

Introducción a la Probabilidad y Estadística - Probabilidad y Estadística

Grisel Britos

2023

ESTIMACIÓN PUNTUAL

Comenzamos el curso con Estadística Descriptiva, en donde a partir de un conjunto de datos se pretendía resumir información del fenómeno. Después construimos modelos y variables aleatorias que nos permitan describir ciertos fenómenos aleatorios. También vimos distribuciones de variables aleatorias que dependían de parámetros conocidos, por ejemplo: p en el caso Binomial, λ para una Poisson, μ y σ^2 para la normal, etc.

La situación ahora será diferente, supondremos que tenemos variables aleatorias con distribución conocida pero el o los parámetros son desconocidos y se pretende dar estimaciones para los mismos.

Los procesos de estimación para los parámetros desconocidos forman parte de la Estadística Paramétrica.

Definición: Sea X_1, \dots, X_n una muestra aleatoria (m.a.) con distribución que depende de θ , parámetro desconocido.

Se dice que $\hat{\theta} = h(X_1, \dots, X_n)$ es un **estimador puntual para θ** si $\hat{\theta}$ es una variable aleatoria.

Considerando X_1, \dots, X_n una m.a. con distribución que depende de θ , daremos algunos criterios para elegir entre dos estimadores puntuales para θ .

Definición: Se dice que un estimador puntual $\hat{\theta} = h(X_1, \dots, X_n)$ es un **estimador insesgado para θ** si $E(\hat{\theta}) = \theta$.

Algunos estimadores insesgados se pueden obtener del siguiente resultado:

Proposición: Sea X_1, \dots, X_n una m.a. tal que $E(X_i) = \mu$ y $V(X_i) = \sigma^2$ para todo $i = 1, \dots, n$. Entonces:

- a) \bar{X} es un estimador insesgado para μ .
- b) S_{n-1}^2 es un estimador insesgado para σ^2 .

Ejercicios:

Consideremos que X_1, \dots, X_n una muestra aleatoria con las diferentes distribuciones planteadas en cada uno de los ítems. Determine si los estimadores presentados en cada caso son insesgados.

- I) Distribución $Be(p)$ con $p \in (0, 1)$. $\hat{p} = \bar{X}$
- II) Distribución $\mathcal{P}(\lambda)$ con $\lambda > 0$. $\hat{\lambda} = \bar{X}$
- III) Distribución $U[0, \theta]$ con $\theta > 0$. $\hat{\theta}_1 = \bar{X}$ y $\hat{\theta}_2 = \max(X_1, \dots, X_n)$
- IV) Distribución $N(\mu, \sigma^2)$ con $\mu \in \mathbb{R}$ y $\sigma^2 > 0$. $\hat{\mu} = \bar{X}$ y $\hat{\sigma}^2 = S_{n-1}^2$

Observación: En el ejercicio III) se encontraron dos estimadores insesgados para el parámetro θ :

$$\hat{\theta}_1^* = 2\bar{X} \quad \text{y} \quad \hat{\theta}_2^* = \frac{n+1}{n} \max(X_1, \dots, X_n)$$

Entonces nos podemos preguntar: ¿cuál elegir entre ellos?

A continuación daremos un criterio que nos permitirá elegir entre estimadores insesgados para el mismo parámetro.

Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son estimadores para θ , vamos a elegir aquel que posea la **menor varianza**.
Por ejemplo, si

$$V(\hat{\theta}_1) \geq V(\hat{\theta}_2)$$

entonces decimos que $\hat{\theta}_2$ es mejor estimador que $\hat{\theta}_1$ para estimar θ .

Volviendo al Ejercicio III), entre los dos estimadores insesgados para θ , ¿cuál elegiría?.

Teorema: Sea X_1, \dots, X_n una muestra aleatoria con distribución $N(\mu, \sigma^2)$. Entonces \bar{X} es el estimador de mínima varianza uniforme entre todos los insesgados para μ (se dice que \bar{X} es IMVU).
Es decir, si $h(X_1, \dots, X_n)$ es un estimador insesgado para μ , entonces $V(h(X_1, \dots, X_n)) \geq V(\bar{X}) \quad \forall \mu \in \mathbb{R}$.

Definición: Sea $\hat{\theta} = h(X_1, \dots, X_n)$ un estimador puntual para θ . Llamaremos **Error Estándar del Estimador** a $\sigma_{\hat{\theta}} = \sqrt{V(\hat{\theta})}$.

Cuando el error estándar del estimador depende de parámetros desconocidos y son reemplazados por estimaciones, se lo llama **Error Estándar Estimado**, denotado por $\hat{\sigma}_{\hat{\theta}} = \sqrt{\widehat{V(\hat{\theta})}}$.

Volvamos a los ejercicios anteriores y calculemos los errores estándar de los estimadores dados.

A continuación mencionaremos dos de los métodos de mayor difusión para obtener estimadores de parámetros. Ellos son el Método de los Momentos (MM) y el Método de Máxima Verosimilitud (MV).

Método de los Momentos

Definición: Sea X_1, \dots, X_n una muestra aleatoria entonces se define el **k -ésimo momento poblacional** a $E(X^k)$ y el **k -ésimo momento muestral** a $\frac{1}{n} \sum_{i=1}^n X_i^k$, $\forall k \in \mathbb{N}$

Descripción del MM:

Sea X_1, \dots, X_n una muestra aleatoria con distribución que depende de m parámetros desconocidos $\underline{\theta} = (\theta_1, \dots, \theta_m)$.

El método consiste en plantear m ecuaciones dadas por:

$$\begin{aligned} E(X) &= \frac{1}{n} \sum_{i=1}^n X_i \\ E(X^2) &= \frac{1}{n} \sum_{i=1}^n X_i^2 \\ &\vdots \\ E(X^m) &= \frac{1}{n} \sum_{i=1}^n X_i^m \end{aligned}$$

Entonces la o las soluciones a este sistema, $\hat{\theta}_1, \dots, \hat{\theta}_m$ son los **Estimadores por el Método de los Momentos** para $\underline{\theta} = (\theta_1, \dots, \theta_m)$.

Ejercicio: Hallar los estimadores por el MM para las distribuciones dadas en los ejercicios I), II), III) y IV).

Método de Máxima Verosimilitud (MV)

Descripción del método de MV: Sean X_1, \dots, X_n una muestra aleatoria con función probabilidad de masa conjunta $p_{\underline{X}}(\cdot, \underline{\theta})$ o función de densidad conjunta $f_{\underline{X}}(\cdot, \underline{\theta})$ según se trate de v.a. discretas o v.a. continuas respectivamente, que dependen de m parámetros desconocidos $\underline{\theta} = (\theta_1, \dots, \theta_m)$. Diremos que $\hat{\underline{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_m)$ es el **Estimador de Máxima Verosimilitud** (MV) si cumple

- Caso discreto: $p_{\underline{X}}(\underline{x}, \hat{\underline{\theta}}) \geq p_{\underline{X}}(\underline{x}, \underline{\theta}) \quad \forall \underline{\theta} \text{ y } \forall \underline{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$
- Caso continuo: $f_{\underline{X}}(\underline{x}, \hat{\underline{\theta}}) \geq f_{\underline{X}}(\underline{x}, \underline{\theta}) \quad \forall \underline{\theta} \text{ y } \forall \underline{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$

O sea, el estimador por MV, $\hat{\theta}_{MV}$, para θ es aquel que maximiza la función de verosimilitud.

Ejercicio: Hallar los estimadores por MV para las distribuciones dadas en los ejercicios I), II), III) y IV).

Los estimadores obtenidos por MV tienen importantes propiedades como son la Propiedad de Invarianza y de ser Asintóticamente insesgados y de menor varianza.

Propiedad de invarianza

Sea X_1, \dots, X_n una muestra aleatoria con $\hat{\theta}$ estimador por MV para θ y $h : \mathbb{R}^m \rightarrow \mathbb{R}$ una función. Entonces $h(\hat{\theta})$ es el estimador por MV para $h(\theta)$.

Ejercicio: Sea X_1, \dots, X_n una m.a con distribución $N(\mu, \sigma^2)$.

- a) Para $p \in (0, 1)$ hallar el estimador por MV para el percentil ($p100$) o cuantil p para una variable aleatoria $N(\mu, \sigma^2)$.

b) *Obtener el estimador por MV para $P(X \leq x_0)$ donde $x_0 \in \mathbb{R}$ fijo y una estimación para $P(X \leq x_0)$ cuando se observa la muestra x_1, \dots, x_n .*

Propiedad asintótica del estimador por MV

Para n suficientemente grande el estimador por MV, $\hat{\theta}$, para θ es aproximadamente insesgado y la varianza es menor o igual que la de cualquier otro estimador.