3.} = 2933.75$ ,  $\bar{y}_{4.} = 2666.25$ , y  $\bar{y}_{..} = 2931.81$

(10 pts)