Lección 7: Inferencia Estadística. Estimación puntual

Problema que planteamos: Tenemos un población y queremos estudiar una cierta característica de ella. Esta, la vamos a leer en términos de una cierta variable aleatoria X. A partir de ahora nos referiremos a X como una población.

- No conocemos la población, en el sentido de que no conocemos la función de distribución de la variable aleatoria.
- Será habitual suponer que la población se adapta a un cierto modelo de probabilidad (binomial, Poisson, normal,...)
- Queremos tomar muestras aleatorias simples de la población que nos dé información sobre la función de distribución.
- Si la variable aleatoria es discreta, sea P_θ su función de masa (tambin denominada funcin de probabilidad), si es continua f_θ su función de densidad, dependiendo de un parámetro θ desconocido, y que estará en el espacio paramétrico Θ, que es el conjunto de los posibles valores del parámetro.
- El parámetro que desconocemos puede ser de una o más dimensiones. Para una Poisson $P(\lambda)$ el parámetro $\theta = \lambda$; para una normal $N(\mu; \sigma)$ el parámetro $\theta = (\mu, \sigma)$.

Siempre que estemos en una situación de este tipo, nos estamos refiriendo a un problema de inferencia estadística.

La inferencia estadística se divide en tras grandes partes, dependiendo de la naturaleza del problema que queremos resolver:

- Estimación puntual.
- Estimación por intervalos de confianza.
- Contraste de hipótesis paramétrica.

Objetivo de la estimación puntual: en función del modelo que estemos aceptando y en función de una muestra concreta $(x_1, x_2, \cdots, x_n)$, cual es el valor $\theta \in \Theta$ que nos parece más pausible.

Definición 1 *Sea $(X_1, X_2, \cdots, X_n)$ una muestra aleatoria simple de una población X con función de masa P_θ, o de densidad f_θ, $\theta \in \Theta$. Un estimador puntual de $g(\theta)$ es una función T que a cada posible muestra $(x_1, x_2, \cdots, x_n)$ le hace corresponder una estimación $T(x_1, x_2, \cdots, x_n)$ de $g(\theta)$.*

$T(X_1, X_2, \cdots, X_n)$ es una variable aleatoria, (o un vector si $g(\theta)$ lo es), y por lo tanto, un estimador puntual es un estadístico.

Definición 2 *Sea X una población con función de masa P_θ, o de densidad f_θ, $\theta \in \Theta$, y $(X_1, X_2, \cdots, X_n)$ una muestra aleatoria simple. El error medio cuadrático de un estimador T, para estimar $g(\theta)$, se define como*

$$\boxed{E_\theta\left[\left(T(X_1, X_2, \cdots, X_n) - g(\theta)\right)^2\right]} \tag{1}$$

Vamos a obtener una expresión sencilla para el error.

$$E_\theta\left[(T - g(\theta))^2\right] = E_\theta\left[(T - E_\theta(T) + E_\theta(T) - g(\theta))^2\right]$$

$$= E_\theta\left[(T - E_\theta(T))^2 + 2(T - E_\theta(T))(E_\theta(T) - g(\theta)) + (E_\theta(T) - g(\theta))^2\right]$$

$$= E_\theta\left[(T - E\theta(T))^2\right] + 2(E\theta(T) - g(\theta))E_\theta\left[T - E_\theta(T)\right] + E_\theta\left[(E_\theta(T) - g(\theta))^2\right],$$

Como $E_\theta\left[T - E_\theta(T)\right] = 0$ y $E_\theta\left[(E_\theta(T) - g(\theta))^2\right] = (E_\theta(T) - g(\theta))^2$, nos queda

$$\boxed{E_\theta\left[(T - g(\theta))^2\right] = V_\theta(T) + \text{sesgo}^2(T)} \tag{2}$$

donde $\text{sesgo}(T) = E_\theta(T) - g(\theta)$.

Definición 3 *Un estimador T es insesgado o centrado, para estimar $g(\theta)$, si*

$$E_\theta(T) = g(\theta), \qquad \forall \theta \in \Theta.$$

Ejemplo 4

Tomamos una población X normal $N(\mu; \sigma)$, con $\theta = (\mu, \sigma)$ desconocido. Sea $(X_1, X_2, \cdots, X_n)$ una muestra aleatoria simple.

1. El estimador $T(X_1, X_2, \cdots, X_n) = \overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$ es insesgado para estimar $g_1(\theta) = \mu$, ya que
$$E_\theta(\overline{X}) = \frac{1}{n}\sum_{i=1}^n E_\theta(X_i) = \frac{1}{n}\sum_{i=1}^n \mu = \mu = g_1(\theta).$$

2. El estimador $T(X_1, X_2, \cdots, X_n) = \hat{s}^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$ es insesgado para estimar $g_2(\theta) = \sigma^2$, ya que
$$E_\theta(\hat{s}^2) = \sigma^2 = g_2(\theta).$$
(esto se demuestra en el siguiente ejercicio)

Ejercicio 5 *Sea X una población con esperanza y varianza μ y σ^2 respectivamente. Sea $(X_1, X_2, \cdots, X_n)$ una muestra aleatoria simple de tamaño n. Consideramos los estadísticos*

$$\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i, \qquad \textit{media muestral,}$$

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 \qquad \textit{varianza muestral,}$$

$$\hat{s}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \qquad \textit{cuasivarianza muestral.}$$

1. *Demostrar que $\hat{s}^2$ es un estimador insesgado, o centrado, para estimar σ^2.*
2. *Demostrar que s^2 es un estimador sesgado, o no centrado, para estimar σ^2.*
3. *Deducir que el error medio cuadrático del estimador $\hat{s}^2$ para estimar σ^2, es menor que el error medio cuadrático del estimador s^2, para estimar σ^2.*

Indicación: Para el tercer apartado utilícese que $V(s^2) = \frac{2(n-1)}{n^2}\sigma^4$.

Observese que si $(X_1, X_2, \cdots, X_n)$ es una muestra aleatoria simple de la población X, entonces

- $X_1, X_2, \cdots, X_n$ sin independientes.
- $E[X_i] = \mu, \quad V(X_i) = E[(X_i - \mu)^2] = \sigma^2, \quad i = 1, 2, \cdots, n.$

1. Tendremos que demostrar que $E[\hat{s}^2 - \sigma^2] = 0 \Longleftrightarrow E[\hat{s}^2] = \sigma^2$.

$$E[\hat{s}^2] = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2\right] = \frac{1}{n-1}E\left[\sum_{i=1}^{n}\left(X_i - \mu + \mu - \overline{X}\right)^2\right]$$

$$= \frac{1}{n-1}E\left[\sum_{i=1}^{n}(X_i - \mu)^2 + 2(\mu - \overline{X})\sum_{i=1}^{n}(X_i - \mu) + n(\mu - \overline{X})^2\right],$$

Como

$$\sum_{i=1}^{n}(X_i - \mu) = \sum_{i=1}^{n}X_i - n\mu = n\left(\frac{1}{n}\sum_{i=1}^{n}X_i - \mu\right) = n(\overline{X} - \mu),$$

entonces

$$E[\hat{s}^2] = \frac{1}{n-1}E\left[\sum_{i=1}^{n}(X_i - \mu)^2 - 2n(\overline{X} - \mu)^2 + n(\mu - \overline{X})^2\right]$$

$$= \frac{1}{n-1}E\left[\sum_{i=1}^{n}(X_i - \mu)^2 - n(\overline{X} - \mu)^2\right]$$

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n}E[(X_i - \mu)^2] - nE[(\overline{X} - \mu)^2]\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}V(X_i) - nV(\overline{X})\right)$$

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n}V(X) - V(X)\right) = \frac{1}{n-1}\left(n\sigma^2 - \sigma^2\right) = \sigma^2.$$

2. $E[s^2] = E\left[\frac{n-1}{n}\hat{s}^2\right] = \frac{n-1}{n}\sigma^2 \neq \sigma^2$, y por lo tanto s^2 es sesgado.

3. <u>Error cuadrático para el estimador varianza para estimar σ^2.</u>

$$E\left[(s^2-\sigma^2)^2\right] = V(s^2) + \text{sesg}^2(s^2) = \frac{2(n-1)}{n^2}\sigma^4 + \left(\frac{n-1}{n}\sigma^2 - \sigma^2\right)^2$$

$$= \frac{2(n-1)}{n^2}\sigma^4 + \frac{\sigma^4}{n^2} = \frac{2n-1}{n^2}\sigma^4.$$

<u>Error cuadrático para el estimador cuasivarianza para estimar σ^2.</u>

$$E\left[(\hat{s}^2-\sigma^2)^2\right] = V(\hat{s}^2) = V\left(\frac{n}{n-1}s^2\right) = \frac{n^2}{(n-1)^2}V(s^2)$$

$$= \frac{n^2}{(n-1)^2}\cdot\frac{2(n-1)}{n^2}\sigma^4 = \frac{2}{n-1}\sigma^4.$$

$$E\left[(\hat{s}^2-\sigma^2)^2\right] > E\left[(s^2-\sigma^2)^2\right] \Longleftrightarrow \frac{2}{n-1} > \frac{2n-1}{n^2},$$

lo cual es cierto ya que

$$\frac{2}{n-1} > \frac{2n-1}{n^2} \Longleftrightarrow 2n^2 > 2n^2 - 3n + 1 \Longleftrightarrow 3n > 1,$$

lo cual es cierto cualquiera que sea $n \in \mathbb{N}$.

Comentario: Aunque $\hat{s}^2$ es insesgado para estimar σ^2, y s^2 es sesgado, se comete menos error cuadrático medio para estimar σ^2 con el estimador segado s^2.

$\square$

Definición 6 *Un estimador T es consistente para estimar $g(\theta)$ si para todo $\theta \in \Theta$ se verifica*

$$\begin{array}{ll} \text{lím}_{n\to\infty} F_T(t) = \text{lím}_{n\to\infty} P(T<t) = 0 & t < g(\theta) \\ \text{lím}_{n\to\infty} F_T(t) = \text{lím}_{n\to\infty} P(T<t) = 0 & t > g(\theta) \end{array}$$

donde F_T es la función de distrubución de la variable aleatoria $T(X_1, X_2, \cdots, X_n)$.

El significado intuitivo de la definición es que, a medida que el tamaño muestral n aumenta, la distribución de la variable aleatoria $T(X_1, X_2, \cdots, X_n)$ se concentra cada vez más alrededor del valor $g(\theta)$ y en consecuencia, las estimaciones $T(x_1, x_2, \cdots, x_n)$ cada vez se acercan más a $g(\theta)$.

Proposición 7 *Si el estimador T satisface*

- $\text{lím}_{n\to\infty} E_\theta(T) = g(\theta)$ *para todo* $\theta \in \Theta$,
- $\text{lím}_{n\to\infty} V(T) = 0$ *para todo* $\theta \in \Theta$,

entonces T es consistente para estimar $g(\theta)$.

Ejercicio 8 *Sea X una población normal $N(\mu;\sigma)$ y $(X_1, X_2, \cdots, X_n)$ una muestra aleatoria simple. Demuéstrese que el estimador $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ para estimar $g(\theta) = g(\mu,\sigma) = \mu$ es consistente.*

Vamos a utilizar la proposición anterior.

- $\lim_{n\to\infty} E_\theta(\overline{X}) = \lim_{n\to\infty} \mu = \mu = g(\theta)$.
- $\lim_{n\to\infty} V(T) = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$.

$\square$

Definición 9 Construcción de estimadores por el método de los momentos. *Sea X una población con función de masa P_θ, o de densidad f_θ, $\theta = (\theta_1, \theta_2, \cdots, \theta_k) \in \Theta$, y $(X_1, X_2, \cdots, X_n)$ una muestra aleatoria simple. El estimador de θ por el método de los momentos es el formado por los valores $(\widetilde{\theta}_1, \widetilde{\theta}_2, \cdots, \widetilde{\theta}_k)$ que se obtiene al resolver en $\theta_1, \theta_2, \cdots, \theta_k$ el sistema*

$$\begin{cases} E_\theta[X] = \frac{1}{n}\sum_{i=1}^{n} X_i \\ E_\theta[X^2] = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \\ \cdots\cdots \\ E_\theta[X^k] = \frac{1}{n}\sum_{i=1}^{n} X_i^k \end{cases}$$

Ejercicio 10 *Una población X sigue una exponencial $Exp(\lambda)$ de parámetro desconocido. Calcular el estimador de $\theta = \lambda$ por el método de los momentos.*

X es continua con función de densidad

$$f_\lambda(x) = \begin{cases} 0 \\ \lambda e^{-\lambda x} \end{cases}$$

Sea $(X_1, X_2, \cdots, X_n)$ una muestra aleatoria de tamaño n.

Supuesto que θ es uno dimensional, el estimador es el valor $\tilde{\lambda}$ obtenido al resolver la ecuación

$$E_\lambda[X] = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Calculamos $E_\lambda[X]$.

$$E_\lambda[X] = \int_{\mathbb{R}} x f_\lambda(x)dx = \int_0^\infty x\lambda e^{-\lambda x}dx = \frac{1}{\lambda}.$$

$$E_\lambda[X] = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \Longleftrightarrow \frac{1}{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

y el estimador de λ por el método de los momentos de la muestra obtenida $(x_1, x_2, \cdots, x_n)$ es el valor

$$\widetilde{\lambda} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} x_i}$$

.

$\square$

Ejercicio 11 *Sea $(X_1, X_2, \cdots, X_n)$ una muestra aleatoria de*

$$f_\theta(x) = \begin{cases} 0 & x \leq 0 \\ e^{-x+\theta} & x > \theta \end{cases} \qquad \theta \in \Theta = \mathbb{R}.$$

1. *Hallar el estimador por el método de los momentos de θ.*
2. *Estudiar si el estimador encontrado en el apartado anterior es insesgado para estimar el parámetro θ.*

Definición 12 Construcción de estimadores por el método de máxima verosimilitud. *Sea X una población con función de masa P_θ, o de densidad f_θ, $\theta = (\theta_1, \theta_2, \cdots, \theta_k) \in \Theta$, y $(X_1, X_2, \cdots, X_n)$ una muestra aleatoria simple. El estimador de máxima verosimilitud de θ es el formado por los valores $(\widehat{\theta}_1, \widehat{\theta}_2, \cdots, \widehat{\theta}_k)$ que maximiza lo que llamamos función de verosimilitud de la muestra $(x_1, x_2, \cdots, x_n)$ obtenida*

$$L(\theta) = L(\theta; x_1, x_2, \cdots, x_n) = \begin{cases} P_\theta(x_1) \cdot P_\theta(x_2) \cdots P_\theta(x_n) & \textit{discreta} \\ f_\theta(x_1) \cdot f_\theta(x_2) \cdots f_\theta(x_n) & \textit{continua} \end{cases}$$

Ejercicio 13 *Un una gran piscifactoría hay una proporción desconocida de peces de una especia A. Para obtener información sobre esa proporción, vamos a ir sacando peces al azar.*

1. *Si la proporción de peces de la especie A es p, ¿cuál es la probabilidad de que el primer pez de la especie A sea el décimo que extraemos?.*
2. *Tres personas realizan, independientemente unas de otras, el proceso de sacar peces al azar hasta encontrarse con el primero del tipo A:*
   - *La primera persona obtiene el primer pez tipo A en la décima exacción.*
   - *La segunda persona obtiene el primer pez tipo A en la decimoquinta extracción.*
   - *La tercera persona obtiene el primer pez tipo A en la decimoctava extracción.*

   Escribir la función de verosimilitud y obtener la estimación de máxima verosimilitud de p.

1. Experimento aleatorio: sacar un pez y ver si es de la clase A.

   Consideramos la variable aleatoria X="número de peces que no son de la clase A y que hemos extraído antes de extraer el primer pez de la clase A". Esta es una variable aleatoria discreta y su rango es $\{0, 1, 2, \cdots\}$. Si la probabilidad de sacar un pez de la clase A es p, la de extraer un pez que no sea de la la clase A es $1 - p$.

$$P(X = 9) = (1 - p)^9 p.$$

2. Consideramos como población el conjunto de peces de la piscifactoría. X la variable aleatoria "número de peces que no son de la clase A y que hemos extraído antes de extraer el primer pez de la clase A". Su función de probabilidad es

$$P_p(X = k) = (1 - p)^k p, \qquad k = 0, 1, 2, \cdots,$$

Queremos estimar el parámetro p. Tomamos una muestra aleatoria de tamaño 3, (X_1, X_2, X_3), donde X_i="número de peces que no son de la clase A y que han sido extraídos por la persona iésima antes de extraer el primer pez de la clase A". Las variables X_i, $i = 1, 2, 3$, son independientes y todas están igualmente distribuidas:

$$P_p(X_i = k) = (1 - p)^k p, \qquad k = 0, 1, 2, \cdots.$$

El estimador de máxima verosimilutud de p es el valor $\hat{p}$ que maximiza en $[0, 1]$ lo que llamamos función de verosimilitud de la muestra $(x_1, x_2, x_3) = (9, 14, 17)$ que está dada por

$$L(p) = P_p(X_1 = 9, X_2 = 14, X_3 = 17) = P_p(X_1 = 9) \cdot P_p(X_2 = 14) \cdot P_p(X_3 = 17)$$

$$= (1 - p)^9 p \cdot (1 - p)^{14} \cdot (1 - p)^{17} p = (1 - p)^{40} p^3, \qquad p \in [0, 1].$$

$L(p)$ es una función continua en el compacto $[0, 1]$ por lo tanto alcanzará su máximo en $[0, 1]$. Ya que $l(p)$ es no negativa y en los extremos de $[0, 1]$ de anula, el máximo lo alcanzará en algún punto crítico de $(0, 1)$.

$$L'(p) = 0 \Longleftrightarrow -40(1 - p)^{39} p^3 + 3(1 - p)^{40} p^2 = 0$$

$$\Longleftrightarrow -40p + 3(1 - p) = 0 \Longleftrightarrow p = \frac{3}{43}.$$

Como solo existe un punto en $(0, 1)$ donde se anula la derivada de $L(p)$, este corresponde a un máximo y $\hat{p} = \frac{3}{43}$.

Ejercicio 14 *Para estudiar la proporción p de caballos afectados por la peste equina se les va a someter a una prueba. Sabemos que la prueba será positiva si el animal está enfermo; si está sano, hay una probabilidad de 0,04 de que la prueba resulte positiva.*

1. *Hallar la relación entre la probabilidad p de estar enfermo y la probabilidad q de dar positivo en la prueba.*

2. *Obtener la estimación de máxima verosimilitud de p si 500 ejemplares son sometidos a la prueba y resulta positiva en 95 casos.*

3. *Si realmente hay un 20 % de caballos afectados por la epidemia, ¿cuál es la probabilidad de que la prueba resulte positiva en, al menos, 95 ejemplares de los 500?*

1. Experimento aleatorio: realizar la prueba en una cierta población de caballos.

   Consideramos los sucesos:

   - P^+=caballos a los que la prueba le ha dado positiva.
   - P^-=caballos a los que la prueba no le ha dado positiva.
   - A^+=caballos afectados por la peste equina.
   - A^-=caballos no afectados por la peste equina.

   Sea $P(A^+) = p$ y $P(P^+) = q$. Sabemos por el enunciado que $P(P^+|A^+) = 1$ y $P(P^+|A^-) = 0{,}04$. Se verifica

   $$q = P(P^+) = P(P^+ \cap A^+) + P(P^+ \cap A^-) = P(P^+|A^+) \cdot P(A^+) + P(P^+|A^-) \cdot P(A^-)$$

   $$= 1 \cdot p + 0,04 \cdot (1-p) \Longrightarrow p = \frac{q - 0,04}{0,96}.$$

2. Vamos a dar una estimación de q y utilizando $p = \frac{q-0{,}04}{0{,}96}$ daremos una estimación de p.

   El universo es la población de caballos. Consideramos la variable aleatoria

   $$X = \begin{cases} 1 & \text{si la prueba ha dado positiva} \\ 0 & \text{en caso contrario} \end{cases}$$

   X es una variable discreta con rango $\{0, 1\}$ y función de masa

   $$P_q(X = n) = q^n (1-q)^{1-n}, \qquad n = 0, 1.$$

   Tomamos un vector aleatorio de tamaño 500, $(X_1, X_2, \cdots, X_{500}\}$. Las componentes del vector aleatorio

   - son independeientes,
   - $P_q(X_i = n) = q^n(1-q)^n, \qquad n = 0, 1, \qquad i = 1, 2, \cdots, 500.$

   La estimación de máxima verosimilitud para q es el valor $\widehat{q}$ que maximiza en $[0, 1]$ la función

   $$L(q) = P_q(X_1 = x_1, X_2 = x_2, \cdots, X_{500} = x_{500})$$
   $$= P_q(X_1 = x_1) \cdot P_q(X_2 = x_2) \cdots P_q(X_{500} = x_{500}),$$

   aqui hay 95 x_i que son 1 y 405 que son cero, entonces

   $$L(q) = q^{95}(1-q)^{405} \qquad q \in [0, 1].$$

$L(q)$ es continua en $[0,1]$ y este intervalo es compacto, entonces alcanzará su valor máximo y mínimo. Como $L \geq 0$ y $L(0) = L(1) = 0$, el máximo se alcanzará en un punto crítico en $(0,1)$.

$$L'(q) = 0 \Longleftrightarrow 95q^{94}(1-q)^{405} - 405q^{95}(1-q)^{404} = 0$$

$$\Longleftrightarrow 95(1-q) - 405q = 0 \Longleftrightarrow q = \frac{95}{500}.$$

Como solo existe un punto crítico en $(0,1)$, este corresponde al máximo y el estimador de máxima verosimilitud es $\widehat{q} = \frac{95}{500}$.

La estimación de máxima verosimilitud para p es

$$\widehat{p} = \frac{\widehat{q} - 0,04}{0,96} = \frac{\frac{95}{500} - 0,04}{0,96}.$$

3. Si realmente hay un 20 % de caballos afectados por la peste equina, entonces $p = 0,2$ y $q = 0,04 + 0,96 \cdot 0,2 = 0,232$.

   Sea X la variable aleatoria que indicael número de pruebas positivas que se dan en 500 realizaciones. $X \to B(n = 500, q = 0,232)$. Si utilizamos el teorema central de limite, aproximamos esta binomial por la normal $N(nq, \sqrt{nq(1-q)}) = N(116; 9,44)$.

   $$P(X \geq 95) \cong 0,9868.$$

$\square$

Ejercicio 15 *El parámetro de una distribución de Poisson puede tomar uno de los cuatro valores siguientes:* $\lambda = 4;\ 4,5;\ 5,5;\ 6$. *Decida cual de ellos puede ser, considerando una muestra aleatoria simple de tamaño dos* (X_1, X_2), *con realización* $(x_1, x_2) = (3,7)$, *y basándose en el principio de máxima verosimilitud.*

La muestra puede proceder de cuatro poblaciones diferentes, según sea el valor que tome el parámetro λ de la distribución de Poisson. De acuerdo con el principio de máxima verosimilitud, elegiremos el valor de λ tal que la probabilidad de obtener la citada muestra sea máxima.

- Para $\lambda = 4,5$,

  $$P(X_1 = 3, X_2 = 7) = P(X_1 = 3) \cdot P(X_2 = 7) = \frac{e^{-4} \cdot 4^3}{3!} \cdot \frac{e^{-4} \cdot 4^7}{7|} = 0,02326.$$

- Para $\lambda = 4,5$

  $$P(X_1 = 3, X_2 = 7) = P(X_1 = 3) \cdot P(X_2 = 7) = \frac{e^{-4,5} \cdot 4,5^3}{3!} \cdot \frac{e^{-4,5} \cdot 4,5^7}{7|} = 0,0278.$$

- Para $\lambda = 5,5$,

$$P(X_1 = 3, X_2 = 7) = P(X_1 = 3) \cdot P(X_2 = 7) = \frac{e^{-5,5} \cdot 5,5^3}{3!} \cdot \frac{e^{-5,5} \cdot 5,5^7}{7|} = 0,02798.$$

- Para $\lambda = 6$,

$$P(X_1 = 3, X_2 = 7) = P(X_1 = 3) \cdot P(X_2 = 7) = \frac{e^{-6} \cdot 6^3}{3!} \cdot \frac{e^{-6} \cdot 6^7}{7|} = 0,02457.$$

Cuando $\lambda = 5,5$ obtenemos la mayor probabilidad de extraer la muestra $(3,7)$, por lo cual decidimos que el valor del parámetros λ es $5,5$.

$\square$

Ejercicio 16 *Dada una muestra aleatoria $(X_1, X_2, \cdots, X_n)$ de tamaño n de una población X, calcular el estimador por el método máxima verosimilitud y por el método de los momentos, en lo siguientes casos:*

1. *$X \sim$Bernoulli de parámetro p.*
2. *$X \sim$Poisson de parámetro λ.*
3. *$X \sim N(\mu, \sigma)$.*

Solución: 1. Estimador por el método de los momentos $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, estimador de máxima verosimilitud $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$; 2. Estimador por el método de los momentos $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, estimador de máxima verosimilitud $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$; 3. Estimador por el método de los momentos para μ $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ y para σ^2 $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$, estimador de máxima verosimilitud para μ $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ y para σ^2 $\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$.

Ejercicio 17 *La distancia X entre una árbol cualquiera y el árbol más próximo a él en un bosque sigue uns distribición de Rayleigh con función de densidad*

$$f_\theta(x) = \begin{cases} 0 & x < 0 \\ 2\theta x e^{-\theta x^2} & x \geq 0 \end{cases} \qquad \theta > 0.$$

1. *Obtener los estimadores de máxima verosimilitud de θ y de $g(\theta) = E_\theta(X)$, basado en muestras de tamaño n.*
2. *Obtener el estimador de θ por el método de los momentos.*

Indicación: $\int_0^\infty e^{-x^2} dx = \frac{\pi}{2}$.

1. Dada la muestra aleatoria simple $(X_1, X_2, \cdots, X_n)$ de una cierta población X confunción de densidad f_θ, el estimador de máxima verosimilitud de θ es el valor $\widehat{\theta}$ que maximiza en $(0, \infty)$ la función de verosimilitud de la muestra $(x_1, x_2, \cdots, x_n)$:

$$L(\theta) = f_\theta(x_1, x_2, \cdots, x_n) = 2\theta x_1 e^{-\theta x_1^2} \cdot 2\theta x_2 e^{-\theta x_2^2} \cdots 2\theta x_n e^{-\theta x_n^2}$$

$$= 2^n\theta^n x_1x_2\cdots x_n e^{-\theta\sum_{i=1}^n x_i^2}, \qquad 0 < x_1, x_2, \cdots, x_n.$$

Considerandola solo como función de θ

$$L(\theta) = 2^n\theta^n x_1x_2\cdots x_n e^{-\theta\sum_{i=1}^n x_i^2}, \qquad \theta \in (0,\infty),$$

vamos a obtener en que punto alcanza su valor máximo.

Ya que $L(\theta)$ es continua en $(0,\infty)$, positiva y $\lim_{\theta\to 0} L(\theta) = \lim_{\theta\to\infty} L(\theta) = 0$, l alcanza su valor máximo en $(0,\infty)$, y lo hará en un punto $\widetilde{\theta}$ que anule la derivada de $L(\theta)$.

$$L'(\theta) = 2^n n\theta^{n-1} x_1x_2\cdots x_n e^{-\theta\sum_{i=1}^n x_i^2} - \sum_{i=1}^n x_i^2 2^n\theta^n x_1x_2\cdots x_n e^{-\theta\sum_{i=1}^n x_i^2},$$

y

$$L'(\theta) = 0 \Longleftrightarrow \theta = \frac{n}{\sum_{i=1}^n x_i^2}.$$

Al solo existir un punto crítico en $(0,\infty)$, este corresponde al valor donde $L(\theta)$ alcanza el máximo y el estimador de maáxima verosimilitud para θ es $\widehat{\theta} = \frac{n}{\sum_{i=1}^n x_i^2}$. El estimador, como estadístico, es

$$T(X_1, X_2, \cdots, X_n) = \frac{n}{\sum_{i=1}^n X_i^2}.$$

Vamos a calcular $g(\theta) = E_\theta[X]$,

$$E_\theta[X] = \int_{-\infty}^{\infty} x f_\theta(x)dx = \int_0^\infty 2\theta x^2 e^{-\theta x^2}dx$$

$$\left(\begin{array}{ll} x = u & dx = du \\ 2\theta x e^{-\theta x^2}dx = dv & v = e^{-\theta x^2}\end{array}\right) = -xe^{-\theta x^2}\Big|_0^\infty + \int_0^\infty e^{-\theta x^2}$$

$$= \left(\begin{array}{l}\sqrt{\theta}x = t \\ dx = \frac{dt}{\sqrt{\theta}}\end{array}\right) = \frac{1}{\sqrt{\theta}}\int_0^\infty e^{-t^2}dt = \frac{1}{2}\sqrt{\frac{\pi}{\theta}}.$$

El estimador de máxima verosimilitud para $g(\theta) = E_\theta[X]$ es

$$\widehat{g(\theta)} = \frac{1}{2}\sqrt{\frac{\pi}{\widehat{\theta}}} = \frac{1}{2}\sqrt{\frac{\pi\sum_{i=1}^n x_i^2}{n}}.$$

El estimador, como estadístico,

$$g(T) = \frac{1}{2}\sqrt{\frac{\pi}{\frac{n}{\sum_{i=1}^n X_i^2}}} = \frac{1}{2}\sqrt{\frac{\pi\sum_{i=1}^n X_i^2}{n}}.$$

2. Para obtener el estimador por el método de los momentos, resolvemos la ecuación

$$E_\theta[X] = \overline{X} \Longleftrightarrow \frac{1}{2}\sqrt{\frac{\pi}{\theta}} = \overline{X} \Longleftrightarrow \theta = \frac{\pi}{4\overline{X}^2},$$

y el estimador por el método de los momentos para estimar θ es $\widetilde{\theta} = \frac{\pi}{4\overline{x}^2}$. Como estadístico

$$W(X_1, X_2, \cdots, X_n) = \frac{\pi}{4\overline{X}^2}.$$

Observamos que obtenemos distinto estimador según utilicemos máxima verosimilitud o el método de los momentos.

□

Ejercicio 18 *La lectura de voltaje dada por un voltímetro conectado a un circuito eléctrico es una variable aleatoria con distribución uniforme en el intervalo $(\theta, \theta+1)$, siendo θ el verdadero valor (desconocido) del voltaje. Sea $(X_1, X_2, \cdots, X_n)$ una muestra aleatoria de lecturas de dicho voltímetro.*

1. *Demostrar que la media muestral $\overline{X}$ es un estimador sesgado de θ y calcular el sesgo.*
2. *Calcular el error cuadrático medio de $\overline{X}$.*
3. *Obtener , a partir de $\overline{X}$, un estimador insesgado de theta.*

Sea X la la variable aleatoria que indica la lectura del voltaje. Su función de densidad es

$$f_\theta(x) = \begin{cases} 1 & x \in (\theta, \theta+1) \\ 0 & \text{en el resto} \end{cases}$$

1. Tendremos que ver que $E[\overline{X}] \neq g(\theta) = \theta$.

$$E[\overline{X}] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\sum_{i=1}^{n} E[X] = E[X]$$

$$= \int_{-\infty}^{\infty} x f_\theta(x) dx = \int_{\theta}^{\theta+1} x dx = \left.\frac{x^2}{2}\right|_{\theta}^{\theta+1} = \frac{(\theta+1)^2 - \theta^2}{2} = \theta + \frac{1}{2},$$

luego efectivamente es sesgado.

$$\text{sesgo de } \overline{X} = E_\theta[\overline{X}] - \theta = \frac{1}{2}.$$

2. Error cuadrático de $\overline{X} = V(\overline{X}) + \text{sesgo}^2(\overline{X})$.

$$V(\overline{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} V(X) = \frac{V(X)}{n}.$$

Calculamos ahora $V(X)$.

$$V(X) = E\left[(X - E[X])^2\right] = E\left[X^2 - 2E[X]X - E[X]^2\right]$$

$$= E\left[X^2\right] - 2E[X] \cdot E[X] + E^[X]^2 = E\left[X^2\right] - E^[X]^2$$

$$= \int_{\mathbb{R}} x^2 f_\theta(x)dx - \left(\theta + \frac{1}{2}\right)^2 = \int_\theta^{\theta+1} x^2 dx - \left(\theta^2 + \theta + \frac{1}{4}\right)$$

$$= \frac{1}{3}\left((\theta+1)^3 - \theta^3\right) - \left(\theta + \frac{1}{2}\right)^2 = \frac{1}{3}(3\theta^2 + \theta + 1) - \left(\theta^2 + \theta + \frac{1}{4}\right) = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

Ontenemos

$$\text{Error cuadrático medio} = \frac{1}{12n} + \frac{1}{4}.$$

3. Ya que el sesgo de $\overline{X}$ es $\frac{1}{2}$, el estimador, como puede comprobar el estudiante,

$$T(X_1, X_2, \cdots, X_n) = \overline{X} - \frac{1}{2} = \frac{1}{n}\sum_{i=1}^{n} X_1 - \frac{1}{2},$$

es insesgado o centrado.

□

Ejercicio 19 *Vamos a clasificar a las personas de un pais según dos características: color de los ojos (claros u oscuros) y sexo (hombre o mujer). Las dos características son independientes.*

1. *Obtenemos una muestra al azar de la población con los siguientes resultados*
   - *200 mujeres con ojos claros.*
   - *150 hombres con ojos claros.*
   - *350 mujeres con ojos oscuros.*
   - *300 hombres con ojos oscuros.*

   Obtener la estimación de máxima verosimilitud de
   - *p= probabilidad de ser hombre.*
   - *q= probabilidad de tener ojos claros.*
2. *Después de muchas horas de intenso trabajo llegamos a saber con exactitud que p=0,4 y q=0,6. Si tomamos 8 personas al azar de ese pais, ¿cuál es la probabilidad de encontrar alguna mujer de ojos oscuros?. Y si la muestra que tomamos es de 200 personas, ¿cuál es la probabilidad de que haya más de 60 mujeres de ojos oscuros?.*

1. Tenemos

- p= probabilidad de ser hombre.
- 1-p= probabilidad de ser mujer.
- q= probabilidad de tener los ojos claros.
- 1-q= probabilidad de tener los ojos oscuros.

y

- q(1-p)= probabilidad de ser mujer con ojos claros.
- pq= probabilidad de ser hombre con ojos claros.
- (1-p)(1-q)= probabilidad de ser mujer con ojos oscuros.
- p(1-q)= probabilidad de hombre con ojos oscuros.

Experimento aleatorio: tomar una persona del pais y comprobar sexo y color de los ojos.

Espacio muestral

$$\Omega = \{\text{mujer con ojos claros}, \ \text{mujer con ojos oscuros},$$

$$\text{hombre con ojos claros}, \ \text{hombre con ojos oscuros}\}$$

Definimos la variable aleatoria $X : \Omega \longmapsto \mathbb{R}$ por

$$\begin{array}{ll} X(\text{mujer con ojos claros}) = 1, & X(\text{mujer con ojos oscuros}) = 2, \\ X(\text{hombre con ojos claros}) = 3, & X(\text{ hombre con ojos oscuros}) = 4. \end{array}$$

Tomamos una muestra de tamaño 1000 $(X_1, X_2, \cdots, X_{1000})$.

El estimador de máxima verosimilitud de (p, q) es $(\hat{p}, \hat{q})$, donde este punto maximiza la función de verosimilitud de la muestra $(x_1, x_2, \cdots, x_{1000})$ obtenida $(x_i = 1, \ i = 1, 2, \cdots, 200, \ x_i = 2, \ i = 201, 202, \cdots, 350, \ x_i = 1, \ i = 351, 352, \cdots, 700, \ x_i = 1, \ i = 701, 702, \cdots, 1000,)$:

$$L(p, q) = P(p, q, x_1, x_2, \cdots, x_{1000}) = P_{p,q}(X_1 = 1) \cdots P_{p,q}(X_{200} = 1) \cdot$$

$$P_{p,q}(X_{201} = 2) \cdots P_{p,q}(X_{350} = 2) \cdot P_{p,q}(X_{351} = 3) \cdots P_{p,q}(X_{700} = 3) \cdot$$

$$P_{p,q}(X_{701} = 4) \cdots P_{p,q}(X_{1000} = 4) =$$

$$[q(1-p)]^{200}[pq]^{150}[(1-p)(1-q)]^{350}[p(1-q)]^{300} = p^{450}(1-p)^{550}q^{350}(1-q)^{650},$$

en $[0, 1] \times [0, 1]$. Como $[0, 1] \times [0, 1]$ es compacto y $L(p, q)$ es continia, esta función alcanza su máximo en $[0, 1] \times [0, 1]$. Al ser $L(p, q)$ no negativa y valer cero en la frontera, el máximo se alcanza en el interior de $[0, 1] \times [0, 1]$. Ya que la función $\log x$ es creciente y $L(p, q) > 0$ en $(0, 1) \times (0, 1)$, $L(p, q)$ y

$$\log(L(p, q)) = 450 \log p 550 \log(1-p) 350 \log q + 650 \log(1-q),$$

alcanzarán su valor máximo en los mismos puntos. Estos tiene solución de

$$\nabla log(L(p,q)) = \left(\frac{\partial \log L}{\partial p}(p,q), \frac{\partial \log L}{\partial q}(p,q) \right) = (0,0).$$

$$\begin{cases} \frac{\partial \log L}{\partial p}(p,q) = \frac{450}{p} - \frac{550}{1-p} = 0 \implies p = 0,45 \\ \frac{\partial \log L}{\partial q}(p,q) = \frac{350}{q} - \frac{650}{1-q} = 0 \implies q = 0,355. \end{cases}$$

La estimación de máxima verosimilitud es $(\hat{p}, \hat{q}) = (0,45; 0,35)$.

2. Definimos, para $i = 1, 2, \cdots, 8$, las variables aleatorias

$$X_i = \begin{cases} 1 & \text{la persona i-ésima elegida es una mujer de ojos oscuros} \\ 0 & \text{en caso contrario} \end{cases}$$

Tenemos

$$P(X_i = 1) = (1 - 0,4) \cdot (1 - 0,6) = 0{,}24.$$

La variable aleatoria

$$X = X_1 + X_2 + \cdots + X_8;$$

me cuantifica el número de mujeres con los ojos oscuros hay entre las 8 personas elegidas al azar. X sigue una binomial $B(8; 0,24)$, y la probabilidad pedida es

$$P(X \geq 1) = 1 - P(X = 0) \cong 0,9.$$

Si la muestra que tomamos es de 200 personas, la variable aleatoria

$$Y = Y_1 + Y_2 + \cdots + Y_{200},$$

donde para $i = 1, 2, \cdots, 200$,

$$Y_i = \begin{cases} 1 & \text{la persona i-ésima elegida es una mujer de ojos oscuros} \\ 0 & \text{en caso contrario} \end{cases},$$

sigue una binomial $B(200; 0,24)$. Como la probabilidad pedida es

$$P(Y > 60) = 1 - P(Y \leq 60) = 1 - \sum_{i=0}^{60} P(Y = i),$$

pero no podemos hacer tantos cáculos. Vamos a paroximar por una Normal $N(200 \cdot 0,24; \sqrt{200 \cdot 0,24 \cdot (1 - 0,24)}) = N(48, 6, 04)$. Entonces

$$P(Y > 60) = P(N(48; 6,04) > 60 - 0,5).$$

Ejercicio 20 *El error (en centígramos) que se comete al pesar un objeto en una balanza puede considerarse como una variable aleatoria X con distribución $N(0; 15)$.*

1. Probabilidad de que el error cometido (en valor absoluto) en una pesada sea inferior a 20 centígramos.

2. Número mínimo de pesadas para que el error medio cometido (en valor absoluto) sea inferior a 5 centígramos con una probabilidad de 0,9.

1. $P(|X| < 20) = P(-20 < X < 20) = P\left(-1,33 < \frac{X}{15} < 1,33\right) = 1-2P\left(\frac{X}{15} > 1,33\right) = 0,8164$.

2. Sea $(X_1, X_2, \cdots, X_n)$ una muestra aleatoria de tamaño, donde X_i, $i = 1, 2, \cdots, n$, la vemos como la pesada i-ésima. Consideramos la media aleatoria $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, que sigue una $N(0, \frac{15}{\sqrt{n}})$. Queremos que

$$P(|\overline{X}| < 5) = 0{,}9 \Longleftrightarrow P(-5 < \overline{X} < 5) = 0{,}5$$

$$\Longleftrightarrow P\left(-\frac{5\sqrt{n}}{15} < \overline{X} < \frac{5\sqrt{n}}{15}\right) = 0{,}9 \Longleftrightarrow 1-2P\left(\overline{X} > \frac{5\sqrt{n}}{15}\right) = 0{,}9$$

$$\Longleftrightarrow P\left(\overline{X} > \frac{5\sqrt{n}}{15}\right) = 0,05 \Longrightarrow \frac{5\sqrt{n}}{15} = 1,64 \Longrightarrow n \geq 25.$$

Ejercicio 21 *Los gastos diarios de una empresa son una variable aleatoria con la siguiente función de densidad*

$$f_a(x) = \begin{cases} 0 & x \leq 0 \\ \frac{x}{a^2}e^{-\frac{x}{a}} & x > 0 \end{cases} \qquad a > 0.$$

Se decide tomar una muestra aleatoria simple de 10 dias en los que el gasto medio fué:

$$12,47 \quad 15,53 \quad 12,80 \quad 11,01 \quad 13,05 \quad 12,63 \quad 13,05 \quad 12,63 \quad 14,85 \quad 14,45$$

Se pide obtener una estimación de máxima verosimilitud del parámetro a y establecer si el estimador es centrado

Sea X una población con función de densidad $f_a(x)$. Tomamos una muestra aleatoria de tamaño 10, $(X_1, X_2, \cdots, X_{10})$, con el fin de obtener un valor aproximado para a. Cada variable aleatoria X_i se distribuye según la función de densidad $f_a(x)$. El estimador de máxima verosimilitud de a es el valor $\hat{a}$ que maximiza en $(0, \infty)$ lo que llamamos función de verosimilitud de la muestra

$$(x_1 = 12,47; \quad x_2 = 15,53; \quad x_3 = 12,80; \quad x_4 = 11,01; \quad x_5 = 13,05;$$

$$x_6 = 12,63; \quad x_7 = 13,05; \quad x_8 = 12,63; \quad x_9 = 14,85; \quad x_{10} = 14,45)$$

obtenida

$$L(a) = f_a(x_1, x_2, \cdots, x_{10}) = f_a(x_1) \cdot f_a(x_2) \cdots f_a(x_{10})$$

$$= \frac{x_1}{a^2}e^{-\frac{x_1}{a}} \cdot \frac{x_2}{a^2}e^{-\frac{x_2}{a}} \cdots \frac{x_{10}}{a^2}e^{-\frac{x_{10}}{a}} = \frac{x_1 x_2 \cdots x_{10}}{a^{20}}e^{-\frac{x_1+x_2+\cdots x_{10}}{a}} = \frac{\prod_{i=1}^{10} x_i}{a^{20}}e^{-\frac{\sum_{i=1}^{10} x_i}{a}}.$$

Ya que $L(a)$ es positiva, continua y acotada en $(0,\infty)$ y $\lim_{a\to 0^+} L(a) = \lim_{a\to\infty} L(a) = 0$, $L(a)$ alcanza el valor máximo en $(0,\infty)$ y lo hará en un punto que anule la primera derivada.

$$L'(a) = -20\frac{\prod_{i=1}^{10} x_i}{a^{21}} e^{-\frac{\sum_{i=1}^{10} x_i}{a}} + \frac{\left(\prod_{i=1}^{10} x_i\right)\cdot\left(\sum_{i=1}^{10} x_i\right)}{a^{22}} e^{-\frac{\sum_{i=1}^{10}}{a}},$$

$$L'(a) = 0 \Longleftrightarrow -20 + \frac{\sum_{i=1}^{10} x_i}{a} \Longleftrightarrow a = \frac{\sum_{i=1}^{10} x_i}{20}.$$

Como solo hay un valor que anule la derivada de $L(a)$, este corresponderá al máximo, y el estimador de másima verosimilitud para a es

$$\hat{a} = \frac{\sum_{i=1}^{10} x_i}{20} = \frac{12,47 + 15,53 + \cdots + 14,45}{20} = 6,6235.$$

Para ver si el estimador es insesgado o centrado tendríamos que ver que la variable aleatoria $T(X_1, X_2, \cdots, X_{10}) = \frac{\sum_{i=1}^{10} X_i}{20}$ satisface que $E_a[T] = a$.

$$E_a[T] = \frac{\sum_{i=1}^{10} E_a[X_i]}{20} = \frac{\sum_{i=1}^{10} E_a[X]}{20} = \frac{1}{2}E_a[X]$$

$$= \frac{1}{2}\int_{\mathbb{R}} x f_a(x)dx = \frac{1}{2a^2}\int_0^\infty x^2 e^{-\frac{x}{a}}dx = \frac{1}{2} = \begin{pmatrix} x^2 = u, & du = 2xdx \\ dv = e^{-\frac{x}{a}}, & v = -ae^{-\frac{x}{a}} \end{pmatrix}$$

$$= \frac{1}{2a^2}\left[-x^2 a e^{-\frac{x}{a}}\Big|_0^\infty + 2a\int_0^\infty x e^{-\frac{x}{a}}dx\right] = \frac{1}{a}\int_0^\infty x e^{-\frac{x}{a}}dx = \begin{pmatrix} x = u, & dx = du \\ dv = e^{-\frac{x}{a}}, & v = -ae^{-\frac{x}{a}} \end{pmatrix}$$

$$= \frac{1}{a}\left[-xee^{-\frac{x}{a}}\Big|_0^\infty + a\int_0^\infty e^{-\frac{x}{a}}dx\right] = \int_0^\infty e^{-\frac{x}{a}} = -ae^{-\frac{x}{a}}\Big|_0^\infty = a.$$

EL estimador calculado es centrado o insesgado.

□

Ejercicio 22 *La duración en minutos de un determinado viaje es una variable aleatoria con distribución normal de media desconocida y desviación típica igual a* 3*. En una muestra aleatoria de diez realizaciones del viaje en cuestión se obtuvieron los siguientes tiempos*

$$x_1 = 10,1 \;\; x_2 = 6,5 \;\; x_3 = 5,5 \;\; x_4 = 7,9 \;\; x_5 = 8,2 \;\; x_6 = 6,5$$

$$x_7 = 7,0 \;\; x_8 = 8,1 \;\; x_9 = 6,9 \;\; x_{10} = 7,7.$$

Se pide:

1. *Estimar por máxima verosimilitud la duración media del viaje.*
2. *Calcular la probabilidad de que, en valor absoluto, la diferencia entre la estimación media estimada y la real sea menor que un minuto.*

Sea X la variable aleatoria que mide la duración del viaje en minutos. X sigue una $N(\mu; 3)$. Tomamos una muestra aleatoria de tamaño 10, $(X_1, X_2, \cdots, X_{10})$.

1. Recordamos que la función de densidad de la variable aleatoria X es

$$f_\mu(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{3\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{18}}.$$

El estimador de máxima verosimilitud de μ es el valor $\hat{\mu}$ que maximiza en $[0, \infty)$lo que llamamos la funciín de verosimilitud de la muestra

$$(x_1 = 10,1; x_2 = 6,5; \cdots x_9 = 6,9; x_{10} = 7,7)$$

obtenida

$$L(\mu) = f_\mu(x_1, x_2, \cdots, x_10) = \frac{1}{3\sqrt{2\pi}} e^{-\frac{(x_1-\mu)^2}{18}} \cdot \frac{1}{2\sqrt{3\pi}} e^{-\frac{(x_2-\mu)^2}{18}} \cdots \frac{1}{3\sqrt{2\pi}} e^{-\frac{(x_{10}-\mu)^2}{18}}$$

$$= \frac{1}{(18\pi)^5} e^{-\frac{\sum_{i=1}^{10}(x_i-\mu)^2}{18}}.$$

Ya que $L(\mu)$ es continua, acotada en $[0, \infty)$ y $\lim_{\mu\to\infty} L(\mu) = 0$, $L(\mu)$ alcanza u valor máximo en un punto donde se anule la derivada.

$$L'(\mu) = \frac{1}{(18\pi)^5} \frac{\sum_{i=1}^{10}(x_i - \mu)}{18} e^{-\frac{\sum_{i=1}^{10}(x_i-\mu)^2}{18}}.$$

$$L'(\mu) = 0 \Longleftrightarrow \sum_{i=1}^{10}(x_i - \mu) = 0 \Longleftrightarrow \mu = \frac{\sum_{i=1}^{10} x_i}{10}.$$

Como solo existe un punto donde se anula la derivada, en este es donde $L(\mu)$ alcanza su valor máximo. El estimador de máxima verosimilitud para μ es

$$\hat{\mu} = \frac{\sum_{i=1}^{10} x_i}{10} = 7,44.$$

2. La media muestral

$$\overline{X} = \frac{1}{10}\sum_{i=1}^{10} X_i \sim N\left(\mu; \frac{3}{\sqrt{10}}\right).$$

El ejercicio nos pide $P(|\overline{X} - \mu| < 1)$.

$$P(|\overline{X} - \mu| < 1) = P(-1 < \overline{X} - \mu < 1) = P\left(-\frac{\sqrt{10}}{3} < \frac{\overline{X} - \mu}{\frac{3}{\sqrt{10}}} < \frac{\sqrt{10}}{3}\right)$$

$$= 1 - 2P\left(\frac{\overline{X} - \mu}{\frac{3}{\sqrt{10}}} > \frac{\sqrt{10}}{3}\right) = 0,70804.$$

$\square$

Ejercicio 23 *El tiempo de vida en dias X de los individuos de una población afectados de una nueva enfermedad es una variable aleatoria continua con función de densidad, dependiendo de un parámetro desconocido $\phi > 0$, dada por*

$$f_\phi(x) = \begin{cases} 0 & x \leq \phi \\ 2\phi^2 x^{-3} & x > \phi \end{cases}$$

Con objeto de estimar el parámetro ϕ, se extrajo una muestra aleatoria simple de dicha población, obteniéndose los siguientes tiempos de vida, en dias dias, de los 1o individuos seleccionados, todos los cuales fallecieron por enfermedad en estudio:

$$x_1 = 398 \;\; x_2 = 356 \;\; x_3 = 615 \;\; x_4 = 265 \;\; x_5 = 650$$

$$x_6 = 325 \;\; x_7 = 400 \;\; x_8 = 223 \;\; x_9 = 368 \;\; x_{10} = 680.$$

Determínese la estimación de máxima verosimilitud de ϕ.

La función de verosimilitud de la variable aleatoria en estudio X, tiempo de vida de los individuos afectados por la enfermedad en estudio, nos indica que dichos individuos contraen la enfermedad en un momento desconocido, ϕ, (puesto que en este punto la función de distribución $F(x)$ empieza a crecer desde cero, o lo que es lo mismo la función de supervivencia $S(x) = 1 - F(x)$ vale 1, lo que quiere decir que todos los individuos están vivos), momento a partir del cual, y por la forma de dicha función de densidad, la probabilidad de sobrevivir va disminuyendo.

Es precisamente el inicio de la enfermedad lo que queremos estimar. El estimador de máxima verosimilitud de ϕ es el valor $\widehat{\phi}$ que maximiza en $(0, \infty)$ la función de verosimilitud de la muestra $(x_1, x_2, \cdots, x_{10})$

$$L(\phi) = f_\phi(x_1, x_2, \cdots, x_{10}) = f_\phi(x_1) \cdot f_\phi(x_2) \cdots f_\phi(x_{10})$$

$$= 2^{10}\phi^{20} \prod_{i=1}^{10} x_i^{-3}, \qquad x_i > \phi \;\; i = 1, 2, \cdots 10.$$

Como siempre, el método de máxima verosimilitud se basa en asignar a ϕ el valor que maximice $L(\phi)$; el problema es que ahora ϕ aparece en el recorrido de la variable, es decir, que $L(\phi)$ toma un valor distinto de cero si $\phi < x_i$ para todo $i == 1, 2, \cdots 10$, y si algún x_i es tal que $X_i \leq \phi$ entonces $L(\phi) = 0$. En la estimación de ϕ habrá que tener en cuenta también el recorrido de ϕ.

La función $L(\phi) = 2^{10}\phi^{20} \prod_{i=1}^{10} x_i^{-3}$ crece al crecer ϕ, por lo que será tanto mayor cuanto mayor sea ϕ, y esto hasta que ϕ sea mayor o igual que algún x_i, que entonces $L(\phi)$ vale cero. Por tanto el valor de ϕ que maximiza $L(\phi)$ es el mínimo de $\{x_1, x_2, \cdots, x_{10}\}$. Este valor corresponde a 223, que será el valor estimado para ϕ.

□

COMPARACIÓN ENTRE EL MÉTODO DE LOS MOMENTOS Y EL DE MÁXIMA VEROSIMILITUD.

El siguiente comentario ha sido tomado del libro, páginas 88 y 89:

-ELEMENTOS DE LA ESTADÍSTICA APLICADA. Teoría de muestras e inferencia estadística. -José Javier Muruzábal Irigoyen. -Ed. Colegio de Ingenieros de Caminos, Canales y Puertos.

Según acabamos de ver, el método de los moméntos y de máxima verosimilitud coinciden en el caso de la estimación de la esperanza y varianza de poblaciones normales, pero no coinciden en el caso de la amplitud del intervalo de una variable aleatoria uniforme. Ello corrobora la consideración anterior acerca de que los resultados de ambos métodos son con frecuencia coincidentes, sin que ello pueda ser elevado a una categiría universal.

Con frecuencia, cuando los dos métodos no coinciden en sus resultados, la estimación por máxima verosimilitud, (más complicada en sus cálculos), conduce a estimadores de menor riesgo que el método de los momentos, (más sencillo desde el punto de vista operativo), sin que pueda enunciarse una regla general en este sentido. Sin embargo, también ha quedado demostrado que en el caso de la amplitud del intervalo de una distribución uniforme, el estimador de máxima verosimilitud no es apropiado porque tiende a subestimar el valor del parámetro, y esto hace que pueda llegar a ser rechazado.

Como ya se ha indicado, a la hora de comparar dos estimadores para elegir el mejor, la regla general es que no hay reglas generales que permitan establecer que un tipo de estimador es siempre mejor que otro. Ello lleva, por tanto, a la necesidad de decidir en cada casoconcreto calculando el riesgo o error de los estimadores para elegir el de menor valor.

Esto quiere decir que no es posible predecir de antemano si los estimadores obtenidos por el método de los momentos son mejores que los obtenidos por máxima verosimilitud o viceversa.