Ayudantía 8 - Mat044

2 de julio de 2021

Estimación Máximo Verosímil

Considerando una muestra aleatoria $x_1, x_2, \ldots, x_n$ iid con función de probabilidad $f(\cdot)$ la cual pertenece a un modelo estadístico $\mathcal{F}_\Theta = \{f(\cdot|\theta), \theta \in \Theta\}$ de manera que $f(\cdot) = f(\cdot|\theta_0)$ donde θ_0 es el parámetro real de donde proviene la muestra.

Se desea inferir el parámetro $\hat{\theta}$ tal que $\hat{\theta}$ sea lo más cercano posible a θ_0 a través de la muestra $\{x_i\}_i$

Se tiene la densidad conjunta de la muestra:
$L(\theta_0|x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n|\theta_0) = \Pi_i f(x_i|\theta_0)$
pues es una muestra de variables independientes.

Donde $L(\theta_0|x_i)$ se le llama función de verosimilitud (función del parámetro vs la densidad conjunta que es función de la muestra)

Considerando la función de verosimilitud

$$L(\theta|x_i) = \Pi_i f(x_i|\theta)$$

se busca el parámetro $\hat{\theta}$ tal que maximice la probabilidad de obtener dicha muestra, entonces se obtiene el estimador máximo verosímil (MLE):

$$\hat{\theta}_{MLE} \in \arg\max_{\theta \in \Theta} L(\theta|x_i)$$

Cosiderando que $L(\theta)$ es una productoria, se tiene que $ln(L(\theta)) = \sum ln(f(\theta|x_i))$, luego, como ln es función monótona creciente, se tiene que

$$\hat{\theta}_{MLE} \in \arg\max_{\theta \in \Theta} L(\theta|x_i) \iff \hat{\theta}_{MLE} \in \arg\max_{\theta \in \Theta} ln(L(\theta|x_i))$$

por lo tanto se define la log-verosimilitud como:

$$\ell(\theta) = ln(L(\theta))$$

también a $\frac{\partial}{\partial \theta}\ell(\theta) = U(\theta)$ se le conoce como función Score.

Entonces el método de estimación máximo verisímil consiste en maximizar la función de logverosimilitud :

$$\hat{\theta}_{MLE} \in \arg\max_{\theta \in \Theta} \ell(\theta)$$

Regresión Lineal Simple

Considerando un modelo lineal simple $y = ax + c$, si tomamos una muestra aleatoria $y_1, y_2, \dots y_n$ donde se tiene que:

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

considerando x_i como variables regresoras (variables conocidas),
$(\beta_0, \beta_1) = \theta$ y $\Theta = \mathbb{R}^2$
(los ϵ_i se les conoce como errores y si $var(\epsilon_i) = \sigma^2 \ \forall i$ entonces se les llama homocedásticos)

Luego, se tiene que cada v.a. y_i tiene una distribución normal:

$$E(y_i) = E(\beta_0 + \beta_1 x_i + \epsilon_i) = \beta_0 + \beta_1 x_i$$

$$Var(y_i) = Var(\beta_0 + \beta_1 x_i + \epsilon_i) = Var(\epsilon_i) = \sigma^2$$

De lo anterior se tiene entonces que:

$$y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$$

por lo tanto la densidad de cada v.a. es

$$f_i(\theta|y_i) = \frac{1}{\sqrt{\pi\sigma^2}} exp\left\{-\frac{1}{2\sigma^2}(y_i - \beta_0 - \beta_1 x_i)^2\right\}$$

entonces, como son todas independientes la distribución conjunta:

$$L(\theta|y_1, y_2, \ldots, y_n) = \Pi_i \frac{1}{\sqrt{\pi\sigma^2}} exp\left\{-\frac{1}{2\sigma 2}(y_i - \beta_0 - \beta_1 x_i)^2\right\}$$

$$L(\theta|y_1, y_2, \ldots, y_n) = (\pi\sigma^2)^{-\frac{n}{2}} exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2\right\}$$

$$L(\theta|y_1, y_2, \ldots, y_n) = (\pi\sigma^2)^{-\frac{n}{2}} exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \right\}$$

Luego, la logverosimilitud:

$$\ell(\theta) = \ell(\beta_0, \beta_1) = -\frac{n}{2} ln(\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

Entonces, para encontrar el estimador máximo verosímil (considerando σ^2 conocido) se debe encontrar el máximo de $\ell(\theta)$:

$$\begin{aligned} \frac{\partial}{\partial\beta_0}\ell(\theta) &= -\tfrac{1}{2\sigma^2} \textstyle\sum_{i=1}^{n} -2(y_i - \beta_0 - \beta_1 x_i) \\ \frac{\partial}{\partial\beta_1}\ell(\theta) &= -\tfrac{1}{2\sigma^2} \textstyle\sum_{i=1}^{n} -2x_i(y_i - \beta_0 - \beta_1 x_i) \end{aligned}$$

se tiene que

$$\frac{\partial}{\partial\beta_0}\ell(\hat{\theta}_{MLE}) = 0$$
$$\frac{\partial}{\partial\beta_1}\ell(\hat{\theta}_{MLE}) = 0$$

$$\frac{\partial}{\partial \beta_0}\ell(\hat{\theta}_{MLE}) \quad = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} -(y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\iff$$

$$\sum_{i=1}^{n} -(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

equivalentemente

$$\frac{\partial}{\partial \beta_1}\ell(\hat{\theta}_{MLE}) \quad = -\frac{1}{2\sigma^2}\sum_{i=1}^{n} -2x_i(y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\iff$$

$$\sum_{i=1}^{n} -x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

por lo tanto se debe resolver

$$\begin{aligned} \sum_{i=1}^{n} x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) &= 0 \\ \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) &= 0 \end{aligned}$$

$$\begin{aligned} \sum_{i=1}^{n} x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) &= 0 \quad (1) \\ \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) &= 0 \quad (2) \end{aligned}$$

reeordendando el sistema:

$$\begin{aligned} \hat{\beta}_0 \sum x_i + \hat{\beta}_1 \sum x_i^2 &= \sum x_i y_i \\ \hat{\beta}_0 n + \hat{\beta}_1 \sum x_i &= \sum y_i \end{aligned}$$

$$\begin{pmatrix} \sum x_i & \sum x_i^2 \\ n & \sum x_i \end{pmatrix} \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}$$

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \frac{1}{(\sum x_i)^2 - n \sum x_i^2} \begin{pmatrix} \sum x_i & -\sum x_i^2 \\ -n & \sum x_i \end{pmatrix} \begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}$$

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \frac{1}{(\sum x_i)^2 - n \sum x_i^2} \begin{pmatrix} \sum x_i \sum x_i y_i - \sum y_i \sum x_i^2 \\ \sum x_i \sum y_i - n \sum x_i y_i \end{pmatrix}$$

Resumeindo y reeordenando se pueden reescribir como:

$$\hat{\beta}_1 = \frac{\sum x_i y_i - \frac{1}{n}\sum x_i \sum y_i}{\sum x_i^2 - \frac{1}{n}(\sum_x i)^2}$$

donde

$$\begin{aligned}\sum x_i y_i - \tfrac{1}{n}\sum x_i \sum y_i \quad &= \sum x_i y_i - n\bar{x}\bar{y} \\ &= \sum x_i y_i - \bar{x}\sum y_i - \bar{y}\sum x_i + n\bar{x}\bar{y} \\ &= \sum (x_i - \bar{x})(y_i - \bar{y}) \\ &= Q_{xy}\end{aligned}$$

por otro lado, se tiene que

$$\sum x_i^2 - \frac{1}{n}(\sum x_i)^2 = \sum (x_i - \bar{x})^2 = Q_{xx}$$

entonces, podemos resumir:

$$\hat{\beta}_1 = \frac{Q_{xy}}{Q_{xx}}$$

Análogamente:

$$\hat{\beta}_0 = \frac{\frac{1}{n}\sum x_i^2 \sum y_i - \frac{1}{n}\sum x_i \sum x_i y_i}{\sum x_i^2 - \frac{1}{n}(\sum x_i)^2}$$

donde es posible simplifcar:

$$\begin{aligned} \sum(x_i - \bar{x})^2 &= Q_{xx} = \sum x_i^2 - \frac{1}{n}(\sum x_i)^2 \\ \implies \sum x_i^2 &= Q_{xx} - \frac{1}{n}(\sum x_i)^2 \end{aligned}$$

entonces, se tiene que

$$\hat{\beta}_0 = \frac{\frac{1}{n}(Q_{xx} - \frac{1}{n}(\sum x_i)^2) \sum y_i - \frac{1}{n}\sum x_i \sum x_i y_i}{Q_{xx}}$$

Desarrollando el denomiandor:

$$= \bar{y} Q_{xx} + \frac{1}{n}\sum y_i \frac{1}{n}(\sum x_i)^2 - \frac{1}{n}\sum x_i \sum x_i y_i$$

$$= \bar{y} Q_{xx} + \bar{x}(\frac{1}{n}\sum x_i \sum y_i - \sum x_i y_i)$$

entonces, se tiene que

$$\hat{\beta}_0 = \frac{\bar{y}Q_{xx} + \bar{x}(\frac{1}{n}\sum x_i \sum y_i - \sum x_i y_i)}{Q_{xx}}$$

$$\hat{\beta}_0 = \frac{\bar{y}Q_{xx} + \bar{x}Q_{xy}}{Q_{xx}} = \bar{y} + \bar{x}\hat{\beta}_1$$

finalmente, vemos que son estimadores insesgados:

$$E(\hat{\beta}_1) = E\left(\frac{Q_{xy}}{Q_{xx}}\right) = \frac{1}{Q_{xx}}E(Q_{xy}) = \frac{1}{Q_{xx}}\sum(x_i - \bar{x})E(y_i - \bar{y})$$

donde $E(y_i) = \beta_0 + \beta_1 x_i$ y $E(\bar{y}) = \frac{1}{n}\sum \beta_0\beta_1 x_i = \beta_0 + \beta_1\bar{x}$ entonces,

$$E(\hat{\beta}_1) = \frac{1}{Q_{xx}}\sum(x_i - \bar{x})(\beta_0 + \beta_i x_i - \beta_0 - \beta_1\bar{x}) = \frac{1}{Q_{xx}}\beta_1\sum(x_i - \bar{x})(x_i - \bar{x})$$

$$= \beta_1\frac{Q_{xx}}{Q_{xx}} = \beta_1$$

por lo tanto, $\hat{\beta}_1$ es insesgado.

Luego, como $\hat{\beta}_0 = \bar{y} + \bar{x}\hat{\beta}_1$

$$E(\hat{\beta}_0) = E(\bar{y} + \bar{x}\hat{\beta}_1) = \beta_0 + \beta_1\bar{x} - \bar{x}\beta_1 = \beta_0$$

entonces, $\hat{\beta}_0$ también es insesgado.

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} \bar{y} + \bar{x}\, Q_{xy}/Q_{xx} \\ Q_{xy}/Q_{xx} \end{pmatrix}$$

donde

$$Q_{xx} = \sum (x_i - \bar{x})^2, \quad Q_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y})$$

y

$$E(\hat{\beta}_0) = \beta_0, \quad E(\hat{\beta}_1) = \beta_1$$