

MAT-206: Método de momentos

Felipe Osorio

`fosorios.mat.utfsm.cl`

Departamento de Matemática, UTFSM



Suponga X_1, \dots, X_n una muestra aleatoria desde el modelo estadístico

$$\mathcal{P} = \{P_\theta : \theta \in \Theta\}.$$

Deseamos entender el mecanismo o **modelo verdadero** (basado en $\theta_0 \in \Theta$) que **generó los datos**.

IDEA:

Desarrollar procedimientos para '**seleccionar**' θ desde los datos (estimación) así como discutir sus propiedades.



Definición 1:

Una función $T : \mathcal{X}^n \rightarrow \Gamma$ es llamado un **estimador** (puntual). Este es usado para estimar $\gamma = g(\theta)$.

El valor $T(x)$ es llamado **estimación** de $g(\theta)$ y corresponde a una realización de la variable aleatoria $T(\mathbf{X})$.

Observación:

Usualmente anotamos un **estimador** (y una estimación) como $\hat{\gamma} = T(X_1, \dots, X_n)$ y distinguimos el método usado como $\hat{\gamma}_{ML}$, $\hat{\gamma}_{MM}$ o $\hat{\gamma}_{LS}$.



Importante:

Este método **no** requiere **conocer** la distribución subyacente X ($\sim P_\theta \in \mathcal{P}$), sino que requiere **asumir** formas específicas para sus momentos.

Considere X_1, \dots, X_n una m.a.(n) desde P_θ (uni-dimensional). Es decir,

$$\mathcal{P} = \{P_\theta^{\otimes n} : \theta \in \Theta\},$$

y sea

$$\mu_k = \mu_k(P_\theta) = E(X^k) = \int x^k dP_\theta = \int x^k f_X(x; \theta) dx.$$

Asumiremos también que el parámetro de interés γ depende de θ a través de los momentos μ_k como:

$$\gamma = h(\mu_1(P_\theta), \dots, \mu_r(P_\theta)),$$

donde h es una función conocida.



Definición 2 (Estimador de momentos):

Suponga X_1, \dots, X_n que sigue el modelo estadístico $\{P_{\theta}^{\otimes n} : \theta \in \Theta\}$. El **estimador de momentos** es definido como

$$\hat{\gamma}_{\text{MM}} = h(m_1, \dots, m_r),$$

donde $m_k = \hat{\mu}_k$ es el **momento empírico** (o muestral) de orden k , dado por

$$m_k = \frac{1}{n} \sum_{i=1}^n x_i^k.$$



Ejemplo:

Considere X_1, \dots, X_n una muestra aleatoria desde P_θ y considere

$$\gamma = \int x f(x; \theta) dx.$$

Usando el método de momentos, tenemos que:

$$\hat{\gamma}_{\text{MM}} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}.$$



Método de momentos

Ejemplo:

Considere que el parámetro de interés es la desviación estándar σ . Note que

$$\sigma^2 = \int (x - \mu_1)^2 f(x; \theta) dx,$$

de este modo,

$$\sigma = \sqrt{\mu_2 - \mu_1^2} = h(\mu_1, \mu_2),$$

cuyo estimador usando el método de momentos adopta la forma:

$$\hat{\sigma}_{\text{MM}} = \sqrt{m_2 - m_1^2}.$$

Note que

$$\begin{aligned} m_2 - m_1^2 &= \frac{1}{n} \sum_{i=1}^n x_i^2 - \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2 = \frac{1}{n} \left(\sum_{i=1}^n x_i^2 - \frac{1}{n} (n\bar{x})^2 \right) \\ &= \frac{1}{n} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2. \end{aligned}$$

De este modo,

$$\hat{\sigma}_{\text{MM}} = \left\{ \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right\}^{1/2}.$$



Ejemplo:

El sesgo de una variable aleatoria X con distribución F es definida como

$$\gamma = \frac{E_F(X - E_F(X))^3}{(\text{var}_F(X))^{3/2}} = \frac{\mu_3 - 3\mu_2\mu_1 + 2\mu_1^3}{(\mu_2 - \mu_1^2)^{3/2}},$$

así el estimador de momentos asume la forma

$$\hat{\gamma}_{\text{MM}} = \frac{m_3 - 3m_2m_1 + 2m_1^3}{(m_2 - m_1^2)^{3/2}} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\}^{3/2}}$$

Ejemplo:

La función de distribución acumulada es definida como

$$F(t) = P_F((-\infty, t]),$$

para t fijo. Un estimador natural para la probabilidad del conjunto $(-\infty, t]$ es la frecuencia relativa,

$$\hat{F}_n(t) = \hat{F}_n(t; \mathbf{x}) = \frac{1}{n} \sum_{i=1}^n I_{(-\infty, t]}(x_i).$$

\hat{F}_n se denomina la **función de distribución empírica**. Note que

$$m_k = \mu_k(\hat{F}_n) = \int x^k d\hat{F}_n,$$

es decir, \hat{F}_n es un estimador de momentos de F .



Ejemplo:

Sea X_1, \dots, X_n una muestra aleatoria desde una distribución log-normal con vector de parámetros $\theta = (\mu, \sigma^2)^\top \in \mathbb{R}_+ \times \mathbb{R}_+$. Es decir, cada X_i tiene densidad

$$f(z; \mu, \sigma^2) = \frac{1}{\sigma z \sqrt{2\pi}} \exp \left\{ -\frac{(\log z - \mu)^2}{2\sigma^2} \right\}, \quad z > 0.$$

En este caso

$$\mu_1 = \exp(\mu + \sigma^2/2), \quad \mu_2 = \exp(\sigma^2) (\exp(\mu + \sigma^2/2))^2,$$

y portanto los estimadores de momentos son dados por:

$$\hat{\mu}_{\text{MM}} = 2 \log m_1 - \frac{1}{2} \log m_2, \quad \hat{\sigma}_{\text{MM}}^2 = \log m_2 - 2 \log m_1.$$



Ejemplo:

Suponga X_1, \dots, X_n una muestra aleatoria desde $\text{Poi}(\lambda)$, $\lambda > 0$. Recuerde que

$$E(X_1) = \text{var}(X_1) = \lambda.$$

De este modo, podemos considerar

$$\hat{\lambda}_{\text{MM}} = \frac{1}{n} \sum_{i=1}^n x_i \quad (= h(m_1))$$

por otro lado, otro estimador puede ser¹

$$\tilde{\lambda}_{\text{MM}} = h(m_1, m_2) = m_2 - m_1^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

¹¿Cuál $\hat{\lambda}_{\text{MM}}$ o $\tilde{\lambda}_{\text{MM}}$ es mejor?

En la práctica, tenemos que θ es vector p -dimensional. Por simplicidad asumiremos que $\gamma = \theta$. De este modo tenemos

$$\mu_1 = g_1(\theta_1, \dots, \theta_p),$$

$$\vdots$$

$$\mu_p = g_p(\theta_1, \dots, \theta_p).$$

Resolviendo para los p -parámetros en función de los momentos, obtenemos

$$\theta_1 = h_1(\mu_1, \dots, \mu_p),$$

$$\vdots$$

$$\theta_p = h_p(\mu_1, \dots, \mu_p).$$

(1)



Método de momentos

El estimador $\hat{\theta}_{\text{MM}}$, puede ser obtenido substituyendo en (1) por los momentos muestrales, es decir:

$$\begin{aligned}\hat{\theta}_1 &= h_1(m_1, \dots, m_p), \\ &\vdots \\ \hat{\theta}_p &= h_p(m_1, \dots, m_p).\end{aligned}$$

Debemos resaltar que el método de momentos requiere resolver el sistema de ecuaciones **no lineales**

$$\begin{aligned}g_1(\theta_1, \dots, \theta_p) - \mu_1 &= 0, \\ &\vdots \\ g_p(\theta_1, \dots, \theta_p) - \mu_p &= 0,\end{aligned}$$



El sistema de ecuaciones puede ser escrito como:

$$\Psi(\theta) = 0, \tag{2}$$

donde $\Psi : \mathbb{R}^p \rightarrow \mathbb{R}^p$ y por tanto $\hat{\theta}_{\text{MM}}$ corresponde a una raíz de la [ecuación de estimación](#) en (2).

En general, para obtener $\hat{\theta}_{\text{MM}}$ se requiere de [métodos iterativos](#), tal como:

$$\theta^{(s+1)} = \theta^{(s)} - \{\dot{\Psi}(\theta^{(s)})\}^{-1} \Psi(\theta^{(s)}), \quad s = 0, 1, \dots,$$

donde $\theta^{(s)}$ representa una estimación para θ en la etapa s -ésima y

$$\dot{\Psi}(\theta) = \frac{\partial \Psi(\theta)}{\partial \theta^\top}.$$



Observación:

Una dificultad evidente del método de momentos es que $\Psi(\theta) = 0$ puede tener **múltiples raíces**.

Importante:

El método de momentos es un caso particular de una **función de inferencia**. Más adelante, se estudiará las **propiedades** de estimadores obtenidos como la raíz de una función de inferencia.

