ITAM - Métodos Estadísticos para C.Pol y R.I.

Assignment 02 - Respuestas

1. - Sabemos que:

   Los estimadores de mínimos cuadrados de α y β son

   $$b = \frac{S_{xy}}{S_{xx}}$$

   y

   $$a = \bar{y} - b\bar{x}$$

   Entonces, para encontrar la recta de predicción de mínimos cuadrados para los datos de la tabla:

   Paso 1:

   $S_{xx} = \sum x_i^2 - \frac{\sum x_i{}^2}{n} = 23634 - \frac{460^2}{10} = 2474$

   $S_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$

   $= 36854 - \frac{460*760}{10} = 1894$

   $\bar{y} = \frac{\sum y_i}{n} = \frac{760}{10} = 76$

   $\bar{x} = \frac{\sum x_i}{n} = \frac{460}{10} = 46$

   Paso 2:

   $b = \frac{S_{xy}}{S_{xx}} = \frac{1894}{2474} = 0.76556$

   y

   $$a = \bar{y} - b\bar{x} = 76 - (0.76556)(46) = 40.78424$$

   Entonces, la recta de regresión de mínimos cuadrados es:

   $$\hat{y} = a + bx$$

   o bien:

   $$\hat{y} = \beta_0 + \beta_1 x = 40.78424 + 0.76556x$$

   -

   - que "hay una relación lineal significativa entre las calificaciones esperadas y la puntuación final del examen.

   - Para estimar el promedio de las calificaciones cuyo aprovechamiento es de 50, con un intervalo de confianza debemos:

   La estimación puntual de $E(y \mid x_0 = 50)$ el promedio de calificación es:

   $\hat{y} = 40.78424 + 0.76556(50) = 79.06$

2. - Sabemos que:
   $S_{xx} = \sum x_i^2 - \frac{\sum x_i{}^2}{n}$
   $S_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$
   $S_{yy} = \sum y_i^2 - \frac{\sum y_i{}^2}{n}$
   Entonces: $S_{xx} = 60.4$ $S_{xy} = 328$
   $S_{yy} = 2610$

   $$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

   Entonces:
   $r = \frac{328}{\sqrt{60.4*2610}} = 0.8261$
   El valor obtenido para r es cerca a 1, indica que hay una relación lineal positiva bastante fuerte entre estatura y peso

   - Para probar si es significativamente diferente de cero, tenemos:
   $H_o : \rho = 0$ contra $H_a : \rho \neq 0$
   El valor del estadístico de prueba es
   $t = r\sqrt{\frac{n-2}{1-r^2}}$
   $= 0.8261 * \sqrt(\frac{10-2}{1-(0.8261)^2}) = 4.15$
   La t observada o el valor de tablas: tiene una distribución, con $n = 10$, de 8 grados de libertad.
   Dado que la t observada es mayor que
   $t_0.005 = 3.355$
   y el $valorp$ es menor a $2(0.005) = 0.01$, entonces rechazamos Ho. Concluimos que la correlación es significativamente diferente de 0

3. • Tenemos que:

$$\hat{\beta}_1 = \frac{S_x y}{S_x x}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Entonces:
$S_{xx} = \sum x_i^2 - \frac{\sum x_i{}^2}{n}$
$S_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$
$S_{yy} = \sum y_i^2 - \frac{\sum y_i{}^2}{n}$

$$\hat{\beta}_1 = \frac{7 - \frac{1}{5}(0)(5)}{10 - \frac{1}{5}(0)^2} = 0.7$$

Ahora:
$\bar{y} = \frac{\sum(y_i)}{n} = 0.7$
$\bar{x} = \frac{\sum(x_i)}{n} = 0$

$$\hat{\beta}_0 = \frac{5}{5} - (0.7)(0) = 1$$

Entonces, la recta de regresión de mínimos cuadrados es:

$$\hat{y} = \bar{y} + \hat{\beta}_0 x$$

$$\hat{y} = 1 + 0.7x$$

• Ahora, para determinar el intervalo de confianza al 90% para E(y) cuando x= 1 tenemos que:

$E(y) = \beta_0 + \beta_1 x$

Para estimar dicho valor fijo, usaremos el estimador insesgado $E(\hat{Y}) = \hat{\beta}_0 + \hat{\beta}_1 x*$ Entonces:

$E(\hat{Y}) = 1 + 0.7x*$ En este caso, $x* = 1$ y como $n = 5$, $\bar{x} = 0$ y $S_x x = 10$

Entonces, el error estándar es

$\frac{1}{n} + \frac{(x-\bar{x})^2}{S_x x} = \frac{1}{5} + \frac{(1-0)^2}{10} = 0.3$

Ahora bien, el valor de la $t_0.05$ de tablas es con n-2= 3 grados de libertad es 2.353

El intervalo está dado por:

$$\hat{\beta}_0 + \hat{\beta}_1 x * \pm t_{\alpha/2} * S * \sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{S_x x}}$$

sustituyendo:

$$((1 + (0.7)(1)) \pm (2.353)(0.606)\sqrt{0.3}$$

$= 1.7 \pm 0.781$ El intervalo está dado por:

$$(0.919, 2.481)$$

Es decir: Con un nivel de confianza de 90% cuando la variable independiente tome el valor de 1, la variable dependi

- Para demostrar que la pendiente difiere de cero dicho ejercicio implica una prueba de hipótesis
  $H_o : \beta_1 = 0$ contra $H_a : \beta_1 \neq 0$
  El valor del estadístico de prueba es
  $t = \frac{\hat{\beta}_1}{s\sqrt{c_{11}}}$
  Sustituyendo tenemos:

$$c_{11} = \frac{1}{S_x x} = 1/10 = 0.1$$

  $t = \frac{0.7-0}{0.606\sqrt{0.1}} = 3.65$
  Ahora, si tomamos con un $\alpha = 0.05$, el valor de t de tablas , $t_{\alpha/2} = t_{0.025}$ con 3 grados de libertad es 3.182
  Bajo la hipótesis alternativa la región de rechazo está dado por:

$$-3.182 \leq t \geq 3.182$$

  Ahora bien, dado que es una prieba de dos colas para calcular el valor p:

$$valorp = 2P(t > 3.65)$$

$$= 0.01 < P(t > 3.65) < 0.025$$

  Entonces:

$$= 0.02 < valorp < 0.05$$

  Dado que el p value es menor a 0.1, rechazamos la hipótesis nula

- Para calcular el intervalo de confianza para β_1 tenemos:
  El valor de tablas para $t_{0.025}$ con 3 grados de libertad es 3.182. Entonces el intervalo de confianza está dado por

$$\hat{\beta}_1 \pm t_{0.025} * s * \sqrt{c_{11}}$$

  siendo $c_{11} = 1/S_{xx}$
  Susituyendo, obtenemos

$$0.7 \pm (3.182)(0.606)\sqrt{0.1}$$

  O bien,

$$0.7 \pm 0.610$$

  El parámetro de β_1 variará entre 0.9 y 1.31 unidades

- Para determinar el pronóstico puntual cuando x = 2 con un nivel de confianza de 0.90. Tenemos:

$$\hat{\beta}_0 = 1$$

$$\hat{\beta}_1 = 0.7$$

De modo que el valor pronósticado para Y cuando $x = 2$ es

$$\hat{\beta}_0 + \hat{\beta}_1 x* = 1 + (0.7)(2) = 2.4$$

Además, el intervalo de predicción está dado por:

$$\hat{\beta}_0 + \hat{\beta}_1 x * \pm t_{\alpha/2} S \sqrt{1 + \frac{1}{n} + \frac{(x * - \bar{x})^2}{S_{xx}}}$$

Sustituyendo, tenemos:

$$2.4 \pm (2.353)(0.606)\sqrt{1 + 0.6}$$

$$2.4 \pm 1.804$$

4. • Para ajustar el modelo de regresión de mínimos cuadrados:
$S_{xx} = \sum x_i^2 - \frac{\sum x_i{}^2}{n}$
$S_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$
$S_{yy} = \sum y_i^2 - \frac{\sum y_i{}^2}{n}$

$$\hat{\beta}_1 = \frac{S_x y}{S_x x}$$
$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Sustituyendo, obtenemos:
$S_{xy} = 8.709 - \frac{1}{6}(8.74)(6.148) = -0.247$
$S_{xx} = 12.965 - \frac{1}{6}(8.74)^2 = 0.234$
$S_{yy} = 6.569 - \frac{1}{6}(6.148)^2 = 0.269$

$$\hat{\beta}_1 = \frac{-0.247}{0.234} = -1.056$$
$$\hat{\beta}_0 = \frac{6.148}{6} - (-1.056)(\frac{8.74}{6}) = 2.563$$

Entonces, la recta de minimos cuadrados está dada por

$$\hat{\beta}_0 + \hat{\beta}_1 x = -1.056 + 2.563x$$

• Prueba de hipótesis para β_1

Bibliografía Wackerly. (2008). Estadística Matemática con Aplicaciones (7.a ed.). Cengage Learning.