

Apuntes Unidad 2

Estimación e interpretación de los parámetros
del modelo de regresión lineal



Curso: Probabilidad y estadística descriptiva

Unidad 2 : Media muestral, dispersión y correlación

Tema: Modelo de regresión lineal

Contenido: Estimación e interpretación de los parámetros del modelo de regresión lineal

MODELO DE REGRESIÓN LINEAL

Los valores \hat{y}_i corresponden a las estimaciones de los valores y_i . Estos se obtienen al reemplazar un valor x_i de la variable x en el modelo de regresión, cuya fórmula matemática es la siguiente:

$$\hat{y}_i = \hat{\beta}_1 \cdot x_i + \hat{\beta}_0$$

Estas estimaciones son las que conforman la recta de regresión lineal. Como los \hat{y}_i fueron estimados, puede que existan diferencias entre estos valores y los de y_i . Estas diferencias corresponden a los errores del modelo.

$$\hat{e}_i = y_i - \hat{y}_i$$

Hasta el momento no hemos hablado de cómo conocer el valor que acompaña a la variable explicativa, es decir, $\hat{\beta}_1$, ni el coeficiente $\hat{\beta}_0$, por lo que a continuación aprenderemos a estimarlos.

ESTIMACIÓN DE LOS PARÁMETROS DEL MODELO DE REGRESIÓN LINEAL

La siguiente igualdad se conoce como regresión lineal simple.

$$y = \beta_1 \cdot x + \beta_0$$

Para determinar la recta de regresión, debemos encontrar valores que estimen los coeficientes β_1 y β_0 del modelo.

Aunque no vamos a verlo en detalle, el método de mínimos cuadrados nos brinda una solución algebraica para estimar el valor de los coeficientes antes mencionados a través de las expresiones:

$$\hat{\beta}_1 = \frac{s_{xy}}{s_x^2} \qquad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \cdot \bar{x}$$

Donde:

- s_{xy} : Covarianza entre x e y
- s_x^2 : Varianza de x
- \bar{x} : Media de x
- \bar{y} : Media de y

Curso: Probabilidad y estadística descriptiva

Unidad 2 : Media muestral, dispersión y correlación

Tema: Modelo de regresión lineal

Contenido: Estimación e interpretación de los parámetros del modelo de regresión lineal

Una vez estimados los valores de los parámetros, construimos el modelo de regresión lineal:

$$\hat{y} = \hat{\beta}_1 \cdot x + \hat{\beta}_0$$

Es importante recalcar que utilizamos un $\hat{\quad}$ para denotar que estos son valores estimados de la variable y y de los parámetros β_0 y β_1 .

SÍNTESIS

- El **modelo de regresión lineal** permite modelar la relación entre dos variables x e y .
- La relación entre estas variables la podemos describir por la ecuación $y = \beta_1 \cdot x + \beta_0$.
- Los coeficientes β_1 y β_0 se conocen como parámetros del modelo y en general son desconocidos.
- Estos parámetros se pueden estimar a partir de datos empíricos. Para ello, se toman datos y , utilizando el *método de los mínimos cuadrados*, se encuentra la recta que se ajusta mejor a los datos
- El método de mínimos cuadrados nos brinda una solución algebraica para estimar el valor de los coeficientes antes mencionados a través de las expresiones:

$$\hat{\beta}_1 = \frac{s_{xy}}{s_x^2} \qquad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \cdot \bar{x}$$

- Donde \bar{x} corresponde a la media de x , \bar{y} corresponde la media de y , s_{xy} es la covarianza entre x e y y s_x^2 es la varianza de x .