

# XS3310 Teoría Estadística

I Semestre 2021

Escuela de Estadística

2021-04-26

## ¿Qué vamos a discutir hoy?

- Mapa de recursos ¿dónde encuentro qué?
- Estimadores de momentos
- Definición 5.2 momento al origen
- Definición 5.3 momento central
- Definición 5.4 momento muestral
- Ejemplo de cálculo (2)
- Estimadores de máxima verosimilitud
- Definición
- Ejemplo simple: Motivación
- Otro ejemplo
- Propiedades de los estimadores Máximo verosímiles
- Asignación de la tarea o examen corto 1

## Vídeos

- Método de momentos
  - Comentarios sobre el método de momentos
  - Método de máxima verosimilitud
  - Otros ejemplos del método de máxima verosimilitud
  - Comentarios sobre el método de máxima verosimilitud
  - Funciones parametrales y máxima verosimilitud
  - Principio de invarianza
- 

## Estimadores de momentos

**Definición 5.2.** Momento al origen. Si  $X$  es una variable aleatoria con distribución conocida, se define el  $k$ -ésimo momento al origen como:  $\mu'_k = E[X^k]$

**Definición 5.3.** Momento central. Si  $X$  es una variable aleatoria con distribución conocida y media  $\mu$ , se define el  $k$ -ésimo momento central como:

$$\mu_k = E[(X - \mu)^k]$$

Para este tema también vamos a definir un nuevo tipo de momento, que llamaremos momento muestral:

**Definición 5.4.** Momento muestral. Si  $X_1, X_2, \dots, X_n$  es una muestra aleatoria de una distribución conocida, entonces se define el  $k$ -ésimo momento muestral como:

$$m'_k = \frac{1}{n} \sum_{j=1}^n X_j^k$$

---

## Estimadores de momentos

Si nos basamos en la idea de que los momentos muestrales son buenos estimadores de los momentos poblacionales al origen, entonces para determinar el estimador de momentos (EM) en relación con un parámetro desconocido  $\theta$ , se resuelve la ecuación generada al igualar el momento (o momentos) poblacional al origen con el momento muestral y resolviendo para  $\theta$ .

**Ejemplo.** Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria tal que  $X_j \sim Unif(0, \theta)$ . Encuentre el estimador de momentos para  $\theta$ .

---

## Estimadores de momentos

**Solución.** Empezando por el momento al origen tenemos:

$$\mu'_1 = E[X^1] = E[X] = \frac{\theta}{2}$$

El primer momento muestral siempre va a ser  $\bar{X}$  por lo que en este caso tenemos la siguiente igualdad:

$$\mu'_1 = m'_1$$

$$\Rightarrow \frac{\theta}{2} = \bar{X}$$

$$\Rightarrow \hat{\theta} = 2\bar{X}$$

Por lo tanto concluimos que  $\hat{\theta} = 2\bar{X}$  es el estimador de momentos para  $\theta$ . Se puede demostrar que este es un estimador insesgado y consistente para  $\theta$ .

---

## Estimadores de momentos

**Ejemplo.** Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria tal que  $X_j \sim \text{Gamma}(\alpha, \beta)$ . Determine un estimador de momentos para  $\alpha$  y para  $\beta$ .

**Solución.** En los casos en que se requiere estimar más de un parámetro se debe utilizar un número de igualdades similar al número de parámetros desconocidos. En este caso tenemos:

$$\begin{aligned}\mu'_1 = E(X) &= \alpha\beta & m'_1 &= \overline{X} \\ \mu'_2 = E(X^2) &= \alpha\beta^2 + (\alpha\beta)^2 & m'_2 &= \frac{1}{n} \sum_{j=1}^n X_j^2 = \overline{X^2}\end{aligned}$$

Ahora debemos resolver el siguiente sistema de ecuaciones:

$$\begin{aligned}\alpha\beta &= \overline{X} \\ \alpha\beta^2 + (\alpha\beta)^2 &= \overline{X^2}\end{aligned}$$

Podemos sustituir  $\alpha\beta$  de la primera ecuación en la segunda y despejar  $\beta$ :

$$\begin{aligned}\overline{X}\beta + \overline{X}^2 &= \overline{X^2} \\ \Rightarrow \hat{\beta} &= \frac{\overline{X^2} - \overline{X}^2}{\overline{X}}\end{aligned}$$

---

## Estimadores de momentos

Sustituyendo este resultado en la primera ecuación podemos obtener que

$$\hat{\alpha} = \frac{\overline{X}^2}{\overline{X^2} - \overline{X}^2}$$

Concluimos que  $\hat{\alpha}$  y  $\hat{\beta}$  son los estimadores de momentos de  $\alpha$  y  $\beta$  respectivamente.

Si se quisieran simplificar un poco estas expresiones se puede demostrar que

$$\hat{\alpha} = \frac{n\overline{X}^2}{(n-1)S^2}$$

$$\hat{\beta} = \frac{(n-1)S^2}{n\overline{X}}$$

---

## Estimadores de máxima verosimilitud

Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria sobre una población con distribución que incluye parámetros  $\theta_1, \theta_2, \dots, \theta_k$ . La función de verosimilitud  $\mathcal{L}(x_1, \dots, x_n | \theta_1, \theta_2, \dots, \theta_k)$ , mejor expresada como  $\mathcal{L}(\theta_1, \theta_2, \dots, \theta_k)$ , se dice que es una función de los parámetros para un específico resultado de la muestra aleatoria.

**Ejemplo.** Suponga que tenemos una función de densidad discreta (que toma valores 1, 2, 3 y 4) que depende de un parámetro  $\theta$  que solo puede tomar tres valores (0,1,2). Su función de probabilidad viene dada en la siguiente tabla:

Función de probabilidad

y	1	2	3	4
$f(y; 0)$	.980	.005	.005	.010
$f(y; 1)$	.100	.200	.200	.500
$f(y; 2)$	.098	.001	.001	.900

---

## Estimadores de máxima verosimilitud

Supongamos que obtuvimos una muestra aleatoria de tamaño tres donde observamos los siguientes datos: 4,4,3,4. ¿Cuál es el valor de  $\theta$  que maximiza la función de verosimilitud?

Solución. Debemos obtener la verosimilitud para la muestra obtenida para cada uno de los valores de  $\theta$ . El valor de  $\theta$  que genere la máxima verosimilitud será el estimador de máxima verosimilitud.

Recordemos que la función de verosimilitud es la probabilidad conjunta de la muestra observada. En este caso viene dada por:

$$\mathcal{L}(\theta) = P(Y = 4|\theta) \cdot P(Y = 4|\theta) \cdot P(Y = 3|\theta) \cdot P(Y = 4|\theta) = P(Y = 4|\theta)^3 \cdot P(Y = 3|\theta)$$

Debemos encontrar este valor para cada uno de los posibles valores de  $\theta$ :

$$\mathcal{L}(0) = (0.010)^3 \cdot (0.005) = 5 \cdot 10^{-09}$$

$$\mathcal{L}(1) = (0.500)^3 \cdot (0.200) = 0.025$$

$$\mathcal{L}(2) = (0.900)^3 \cdot (0.001) = 0.00729$$

## Estimadores de máxima verosimilitud

En este caso obtenemos la mayor verosimilitud cuando  $\theta = 1$ , por lo tanto el estimador de máxima verosimilitud sería  $\hat{\theta} = 1$ .

En el caso en que  $\theta$  sea una variable cuyo dominio es continuo, para encontrar el estimador de  $\theta$  basta con optimizar la verosimilitud siempre que existan los máximos correspondientes.

**Ejemplo.** Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria tal que  $X_j \sim \text{Bernoulli}(p)$ . Determinar el estimador de máxima verosimilitud (EMV) para  $p$ .

**Solución.** Ya anteriormente habíamos obtenido la función de verosimilitud de una muestra aleatoria Bernoulli,

$$\mathcal{L}(p) = p^u (1-p)^{n-u} \quad \text{donde} \quad u = \sum_{j=1}^n x_j$$

## Estimadores de máxima verosimilitud

Derivamos esta función con respecto a  $p$  e igualamos a cero. No obstante podemos notar como esto va a ser un proceso tedioso pues existe un producto de distintas funciones de  $p$ . Para facilitar este cálculo podemos utilizar una propiedad matemática que establece que si  $f(x)$  tiene un punto extremo en  $x_0$  y  $f(x) > 0$  entonces  $\ln f(x)$  también tiene un valor extremo en  $x_0$ . Por lo tanto podemos hacer uso de la log-verosimilitud:

$$\ell(p) = \ln \mathcal{L}(p) = u \ln p + (n - u) \ln(1 - p)$$

$$\Rightarrow \frac{\partial \ln \mathcal{L}}{\partial p} = \frac{u}{p} - \frac{n - u}{1 - p} = 0 \Rightarrow \frac{u(1 - p) - p(n - u)}{p(1 - p)} = 0$$

$$\Rightarrow u - up - np + up = 0 \Rightarrow u = np \Rightarrow \hat{p} = \frac{u}{n} = \frac{\sum_{j=1}^n x_j}{n}$$

El valor crítico es  $\hat{p} = \frac{\sum_{j=1}^n x_j}{n}$  pero falta demostrar que efectivamente sea un máximo,

$$\frac{\partial^2 \ln \mathcal{L}}{\partial p^2} = \frac{-u}{p^2} - \frac{n - u}{(1 - p)^2} < 0$$

Por lo tanto, concluimos que  $\hat{p}$  es el estimador de máxima verosimilitud.

---

## Estimadores de máxima verosimilitud

**Ejemplo.** Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria tal que  $X_j \sim N(\mu, \sigma^2)$ . Encuentre los estimadores máximo-verosímiles para  $\mu$  y  $\sigma^2$ .

**Solución.** Empezemos por obtener la función de verosimilitud:

$$\mathcal{L}(\mu, \sigma^2) = \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_j - \mu)^2}{2\sigma^2}} = (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{n}{2}} e^{-\frac{\sum (x_j - \mu)^2}{2\sigma^2}}$$

Obteniendo la log-verosimilitud:

$$\ell(\mu, \sigma^2) = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{\sum (x_j - \mu)^2}{2\sigma^2}$$

Derivando con respecto a  $\mu$ :

$$\frac{\partial \ell(\mu, \sigma^2)}{\partial \mu} = \frac{\sum (x_j - \mu)}{\sigma^2} = 0 \Rightarrow \sum (x_j - \mu) = 0 \Rightarrow \hat{\mu} = \bar{X}$$

$$\frac{\partial^2 \ell(\mu, \sigma^2)}{\partial \mu^2} = -\frac{n}{\sigma^2} < 0$$

---

## Estimadores de máxima verosimilitud

Ahora procedemos a derivar con respecto a  $\sigma^2$ . Para no confundirnos llamemos  $\omega = \sigma^2$ .

$$\frac{\partial \ell(\mu, \omega)}{\partial \omega} = -\frac{n}{2\omega} + \frac{\sum (x_j - \mu)^2}{2\omega^2} = \frac{-n\omega + \sum (x_j - \mu)^2}{2\omega^2} = 0$$

$$\Rightarrow \sum (x_j - \mu)^2 = n\omega \Rightarrow \hat{\omega} = \hat{\sigma}^2 = \frac{\sum (x_j - \mu)^2}{n}$$

$$\frac{\partial^2 \ell(\mu, \sigma^2)}{\partial \omega^2} = \frac{n}{2\omega^2} - \frac{\sum (x_j - \mu)^2}{\omega^3} < 0 \quad \text{cuando} \quad \omega = \frac{\sum (x_j - \mu)^2}{n}$$

---

## Estimadores de máxima verosimilitud

Nótese que este estimador está en términos de un parámetro desconocido,  $\mu$ , por lo que podemos sustituirlo por su respectivo estimador de máxima verosimilitud. Por lo tanto obtenemos el estimador

$$\hat{\sigma}^2 = \frac{\sum (x_j - \bar{X})^2}{n} = \frac{(n-1)}{n} S^2$$



Concluimos que  $\overline{X}$  y  $\frac{(n-1)}{n}S^2$  son los estimadores de máxima verosimilitud de  $\mu$  y  $\sigma^2$ , respectivamente.

---

## Estimadores de máxima verosimilitud

**Ejemplo.** Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria tal que  $X_j \sim Unif(0, \theta)$ . Encuentre el estimador de máxima verosimilitud de  $\theta$ .

**Solución.** Empezemos por obtener la función de verosimilitud y log-verosimilitud:

$$\mathcal{L}(\theta) = \prod_{j=1}^n \frac{1}{\theta} = \theta^{-n}$$

$$\ell(\theta) = -n \ln \theta$$

Derivando con respecto a  $\theta$ :

$$\frac{\partial \ell(\theta)}{\partial \theta} = \frac{-n}{\theta} < 0$$

---

## Estimadores de máxima verosimilitud

Esto significa que  $\ell(\theta)$  no tiene puntos críticos pero sí podemos encontrar un máximo local. Como es una función decreciente sabemos que el máximo se encuentra en el valor mínimo que  $\theta$  puede tomar. Recordemos que la verosimilitud es una función de  $\theta$  para un resultado específico de la muestra aleatoria, por lo que los valores de  $\theta$  van a depender de los valores de la muestra aleatoria. Siendo  $\theta$  el máximo poblacional es lógico pensar que para una muestra aleatoria dada lo mínimo que puede ser este valor es el máximo muestral  $X_{(n)}$ . Por lo tanto, concluimos que el estimador de máxima verosimilitud de  $\theta$  es  $X_{(n)}$ .

---

## Estimadores de máxima verosimilitud

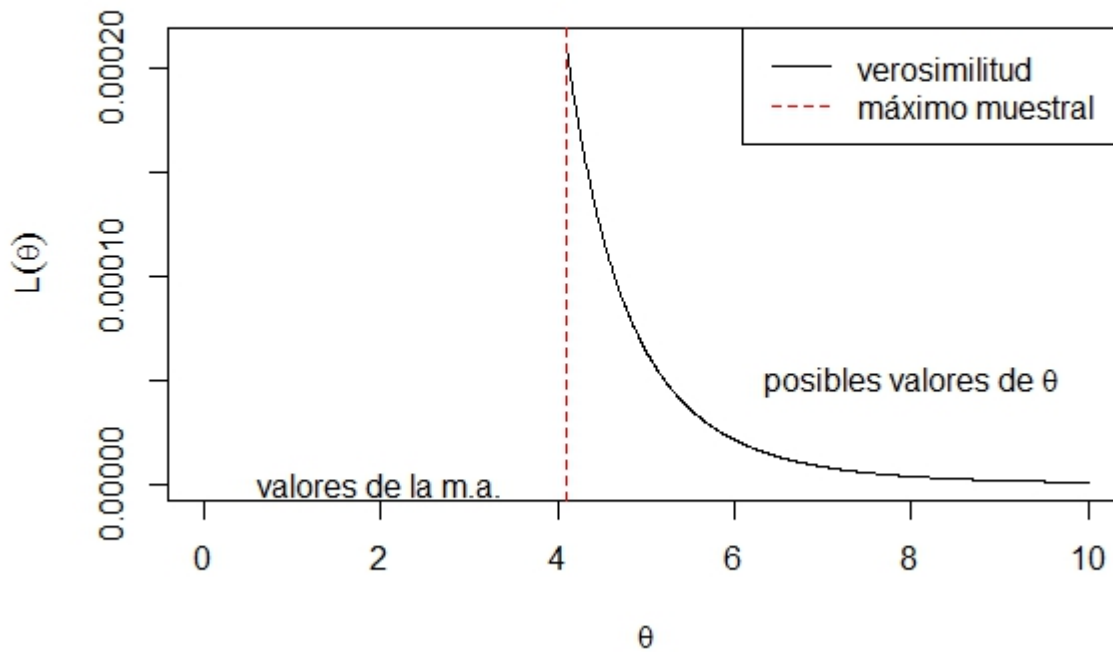


Figure 1: Figura 4. Verosimilitud de una  $Unif(0, \theta)$  a partir de una muestra aleatoria.

---

## Propiedades de los estimadores de máxima verosimilitud

- Si  $\mathcal{L}(\theta)$  es función de verosimilitud entonces para un estadístico suficiente  $U = T(X_1, X_2, \dots, X_n)$  para estimar  $\theta$  se cumple,

$$\begin{aligned}\mathcal{L}(\theta) &= g(u, \theta) \cdot h(x_1, x_2, \dots, x_n) \\ \Rightarrow \ell(\theta) &= \ln(g(u, \theta)) + \ln(h(x_1, x_2, \dots, x_n)) \\ \Rightarrow \frac{\partial \ell(\theta)}{\partial \theta} &= \frac{g'(u, \theta)}{g(u, \theta)}\end{aligned}$$

Con lo anterior se demuestra que el EMV de  $\theta$  debe ser función de  $U$ . Por el Teorema de Rao-Blackwell se cumple que son estimadores de variancia mínima.

- Se puede demostrar que los EMV son asintóticamente insesgados.
  - **Principio de Invariancia:** Si  $\hat{\theta}$  es EMV de  $\theta$  y  $t(\theta)$  es una función inyectiva de  $\theta$  entonces  $t(\hat{\theta})$  es el EMV de  $t(\theta)$ .
- 

## Propiedades de los estimadores de máxima verosimilitud

**Ejemplo.** Sea  $X_1, X_2, \dots, X_n$  una muestra aleatoria Normal con  $\mu$  y  $\sigma^2$  desconocidos. Encuentre el EMV de  $\sigma$ .

**Solución.** Se demostró con anterioridad que  $\hat{\sigma}^2 = \frac{(n-1)}{n} S^2$  es el EMV para  $\sigma^2$ . Sabiendo que la función  $t(x) = \sqrt{x}$  es continua en los reales positivos entonces por la propiedad de Invariancia se cumple que  $\sqrt{(\hat{\sigma}^2)} = \sqrt{\frac{(n-1)}{n}} S$  es el EMV de  $\sigma$ .

## Ejercicios

9.80 Suponga que  $Y_1, Y_2, \dots, Y_n$  denotan una muestra aleatoria de la distribución de Poisson con media  $\lambda$ .

- Encuentre el MLE  $\hat{\lambda}$  para  $\lambda$ .
- Encuentre el valor esperado y la varianza de  $\hat{\lambda}$ .
- Demuestre que el estimador del inciso a es consistente para  $\lambda$ .
- ¿Cuál es el MLE para  $P(Y = 0) = \exp(-\lambda)$ ?

9.81 Suponga que  $Y_1, Y_2, \dots, Y_n$  denotan una muestra aleatoria de una población distribuida exponencialmente con media  $\theta$ . Encuentre el MLE de la varianza poblacional  $\theta^2$ . [Sugerencia: recuerde el Ejemplo 9.9. de Mendenhall]

---

class: center, middle

## ¿Qué discutimos hoy?

Método de momentos. Método de máxima verosimilitud. Principio de invariancia.