

Regresión: Tarea por Grupos

Cuervo Fernández, Esther

García Prado, Sergio

Martín Villares, Pablo

26 de noviembre de 2017

Parte I

Ejercicios Montgomery:

- 2.1. En la tabla B.1 del apéndice aparecen datos sobre el desempeño de los 26 equipos de la Liga Nacional de Fútbol en 1976. Se cree que la cantidad de yardas ganadas por tierra por los contrarios (x_8) tiene un efecto sobre la cantidad de juegos que gana un equipo (y).
- a. Ajustar un modelo de regresión lineal simple que relacione los juegos ganados, y , con las yardas ganadas por tierra por los contrarios, x_8 .

$$y_i = \beta_0 + \beta_1 x_{8i} + \varepsilon_i \quad (1)$$

Analizaremos el modelo de la ecuación 1 mediante un estudio de regresión lineal simple, con y como variable dependiente, x_8 como variable independiente, β_0 el intercepto, y β_1 la pendiente. ε es el error aleatorio.

El procedimiento *REG* permite estimar los valores del intercepto y la pendiente, y analizar las hipótesis nulas:

$$H_0 : \beta_0 = 0 \quad (2)$$

$$H'_0 : \beta_1 = 0 \quad (3)$$

La hipótesis más importante en la regresión es la que incumbe a β_1 , ya que si la pendiente de la recta es 0, no existe regresión, y se trata de una población simple sobre la que x_8 no tiene efecto.

Parameter Estimates						
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	21.78825	2.69623	8.08	<.0001
x8	x8	1	-0.00703	0.00126	-5.58	<.0001

Figura 1: Estimadores y p-valores para la regresión.

Como podemos ver, el *p-valor* para β_1 está por debajo de 0,05, por lo que podemos rechazar la hipótesis nula con una confianza del 95 %. El *p-valor* para el intercepto también permite rechazar la hipótesis nula.

Por tanto, los estimadores son:

$$\hat{\beta}_0 = 21,78825 \quad (4)$$

$$\hat{\beta}_1 = -0,00703 \quad (5)$$

Con la recta de regresión:

$$y_i = 21,78825 - 0,00703x_{8i} + \varepsilon_i \quad (6)$$

La interpretación de estos parámetros es que con 0 yardas ganadas por el contrario, un equipo gana de media 21,78825 juegos, y que por cada yarda que gana el contrario, el equipo pierde de media 0,00703 juegos.

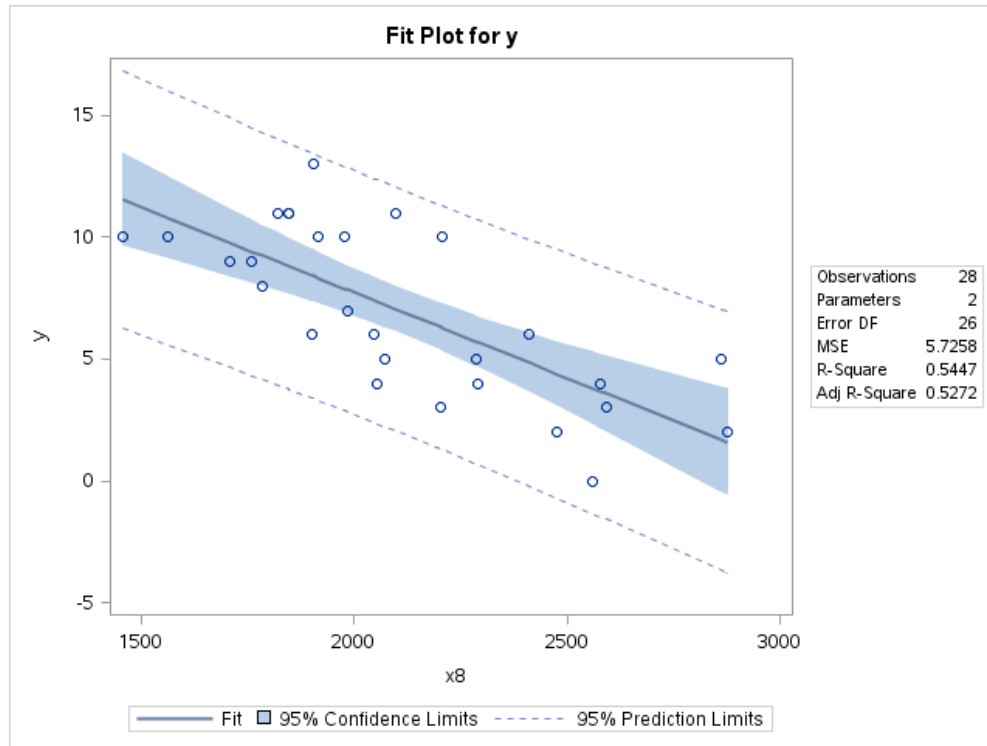


Figura 2: Fit plot del modelo de regresión lineal simple

- b. Formar la tabla de análisis de varianza y probar el significado de la regresión.
 - c. Determinar un intervalo de confianza de 95 % para la pendiente.
 - d. ¿Qué porcentaje de variabilidad total da y , y explica este modelo?
 - e. Determinar un intervalo de confianza de 95 % para la cantidad promedio de juegos ganados, si la distancia ganada por tierra por los contrarios se limita a 2000 yardas.
- 2.2. Supóngase que se quiere usar el modelo desarrollado en el problema 2.1 para pronosticar la cantidad de juegos que ganará un equipo si puede limitar los avances por tierra de sus contrarios a 1800 yardas. Determinar un estimado de punto de la cantidad de juegos ganados cuando $x_8 = 1800$. Determinar un intervalo de predicción de 90 % para la cantidad de juegos ganados.

Parte II

Código Fuente