Técnicas Cuantitativas II

Estimación puntual de parámetros

Concepto de estimador paramétrico
Obtención de estimadores puntuales
Propiedades deseables de un estimador paramétrico

Índice

Introducción a la estimación: concepto de estimador de un parámetro

Supongamos que el número medio de mensajes multimedia recibidos en el móvil semanalmente por los estudiantes de la Universidad de Granada puede considerarse como una variable aleatoria con distribución Poisson de parámetro λ. Un problema fundamental en estadística es conocer el valor del parámetro desconocido λ, para lo cual existen distintas técnicas.

En el presente tema se presenta la estimación puntual, que consiste en obtener un único número calculado a partir de las observaciones muestrales y que es utilizado como estimación del valor del parámetro desconocido. Así, se entiende por **estimador** un estadístico muestral usado para calcular una aproximación numérica de un parámetro desconocido de la población. Esto es:

Definición 1 Sea θ el parámetro desconocido de una determinada distribución de probabilidad. Un estimador de θ, que se denota como $\hat{\theta}(x_1, \ldots, x_n)$, será un estadístico muestral que aproxime el verdadero valor de θ a partir de la muestra.

Téngase en cuenta que el estimador depende sólo de la muestra y, en ningún caso, del parámetro que se desea estimar. Por tanto, el estimador será un valor numérico distinto dependiendo de la muestra considerada. Además, para cada parámetro pueden existir varios estimadores diferentes. En general, escogeremos aquel que posea mejores propiedades.

Ejemplo 1 La inversión en publicidad de seis empresas automovilistas españolas en millones de euros durante el año 2008 fue 86’6, 69’7, 45’1, 44, 42 y 40’1. Suponiendo que los datos se distribuyen según una normal, se pide obtener estimaciones puntuales de la inversión media en publicidad y de la proporción de estas cuya inversión en publicidad fue inferior a 45.

Como veremos, un estimador de la media poblacional sería la media muestral:

$$\hat{\mu}(x_1, \ldots, x_6) = \bar{X} = \frac{86'6 + 69'7 + 45'1 + 44 + 42 + 40'1}{6} = 54'58.$$

Considerando la proporción muestral como estimador puntual de la proporción poblacional. Dado que tres de las seis empresas no superaron los 45 millones de euros en inversion en publicidad, a partir de

$$A_i = \begin{cases} 1 & \text{, la empresa } i \text{ tiene inversión inferior a 45} \\ 0 & \text{, en otro caso} \end{cases}$$

con $i = 1, \ldots, 6$, se tiene que $\hat{p} = \frac{3}{6} = 0'5$. $\sharp$

Método de máxima verosimilitud para la obtención de estimadores puntuales

En esta sección empezamos a estudiar los métodos de obtención de estimadores puntuales. Aunque se pueden obtener estimadores de forma simultánea para todos los parámetros poblacionales desconocidos de una distribución, nos centraremos en estudiar el caso en el que la distribución de probabilidad depende de un único parámetro desconocido.

El **método de máxima verosimilitud** consiste en escoger como estimador puntual aquel valor del parámetro que maximice la probabilidad de aparición de los valores muestrales obtenidos, es decir, aquel que maximice la función de probabilidad.

Ejemplo 2 Supongamos que el número de inundaciones por año sigue una distribución de Poisson con parámetro desconocido λ. En una determinada provincia han ocurrido dos inundaciones durante un año, ¿cómo se puede estimar el valor de λ si solo conocemos una observación?

Si los posibles valores de λ fueran 1, 1'5, 2, 2'5 y 3, la probabilidad de obtener la observación dos sería, respectivamente, 0'1839, 0'2510, 0'2770, 0'2565 y 0'2250.

Observando las tablas correspondientes a la distribución de Poisson se aprecia que el valor de λ que maximiza la probabilidad de aparición del valor observado es 2. $\sharp$

Sea X una variable aleatoria (discreta o continua) cuya distribución de probabilidad depende de un único parámetro desconocido θ. Dada una muestra aleatoria simple, $X_1, \ldots, X_n$, de X, se define la función de verosimilitud como

$$L(\theta; x_1, \ldots, x_n) = \begin{cases} \displaystyle\prod_{i=1}^{n} P[X_i = x_i; \theta] & \text{, caso discreto} \\ \displaystyle\prod_{i=1}^{n} f(x_i; \theta) & \text{, caso continuo} \end{cases},$$

donde $p_i = P[X_i = x_i; \theta]$ denota la función de cuantía y f la función de densidad, dependiendo de la naturaleza de la variable.

Adviértase, que mientras que las distribuciones de probabilidad dependen de la muestra, la función de verosimilitud depende del parámetro desconocido, θ.

Entonces, el método para encontrar el estimador de θ se convierte en el problema matemático de maximizar la función $L(\theta; x_1, x_2, \ldots, x_n)$. Es decir, hallar aquellos valores que anulan la primera derivada de $L(\theta; x_1, x_2, \ldots, x_n)$ y hacen negativa a la segunda.

Luego $\widehat{\theta} = \widehat{\theta}(x_1, x_2, \ldots, x_n)$ será el estimador máximo verosímil del parámetro θ si verifica que:

$$\begin{aligned} \frac{\partial}{\partial\theta} L(\widehat{\theta}; x_1, x_2, \ldots, x_n) &= 0, \\ \frac{\partial^2}{\partial\theta^2} L(\widehat{\theta}; x_1, x_2, \ldots, x_n) &< 0. \end{aligned}$$

Puesto que al derivar directamente la función de verosimilitud se suelen obtener expresiones complicadas, normalmente se maximiza el logaritmo neperiano de la función de verosimilitud, ya que la solución obtenida es la misma al ser monótona creciente la función logaritmo.

Por tanto, el estimador máximo verosímil será aquel valor $\widehat{\theta}$ que verifique que:

$$\begin{aligned} \frac{\partial}{\partial\theta} \ln L(\widehat{\theta}; x_1, x_2, \ldots, x_n) &= 0, \\ \frac{\partial^2}{\partial\theta^2} \ln L(\widehat{\theta}; x_1, x_2, \ldots, x_n) &< 0. \end{aligned}$$

A modo de resumen, para obtener un estimador por máxima verosimilitud, los pasos a seguir son los siguientes:

- Calcular la función de verosimilitud de la muestra teniendo en cuenta la distribución de probabilidad de los datos obtenidos.
- Determinar el logaritmo de la función de verosimilitud. Se puede demostrar que si una función alcanza un máximo en un punto, su logaritmo neperiano alcanza el máximo en ese mismo punto. Esto nos facilitará buscar el estimador máximo-verosímil, ya que generalmente es más fácil determinar el máximo del logaritmo de una función que directamente el de la función.
- Derivar el logaritmo neperiano con respecto al parámetro que se pretende estimar e igualar a cero para determinar el valor que maximiza dicha función. De esta expresión despejaremos el estimador, comprobando que la segunda derivada evaluada en dicho punto es negativa.

A continuación vamos a obtener el estimador máximo verosímil del parámetro λ de la distribución de Poisson para las siguientes muestras correspondientes al número de visitas semanales recibidas por un profesor en tutorías:

Muestra 1: $3, 3, 6, 2, 1, 4, 0, 2, 2, 5.$

Muestra 2: $1, 3, 1, 4, 5, 3, 4, 6, 1, 3.$

Dada una muestra aleatoria simple $X_1, \ldots, X_n$ procedente de una $P(\lambda)$, sabemos que la función de probabilidad en este caso es

$$P[X_i = x_i; \lambda] = e^{-\lambda} \cdot \frac{\lambda^{x_i}}{x_i!},$$

por lo que la función de verosimilitud se expresará como:

$$\begin{aligned} L(\lambda; x_1, x_2, \ldots, x_n) &= \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} \\ &= e^{-\lambda} \frac{\lambda^{x_1}}{x_1!} \cdot e^{-\lambda} \frac{\lambda^{x_2}}{x_2!} \cdots e^{-\lambda} \frac{\lambda^{x_n}}{x_n!} = e^{-n\lambda} \cdot \frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}. \end{aligned}$$

Método de máxima verosimilitud: Ejemplo 1

Entonces, tomando logaritmos

$$\ln L(\lambda; x_1, x_2, \ldots, x_n) = \left(\sum_{i=1}^{n} x_i\right) \cdot \ln \lambda - \ln\left(\prod_{i=1}^{n} x_i!\right) - n\lambda,$$

y derivando la expresión anterior

$$\frac{\partial}{\partial \lambda} \ln L(\lambda; x_1, x_2, \ldots, x_n) = \frac{\sum\limits_{i=1}^{n} x_i}{\lambda} - n.$$

Igualando a cero la derivada anterior se obtiene la igualdad

$$\frac{\sum\limits_{i=1}^{n} x_i}{\lambda} - n = 0,$$

cuya solución es el posible estimador puntual buscado

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

Método de máxima verosimilitud: Ejemplo 1

Finalmente, ya tan sólo queda calcular la segunda derivada

$$\frac{\partial^2}{\partial\lambda^2}\ln L(\lambda; x_1, x_2, \ldots, x_n) = -\frac{\sum\limits_{i=1}^{n} x_i}{\lambda^2}.$$

y evaluarla en $\widehat{\lambda}$ para estudiar su signo. En este caso es evidente que

$$\frac{\partial^2}{\partial\lambda^2}\ln L(\widehat{\lambda}; x_1, x_2, \ldots, x_n) = -\frac{\sum\limits_{i=1}^{n} x_i}{\widehat{\lambda}^2} < 0,$$

ya que $\widehat{\lambda}^2 > 0$ y, puesto que la distribución de Poisson solo toma valores positivos, es decir, $x \in \{0, 1, \ldots\}$, es claro que $\sum\limits_{i=1}^{n} x_i > 0$.

Por tanto, para obtener el valor del estimador máximo verosímil para cada una de las muestras consideradas tan sólo hay que calcular la media aritmética de cada colección de datos:

$$\widehat{\lambda}_1 = \frac{3+3+\cdots+2+5}{10} = 2'8, \quad \widehat{\lambda}_2 = \frac{1+3+\cdots+1+3}{10} = 3'1.$$

Método de máxima verosimilitud: Ejemplo 2

Una variable aleatoria con distribución gamma, con parámetro α conocido, tiene función de densidad:

$$f(x;\theta) = \frac{1}{\Gamma(\alpha)\theta^{\alpha}} \cdot x^{\alpha-1} \cdot \exp\left\{-\frac{x}{\theta}\right\}, \quad x > 0,$$

con valor esperado $E[X] = \alpha\theta$ y varianza $Var(X) = \alpha\theta^2$.
Para obtener el estimador puntual para θ mediante el método de máxima verosimilitud hay que obtener en primer lugar la función de verosimilitud muestral. Es decir

$$L(\theta; x_1, x_2, \ldots, x_n) = \frac{1}{\Gamma(\alpha)^n \theta^{n\alpha}} \cdot \prod_{i=1}^{n} x_i^{\alpha-1} \cdot \exp\left\{-\frac{\sum_{i=1}^{n} x_i}{\theta}\right\}.$$

Tomando logaritmo en la verosimilitud

$$\ln L(\theta; x_1, x_2, \ldots, x_n) = -n\alpha \ln\theta - n\ln\Gamma(\alpha) + (\alpha-1)\sum_{i=1}^{n} \ln x_i - \frac{\sum_{i=1}^{n} x_i}{\theta},$$

derivando con respecto a θ

$$\frac{\partial}{\partial\theta}\ln L(\theta; x_1, x_2, \ldots, x_n) = -\frac{n\alpha}{\theta} + \frac{\sum_{i=1}^{n} x_i}{\theta^2},$$

e igualando a cero

$$-\frac{n\alpha}{\theta} + \frac{\sum_{i=1}^{n} x_i}{\theta^2} = 0,$$

se obtiene la solución

$$\widehat{\theta}(x_1, \ldots, x_n) = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i}{\alpha} = \frac{\overline{x}}{\alpha}.$$

Para confirmar que este valor es un máximo, hay que comprobar que la derivada segunda es negativa cuando se sustituye en ella θ por la solución obtenida.

En efecto, puesto que

$$\frac{\partial^2}{\partial\theta^2}\ln L(\theta; x_1, x_2, \ldots, x_n) = \frac{n\alpha}{\theta^2} - \frac{2\theta\sum_{i=1}^{n} x_i}{\theta^4} = \frac{n\alpha}{\theta^2} - \frac{2\sum_{i=1}^{n} x_i}{\theta^3},$$

si la evaluamos en $\widehat{\theta}(x_1, \ldots, x_n)$ se obtiene que

$$\begin{aligned}\frac{\partial^2}{\partial\theta^2}\ln L(\widehat{\theta}; x_1, x_2, \ldots, x_n) &= \frac{n\alpha}{\frac{\overline{x}^2}{\alpha^2}} - \frac{2\sum_{i=1}^{n} x_i}{\frac{\overline{x}^3}{\alpha^3}} = \frac{n\alpha^3}{\overline{x}^2} - \frac{2n\overline{x}\alpha^3}{\overline{x}^3} \\ &= \frac{n\alpha^3}{\overline{x}^2} - \frac{2n\alpha^3}{\overline{x}^2} = -\frac{n\alpha^3}{\overline{x}^2},\end{aligned}$$

donde se ha usado que $n\overline{x} = \sum_{i=1}^{n} x_i$.

Puesto que $n, \alpha, \overline{x}^2 > 0$, la segunda derivada evaluada en $\widehat{\theta}$ es negativa, por lo que $\widehat{\theta}$ es el estimador máximo verosímil de θ.

Método de los momentos para la obtención de estimadores puntuales

Este método se basa en la relación existente entre los parámetros poblacionales desconocidos y los momentos poblacionales. Más concretamente, consiste en trasladar esa relación sustituyendo los momentos poblacionales por los muestrales, para así estimar los primeros.

Dada una variable aleatoria, X, cuya distribución de probabilidad depende de un único parámetro desconocido, el método de los momentos consiste en enfrentar características poblacionales, $E[X^r]$, a sus homólogas en la muestra

$$\sum_{i=1}^{n} x_i^r \cdot n_i.$$

En el caso que se estudia un único parámetro desconocido, habrá que enfrentar un único parámetro poblacional a un único parámetro muestral ($r = 1$). Es decir, $E[X] = \bar{X}$, de forma que se plantea una ecuación, cuya solución es el estimador puntual buscado.

A continuación se van a obtener los estimadores puntuales por el método de los momentos de las distribuciones de Poisson de parámetro λ y binomial de parámetro p.
Adviértase que la igualdad a plantear es $E[X] = \bar{X}$.
En el caso de la binomial, $E[X] = np$, luego entonces $np = \bar{X}$, y por tanto, el estimador puntual buscado es

$$\widehat{p} = \frac{\bar{X}}{n}.$$

Mientras que para la Poisson, $E[X] = \lambda$, luego directamente se obtiene que $\widehat{\lambda} = \bar{X}$.
Obsérvese que los estimadores obtenidos coinciden con los del método anterior. Esto se debe a que, bajo ciertas condiciones, ambos métodos coinciden.

Método de los momentos: Ejemplo 2

Para estimar el parámetro θ de una variable aleatoria con distribución gamma, con parámetro α conocido, por el método de los momentos, simplemente habría que igualar la media poblacional a la media muestral y despejar θ. Esto es

$$E[X] = \overline{X} \Rightarrow \alpha\theta = \overline{X} \Rightarrow \theta = \frac{\overline{X}}{\alpha}.$$

Puede parecer que este método es más sencillo de aplicar que el anterior, pero nada más lejos de la realidad. Supongamos que se ignora en este caso que $E[X] = \alpha\theta$, entonces tendría que calcular dicho valor esperado y, por tanto, se empieza a complicar este método.

Propiedades deseables para un estimador paramétrico: insesgadez, consistencia, eficiencia y suficiencia

Tras estudiar métodos enfocados a la obtención de estimadores, a continuación vamos a preguntarnos por la calidad de las estimaciones que estos proporcionan. Necesitamos alguna medida que nos permita seleccionar el mejor estimador, ya que recordemos que para un mismo parámetro desconocido es posible que exista más de un estimador puntual.

De esta forma, desearemos que los valores proporcionados por el estimador para cada muestra se aproximen lo más posible al verdadero valor del parámetro (insesgadez), siendo las posibles diferencias existentes lo más pequeñas posibles (eficiencia). Además, conforme aumente la muestra menor ha de ser el error que se cometa en la estimación[1] (consistencia) y el estimador obtenido debe contener toda la información existente en la muestra (suficiencia).

A continuación se estudiará como comprobar si un estimador verifica cada una de estas propiedades.

[1] O lo que es lo mismo, conforme aumente la muestra la aproximación numérica proporcionada por el estimador debe tender al verdadero valor del parámetro.

Dada una variable aleatoria, X, cuya distribución de probabilidad depende de un parámetro desconocido, θ, se dice que el estimador $\widehat{\theta}(X_1, \ldots, X_n)$ de θ es insesgado si $E[\widehat{\theta}] = \theta$. Si no se verifica la condición anterior, se dice que no es insesgado o que es sesgado, y que tiene un sesgo $d\left(\widehat{\theta}\right) = E\left[\widehat{\theta}\right] - \theta$.

Ejemplo 3 Sea X una variable aleatoria cuya distribución de probabilidad depende de un parámetro desconocido θ tal que $E[X] = \theta$ y sea $\widehat{\theta}(X_1, \ldots, X_n) = \bar{X}$ un estimador puntual de dicho parámetro.
Puesto que se ha demostrado que $E[\overline{X}] = E[X]$, entonces se verifica que $E[\widehat{\theta}] = E[\overline{X}] = \theta$, por lo que $\widehat{\theta}$ es un estimador insesgado de θ, es decir, que la media aritmética siempre es estimador insesgado de la media poblacional.
Por tanto, en el caso en el que X se distribuya según una Normal, una Poisson o una Bernoulli ($\theta = \mu$, $\theta = \lambda$ y $\theta = p$, respectivamente), puesto que en estos casos se verifica que $\widehat{\theta}(X_1, \ldots, X_n) = \bar{X}$, entonces $\widehat{\theta}$ es un estimador puntual insesgado de θ de forma inmediata. $\sharp$

Ejemplo 4 Dada una muestra aleatoria simple $X_1, \ldots, X_n$ procedente de una variable aleatoria, X, distribuida según una normal de parámetros μ y σ^2, entonces se verifica que la media poblacional es $E[X] = \mu$ y la varianza poblacional es $Var(X) = \sigma^2$.
Por tanto, en tal caso hemos demostrado que se verifica que $E[S^2_{n-1}] = \sigma^2$ y $E[S^2_n] = \frac{n-1}{n}\sigma^2$, es decir, la cuasivarianza muestral es un estimador insesgado de σ^2, mientras que la varianza muestral lo es segado. $\sharp$

Ejemplo 5 Una vez obtenido el estimador para el parámetro θ de una distribución gamma con α conocido, $\widehat{\theta}(x_1, \ldots, x_n) = \frac{\overline{x}}{\alpha}$, para estudiar si es insesgado hemos de comprobar que se verifica que $E[\widehat{\theta}] = \theta$. En efecto,

$$\begin{aligned} E[\widehat{\theta}] &= E\left[\frac{\overline{X}}{\alpha}\right] = \frac{1}{\alpha}E[\overline{X}] = \frac{1}{\alpha}E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n\alpha}E\left[\sum_{i=1}^{n} X_i\right] \\ &= \frac{1}{n\alpha}\sum_{i=1}^{n} E[X_i] = \frac{1}{n\alpha}\sum_{i=1}^{n} \alpha\theta = \frac{1}{n\alpha}n\alpha\theta = \theta. \end{aligned}$$

$\sharp$

Dada una variable aleatoria, X, cuya distribución de probabilidad depende de un parámetro desconocido, θ, y siendo $\widehat{\theta}$ un estimador puntual de θ. Bajo condiciones de regularidad, los estimadores de un parámetro θ verifican siempre la siguiente desigualdad

$$Var\left(\widehat{\theta}\right) \geq \frac{1}{n \cdot E\left[\left(\frac{\partial}{\partial\theta} \ln L(\theta; x)\right)^2\right]}.$$

Entonces, un estimador es eficiente si es insesgado[2] y se verifica la igualdad en la expresión anterior.

La expresión

$$\frac{1}{n \cdot E\left[\left(\frac{\partial}{\partial\theta} \ln L(\theta; x)\right)^2\right]},$$

recibe el nombre de cota de Frechet-Cramer-Rao.

[2] Por tanto, los estimadores sesgados no son eficientes.

De igual forma, para estudiar si es eficiente el estimador para el parámetro θ de una distribución gamma con α conocido, hay que comprobar que $Var\left(\widehat{\theta}\right)$ coincide con la cota de Frechet-Cramer-Rao. En efecto,

$$\natural Var\left(\widehat{\theta}\right) = Var\left(\frac{\overline{X}}{\alpha}\right) = \frac{1}{\alpha^2}Var(\overline{X}) = \frac{1}{\alpha^2}\cdot\frac{Var(X)}{n} = \frac{1}{\alpha^2}\cdot\frac{\alpha\theta^2}{n} = \frac{\theta^2}{\alpha n}.$$

$$\natural L(\theta;x) = \frac{1}{\theta^\alpha}\cdot\exp\left\{-\frac{x}{\theta}\right\} \Rightarrow \ln L(\theta;x) = -\alpha\ln\theta - \frac{x}{\theta} \Rightarrow$$

$$\Rightarrow \frac{\partial}{\partial\theta}\ln L(\theta;x) = -\frac{\alpha}{\theta} + \frac{x}{\theta^2} = \frac{x-\alpha\theta}{\theta^2} \Rightarrow$$

$$\Rightarrow E\left[\left(\frac{\partial}{\partial\theta}\ln L(\theta;x)\right)^2\right] = E\left[\left(\frac{X-\alpha\theta}{\theta^2}\right)^2\right] = \frac{1}{\theta^4}E\left[(X-\alpha\theta)^2\right] =$$

$$= \frac{1}{\theta^4}E\left[(X-E[X])^2\right] = \frac{1}{\theta^4}Var(X) = \frac{\alpha\theta^2}{\theta^4} = \frac{\alpha}{\theta^2} \Rightarrow$$

$$\Rightarrow \frac{1}{nE\left[\left(\frac{\partial}{\partial\theta}\ln L(\theta;x)\right)^2\right]} = \frac{\theta^2}{\alpha n}.$$

Estudiemos a continuación si el estimador puntual $\widehat{\lambda}(X_1, \ldots, X_n) = \bar{X}$ del parámetro desconocido λ de una variable aleatoria discreta, X, distribuida según una Poisson, que sabemos que es insesgado, es también eficiente.
En este caso, se tiene que la función de cuantía es:

$$P[X = x; \lambda] = e^{-\lambda} \cdot \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \ldots, \quad \lambda > 0,$$

con $E[X] = \lambda = Var(X)$.
Con dicha información:

$$Var\left(\widehat{\lambda}\right) = Var\left(\overline{X}\right) = \frac{Var\,(X)}{n} = \frac{\lambda}{n}.$$

Mientras que partiendo de la función de verosimilitud para una muestra de tamaño uno, que coincidirá, por tanto, con la función de cuantía anterior:

$$L(\lambda; x) = e^{-\lambda} \cdot \frac{\lambda^x}{x!}.$$

Tomando logaritmos en la expresión anterior:

$$\begin{aligned}\ln L(\lambda;x) &= \ln\left(e^{-\lambda}\cdot\frac{\lambda^x}{x!}\right)=\ln e^{-\lambda}+\ln\frac{\lambda^x}{x!}\\ &= -\lambda\ln e+\ln\lambda^x-\ln x! = -\lambda + x\ln\lambda-\ln x!,\end{aligned}$$

donde se han usado las siguientes propiedades básicas del logaritmo neperiano: $\ln(a\cdot b)=\ln\ a+\ln\ b,\quad \ln a^b=b\ln\ a,\quad \ln\frac{a}{b}=\ln\ a-\ln\ b,\quad \ln e=1.$
El siguiente paso será derivar con respecto al parámetro desconocido:

$$\frac{\partial}{\partial\lambda}\ln L(\lambda;x)=-1+\frac{x}{\lambda}=\frac{x-\lambda}{\lambda}.$$

Entonces, elevando al cuadrado y considerando la esperanza matemática en la expresión anterior:

$$E\left[\left(\frac{\partial}{\partial\lambda}\ln L(\lambda;x)\right)^2\right]=E\left[\left(\frac{X-\lambda}{\lambda}\right)^2\right]=\frac{E[(X-E[X])^2]}{\lambda^2}=\frac{Var(X)}{\lambda^2}=\frac{1}{\lambda},$$

donde se ha tenido en cuenta la definición de varianza y que $E[X]=\lambda=Var(X)$.

Finalmente, la cota de Frechet-Cramer-Rao queda

$$\frac{1}{nE\left[\left(\frac{\partial}{\partial p}\ln L(\lambda;x)\right)^2\right]}=\frac{\lambda}{n}.$$

Entonces, se deduce que el estimador $\widehat{\lambda}$ es eficiente ya que

$$Var\left(\widehat{\lambda}\right)=\frac{\lambda}{n}=\frac{1}{n\cdot E\left[\left(\frac{\partial}{\partial\theta}\ln L(\lambda;x)\right)^2\right]}.$$

Dada una variable aleatoria, X, cuya distribución de probabilidad depende de un parámetro desconocido, θ, y sea $\widehat{\theta}(X_1, \ldots, X_n)$ un estimador puntual de dicho parámetro. Se dice que dicho estimador es consistente cuando

$$\lim_{n\to\infty} P\left[|\widehat{\theta}(X_1, \ldots, X_n) - \theta| > \epsilon\right] = 0, \qquad \forall \epsilon > 0.$$

Puesto que la condición dada anteriormente no siempre es fácil de comprobar, en la práctica se trabajará con la siguiente condición suficiente:

- El estimador es asintóticamente insesgado: $\lim_{n\to\infty} E[\widehat{\theta}] = \theta$.
- El estimador tiene asintóticamente varianza cero: $\lim_{n\to\infty} Var[\widehat{\theta}] = 0$.

En el caso en el que tales condiciones no se cumplen por un estimador no podemos concluir que dicho estimador no es consistente.
Otra opción para estudiar si el estimador es consistente es usar la desigualdad de Tchebychev

$$P\left[|X - E[X]| > \epsilon\right] \leq \frac{Var(X)}{\epsilon^2}, \quad \forall \epsilon > 0.$$

Consistencia: Ejemplo 1

Dada una variable aleatoria continua con función de densidad

$$f(x, \lambda) = \frac{2}{\lambda^2}(\lambda - x), \qquad 0 < x < \lambda,$$

obtener el estimador del parámetro λ por el método de los momentos y comprobar que es consistente.

Usando el método de los momentos, el estimador para una muestra de tamaño n saldrá de la igualdad que se obtiene al igualar el momento poblacional de orden uno con su homólogo en la muestra

$$E[X] = \overline{x},$$

donde

$$E[X] = \int_0^\lambda x \cdot f(x)dx = \cdots = \frac{\lambda}{3}.$$

Por tanto,

$$\frac{\lambda}{3} = \overline{x} \Longrightarrow \widehat{\lambda} = 3\overline{x}.$$

Para estudiar la consistencia del estimador obtenido veremos si es asintóticamente insesgado y con varianza cero en el límite:

- $E[\widehat{\lambda}] = 3E[\overline{X}] = \frac{3}{n}\sum_{i=1}^{n} E[X_i] = \frac{3}{n}\sum_{i=1}^{n} \frac{\lambda}{3} = \lambda \Longrightarrow \lim_{n\to\infty} E[\widehat{\lambda}] = \lim_{n\to\infty} \lambda = \lambda$,
- $Var(\widehat{\lambda}) = 9Var(\overline{X}) = \frac{9}{n}Var(X) = \frac{9}{n}\cdot\frac{\lambda^2}{18} = \frac{\lambda^2}{2n} \Longrightarrow \lim_{n\to\infty} Var(\widehat{\lambda}) = \lim_{n\to\infty} \frac{\lambda^2}{2n} = 0$,

donde se ha usado que

$$E[X^2] = \int_0^{\lambda} x^2 \cdot f(x)dx = \cdots = \frac{\lambda^2}{6},$$

con lo que

$$Var(X) = E[X^2] - (E[X])^2 = \frac{\lambda^2}{6} - \frac{\lambda^2}{9} = \frac{\lambda^2}{18}.$$

Luego se ha demostrado que el estimador obtenido es consistente.

Comprobemos si el estimador puntual para el parámetro θ de una distribución gamma con α conocido es consistente. Puesto que sabemos que $\widehat{\theta}$ es una variable aleatoria, a partir de la desigualdad de Tchebychev se verifica que

$$P\left[|\widehat{\theta} - E[\widehat{\theta}]| > \epsilon\right] \leq \frac{Var(\widehat{\theta})}{\epsilon^2}.$$

Teniendo en cuenta que $\widehat{\theta}$ es insesgado y que $Var(\widehat{\theta}) = \frac{\theta^2}{\alpha n}$

$$P\left[|\widehat{\theta} - \theta| > \epsilon\right] \leq \frac{\theta^2}{\alpha n \epsilon^2},$$

y entonces considerando límites

$$\lim_{n\to\infty} P\left[|\widehat{\theta} - \theta| > \epsilon\right] \leq \lim_{n\to\infty} \frac{\theta^2}{\alpha n \epsilon^2} = 0.$$

Puesto que una probabilidad no puede ser negativa, realmente se ha obtenido que

$$\lim_{n\to\infty} P\left[|\widehat{\theta} - \theta| > \epsilon\right] = 0,$$

por lo que el estimador puntual $\widehat{\theta}$ es consistente.

Sea $X_1, \ldots, X_n$ una muestra aleatoria simple procedente de una variable aleatoria X cuya distribución de probabilidad depende de un parámetro desconocido θ. Se dice que el estimador $\hat{\theta}$ de θ es suficiente si, y solamente si, la función de verosimilitud de la muestra, $L(\theta; x_1, \ldots, x_n)$, se puede descomponer como producto de dos factores no negativos

$$L(\theta; x_1, \ldots, x_n) = h_1(x_1, \ldots, x_n) \cdot h_2(\widehat{\theta}; \theta),$$

donde $h_1(x_1, \ldots, x_n)$ no depende del parámetro θ y $h_2(\widehat{\theta}; \theta)$ es una función que depende solo de θ y de la muestra a partir del estimador.

Ejemplo 6 El estimador puntual para el parámetro θ de una distribución gamma con α conocido es $\widehat{\theta} = \frac{\overline{X}}{\alpha}$.
Puesto que la función de verosimilitud puede factorizarse según el teorema de Neyman se puede decir que $\hat{\theta}$ es un estadístico suficiente:

$$L(\theta; x_1, ..., x_n) = h_1(x_1, \ldots, x_n) \cdot h_2(\widehat{\theta}; \theta),$$

siendo $h_1(x_1, \ldots, x_n) = \prod_{i=1}^{n} x_i$ y $h_2(\widehat{\theta}; \theta) = \frac{e^{(-1/\theta) \cdot \sum_{i=1}^{n} x_i}}{\theta^{\alpha n}} = \frac{e^{(-1/\theta) \cdot \alpha n \widehat{\theta}}}{\theta^{\alpha n}}$. ♯

Ejercicios Propuestos

2.1 El departamento de calidad de cierta empresa realiza diez llamadas diarias para conocer la satisfacción del cliente. Imaginemos que el número de clientes que se declaran insatisfechos diariamente se distribuye según una distribución binomial. Si de los últimos estudios se desprende que el número medio y más frecuente de clientes insatisfechos diariamente es 3. ¿Cuál sería el valor del estimador del parámetro p siguiendo el razonamiento intuitivo del método de estimación de maxima verosimilitud? ¿Y usando el método de los momentos?

2.2 En una distribución normal de media desconocida y varianza 25, se toman muestras aleatorias simples de tamaño 2, considerándose los siguientes estimadores de la media poblacional:

$$A = 0'65X_1 + 0'35X_2, \quad B = X_1 + X_2, \quad C = \frac{X_1 + X_2}{2}.$$

Determine cual de los tres es el mejor desde el punto de vista de sesgo y eficiencia.

2.3 Erase una vez un día, del mes de agosto de cualquier año, que estaba aburrido/a en mi casa, por lo que me asomé a la ventana y me dediqué a contar los coches que pasaban en intervalos de 2 minutos. Obtuve los siguientes datos: $12, 11, 9, 11, 10, 9, 11, 9, 10, 10, 8$. Sabiendo que la distribución que sigue la variable aleatoria *número de coches que pasan por debajo de la ventana de mi casa cada 2 minutos* es una distribución de Poisson con parámetro λ, se pide:

(a) Estimar, por el método de los momentos, el parámetro poblacional desconocido de la distribución anterior. Obtener un valor numérico a partir de la muestra anterior.

(b) Teniendo en cuenta el valor del apartado anterior para el parámetro poblacional desconocido, calcular la probabilidad de que por dicho punto pasen exactamente 9 coches cada 2 minutos.

2.4 Sea una población distribuida según una binomial de parámetros $n = 15$ y p, obtener el estimador de máxima verosimilitud utilizando una muestra aleatoria de tamaño 3.

2.5 En la siguiente tabla se recogen las previsiones realizadas en abril de 2007 por el Instituto Nacional de Estadística (INE) y la Fundación de las Cajas de Ahorro (FUNCAS) para la tasa de variación anual del IPC en tanto por cien para los distintos meses del año 2008:

	E	F	M	A	M	J	J	A	S	O	N	D
Previsiones	3'1	3'1	2'9	2'7	2'8	2'8	2'8	2'8	2'8	2'8	2'8	2'8
Observado	4'3	4'4	4'5	4'2	4'6	5	5'3	4'9	4'5	3'6	2'4	1'4
Errores	1'2	1'3	1'6	1'5	1'8	2'2	2'5	2'1	1'7	0'8	-0'4	-1'4

Suponiendo que los errores en la predicción de la tasa de variación anual del IPC se distribuyen normalmente, calcule una estimación para el error medio y su varianza.

2.6 Sea una población $N(10, \sigma^2)$ donde σ^2 es desconocida. Se pide obtener el estimador de máxima verosimilitud de σ^2 para una muestra de tamaño 20.

2.7 El número de suspensiones de pago que suceden en una semana en determinada ciudad, es una variable aleatoria que sigue una distribución de Poisson. Comprobar que el estimador del número medio de suspensiones de pagos semanales, parámetro λ, es suficiente.

2.8 Supongamos que el número de siniestros de responsabilidad civil declarados mensualmente por una determinada empresa de transporte de mercancías por carretera es una variable aleatoria con distribución de Poisson de parámetro λ. Se realiza un estudio de los últimos 7 meses y se obtienen los siguientes resultados: 3, 5, 2, 1, 2, 3 y 4 siniestros, respectivamente. Determine la mejor estimación que se puede obtener para el parámetro λ a partir de estos datos y explique las razones teóricas que le han hecho tomar esa decisión.

2.9 Según un estudio de *Dun & Braderstreet* presentado el 30 de enero de 2009, el número de procesos de quiebra e insolvencia en Portugal durante el año 2008 fueron 3344, lo que supone aproximadamente 279 procesos mensuales de media. Suponiendo que el número de quiebras mensuales se ajusta a una distribución de Poisson, se pide obtener una estimación máximo verosímil del parámetro λ.

2.10 Suponga que una empresa produce cierto material de manera que el peso del producto final sigue una distribución normal con desviación típica igual a $0'7$. Deduzca el estimador de máxima verosimilitud para el peso medio del producto final a partir de una muestra de tamaño 10.

2.11 La proporción de hombres con cargo directivo se considera una variable aleatoria con función de densidad:

$$f(x) = \frac{1}{\lambda} x^{\frac{1}{\lambda}-1},$$

para todo $x \in (0,1)$ y $\lambda > 0$.

(a) Determine el estimador del parámetro λ por el método de máxima verosimilitud.

(b) Si suponemos una muestra de 16 empresas elegidas al azar, la proporción de hombres con cargo de directivo es

$$0'36, 0'71, 0'65, 0'94, 0'83, 0'75, 0'76, 0'64,$$

$$0'61, 0'56, 0'84, 0'67, 0'79, 0'52, 0'94, 0'73.$$

Se pide obtener una estimación del parámetro λ.

2.12 Comprobar si el estimador de máxima verosimilitud en una distribución $B(2,p)$ es insesgado.