Universidad de El Salvador
Facultad de Ciencias Naturales y Matemática
Maestría en Estadística y Ciencia de Datos
Inferencia Estadística y Regresión

Desarrollo de la tarea 2

Presentado por: Salvador Enrique Rodríguez Hernández (rh06006)
Fecha de entrega: 04 de diciembre de 2024

1. *Estimar los parámetros β_0 y β_1 de un modelo de regresión lineal simple por el método de máxima verosimilitud y comparar los resultados obtenidos con el método de mínimos cuadrados.*

Solución

<u>Estimación por Mínimos Cuadrados Ordinarios (OLS)</u>

El método de mínimos cuadrados ordinarios (OLS) minimiza la suma de los errores cuadráticos:

$$SSE = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

Para encontrar las estimaciones $\widehat{\beta_0}$ y $\widehat{\beta_1}$, se deriva SSE respecto a β_0 y β_1 y se igualan a cero:

$$\frac{\partial SSE}{\partial \beta_0} = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0 \quad \ldots\ldots (1)$$
$$\frac{\partial SSE}{\partial \beta_1} = -2\sum_{i=1}^{n}x_i(y_i - \beta_0 - \beta_1 x_i) = 0 \ldots\ldots (2)$$

Resolviendo la ecuación (1):

$$\sum_{i=1}^{n} y_i = n\beta_0 + \beta_1 \sum_{i=1}^{n} x_i$$

Dividiendo entre n:

$$\bar{y} = \beta_0 + \beta_1 \bar{x}$$

Así:

$$\beta_0 = \bar{y} - \beta_1 \bar{x}$$

Despejando la ecuación (2) se tiene que:

$$\sum_{i=1}^{n} x_i y_i = \beta_0 \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2$$

Usando $\beta_0 = \bar{y} - \beta_1 \bar{x}$:

$$\sum_{i=1}^{n} x_i y_i = (\bar{y} - \beta_1 \bar{x}) \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2$$
$$\sum_{i=1}^{n} x_i y_i = \bar{y} \sum_{i=1}^{n} x_i - \beta_1 \bar{x} \sum_{i=1}^{n} x_i + \beta_1 \sum_{i=1}^{n} x_i^2$$
$$\sum_{i=1}^{n} x_i y_i = \bar{y} \sum_{i=1}^{n} x_i + \beta_1 \left(\sum_{i=1}^{n} x_i^2 - \bar{x} \sum_{i=1}^{n} x_i\right)$$

Despejando β_1:

$$\beta_1 = \frac{\sum_{i=1}^{n} x_i y_i - \bar{y}\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 - \bar{x}\sum_{i=1}^{n} x_i}$$

Multiplicando le numerador y denominador por $\frac{1}{n}$:

$$\beta_1 = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\bar{y}\sum_{i=1}^{n} x_i}{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\bar{x}\sum_{i=1}^{n} x_i}$$

$$\beta_1 = \frac{\frac{\sum_{i=1}^{n} x_i y_i}{n} - \bar{y}\bar{x}}{\frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2} \text{ o } \beta_1 = \frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

Expresando el numerador como la covarianza y el denominador como la varianza se tiene:

$$\beta_1 = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}$$

Finalmente, se sustituye β_1 en $\beta_0 = \bar{y} - \beta_1\bar{x}$.

$$\beta_0 = \bar{y} - \beta_1\bar{x}$$

$$\beta_0 = \bar{y} - \left(\frac{\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}\right)\bar{x}$$

$$\beta_0 = \frac{\bar{y}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right) - \bar{x}\left(\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}\right)}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

$$\beta_0 = \frac{\bar{y}\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\bar{y} - \bar{x}\sum_{i=1}^{n} x_i y_i + n\bar{x}^2\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

$$\beta_0 = \frac{\sum_{i=1}^{n} x_i^2\bar{y} - \sum_{i=1}^{n} x_i y_i\bar{x}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

Estimación por Máxima Verosimilitud (MLE)

Dado el modelo de regresión:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad \varepsilon_i \sim N(0,\sigma^2)$$

La función de verosimilitud es:

$$L(\beta_0,\beta_1,\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}\right)$$

Tomando el logaritmo natural se tiene que:

$$\mathrm{l}(\beta_0,\beta_1,\sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}$$

Para maximizar l, se deriva respecto a β_0, β_1 y σ^2, y se iguala a cero:
Respecto a β_0 y β_1:

$$\frac{\partial \mathrm{l}}{\partial \beta_0} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\frac{\partial \mathrm{l}}{\partial \beta_1} = \frac{1}{\sigma^2}\sum_{i=1}^{n} x_i(y_i - \beta_0 - \beta_1 x_i) = 0$$

Estas ecuaciones son equivalentes a las ecuaciones normales de OLS. Por lo tanto:

$$\beta_1 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}, \beta_0 = \frac{\sum_{i=1}^{n} x_i^2\bar{y} - \sum_{i=1}^{n} x_i y_i \bar{x}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

Respecto a σ^2:

$$\frac{\partial \mathrm{l}}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^{n}(y_i-\beta_0-\beta_1 x_i)^2}{2\sigma^4} = 0$$

Resolviendo para σ^2:

$$\widehat{\sigma^2} = \frac{\sum_{i=1}^{n}\left(y_i-\widehat{\beta_0}-\widehat{\beta_1}x_i\right)^2}{n}$$

<u>Comparación entre OLS y MLE</u>

- Ambos métodos proporcionan las mismas estimaciones para β_0 y β_1 bajo la suposición de normalidad en los residuos.
- MLE además estima σ^2, la varianza de los residuos, lo que no es calculado explícitamente por OLS.

2. *En la regresión lineal múltiple, la idea es minimizar la suma de los cuadrados de los residuos* (e) *definida como:*

$$S(\beta) = \sum_{i=1}^{n} e_i^2 = e'e = (Y - X\beta)'(Y - X\beta)$$

Donde:
Y: *Vector de respuestas (variable dependiente).*
X*: Matriz de diseño de las variables independientes.*
β*: Vector de parámetros del modelo (coeficientes).*
$e = Y - X\beta$: *Vector de residuos.*
Al desarrollar $S(\beta)$*:*

$$S(\beta) = (Y - X\beta)'(Y - X\beta)$$

Usamos las propiedades del álgebra matricial para expandir el producto:

$$S(\beta) = Y'Y - \beta'X'Y - Y'X\beta + \beta'X'X\beta$$

Demostrar que $-\beta'X'Y$ *y* $-Y'X\beta$ *son escalarmente equivalentes porque* $Y'X\beta$ *es un escalar* (1×1)*, y la transposición no afecta el valor escalar y, por lo tanto, podemos combinarlos:*

$$S(\beta) = Y'Y - 2\beta'X'Y + \beta'X'X\beta$$

Solución

Para justificar que $-\beta' X' Y$ y $-Y' X\beta$ son escalares y equivalentes, se comienza analizando sus dimensiones y propiedades fundamentales.

Dado que:

- β es un vector columna de dimensión $p \times 1$,
- X es una matriz de diseño de dimensión $n \times p$,
- Y es un vector columna de dimensión $n \times 1$,

se tiene que X' siendo la transpuesta de X, tiene dimensión $p \times n$. Al multiplicar X' por Y, el resultado es un vector columna de dimensión $p \times 1$. Posteriormente, al multiplicar este resultado por β', que tiene dimensión $1 \times p$, se obtiene un escalar de dimensión 1×1. Por lo tanto, el término $\beta' X' Y$ es un escalar.

De manera similar, en el término $Y' X\beta$, el vector Y', siendo la transpuesta de Y, tiene dimensión $1 \times n$. Al multiplicarlo por X de dimensión $n \times p$, se genera un vector fila de dimensión $1 \times p$. Este vector fila, al multiplicarse por β, de dimensión $p \times 1$, produce nuevamente un escalar de dimensión 1×1. Por consiguiente, $Y' X\beta$ también es un escalar.

Dado que $\beta' X' Y$ es un escalar, su traspuesto es igual a sí mismo (el escalar es una matriz de 1×1), es decir, $\beta' X' Y = (\beta' X' Y)'$, lo que implica que:

$$
\begin{aligned}
(\beta' X' Y)' &= (X' Y)' \beta \\
&= Y' X \beta
\end{aligned}
$$

En consecuencia, se tiene que:

$$\beta' X' Y = (\beta' X' Y)' = Y' X \beta$$

Es decir, $\beta' X' Y$ y $Y' X\beta$ son escalares equivalentes porque ambos representan el mismo valor. Esta equivalencia permite combinarlos directamente en la expansión cuadrática de $e'e$.

Sustituyendo la equivalencia escalar en la expansión original, se tiene que:

$$S(\beta) = Y' Y - 2\beta' X' Y + \beta' X' X$$