

1

Regresión lineal simple

1.1 Introducción

Fue introducido por Francis Galton (1908). El modelo de regresión lineal simple está formado típicamente por:

$$y = \beta_0 + \beta_1 x + \epsilon.$$

Donde:

- y = variable dependiente o variable de respuesta.
- x = variable independiente o explicativo o predictor.
- β_0 = intercepto y .
- β_1 = pendiente.
- ϵ = error aleatorio.

Una presentación más general de un modelo de regresión sería:

$$y = E(y) + \epsilon,$$

Donde: $E(y)$ es la esperanza matemática de la variable respuesta. Cuando $E(y)$ es una combinación lineal de las variables explicativas x_1, x_2, \dots, x_k la regresión es una regresión lineal. Con $E(\epsilon_i) = 0$ y $Var(\epsilon_i) = \sigma^2$. Todos los ϵ_i son independientes.

Ahora debemos hallar buenos estimadores para β_0 y β_1 .

1.2 Estimaciones por mínimos cuadrados

El principal objetivo de los mínimos cuadrados para un modelo de regresión lineal simple es hallar los estimadores b_0 y b_1 tales que la suma de la distancia al cuadrados de la respuesta real y_i y las respuesta de las pronosticadas $\hat{y}_i = \beta_0 + \beta_1 x_i$ alcanza el mínimo entre todas las opciones posibles de coeficientes de regresión β_0 y β_1 . Es decir,

$$(b_0, b_1) = \arg \min_{(\beta_0, \beta_1)} \sum_{i=1}^n [\beta_0 + \beta_1 x_i]^2.$$

Matemáticamente, las estimaciones de mínimos cuadrados de la regresión lineal simple se obtienen resolviendo el siguiente sistema:

$$\frac{\partial}{\partial \beta_0} \sum_{i=1}^n [y_i - (\beta_0 + \beta_1 x_i)]^2 = 0 \quad (1.1)$$

$$\frac{\partial}{\partial \beta_1} \sum_{i=1}^n [y_i - (\beta_0 + \beta_1 x_i)]^2 = 0 \quad (1.2)$$

Supongamos que b_0 y b_1 son soluciones del sistema de arriba, podemos describir la relación entre x e y por la regresión lineal $\hat{y} = b_0 + b_1 x$, el cual es llamado la **recta de regresión ajustada**. Es más conveniente resolver para b_0 y b_1 usando el modelo lineal centralizado:

$$y_i = \beta_0 + \beta_1 \bar{x} - \beta_1 \bar{x} + \beta_1 x_i + \epsilon_i \Rightarrow y_i = \beta_0^* + \beta_1 (x_i - \bar{x}) + \epsilon_i,$$

donde $\beta_0 = \beta_0^* - \beta_1 \bar{x}$. Necesitamos resolver para

$$\frac{\partial}{\partial \beta_0^*} \sum_{i=1}^n [y_i - (\beta_0^* + \beta_1 (x_i - \bar{x}))]^2 = 0$$

$$\frac{\partial}{\partial \beta_1} \sum_{i=1}^n [y_i - (\beta_0^* + \beta_1 (x_i - \bar{x}))]^2 = 0$$

Realizando la derivada parcial para β_0 y β_1 tenemos

$$\sum_{i=1}^n [y_i - (\beta_0^* + \beta_1 (x_i - \bar{x}))] = 0$$

$$\sum_{i=1}^n [y_i - (\beta_0^* + \beta_1 (x_i - \bar{x}))] (x_i - \bar{x}) = 0$$

Notemos que

$$\sum_{i=1}^n y_i = n\beta_0^* + \sum_{i=1}^n \beta_1 (x_i - \bar{x}) = n\beta_0^* \quad (1.3)$$

Por lo tanto, tenemos

$$\beta_0^* = \frac{1}{n} \sum_{i=1}^n y_i = \bar{y}.$$

Luego, sustituyendo β_0^* por \bar{y} en (2.3) obtenemos

$$\sum_{i=1}^n [y_i - (\bar{y} + \beta_1 (x_i - \bar{x}))] (x_i - \bar{x}) = 0$$

Después denotamos b_0 y b_1 las soluciones de los sistemas (2.1) y (2.2). Ahora, es fácil ver que

$$b_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} = \frac{S_{xy}}{S_{xx}} \quad (1.4)$$

y

$$b_0 = b_0^* - b_1 \bar{x} = \bar{y} - b_1 \bar{x} \quad (1.5)$$

El valor ajustado de la regresión lineal simple es definida como $\hat{y}_i = b_0 + b_1 x_i$. La diferencia entre y_i y el valor ajustado \hat{y}_i es $e_i = y_i - \hat{y}_i$, que se refiere al residuo de la regresión. Los residuos de regresión se pueden calcular a partir de las respuestas observadas y_i y los valores ajustados \hat{y}_i , por lo tanto, los residuos son observables. Cabe señalar que el término de error ϵ_i en el modelo de regresión no es observable. El error de regresión es la cantidad por la cual una observación difiere de su valor esperado; este último se basa en la población total de la que se eligió aleatoriamente la unidad estadística. El valor esperado, el promedio de toda la población, normalmente no es observable.

Un residual, por otro lado, es una estimación observable de un error no observable. El caso más simple implica una muestra aleatoria de n hombres cuyas alturas se miden. El promedio de la muestra se utiliza como una estimación del promedio de la población. Entonces, la diferencia entre la altura de cada hombre de la muestra y el promedio de la población no observable es un error, y la diferencia entre la altura de cada hombre de la muestra y el promedio de la muestra observable es un residuo. Dado que los residuales son observables, podemos usar los residuales para estimar el error del modelo no observable. La discusión detallada se proporcionará más adelante.

1.3 Propiedades estadísticas de la estimación por mínimos cuadrados

Primero discutiremos las propiedades estadísticas sin el supuesto de distribución del término de error. Pero asumiremos que $E(\epsilon_i) = 0$, $\text{Var}(\epsilon_i) = \sigma^2$ y ϵ_i para $i = 1, 2, \dots, n$ son independientes.

Teorema 1.1 El estimador de mínimos cuadrados b_0 es un estimador insesgado de β_0 .

Demostración.-

$$\begin{aligned}
 E(b_0) &= E(\bar{y} - b_1 \bar{x}) \\
 &= E\left(\frac{1}{n} \sum_{i=1}^n y_i\right) - E(b_1 \bar{x}) \\
 &= \frac{1}{n} \sum_{i=1}^n E(y_i) - \bar{x} E(b_1) \\
 &= \frac{1}{n} \sum_{i=1}^n (\beta_0 + \beta_1 x_i) - \beta_1 \bar{x} \\
 &= \frac{1}{n} \sum_{i=1}^n \beta_0 + \beta_1 \frac{1}{n} \sum_{i=1}^n x_i - \beta_1 \bar{x} \\
 &= \frac{1}{n} \sum_{i=1}^n \beta_0 + \beta_1 \frac{1}{n} \sum_{i=1}^n x_i - \beta_1 \frac{1}{n} \sum_{i=1}^n x_i \\
 &= \frac{n\beta_0}{n} \\
 &= \beta_0.
 \end{aligned}$$



Teorema El estimador de mínimos cuadrados b_1 es un estimador insesgado de β_1 .

1.2

Demostración.-

$$\begin{aligned}
 E(b_1) &= E\left(\frac{S_{xy}}{S_{xx}}\right) \\
 &= \frac{1}{S_{xx}} E\left[\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})\right] \\
 &= \frac{1}{S_{xx}} \frac{1}{n} E\left[\sum_{i=1}^n (x_i - \bar{x})y_i - \sum_{i=1}^n (x_i - \bar{x})\bar{y}\right] \\
 &= \frac{1}{S_{xx}} \frac{1}{n} E\left[\sum_{i=1}^n (x_i - \bar{x})y_i - \bar{y} \sum_{i=1}^n (x_i - \bar{x})\right] \quad \text{ya que } \bar{y} \text{ es constante.}
 \end{aligned}$$

Sabemos que $\sum_{i=1}^n (x_i - \bar{x}) = \sum_{i=1}^n x_i - n\bar{x} = \sum_{i=1}^n x_i - n\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = 0$, por lo que

$$\begin{aligned}
 &= \frac{1}{S_{xx}} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) E(y_i) \\
 &= \frac{1}{S_{xx}} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) (\beta_0 + \beta_1 x_i) \\
 &= \frac{1}{S_{xx}} \frac{1}{n} \left[\beta_0 \sum_{i=1}^n (x_i - \bar{x}) + \sum_{i=1}^n (x_i - \bar{x}) \beta_1 x_i \right] \\
 &= \frac{1}{S_{xx}} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) \beta_1 x_i \\
 &= \frac{1}{S_{xx}} \frac{1}{n} \left[\sum_{i=1}^n (x_i - \bar{x}) \beta_1 x_i - \sum_{i=1}^n (x_i - \bar{x}) \beta_1 \bar{x} \right] \\
 &= \frac{1}{S_{xx}} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) \beta_1 (x_i - \bar{x}) \\
 &= \frac{1}{S_{xx}} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \beta_1 \\
 &= \frac{S_{xx}}{S_{xx}} \beta_1 \\
 &= \beta_1.
 \end{aligned}$$

■

Teorema 1.3 $\text{Var}(b_1) = \frac{\sigma^2}{nS_{xx}}.$

Demostración.- Sea X_1, X_2, \dots, X_n IDD, con $\text{Var}(X_i) = \sigma_i^2$ para $i = 1, 2, \dots, n$. Si $\sum_{i=1}^n a_i X_i$. Entonces,

$$\text{Var} \left(\sum_{i=1}^n a_i X_i \right) = \sum_{i=1}^n a_i^2 \sigma_i^2.$$

Por lo tanto,

$$\begin{aligned} \text{Var}(b_1) &= \text{Var} \left(\frac{S_{xy}}{S_{xx}} \right) \\ &= \left(\frac{1}{S_{xx}} \right)^2 \text{Var} \left[\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x}) \right] \\ &= \frac{1}{S_{xx}^2} \text{Var} \left[\frac{1}{n} \sum_{i=1}^n y_i (x_i - \bar{x}) \right] \\ &= \frac{1}{S_{xx}^2} \frac{1}{n^2} \sum_{i=1}^n (x_i - \bar{x})^2 \text{Var}(y_i) \\ &= \frac{1}{S_{xx}^2} \frac{1}{n^2} \sum_{i=1}^n (x_i - \bar{x})^2 \sigma^2 \\ &= \frac{\sigma^2}{nS_{xx}}. \end{aligned}$$

■

Teorema 1.4 El estimador de mínimos cuadrados b_1 e \bar{y} no están correlacionados. Bajo el supuesto de normalidad de y_i para $i = 1, 2, \dots, n$, b_1 e \bar{y} se distribuyen normalmente y son independientes.

Demostración.-

$$\begin{aligned} \text{Cov}(b_1, \bar{y}) &= \text{Cov} \left(\frac{S_{xy}}{S_{xx}}, \bar{y} \right) \\ &= \frac{1}{S_{xx}} \text{Cov}(S_{xy}, \bar{y}) \\ &= \frac{1}{nS_{xx}} \text{Cov} \left[\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}), \bar{y} \right] \\ &= \frac{1}{n^2 S_{xx}} \text{Cov} \left[\sum_{i=1}^n (x_i - \bar{x}) y_i, \sum_{i=1}^n y_i \right] \\ &= \frac{1}{n^2 S_{xx}} \sum_{i=1}^n (x_i - \bar{x}) \text{Cov}(y_i, y_j) \end{aligned}$$

Notemos que $E(\epsilon_i) = 0$ y ϵ_i son independientes. De donde por definición de covarianza, podemos

escribir

$$\text{Cov}(y_i, y_j) = E \{ [y_i - E(y_i)][y_j - E(y_j)] \} = E(\epsilon_i, \epsilon_j) = \begin{cases} \sigma^2 & \text{si } i = j \\ 0 & \text{si } i \neq j. \end{cases}$$

Concluimos que

$$\text{Cov}(b_1, \bar{y}) = \frac{1}{n^2} S_{xx} \sum_{i=1}^n (x_i - \bar{x}) \sigma^2 = 0.$$

Recuerde que la correlación cero es equivalente a la independencia entre dos variables normales. Por lo tanto, concluimos que b_1 e \bar{y} son independientes. ■

Teorema 1.5 $\text{Var}(b_0) = \left(\frac{1}{n} + \frac{\bar{x}^2}{nS_{xx}} \right) \sigma^2.$

Demostración.- Sea, X_1, X_2, \dots, X_n IDD, y $E(X) = \mu$ y $\text{Var}(X) = \sigma^2$. Entonces, la media muestral \bar{X} es normal con media μ y varianza $\frac{\sigma^2}{n}$. Por lo tanto,

$$\begin{aligned} \text{Var}(b_0) &= \text{Var}(\bar{y} - b_1 \bar{x}) \\ &= \text{Var}(\bar{y}) + (\bar{x})^2 \text{Var}(b_1) \\ &= \frac{\sigma^2}{n} + \bar{x}^2 \frac{\sigma^2}{nS_{xx}} \\ &= \left(\frac{1}{n} + \frac{\bar{x}^2}{nS_{xx}} \right) \sigma^2 \end{aligned}$$

■

Las varianzas de b_0 y b_1 son importantes cuando queremos hacer inferencias estadísticas sobre la intersección y la pendiente de la regresión.

Dado que las varianzas de los estimadores de mínimos cuadrados b_0 y b_1 involucran la varianza del término de error en el modelo de regresión simple. Esta variación de error es desconocida para nosotros. Por lo tanto, necesitamos estimarlo. Ahora discutimos cómo estimar la varianza del término de error en el modelo de regresión lineal simple. Sea y_i la variable de respuesta observada y $\hat{y}_i = b_0 + b_1 x_i$, el valor ajustado de la respuesta. Tanto y_i como \hat{y}_i están disponibles para nosotros. El verdadero error σ_i en el modelo no es observable y nos gustaría estimarlo. La cantidad $y_i - \hat{y}_i$ es la versión empírica del error ϵ_i . Esta diferencia es un residuo de regresión que juega un papel importante en el diagnóstico del modelo de regresión. Proponemos la siguiente estimación de la varianza del error basada en e_i :

$$s^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Tenga en cuenta que en el denominador es $n-2$. Esto hace que s^2 sea un estimador insesgado de la varianza del error σ^2 . El modelo lineal tiene dos parámetros, por lo que, $n-2$ puede verse como n - números de parámetros simples. En particular, en un modelo de regresión lineal múltiple con p parámetros, el denominador debe ser $n-p$ para construir un estimador insesgado de la varianza del error σ^2 .

El estimador insesgado s^2 para la regresión lineal simple será demostrado en las siguientes derivaciones.

$$y_i - \hat{y}_i = y_i - b_0 - b_1 x_i = y_i - (\bar{y} - b_1 \bar{x}) - b_1 x_i$$

Estamos suponiendo que $E(\epsilon) = 0$; de lo que se sigue,

$$\sum_{i=1}^n (y_i - \hat{y}_i) = \sum_{i=1}^n (y_i - \bar{y}) = \sum_{i=1}^n (y_i - \bar{y}) - b_i \sum_{i=1}^n (x_i - \bar{x}) = 0.$$

Demostremos este supuesto. Note que $(y_i - \hat{y}_i) x_i = [(y_i - \bar{y}) - b_i (x_i - \bar{x})] x_i$, de donde

$$\begin{aligned} \sum_{i=1}^n (y_i - \hat{y}_i) x_i &= \sum_{i=1}^n [(y_i - \bar{y}) x_i - b_i (x_i - \bar{x})] x_i \\ \sum_{i=1}^n (y_i - \hat{y}_i) &= \sum_{i=1}^n [(y_i - \bar{y}) x_i - b_i (x_i - \bar{x})] (x_i - \bar{x}) \\ &= \sum_{i=1}^n (y_i - \bar{y}) (x_i - \bar{x}) - b_i \sum_{i=1}^n (x_i - \bar{x})^2 \\ &= n (S_{xy} - b_1 S_{xx}) \\ &= n \left(S_{xy} - \frac{S_{xy}}{S_{xx}} S_{xx} \right) \\ &= 0. \end{aligned}$$

Para demostrar que s^2 es un estimador insesgado de la varianza del error, primero veamos que

$$(y_i - \hat{y}_i)^2 = [(y_i - \bar{y}) - b_i (x_i - \bar{x})]^2,$$

Por lo que

$$\begin{aligned} \sum_{i=1}^n (y_i - \hat{y}_i)^2 &= \sum_{i=1}^n [(y_i - \bar{y}) - b_i (x_i - \bar{x})]^2 \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 - 2b_i \sum_{i=1}^n (x_i - \bar{x}) (y_i - \bar{y}) + b_i^2 \sum_{i=1}^n (x_i - \bar{x})^2 \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 - 2nb_i S_{xy} + nb_i^2 S_{xx} \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 - 2n \frac{S_{xy}}{S_{xx}} S_{xy} + n \frac{S_{xy}^2}{S_{xx}^2} S_{xx} \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 - n \frac{S_{xy}^2}{S_{xx}} \end{aligned}$$

Después, ya que

$$\begin{aligned} (y_i - \bar{y})^2 &= [\beta_0 + \beta_1 x_i + \epsilon_i - (\beta_0 + \beta_1 \bar{x} + \bar{\epsilon})]^2 \\ &= [\beta_1 (x_i - \bar{x}) + (\epsilon_i - \bar{\epsilon})]^2 \\ &= \beta_1^2 (x_i - \bar{x})^2 + (\epsilon_i - \bar{\epsilon})^2 + 2\beta_1 (x_i - \bar{x}) (\epsilon_i - \bar{\epsilon}) \end{aligned}$$

Entonces, en vista que $E(\epsilon_i - \bar{\epsilon}) = 0$ tenemos

$$\begin{aligned}
E(y_i - \bar{y})^2 &= \beta_1^2 (x_i - \bar{x})^2 + E(\epsilon_i - \bar{\epsilon})^2 \\
&= \beta_1^2 (x_i - \bar{x})^2 + E(\epsilon_i - \bar{\epsilon})^2 - E^2(\epsilon_i - \bar{\epsilon}) + E^2(\epsilon_i - \bar{\epsilon}) \\
&= \beta_1^2 (x_i - \bar{x})^2 + \text{Var}(\epsilon_i - \bar{\epsilon}) \\
&= \beta_1^2 (x_i - \bar{x})^2 + \text{Var}(\epsilon_i) + \text{Var}(\bar{\epsilon}) - 2\text{Cov}(\epsilon_i, \bar{\epsilon}) \\
&= \beta_1^2 (x_i - \bar{x})^2 + \sigma^2 + \frac{\sigma^2}{n} - 2\frac{\sigma^2}{n} \\
&= \beta_1^2 (x_i - \bar{x})^2 + \frac{n-1}{n}\sigma^2,
\end{aligned}$$

y

$$\begin{aligned}
\sum_{i=1}^n E(y_i - \bar{y})^2 &= n\beta_1^2 S_{xx} + \sum_{i=1}^n \frac{n-1}{n}\sigma^2 \\
&= n\beta_1^2 S_{xx} + (n-1)\sigma^2.
\end{aligned}$$

Además, se tiene

$$\begin{aligned}
E(S_{xy}) &= E\left[\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})\right] \\
&= \frac{1}{n} E\left[\sum_{i=1}^n (x_i - \bar{x}) y_i\right] \\
&= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) E(y_i) \\
&= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) (\beta_0 + \beta_1 x_i) \\
&= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) \beta_0 + \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) \beta_1 x_i \\
&= \frac{1}{n} \beta_0 \sum_{i=1}^n x_i - \frac{1}{n} \beta_0 \sum_{i=1}^n x_i + \frac{1}{n} \beta_1 \sum_{i=1}^n (x_i - \bar{x}) x_i \\
&= \frac{1}{n} \beta_0 \left[\sum_{i=1}^n (x_i - \bar{x}) x_i - \sum_{i=1}^n (x_i - \bar{x}) \bar{x} \right] \\
&= \frac{1}{n} \beta_0 \sum_{i=1}^n (x_i - \bar{x})^2 \\
&= \frac{1}{n} \beta_0 S_{xx}
\end{aligned}$$

También; sea X_1, X_2, \dots, X_n IDD, con $\text{Var}(X_i) = \sigma_i^2$ para $i = 1, 2, \dots, n$. Si $\sum_{i=1}^n a_i X_i$. Entonces,

$$\text{Var} \left(\sum_{i=1}^n a_i X_i \right) = \sum_{i=1}^n a_i^2 \sigma_i^2.$$

Por lo tanto,

$$\begin{aligned} \text{Var}(S_{xy}) &= \text{Var} \left[\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}) y_i \right] \\ &= \frac{1}{n^2} \sum_{i=1}^n (x_i - \bar{x})^2 \text{Var}(y_i) \\ &= \frac{1}{n} S_{xx} \sigma^2. \end{aligned}$$

Así, podemos escribir

$$\begin{aligned} \text{Var}(S_{xy}^2) &= E(S_{xy}^2) - E^2(S_{xy}) \\ E(S_{xy}^2) &= \text{Var}(S_{xy}) + E^2(S_{xy}) \\ &= \frac{1}{n} S_{xx} \sigma^2 + \beta_1^2 S_{xx}^2. \end{aligned}$$

y

$$E(S_{xy}^2) = \frac{1}{n} S_{xx} (\sigma^2 + n \beta_1^2 S_{xx})$$

Dado que $E(S_{xx}) = S_{xx}$, entonces

$$E \left(\frac{n S_{xy}^2}{S_{xx}} \right) = \sigma^2 + n \beta_1^2 S_{xx}.$$

Finalmente, $E(\hat{\sigma}^2)$ es dado por:

$$\begin{aligned} E \left[\sum_{i=1}^n (y_i - \hat{y}_i)^2 \right] &= E \left[\sum_{i=1}^n (y_i - \bar{y})^2 \right] - E \left[n \frac{S_{xy}^2}{S_{xx}} \right] \\ &= n \beta_1^2 S_{xx} + (n-1) \sigma^2 - n \beta_1^2 S_{xx} - \sigma^2 \\ &= (n-2) \sigma^2. \end{aligned}$$

En otras palabras, probamos que

$$E(s^2) = E \left[\frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right] = \sigma^2.$$

Por lo tanto, s^2 , la estimación de la varianza del error, es un estimador insesgado de la varianza del error σ^2 en la regresión lineal simple. Otra vista de elegir $n-2$ es que en el modelo de regresión lineal simple hay n observaciones y dos restricciones sobre estas observaciones:

$$\text{i) } \sum_{i=1}^n (y_i - \hat{y}) = 0,$$

$$\text{ii) } \sum_{i=1}^n (y_i - \hat{y}) x_i = 0.$$

Por lo tanto, la estimación de la varianza del error tiene $n - 2$ grados de libertad, que también es el número total de observaciones - el número total de parámetros en el modelo. Veremos características similares en la regresión lineal múltiple.

1.4 Estimación de máxima verosimilitud

El estimador de máxima verosimilitud de una regresión lineal simple puede ser desarrollado si se supone que la variable dependiente y_i tiene una distribución normal $y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$. La función de similitud para (y_1, y_2, \dots, y_n) es dada por:

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

Los estimadores de β_0 y β_1 que maximiza la función de similitud L son equivalentes a los estimadores que minimizan la parte exponencial de la función de verosimilitud lo que produce los mismos estimadores que los estimadores de mínimos cuadrados de la regresión lineal. Por lo tanto, bajo el supuesto de normalidad del término de error, los MLE de β_0 y β_1 y los estimadores de mínimos cuadrados de β_0 y β_1 son exactamente iguales.

Después de obtener b_0 y b_1 , los valores MLE de los parámetros β_0 y β_1 , podemos calcular el valor ajustado \hat{y}_i y la función de probabilidad en términos de los valores ajustados.

$$L = \prod_{i=1}^n f(y_i) = \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} e^{\left(-\frac{1}{2\sigma^2}\right) \sum_{i=1}^n (y_i - \hat{y}_i)^2}.$$

Luego, tomamos la derivada parcial con respecto a σ^2 en la función logarítmica de verosimilitud $\log(L)$ y la igualamos a cero:

$$\frac{\partial \log(L)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (y_i - \hat{y}_i)^2 = 0$$

La estimación de máxima verosimilitud MLE de σ^2 es $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$. Notemos que es un estimador sesgado de σ^2 . No así,

$$s^2 = \frac{1}{n-2} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

que es un estimador insesgado del error de la varianza σ^2 . $\frac{n}{n-2} \hat{\sigma}^2$ es un estimador insesgado de σ^2 . Observe también que $\hat{\sigma}^2$ es una estimación asintóticamente insesgada de σ^2 , lo que coincide con la teoría clásica de MLE.

1.5 Intervalo de confianza sobre la media de regresión y la predicción de regresión

Los modelos de regresión suelen construirse basándose en determinadas condiciones que deben verificarse para que el modelo se ajuste bien a los datos y pueda predecir la respuesta para un determinado regresor con la mayor precisión posible. Uno de los principales objetivos del análisis de regresión es utilizar el modelo de regresión ajustado para realizar predicciones.

El intervalo de confianza de la predicción de regresión permite evaluar la calidad de la predicción. A menudo interesan los siguientes intervalos de confianza de la predicción de regresión:

- Un intervalo de confianza para una sola línea de regresión.
- Un intervalo de confianza para un solo valor futuro de y correspondiente a un valor elegido de x .
- Una región de confianza para la línea de regresión como un todo.

Para analizar el intervalo de confianza para la línea de regresión, consideramos el valor ajustado de la línea de regresión en $x = x_0$, que es $\hat{y}(x_0) = b_0 + b_1 x_0$ y el valor medio de $x = x_0$ es $E(\hat{y}|x_0) = \beta_0 + \beta_1 x_0$. Tenga en cuenta que b_1 es independiente de \bar{y} de donde,

$$\begin{aligned} \text{Var}[\hat{y}(x_0)] &= \text{Var}(b_0 + b_1 x_0) \\ &= \text{Var}[\bar{y} - b_1(x_0 - \bar{x})] \\ &= \text{Var}(\bar{y}) + (x_0 - \bar{x})^2 \text{Var}(b_1) \\ &= \frac{1}{n} \sigma^2 + (x_0 - \bar{x})^2 \frac{1}{S_{xx}} \sigma^2 \\ &= \sigma^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right] \end{aligned}$$

Reemplazando σ por s , el error estandar de la predicción de regresión en x_0 está dado por:

$$s_{\hat{y}}(x_0) = s \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$$

Si $\epsilon \sim N(0, \sigma^2)$ el $(1 - \alpha)100\%$ del intervalo de confianza en $E(\hat{y}|x_0) = \beta_0 + \beta_1 x_0$ puede escribirse como:

$$\hat{y}(x_0) \pm t_{\alpha/2, n-2} s \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}.$$

Ahora, analicemos el intervalo de confianza en la predicción de la regresión. Denotemos la predicción de la regresión en x_0 por y_0 y supongamos que y_0 es independiente de $\hat{y}(x_0)$, donde $y(x_0) = b_0 + b_1 x_0$, y $E[y - \hat{y}(x_0)] = 0$. Ya que $\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - 2\text{Cov}(X, Y)$ con $\text{Cov}(X, Y) = 0$. Entonces,

$$\begin{aligned} \text{Var}[y_0 - \hat{y}(x_0)] &= \sigma^2 + \sigma^2 \left[\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right] \\ &= \sigma^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right] \end{aligned}$$

Bajo el supuesto de normalidad del término de error

$$\frac{y_0 - \hat{y}(x_0)}{\sigma \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}} \sim N(0, 1).$$

Luego, sustituyendo σ con s se tiene

$$\frac{y_0 - \hat{y}(x_0)}{s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}} \sim t_{n-2}.$$

Así, el $(1 - \alpha)100\%$ del intervalo de confianza en la predicción de la regresión en y_0 puede ser expresada como

$$\hat{y}_0 \pm t_{\alpha/2, n-2} s \sqrt{1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}$$

1.6 Inferencia estadística sobre parámetros de regresión

Para dividir la varianza total $\sum_{i=1}^n (y_i - \bar{y})^2$, consideremos la ecuación de regresión ajustada $\hat{y}_i = b_0 + b_1 x_i$, donde $b_0 = \bar{y} - b_1 \bar{x}$ y $b_1 = \frac{S_{xy}}{S_{xx}}$. Podemos escribir

$$\begin{aligned} \bar{\hat{y}} &= \frac{1}{n} \sum_{i=1}^n \hat{y}_i \\ &= \frac{1}{n} \sum_{i=1}^n [(\bar{y} - b_1 \bar{x}) + b_1 x_i] \\ &= \frac{1}{n} \sum_{i=1}^n [\bar{y} + b_1 (x_i - \bar{x})] \quad \text{ya que } \frac{1}{n} \sum x_i - \frac{1}{n} \sum \bar{x} = 0 \\ &= \bar{y}. \end{aligned}$$

Para la respuesta de la regresión y_i , la varianza total es $\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$. Note que la varianza total puede ser dividida en dos partes:

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 &= \frac{1}{n} \sum_{i=1}^n [(y_i - \hat{y}_i) + (\hat{y}_i - \bar{y})]^2 \\ &= \frac{1}{n} \sum_{i=1}^n [(y_i - \hat{y}_i)^2 + (\hat{y}_i - \bar{y})^2 + 2(\hat{y}_i - \bar{y})(y_i - \hat{y}_i)] \\ &= \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \\ &= SS_{Reg} + SS_{Res} \\ &= \text{Varianza explicada por la regresión} + \text{Varianza no explicada (del residuo)}. \end{aligned}$$

Se puede demostrar que $[2(y_i - \hat{y}_i)(\hat{y}_i - \bar{y})]$ es cero. Usando el hecho de que $\sum_{i=1}^n (y_i - \hat{y}_i) = 0$, tenemos

$$\begin{aligned}
\sum_{i=1}^n (\hat{y}_i - \bar{y}) (y_i - \hat{y}_i) &= \sum_{i=1}^n \hat{y}_i (y_i - \hat{y}_i) \\
&= \sum_{i=1}^n [(b_0 + b_1 x_i) (y_i - \hat{y}_i)] \\
&= b_0 \sum_{i=1}^n (y_i - \hat{y}_i) + b_1 \sum_{i=1}^n x_i (y_i - \hat{y}_i) \\
&= b_1 \sum_{i=1}^n x_i (y_i - \hat{y}_i) \\
&= b_1 \sum_{i=1}^n x_i (y_i - b_0 - b_1 x_i) \\
&= b_1 \sum_{i=1}^n x_i [(y_i - \bar{y}) - b_1 (x_i - \bar{x})] \\
&= b_1 \left[\sum_{i=1}^n x_i (y_i - \bar{y}) - \bar{x} (y_i - \bar{y}) - b_1 x_i (x_i - \bar{x}) - b_1 \bar{x} (x_i - \bar{x}) \right] \\
&= b_1 \left[\sum_{i=1}^n [(x_i - \bar{x}) (y_i - \bar{y})] - b_1 \sum_{i=1}^n (x_i - \bar{x})^2 \right] \\
&= b_1 (S_{xy} - b_1 S_{xx}) \\
&= b_1 \left[S_{xy} - \left(\frac{S_{xy}}{S_{xx}} \right) S_{xx} \right] \\
&= 0.
\end{aligned}$$

Los grados de libertad para SS_{Reg} y SS_{Res} se muestran a continuación:

$$\begin{aligned}
SS_{Total} &= SS_{Reg} + SS_{Res} \\
n - 1 &= 1 + n - 2
\end{aligned}$$

Para probar la hipótesis $H_0 : \beta_1 = 0$ contra $H_1 : \beta_1 \neq 0$ es necesario asumir que $\epsilon_i \sim N(0, \sigma^2)$. La siguiente tabla contiene las distribuciones de SS_{Reg} , SS_{Res} y SS_{Total} bajo la hipótesis H_0 .

SS	df	distribucion
SS_{Reg}	1	$\sigma^2 \chi_1^2$
SS_{Res}	$n - 2$	$\sigma^2 \chi_{n-2}^2$
SS_{Total}	$n - 1$	$\sigma^2 \chi_{n-1}^2$

El test estadístico está dado por:

$$F = \frac{\frac{SS_{Reg}}{1}}{\frac{SS_{Res}}{n-2}} \sim F_{1,n-2},$$

que es una prueba F unilateral superior. La tabla de abajo muestra un típico análisis de varianza (anova) de regresión

<i>Fuente</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>
<i>Regresion</i>	SS_{Reg}	1	$SS_{Reg}/1$	$\frac{SS_{Reg}/1}{s^2}$
<i>Residuo</i>	SS_{Res}	$n-2$	$SS_{Res}/(n-2) = s^2$	
<i>Total</i>	SS_{Total}	$n-1$		

Donde:

$$\begin{aligned} SS &= \text{Suma de cuadrados} \\ df &= \text{Grados de libertad} \\ MS &= \text{Media de cuadrados} \\ F &= \text{Estadístico } F. \end{aligned}$$

Para comprobar la pendiente de regresión β_1 . Por los teoremas 2.2 y 2.3, se observa que b_1 sigue la distribución normal

$$b_1 \sim N\left(\beta_1, \frac{\sigma^2}{SS_{xx}}\right)$$

y

$$\left(\frac{b_1 - \beta_1}{s/\sqrt{S_{xx}}}\right) = \left(\frac{b_1 - \beta_1}{s}\right) \sqrt{S_{xx}} \sim t_{n-2}$$

que puede utilizarse para comprobar $H_0 : \beta_1 = \beta_{1_0}$ contra $H_1 : \beta_1 \neq \beta_{1_0}$. Se puede utilizar un enfoque similar para comprobar el intercepto de la regresión. Bajo el supuesto de normalidad normalidad del término de error

$$b_0 \sim N\left[\beta_0, \sigma^2 \left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)\right].$$

Así, podemos usar el estadístico t test con $H_0 : \beta_0 = \beta_{0_0}$ contra $H_1 : \beta_0 \neq \beta_{0_0}$.

$$t = \frac{b_0 - \beta_0}{s \sqrt{\frac{1}{n} + \left(\frac{\bar{x}^2}{S_{xx}}\right)}} \sim t_{n-2}$$

Es sencillo usar las distribuciones de b_0 y b_1 para obtener los $(1 - \alpha)100\%$ intervalos de confianza de β_0 y β_1 :

$$b_0 \pm t_{\alpha/2, n-2} s \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}},$$

y

$$b_1 \pm t_{\alpha/2, n-2} s \sqrt{\frac{1}{S_{xx}}}.$$

Supongamos que la línea de regresión pasa a través de $(0, \beta_0)$. Es decir, el intercepto en y es una constante conocida β_0 . El modelo es dado por $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$ con una constante conocida β_0 . Usando el principio de mínimos cuadrados podemos estimar β_1 :

$$b_1 = \frac{\sum x_i y_i}{\sum x_i^2}$$

En consecuencia, el siguiente estadístico t test se puede utilizar para probar $H_0 : \beta_1 = \beta_{1_0}$ contra $H_1 : \beta_1 \neq \beta_{1_0}$. Bajo el supuesto de normalidad sobre ϵ_i

$$t = \frac{b_1 - \beta_{1_0}}{s} \sqrt{\sum_{i=1}^n x_i^2} \sim t_{n-1}$$

Tenga en cuenta que solo tenemos un parámetro para el modelo de regresión del intercepto y donde la estadística de t test tiene $n - 1$ grados de libertad, que es diferente del modelo lineal simple con 2 parámetros.

La cantidad R^2 , definida a continuación, es una medida de ajuste de regresión:

$$R^2 = \frac{SS_{Reg}}{SS_{Total}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{SS_{Res}}{SS_{Total}}$$

Notemos que $0 \leq R^2 \leq 1$ que representa la proporción de variación total explicada para el modelo de regresión.

la cantidad $CV = \frac{s}{\bar{y}} \times 100$ se llama coeficiente de variación, que también es una medida de la calidad del ajuste y representa la dispersión del ruido alrededor de la línea de regresión.

Ahora discutimos la inferencia simultánea en la regresión lineal simple. Tenga en cuenta que hasta ahora hemos discutido la inferencia estadística sobre β_0 y β_1 individualmente. La prueba individual significa que cuando probamos $H_0 : \beta_0 = \beta_{0_0}$ solo probamos esta H_0 independientemente de los valores de β_1 . Asimismo, cuando probamos $H_0 : \beta_1 = \beta_{1_0}$ solo probamos H_0 independientemente de los valores de β_0 . Si quisiéramos probar si una línea de regresión cae o no en cierta región, necesitamos probar la hipótesis múltiple:

$$H_0 : \beta_0 = \beta_{0_0}, \quad \beta_1 = \beta_{1_0}$$

simultáneamente. Esto cae en el ámbito de la inferencia múltiple. Para la inferencia múltiple sobre β_0 y β_1 notamos que

$$(b_0 - \beta_0, b_1 - \beta_1) \begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix} \begin{pmatrix} b_0 - \beta_0 \\ b_1 - \beta_1 \end{pmatrix} \sim 2s^2 F_{2, n-2}$$

Así, la región de confianza $(1 - \alpha)100\%$ de β_0 y β_1 es dado por

$$(b_0 - \beta_0, b_1 - \beta_1) \begin{pmatrix} n & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix} \begin{pmatrix} b_0 - \beta_0 \\ b_1 - \beta_1 \end{pmatrix} \leq 2s^2 F_{2, n-2},$$

donde $F_{\alpha, 2, n-2}$ es a cola superior del α -ésimo punto porcentual de la distribución F . Tenga en cuenta que esta región de confianza es una elipse.

1.7 Análisis residual y diagnóstico del modelo

Una forma de verificar el comportamiento de un modelo de regresión es a través del residuo de regresión, es decir, $e_i = y_i - \hat{y}_i$. Para la regresión lineal simple, un diagrama de dispersión de e_i frente a x_i proporciona un buen diagnóstico gráfico del modelo de regresión. Una distribución uniforme de los residuos en torno a la media cero indica un buen ajuste del modelo de regresión.

A continuación analicemos las características de los residuos si un modelo de regresión está mal especificado. Supongamos que el modelo correcto debería adoptar la forma cuadrática:

$$y_i = \beta_0 + \beta_1 (x_i - \bar{x}) + \beta_2 x_i^2 + \epsilon_i$$

con $E(\epsilon_i) = 0$. Luego, sea un modelo de regresión lineal especificado incorrectamente como:

$$y_i = \beta_0 + \beta_1 (x_i + \bar{x}) + \epsilon_i^*$$

Entonces, $\epsilon_i^* = \beta_2 x_i^2 + \epsilon_i$, el cual es desconocido. Ahora, la media del error es distinto de cero y es una función de x_i . Del modelo cuadrático se tiene:

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n \beta_0 + \frac{1}{n} \sum_{i=1}^n [\beta_1 (x_i - \bar{x})] + \frac{1}{n} \sum_{i=1}^n (\beta_2 x_i^2) + \frac{1}{n} \sum_{i=1}^n \epsilon_i \Rightarrow \bar{y} = \beta_0 + \beta_2 \bar{x}^2 + \bar{\epsilon}$$

tenemos:

$$b_0 = \bar{y} - b_1 (x_i - \bar{x}) = \bar{y} = \beta_0 + \beta_2 \bar{x}^2 + \bar{\epsilon}$$

y

$$\begin{aligned} \frac{S_{xy}}{S_{xx}} = b_1 &= \frac{\sum_{i=1}^n (x_i - \bar{x}) [\beta_0 + \beta_1 (x_i - \bar{x}) + \beta_2 x_i^2 + \epsilon_i]}{S_{xx}} \\ b_1 &= \beta_1 + \beta_2 \frac{\sum_{i=1}^n (x_i - \bar{x}) x_i^2}{S_{xx}} + \frac{\sum_{i=1}^n (x_i - \bar{x}) \epsilon_i}{S_{xx}} \end{aligned}$$

Luego, es fácil ver que

$$E(b_0) = \beta_0 + \beta_2 \bar{x}^2$$

y

$$E(b_1) = \beta_1 + \beta_2 \frac{\sum_{i=1}^n (x_i - \bar{x}) x_i^2}{S_{xx}}.$$

Por lo tanto, los estimadores b_0 y b_1 son estimaciones sesgadas de β_0 y β_1 . Supongamos que ajustamos el modelo por $\hat{y}_i = b_0 + b_1(x_i - \bar{x})$, el residuo de regresión esperado viene dado por

$$\begin{aligned} E(\epsilon_i) &= E(y_i - \hat{y}_i) = [\beta_0 + \beta_1(x_i - \bar{x}) + \beta_2 x_i^2] - [E(b_0) + E(b_1)(x_i - \bar{x})] \\ &= [\beta_0 + \beta_1(x_i - \bar{x}) + \beta_2 x_i^2] - [\beta_0 + \beta_2 \bar{x}^2] - \left[\beta_1 + \beta_2 \frac{\sum_{i=1}^n (x_i - \bar{x}) x_i^2}{S_{xx}} \right] (x_i - \bar{x}) \\ &= \beta_2 \left[(x_i^2 - \bar{x}^2) - \frac{\sum_{i=1}^n (x_i - \bar{x}) x_i^2}{S_{xx}} \right] \end{aligned}$$

Si $\beta_2 = 0$, entonces el modelo ajustado es correcto y $E(y_i - \hat{y}_i) = 0$. De lo contrario, el valor esperado de residual toma la forma cuadrática de x_i . Como resultado, la gráfica de residuos contra x_i debe tener una curvatura de apariencia cuadrática.

La inferencia estadística en el modelo de regresión se basa en el supuesto de normalidad del término de error. Los estimadores de mínimos cuadrados y los MLE de los parámetros de regresión son exactamente idénticos solo bajo el supuesto de normalidad del término de error. Ahora, la pregunta es ¿cómo verificar la normalidad del término de error? Considere el residuo $y_i - \hat{y}_i$: tenemos $E(y_i - \hat{y}_i) = 0$ y

$$\begin{aligned} \text{Var}(y_i - \hat{y}_i) &= \text{Var}(y_i) + \text{Var}(\hat{y}_i) - 2 \text{Cov}(y_i, \hat{y}_i) \\ &= \sigma^2 + \sigma^2 \left[\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right] - 2 \text{Cov}[y_i, \bar{y} + b_1(x_i - \bar{x})] \end{aligned}$$

Calculando $\text{Cov}[y_i, \bar{y} + b_1(x_i - \bar{x})]$, tenemos

$$\begin{aligned} \text{Cov}[y_i, \bar{y} + b_1(x_i - \bar{x})] &= \text{Cov}(y_i, \bar{y}) + (x_i - \bar{x}) \text{Cov}(y_i, b_1) \\ &= \frac{\sigma^2}{n} + (x_i - \bar{x}) \text{Cov}\left(y_i, \frac{S_{xy}}{S_{xx}}\right) \\ &= \frac{\sigma^2}{n} + (x_i - \bar{x}) \frac{1}{S_{xx}} \text{Cov}\left[\sum_{i=1}^n (x_i - \bar{x}) (y_i - \bar{y})\right] \\ &= \frac{\sigma^2}{n} + (x_i - \bar{x}) \frac{1}{S_{xx}} \left[\sum_{i=1}^n (x_i - \bar{x}) y_i \right] \\ &= \frac{\sigma^2}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \sigma^2. \end{aligned}$$

Así, la varianza del residuo es

$$\text{Var}(\epsilon_i) = \text{Var}(y_i - \hat{y}_i) = \sigma^2 \left[1 - \left(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right) \right],$$

que se puede estimar por

$$s_{e_i} = s \left[1 - \left(\frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}} \right) \right].$$

Si el término de error en la regresión lineal simple se especifica correctamente, es decir, el error se distribuye normalmente, los residuos estandarizados deberían comportarse como la variable aleatoria normal estándar. Por lo tanto, el cuantil de los residuos estandarizados en la regresión lineal simple será similar al cuantil de la variable aleatoria normal estandarizada. Así, la gráfica del cuantil de los residuos estandarizados versus el cuantil normal debe seguir una línea recta en el primer cuadrante si la suposición de normalidad en el término de error es correcta. Por lo general, se denomina gráfico normal y se ha utilizado como una herramienta útil para verificar la normalidad del término de error en la regresión lineal simple. Específicamente, podemos

- (1) Trazar el residuo ordenado $\frac{y_i - \hat{y}_i}{s}$ frente al cuantil de la normal $Z \left(\frac{i - 0.375}{n + 0.25} \right)$.
- (2) Trace el residuo estandarizado ordenado $\frac{y_i - \hat{y}_i}{s_{e_i}}$ contra el cuantil normal $Z \left(\frac{i - 0.375}{n + 0.25} \right)$.