Estimación Puntual

Introducción

- En **Probabilidades** se conocían todos los datos sobre la v.a.
  - Ejemplo: $X \sim B(25, 0.1)$ y podíamos calcular $P(X=5)$.
- En **Estadística** desconocemos las características de X total o parcialmente.
  - Ejemplo: sabemos que X tiene distribución binomial pero ***desconocemos p*** y a partir de una muestra de 25 artículos trataremos de hallar información sobre p.

Inferencia Estadística

- El campo de la ***inferencia estadística*** está formado por los métodos utilizados para tomar decisiones o para obtener conclusiones sobre el o los parámetros de una población a partir de la información contenida en una muestra.
- La inferencia estadística puede dividirse en dos grandes áreas:
  - Estimación de Parámetros.
  - Pruebas de hipótesis.

Estimación Puntual

- Dado un parámetro de interés, como una media poblacional μ o una proporción p de la población, el objetivo de la estimación puntual es usar una muestra para calcular un número que representa, en cierto sentido, una buena estimación del valor real del parámetro.
- El número resultante se llama ***estimación puntual***.
- La estimación se realiza a través de algún ***estadístico*** (ej: media muestral).

Ejemplo 7.1

- Suponga que el parámetro de interés es μ, el tiempo de vida útil promedio de baterías de cierto tipo.
- Ej: muestra de n=3 baterías. Para c/u se indica el tiempo de vida útil observado (horas)

  x_1=5.0, x_2 = 6.4 y x_3 = 5.9. El valor calculado del tiempo de vida útil promedio de la muestra es

$$\bar{x} = 5.77$$

Podría considerarse este valor como la mejor estimación de μ a partir de la información de la muestra.

Estimador Puntual

- Cuando un estadístico se utiliza para estimar un parámetro desconocido se lo llama ***estimador puntual.***
- Es habitual simbolizar en forma genérica a un parámetro con la letra θ y al estadístico que se utiliza como estimador puntual de θ , simbolizarlo con $\hat{\Theta}$.
- Al medir la muestra aleatoria se obtienen $x_1, x_2, \ldots, x_n$ *y entonces* ***el valor que toma*** $\hat{\Theta}$ ***es***

$$\hat{\theta} = h(x_1, x_2, \ldots, x_n)$$

y se denomina ***estimación puntual de*** $\boldsymbol{\theta}$

Estimadores usuales

- Sea X_1, X_2, … ,X_n una muestra aleatoria de una v.a. X donde $E(X)=\mu$ y $V(X)=\sigma^2$.
- Si se desconoce μ , un estadístico que se utiliza para estimar ese parámetro es la **media o promedio muestral**

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Estimadores usuales

- Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una v.a. X donde $E(X)=\mu$ y $V(X)=\sigma^2$.
- Si se desconoce σ^2, un estadístico que se utiliza para estimar ese parámetro es la **varianza muestral**

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

Luego veremos porque utiliza (n-1) en lugar de n

Estimadores usuales

- Si se busca estimar la proporción p de objetos de una población que cumplen una determinada característica, el estimador puntual sería

$$\hat{p} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

donde X_i vale 1 si la i-ésima observación tiene la característica de interés y 0 si no, para i=1,2,…, n.

Más de un estimador

- Puede ocurrir que se tenga más de un estimador para un parámetro.
- Ejemplo: Para estimar el valor de μ puede utilizarse la media muestral o también la semisuma entre X_1 y X_n

$$\hat{\mu} = \frac{X_1 + X_n}{2}$$

- En estos casos necesitamos de algún ***criterio para decidir cuál es mejor estimador*** de μ

Criterios para evaluar estimadores

- En una situación ideal, se podría encontrar un estimador $\hat{\Theta}$ para el cual siempre se cumpla que $\hat{\theta} = \theta$.
- Sin embargo, $\hat{\Theta}$ es una función de las X_i muestrales y por lo tanto, una v.a.

$$\therefore \quad \hat{\theta} = \theta + \underbrace{error\ de\ estimacion}$$

Lo que se necesita es un estimador que tenga las propiedades de *insesgamiento* y *varianza mínima*.

Estimador insesgado

- Se dice que un estimador puntual $\hat{\Theta}$ es ***estimador insesgado*** del parámetro θ si

$$E(\hat{\Theta}) = \theta$$

cualquiera sea el valor verdadero de θ.

- **Sesgo del estimador** : $E(\hat{\Theta}) - \theta$
- Note que si el estimador es insesgado, su sesgo es cero.

Ejemplo 7.2

- Sea X_1, X_2, … ,X_n una muestra aleatoria de una v.a. X donde $E(X)=\mu$ y $V(X)=\sigma^2$.
- Veamos si la media muestral

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

es un estimador insesgado de μ. Debemos ver que

$$E(\overline{X}) = \mu$$

Ejemplo 7.2

$$E(\overline{X}) \overset{?}{=} \mu$$

- Usando la prop.de linealidad de la esperanza

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

- Por tratarse de componentes de una muestra aleatoria

$$E(X_i) = E(X) = \mu \qquad \forall i = 1,2,\ldots,n$$

- Luego

$$E(\overline{X}) = \frac{1}{n} n\mu = \mu$$

Ejemplo 7.3

- Sea X_1, X_2, … ,X_n una muestra aleatoria de una v.a. X donde $E(X)=\mu$ y $V(X)=\sigma^2$.
- Veamos si la varianza muestral

$$\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

es un estimador insesgado de σ^2. Debemos ver que

$$E(S^2) = \sigma^2$$

Ejemplo 7.3

$$\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

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

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

Reescribamos la suma sumando y restando μ
y desarrollando el cuadrado

Ejemplo 7.3

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

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

$$= \sum_{i=1}^{n}[X_i - \mu]^2 + \sum_{i=1}^{n}2[X_i - \mu][\mu - \overline{X}] + \sum_{i=1}^{n}[\mu - \overline{X}]^2) =$$

Ejemplo 7.3

$$=\sum_{i=1}^{n}[X_i-\mu]^2+\sum_{i=1}^{n}2[X_i-\mu][\mu-\overline{X}]+\sum_{i=1}^{n}[\mu-\overline{X}]^2)=$$

$$=\sum_{i=1}^{n}[X_i-\mu]^2+2[\mu-\overline{X}]\sum_{i=1}^{n}[X_i-\mu]+n[\mu-\overline{X}]^2)=$$

$$\sum_{i=1}^{n}[X_i-\mu]=\sum_{i=1}^{n}X_i-n\mu=n(\overline{X}-\mu)$$

Ejemplo 7.3

$$= \sum_{i=1}^{n} [X_i - \mu]^2 + \sum_{i=1}^{n} 2[X_i - \mu][\mu - \overline{X}] + \sum_{i=1}^{n} [\mu - \overline{X}]^2) =$$

$$= \sum_{i=1}^{n} [X_i - \mu]^2 + 2[\mu - \overline{X}] \sum_{i=1}^{n} [X_i - \mu] + n[\mu - \overline{X}]^2) =$$

$$= \sum_{i=1}^{n} [X_i - \mu]^2 + 2[\mu - \overline{X}] n[\overline{X} - \mu] + n[\mu - \overline{X}]^2) =$$

$$= \sum_{i=1}^{n} [X_i - \mu]^2 - 2n[\mu - \overline{X}]^2 + n[\mu - \overline{X}]^2) =$$

$$\sum_{i=1}^{n}[X_i - \overline{X}]^2 = \sum_{i=1}^{n}[X_i - \mu]^2 - 2n[\mu - \overline{X}]^2 + n[\mu - \overline{X}]^2) =$$

- Entonces

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

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

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

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

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

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

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

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

- Usando

$$V(X_i) = V(X) = \sigma^2 \quad \forall i = 1,2,\ldots,n$$

y que

$$V(\overline{X}) = \frac{\sigma^2}{n}$$

$$E(S^2) = \frac{1}{n-1}\left(n\sigma^2 - n\frac{\sigma^2}{n}\right) \Rightarrow \boxed{E(S^2) = \sigma^2}$$

Ejemplo 7.4

- La definición de varianza estándar es un estimador sesgado de la varianza

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

Por ser negativo, este estimador tiende a subestimar σ^2

Resumen

Sea X_1, X_2, … ,X_n una muestra aleatoria de una v.a. X donde $E(X)=\mu$ y $V(X)=\sigma^2$

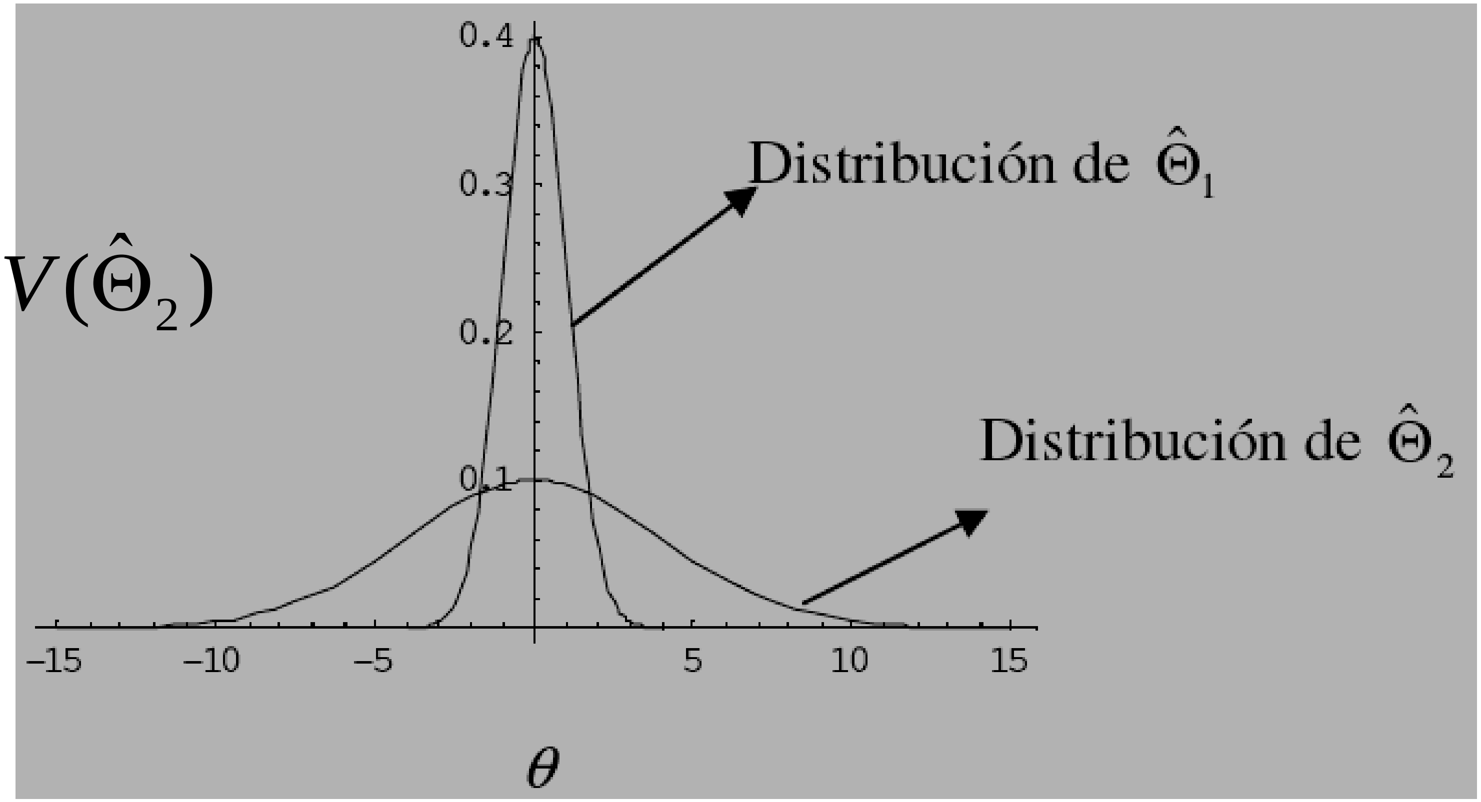
$\overline{X}$	$\frac{1}{n}\sum_{i=1}^{n} X_i$	$E\left(\overline{X}\right)=\mu$ $V\left(\overline{X}\right)=\frac{\sigma^2}{n}$
S^2	$\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$	$E\left(S^2\right)=\sigma^2$
S^2_{EMV}	$\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$	$E\left(S^2_{EMV}\right)=\frac{n-1}{n}\sigma^2$

Criterios para evaluar estimadores

- En ocasiones hay más de un estimador insesgado de un parámetro θ
- Por lo tanto necesitamos un método para seleccionar un estimador entre varios estimadores insesgados.
  - Varianza
  - Error Cuadrático medio

Varianza de un estimador

$$V(\hat{\Theta}_1) < V(\hat{\Theta}_2)$$



Si tenemos dos estimadores insesgados debemos elegir el de menor varianza

Ejemplo 7.5

- Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una v.a. X donde $E(X)=\mu$ y $V(X)=\sigma^2$
- Para estimar el parámetro μ podría utilizarse

$$\hat{\mu}_1 = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad\qquad \hat{\mu}_2 = \frac{X_1 + X_2}{2}$$

Ya vimos que es insesgado

elegi

Veamos si es insesgado

Ejemplo 7.5

$$E(\hat{\mu}_2) \overset{?}{=} \mu$$

$$E(\hat{\mu}_2) = E\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{2}(E(X_1) + E(X_2))$$

$$= \frac{1}{2}(\mu + \mu) = \frac{1}{2} 2\mu = \mu$$

Por lo tanto μ_2 también es insesgado

Entonces ¿Cuál de los dos estimadores es mejor?

Ejemplo 7.5

- Calculamos la varianza de cada uno utilizando las propiedades de la varianza.

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

$$= \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

Las X_i son v.a. independientes

Todas las X_i tienen la misma varianza

Ejemplo 7.5

$$V(\overline{X}) = \frac{\sigma^2}{n}$$

$$V(\hat{\mu}_2) = V\left(\frac{X_1 + X_2}{2}\right) = \frac{1}{4}(V(X_1) + V(X_2))$$

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

Si **n>2** el mejor estimador es $\hat{\mu}_1$ porque

$$V(\hat{\mu}_1) < V(\hat{\mu}_2)$$

Criterios para evaluar estimadores

- Supongamos ahora que $\hat{\Theta}_1$ y $\hat{\Theta}_2$ son dos estimadores de un parámetro θ y ***alguno de ellos no es insesgado.***
- A veces es necesario utilizar un estimador sesgado. En esos casos puede ser importante el ***error cuadrático medio del estimador.***

Error cuadrático medio (ECM)

- El error cuadrático medio de un estimador $\hat{\Theta}$ de un parámetro θ está definido como

$$ECM(\hat{\Theta}) = E\left[\left(\hat{\Theta} - \theta\right)^2\right]$$

- También puede escribirse como

$$ECM(\hat{\Theta}) = V(\hat{\Theta}) + \left(b(\hat{\Theta})\right)^2$$

- siendo

$$b(\hat{\Theta}) = E(\hat{\Theta}) - \theta$$

El ECM es un criterio para comparar estimadores

Eficiencia relativa

- Si $\hat{\Theta}_1$ y $\hat{\Theta}_2$ son dos estimadores de un parámetro θ, la eficiencia relativa de $\hat{\Theta}_2$ con respecto a $\hat{\Theta}_1$ se define como

$$\frac{ECM(\hat{\Theta}_1)}{ECM(\hat{\Theta}_2)}$$

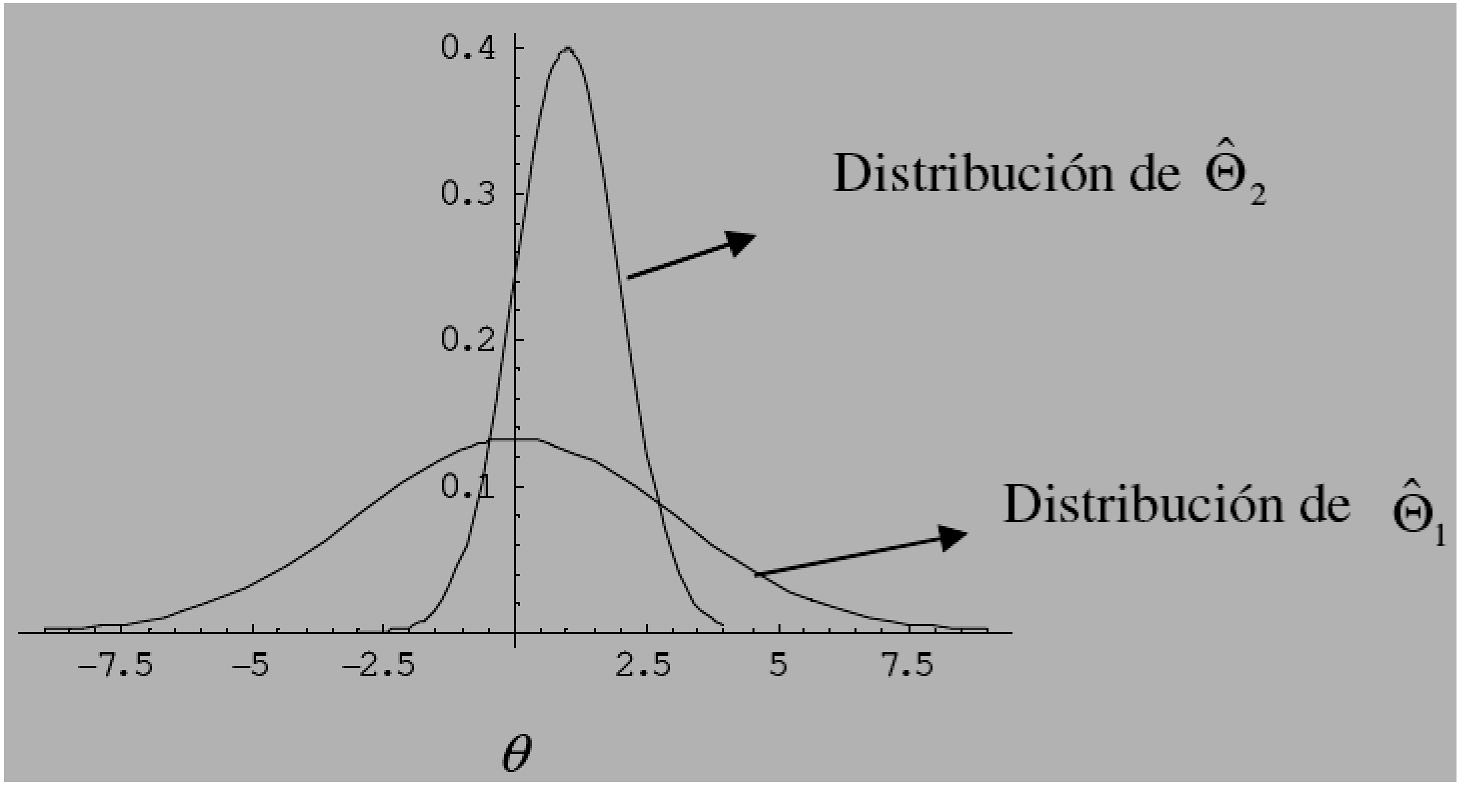
- Si es un valor menor que 1, $\hat{\Theta}_1$ tiene menor ECM que $\hat{\Theta}_2$ y **por lo tanto $\hat{\Theta}_1$ es más eficiente que $\hat{\Theta}_2$**

Criterios para evaluar estimadores

- Si $\hat{\Theta}$ es un estimador insesgado de θ entonces

$$ECM(\hat{\Theta}) = V(\hat{\Theta})$$

- En el ECM se consideran la varianza y el sesgo del estimador.
- Si $\hat{\Theta}_1$ y $\hat{\Theta}_2$ son dos estimadores de θ tales que $E(\hat{\Theta}_1) = \theta$; $E(\hat{\Theta}_2) \neq \theta$ y $V(\hat{\Theta}_2) < V(\hat{\Theta}_1)$, habría que calcular el ECM de cada uno y tomar el menor



- Habría que calcular el ECM de cada uno, y elegir el que tenga el menor valor. Aunque $\hat{\Theta}_2$ sea sesgado, al tener menor varianza podría llegar a tomar valores mas cercanos al verdadero parámetro $\hat{\Theta}_1$.

Ejemplo 7.6

- Sean $\hat{\Theta}_1$, $\hat{\Theta}_2$ y $\hat{\Theta}_3$ tres estimadores de un parámetro θ tales que

$$E(\hat{\Theta}_1) = E(\hat{\Theta}_2) = \theta; E(\hat{\Theta}_3) \neq \theta$$

$$V(\hat{\Theta}_1) = 10, V(\hat{\Theta}_2) = 6, E((\hat{\Theta}_3 - \theta)^2) = 4$$

Haga una comparación de estos estimadores.
¿Cuál prefiere y por qué?

Ejemplo 7.6

- Calculamos el ECM de cada estimador

$$ECM(\hat{\Theta}_1) = V(\hat{\Theta}_1) = 10$$

$$ECM(\hat{\Theta}_2) = V(\hat{\Theta}_2) = 6$$

$$ECM(\hat{\Theta}_3) = E((\hat{\Theta}_3 - \Theta)^2) = 4$$

El mejor estimador es $\hat{\Theta}_3$ porque tiene menor error cuadrático medio.

Consistencia de estimadores puntuales

- Sea $\hat{\Theta}_n$ un estimador del parámetro θ basado en una muestra aleatoria
  (X_1, X_2, …, X_n), si

$$\lim_{n\to\infty} E(\hat{\Theta}_n) = \theta \quad \text{y} \quad \lim_{n\to\infty} V(\hat{\Theta}_n) = 0$$

entonces $\hat{\Theta}_n$ es un **estimador consistente** de θ

Ejemplo 7.7

- La media muestral es un estimador consistente?
- Sabemos que

$$a)\, E(\overline{X}) = \mu = E(X) \qquad \forall n$$

$$b) V(\overline{X}) = \frac{\sigma^2}{n} = \frac{V(X)}{n} \qquad \forall n$$

Vemos que tiende a cero cuando n $\rightarrow \infty$;

por lo tanto el estimador es consistente

Ejercicio 7.1

- Dados los siguientes estimadores para la media muestral

$$\hat{\mu}_1 = \frac{x_1 + x_n}{2} \qquad \hat{\mu}_2 = \frac{x_1 + x_2 + x_3}{3}$$

  - Calcule para cada uno: sesgo, varianza y ECM.
  - Son estimadores consistentes?
  - Utilice la eficiencia relativa para indicar cual de los dos debería ser utilizado.

Ejercicio 7.1

- Sesgo del estimador : $E(\hat{\Theta}) - \theta$

$$\hat{\mu}_1 = \frac{x_1 + x_n}{2}$$

$$E(\hat{\mu}_1) = E\left(\frac{x_1 + x_n}{2}\right) = \frac{1}{2}E\left(x_1 + x_n\right)$$

Ejercicio 7.1

- Sesgo del estimador : $E(\hat{\Theta}) - \theta$

$$\hat{\mu}_1 = \frac{x_1 + x_n}{2}$$

$$E(\hat{\mu}_1) = \mu$$

$$\therefore \; sesgo = 0$$

$$E(\hat{\mu}_1) = \frac{1}{2} E\left(x_1 + x_n\right) = \frac{1}{2} = \frac{1}{2} * 2 * \mu = \mu$$

Ejercicio 7.1

- Varianza

$$V(\hat{\mu}_1) = V\left(\frac{x_1 + x_n}{2}\right) = \frac{1}{4}V\left(x_1 + x_n\right)$$

$$= \frac{1}{4}\left[V\left(x_1\right) + V\left(x_n\right)\right] = \frac{1}{4} * 2 * V\left(x\right)$$

$$V(\hat{\mu}_1) = V(X)/2$$

Ejercicio 7.1

$$ECM(\hat{\mu}_1) = V(\hat{\mu}_1) + \left(b(\hat{\mu}_1)\right)^2$$

- Por ser un estimador insesgado

$$ECM(\hat{\mu}_1) = V(\hat{\mu}_1) = V(X)/2$$

Ejercicio 7.1

- De la misma forma puede verse que
  - $\hat{\mu}_2$ también es insesgado
  - $V(\hat{\mu}_2) = V(X)/3$
  - $ECM(\hat{\mu}_2) = V(\hat{\mu}_2) = V(X)/3$
- Eficiencia relativa

$$\frac{ECM(\hat{\mu}_1)}{ECM(\hat{\mu}_2)} = \frac{V(X)}{2} : \frac{V(X)}{3} = \frac{3}{2} > 1$$

Elegimos $\hat{\mu}_2$

Métodos de estimación puntual

- Se analizan dos métodos constructivos para obtener estimadores puntuales
  - Método de momentos
  - Método de máxima verosimilitud
    - Los estimadores obtenidos de esta forma suelen ser más eficientes pero requieren más cálculo.

Método de momentos

- La idea básica es igualar ciertas características muestrales, como la media, con los valores esperados correspondientes.
- Esto da lugar a un conjunto de ecuaciones cuya solución, para valores de parámetros desconocidos, producen los estimadores.

Método de Momentos

- Sean $X_1, \dots, X_n$ una muestra aleatoria de una función de probabilidad f(x).
- Se definen
  - El **k-ésimo momento de la distribución** *f(x)* como

$$\mu_k = E(X^k) \qquad k=1, 2, 3, \dots$$

  - El **k-ésimo momento muestral** como

$$M_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k \qquad k=1, 2, 3, \dots$$

Método de Momentos

- Sean $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una distribución con función de probabilidad

  $p(x_i, \theta_1, \theta_2, \ldots, \theta_r)$ si X es discreta o

  $f(x_i, \theta_1, \theta_2, \ldots, \theta_r)$ *si* X es continua, donde

  $\theta_1, \theta_2, \ldots, \theta_r$ son parámetros que se desconocen.

- Entonces los estimadores de momento $\hat{\Theta}_1, \hat{\Theta}_2, \ldots, \hat{\Theta}_r$ se obtienen igualando los primeros *r* momentos muestrales con los primeros *r* momentos de la distribución y resolviendo para $\theta_1, \theta_2, \ldots, \theta_r$.

Método de Momentos

- Sean X_1, X_2, …, X_n una muestra aleatoria de una distribución cuya función de probabilidad tiene parámetros θ_1, θ_2, …, θ_r desconocidos, deben resolverse la siguientes ecuaciones para dichos parámetros

$$E(X^k) = \frac{1}{n}\sum_{i=1}^{n} X_i^k \qquad k=1,\ 2,\ \dots,r$$

$$\mu_k = E(X^k) = \begin{cases} \displaystyle\sum_{x_i \in R_X} x_i^k . p(x_i, \theta_1, \theta_2, \dots, \theta_r) & \text{(discreta)} \\ \displaystyle\int_{-\infty}^{\infty} x_i^k . f(x_i, \theta_1, \theta_2, \dots, \theta_r) & \text{(continua)} \end{cases}$$

Ejemplo 7.8

- Sea (X_1, X_2, …, X_n) una muestra aleatoria de tamaño n de una v.a. donde $X \sim U[0,\theta]$,

  con θ desconocido. Hallar el estimador de θ por el método de los momentos.

$\therefore$ debe igualarse el 1er. momento muestral con el 1er. momento de la distribución

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \therefore \quad E(X) = \overline{X}$$

Ejemplo 7.8

$$E(X) = \frac{0+\theta}{2} = \frac{\theta}{2} \quad \Rightarrow \quad \frac{\theta}{2} = \overline{X} \quad \Rightarrow \quad \hat{\Theta} = 2\overline{X}$$

- Verifique que este estimador es consistente

$$E(\hat{\Theta}) = E(2\overline{X}) = 2E(\overline{X}) = 2\frac{\theta}{2} = \theta$$

$$V(\hat{\Theta}) = V(2\overline{X}) = 4V(\overline{X}) = 4\frac{(\theta-0)^2}{12n} = \frac{\theta^2}{3n} \underset{n\to\infty}{\longrightarrow} 0$$

Ejemplo 7.9

- Estimar por el método de los momentos los parámetros μ y σ^2 de una distribución normal.

- Como son dos parámetros utilizaremos los dos primeros momentos de la distribución normal, que están dados por:

$$\mu_1 = E(X) = \mu$$

$$\mu_2 = E(X^2) = \sigma^2 + \mu^2$$

Ejemplo 7.9

- Luego, igualando los dos primeros momentos de la distribución normal con sus respectivos momentos muestrales tenemos que:

$$\mu_1 = M_1 \Rightarrow \mu = \overline{X} \Rightarrow \hat{\mu} = \overline{X}$$

$$\mu_2 = M_2 \Rightarrow \sigma^2 + \mu^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\therefore \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2$$

Ejercicio de la práctica

- Sea t_1, t_2, …, t_n una muestra aleatoria de n observaciones de una variable aleatoria discreta T, con función de frecuencia de probabilidad p(t) ; ($0< \theta < 1$)

t	0	1	2
p(t)	θ^2	$2\,\theta(1-\theta)$	$(1-\theta)^2$

- Halle un estimador de θ utilizando el método de los momentos.

Ejercicio de la práctica

- Para calcular un estimador de θ utilizando el método de los momentos hay que igualar el primer momento de la distribución, E(T), con el primer momento muestral (la media muestral)

$$\underbrace{E(T)} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Usemos la tabla de la distribución de probabilidad de T para calcular E(T)

Ejercicio de la práctica

- Dada p(t), calcular E(T)

t	0	1	2
p(t)	θ^2	$2\,\theta(1-\theta)$	$(1-\theta)^2$

- $E(T) = 0^*\,\theta^2 + 1^*\ (2 * \theta * (1-\theta)) + 2 * (1-\theta)^2$

  $= 2\,\theta - 2\,\theta^2 + 2 - 4\,\theta + 2\,\theta^2$

- $E(T) = -2\,\theta + 2$

Luego igualamos E(T) al primer momento muestral

Ejercicio de la práctica

- Para calcular un estimador de θ utilizando el método de los momentos hay que igualar el primer momento de la distribución, E(T), con el primer momento muestral (la media muestral)

$$E(T) = \frac{\sum_{i=1}^{n} x_i}{n} \quad \therefore \quad -2\theta + 2 = \frac{\sum_{i=1}^{n} x_i}{n} \quad \therefore$$

$$\theta = \frac{1}{2}\left(2 - \frac{\sum_{i=1}^{n} x_i}{n}\right) \quad \therefore \quad \hat{\Theta} = 1 - \frac{\overline{X}}{2}$$

Ejercicio 7.2

- El tiempo de falla *T* de una componente tiene una distribución exponencial con parámetro λ. *T*~ *Exp*(λ) , es decir la fdp es

$$f(t;\lambda) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & si\ no \end{cases}$$

- Recordemos que $E(T) = 1/\lambda$ $V(T) = 1/\lambda^2$
- Calculemos el estimador por el método de momentos del parámetro λ para una muestra de tamaño n.

Ejercicio 7.2

$$E(T) = 1/\lambda$$

- Debe igualarse

$$\boxed{E(T)} = \boxed{\frac{\sum_{i=1}^{n} t_i}{n}}$$

1er. momento de la distribución

1er. momento muestral

Ejercicio 7.2

$$E(T) = 1/\lambda$$

- Debe igualarse

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

$$\therefore \quad \lambda = n \Big/ \sum_{i=1}^{n} t_i \quad \Rightarrow \quad \hat{\lambda} = \frac{1}{\overline{T}}$$

Método de Máxima Verosimilitud

- Sea *X es una v.a. discreta con fdp p(x,θ) , donde* θ es un parámetro desconocido.
- Sean $x_1, x_2, \ldots x_n$ *los valores observados de una muestra aleatoria* de tamaño *n.*
- Se define la ***función de verosimilitud como la función de distribución conjunta de las observaciones***

$$L(x_1, x_2, \ldots, x_n, \theta) = P(X_1 = x_1)P(X_2 = x_2)\ldots P(X_n = x_n)$$
$$= p(x_1, \theta)\, p(x_2, \theta)\ldots p(x_n, \theta)$$

Método de Máxima Verosimilitud

- Sea *X es una v.a. continua con fdp f(x,θ) , donde* θ es un parámetro desconocido.
- Sean $x_1, x_2, \ldots x_n$ *los valores observados de una muestra aleatoria* de tamaño *n.*
- Se define la ***función de verosimilitud como la función de distribución conjunta de las observaciones***

$$L(x_1, x_2, \ldots, x_n, \theta) = f(x_1, \theta)\, f(x_2, \theta) \ldots f(x_n, \theta)$$

Método de Máxima Verosimilitud

- La función de Verosimilitud

$$L(x_1, x_2, ..., x_n, \theta) = p(x_1, \theta)\, p(x_2, \theta) ... p(x_n, \theta)$$

$$L(x_1, x_2, ..., x_n, \theta) = f(x_1, \theta)\, f(x_2, \theta) ... f(x_n, \theta)$$

es una función de θ.

El **estimador de máxima verosimilitud (EMV)** de θ es aquel valor de θ que maximiza la función de verosimilitud

Ejemplo 7.10

- Sea $(X_1, X_2, \ldots, X_n)$ una muestra aleatoria de tamaño n de una v.a. $X \sim B(1, p)$.
- Por ejemplo, se eligen al azar n objetos de una línea de producción, y cada uno se clasifica como defectuoso (en cuyo caso $x_i=1$) o no defectuoso (en cuyo caso $x_i=0$).
- Entonces $p=P(X_i=1)$, es decir es la verdadera proporción de objetos defectuosos en la producción total.
- Queremos hallar el EMV de p.

Ejemplo 7.10

- La función de verosimilitud será

$$L(x_1, x_2, \ldots, x_n, \theta) = p(x_1, \theta)\, p(x_2, \theta) \ldots p(x_n, \theta) =$$
$$\left[p^{x_1}(1-p)^{1-x_1}\right]\left[p^{x_2}(1-p)^{1-x_2}\right] \ldots \left[p^{x_n}(1-p)^{1-x_n}\right]$$

- Esto puede escribirse

$$L(x_1, x_2, \ldots, x_n, \theta) = p^{\sum_{i=1}^{n} x_i}\,(1-p)^{n-\sum_{i=1}^{n} x_i}$$

Ejemplo 7.10

- Aplicando logaritmo natural de L

$$\ln(L(x_1, x_2, \ldots, x_n, \theta)) = \left(\sum_{i=1}^{n} x_i\right)\ln p + \left(n - \sum_{i=1}^{n} x_i\right)\ln(1-p)$$

- Derivando respecto de p e igualando a cero

$$\frac{\partial \ln(L(x_1, x_2, \ldots, x_n, \theta))}{\partial p} = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n - \sum_{i=1}^{n} x_i}{1-p} = 0$$

$$\frac{\partial \ln(L(x_1, x_2, \ldots, x_n, \theta))}{\partial p} = \frac{\sum_{i=1}^{n} x_i}{p} - \frac{n - \sum_{i=1}^{n} x_i}{1-p} = 0$$

$$\frac{\sum_{i=1}^{n} x_i}{p} = \frac{n - \sum_{i=1}^{n} x_i}{1-p} \Rightarrow (1-p)\sum_{i=1}^{n} x_i = p\left(n - \sum_{i=1}^{n} x_i\right)$$

$$\sum_{i=1}^{n} x_i - p\sum_{i=1}^{n} x_i = pn - p\sum_{i=1}^{n} x_i \Rightarrow \sum_{i=1}^{n} x_i = pn$$

$$p = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x}$$

la pr

$$\hat{p} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

sos

de la

Ejemplo 7.11

- El tiempo de falla *T* de una componente tiene una distribución exponencial con parámetro λ. *T*~ *Exp*(λ) , es decir la fdp es

$$f(t;\lambda) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & si\ no \end{cases}$$

- Recordemos que $E(T) = \dfrac{1}{\lambda}$ $\qquad V(T) = \dfrac{1}{\lambda^2}$
- Calculemos el EMV del parámetro λ para una muestra de tamaño n.

Ejemplo 7.11

- La función de verosimilitud será

$$L(t_1,t_2,\ldots,t_n;\lambda) = f(t_1;\lambda)\,f(t_2;\lambda)\ldots f(t_n;\lambda) =$$
$$= \left[\lambda\, e^{-\lambda t_1}\right]\left[\lambda\, e^{-\lambda t_2}\right]\ldots\left[\lambda\, e^{-\lambda t_n}\right]$$

- Esto puede escribirse

$$L(t_1,t_2,\ldots,t_n;\lambda) = (\lambda)^n\, e^{-\lambda \sum_{i=1}^{n} t_i}$$

Ejemplo 7.11

- Aplicando logaritmo natural de L

$$\ln(L(t_1,t_2,\dots,t_n;\lambda)) = \ln\left(\lambda^n e^{-\lambda\sum_{i=1}^{n} t_i}\right) = n\ln(\lambda) - \lambda\sum_{i=1}^{n} t_i$$

- Derivando respecto de λ e igualando a cero

$$\frac{\partial \ln(L(t_1,t_2,\dots,t_n;\lambda))}{\partial\lambda} = n\frac{1}{\lambda} - \sum_{i=1}^{n} t_i = 0$$

$$\lambda = \frac{n}{\sum_{i=1}^{n} t_i} = \frac{1}{\bar{t}} \Rightarrow \boxed{\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} T_i} = \frac{1}{\bar{T}}}$$

Método de máxima verosimilitud

- Características
  - Algunas veces, resulta difícil maximizar la función de verosimilitud debido a que la ecuación obtenida a partir de

$$\frac{\partial L(\theta)}{\partial \theta} = 0$$

    no resulta fácil de resolver.
  - También puede ocurrir que los métodos de cálculo para maximizar *$L(\theta)$ no sean aplicables*

Ejemplo 7.12

- Sea $(X_1, X_2, \ldots, X_n)$ una muestra aleatoria de tamaño n de una v.a. $X \sim U(0, \theta)$, con θ desconocido. Hallar el estimador de θ por el método de máxima verosimilitud.

- La *fdp* de X es

$$f(x) = \begin{cases} \dfrac{1}{\theta} & 0 < x < \theta \\ 0 & si\ no \end{cases}$$

Ejemplo 7.12

- La función de verosimilitud es

$$L(x_1, x_2, \ldots, x_n; \theta) = \begin{cases} \dfrac{1}{\theta^n} & si \quad \max\limits_i (x_i) < \theta \\ 0 & si\ no \end{cases}$$

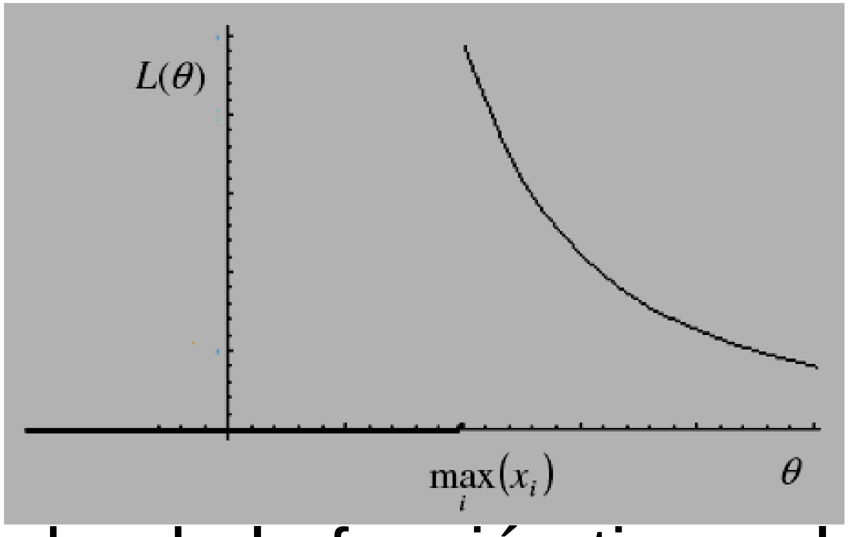
- Si derivamos con respecto a θ obtenemos

$$\frac{d}{d\theta}\theta^{-n} = -\frac{n}{\theta^{n+1}}$$

Siempre es menor que cero

Ejemplo 7.12

- La función de máxima verosimilitud es una función decreciente para todos los $\theta > \max_i (x_i)$



- Vemos que donde la función tiene el máximo hay una discontinuidad no evitable. Por lo tanto

$$\hat{\Theta} = \max_i (x_i)$$

Método de máxima verosimilitud

- El método de máxima verosimilitud puede emplearse en el caso donde hay más de un parámetro desconocido para estimar.
- En ese caso la función de verosimilitud es una función de varias variables.

$$L(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_k)$$

Método de máxima verosimilitud

- Si

$$L(x_1, x_2, ..., x_n; \theta_1, \theta_2, ..., \theta_k)$$

los estimadores de máxima verosimilitud $\hat{\Theta}_1, \hat{\Theta}_2, ..., \hat{\Theta}_k$ se obtienen al plantear (si existen las derivadas parciales) y resolver el sistema de k ecuaciones con k incógnitas $\theta_1, \theta_2, \ \ldots, \theta_k$

$$\frac{d}{d\theta_i} L(x_1, x_2, ..., x_n; \theta_1, \theta_2, ..., \theta_k) = 0 \qquad \text{i=1,2,…,k}$$

Ejercicio 7.3

;

- Encontrar el estimador de máxima verosimilitud para el parámetro ***p*** de una distribución geométrica para una muestra de tamaño n.

- Recuerde que si $X \sim G(p)$

$$p(X = k) = (1 - p)^{k-1} p \qquad k=1,2,\ldots$$

Ejercicio 7.3

- La función de verosimilitud será

$$L(x_1, x_2,\ldots, x_n; \lambda) = p(x_1; p)\,p(x_2; p)\ldots p(x_n; p) =$$
$$= \left[(1-p)^{x_1-1}p\right]\left[(1-p)^{x_2-1}p\right]\ldots\left[(1-p)^{x_n-1}p\right]$$

- Esto puede escribirse

$$L(x_1, x_2,\ldots, x_n; p) = (1-p)^{\sum_{i=1}^{n}(x_i-1)}\,p^n$$

Ejercicio 7.3

- Aplicando logaritmo natural de L

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

$$= \ln\left((1-p)^{\sum_{i=1}^{n}(x_i - 1)} \right) + \ln\left(p^n\right)$$

$$= \sum_{i=1}^{n}(x_i - 1)\ln\left(1-p\right) + n\ln\left(p\right)$$

Ejercicio 7.3

- Por lo tanto,

$$\ln(L(x_1,\ldots,x_n;p)) = \sum_{i=1}^{n}(x_i - 1)\ln\left(1-p\right) + n\ln\left(p\right)$$

- Derivando respecto de λ e igualando a cero

$$\frac{\partial \ln(L(x_1,..,x_n;p))}{\partial p} = -\frac{\sum_{i=1}^{n}(x_i - 1)}{1-p} + \frac{n}{p} = 0$$

- Por lo tanto,

$$\frac{n}{p} = \frac{\sum_{i=1}^{n}(x_i - 1)}{1-p}$$

Ejercicio 7.3

- Por lo tanto,

$$\frac{n}{p} = \frac{\sum_{i=1}^{n}(x_i - 1)}{1-p} \Rightarrow (1-p)n = p\sum_{i=1}^{n}(x_i - 1)$$

$$n - np = p\sum_{i=1}^{n} x_i - p\sum_{i=1}^{n} 1 \Rightarrow n - np = p\sum_{i=1}^{n} x_i - np$$

$$n = p\sum_{i=1}^{n} x_i \Rightarrow \frac{n}{\sum_{i=1}^{n} x_i} = p \Rightarrow \boxed{\hat{p} = \frac{1}{\overline{X}}}$$

Propiedades de los EMV

- **Propiedad 1**

  Los EMV pueden ser ***sesgados***, *pero en general si* $\hat{\Theta}$ *es el EMV de un parámetro* θ *basado en* una muestra de tamaño *n, entonces*

$$\lim_{n\to\infty} E(\hat{\Theta}) = \theta$$

es decir son ***asintóticamente insesgados***

Propiedades de los EMV

- **Propiedad 2**
  - Bajo condiciones bastantes generales se puede probar que los EMV son ***asintóticamente consistentes***
- **Propiedad 3**
  - Bajo condiciones bastantes generales se puede probar que los EMV ***asintóticamente tienen varianza mínima***

Propiedades de los EMV

- **Propiedad 4**
  - Los EMV cumplen la ***propiedad de invarianza***.
  - *E*s decir que si $\hat{\Theta}$ es un EMV de un parámetro θ, el EMV de $g(\theta)$ *es* $g(\hat{\Theta})$, *si g(x) es una* $\underbrace{\textit{función inyectiva.}}$

A cada valor de la imagen le corresponde un único valor en el dominio

Ejemplo 7.14

- En el Ejemplo 7.11 teníamos una v.a. T con distribución exponencial, *T~ Exp(λ).* Si queremos el EMV de la varianza poblacional, podemos calcularlo recordando que

$$V(T) = 1/\lambda^2 \quad \text{es decir que} \quad V(T) = g(\lambda) = 1/\lambda^2$$

- Vimos que

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} T_i} = \frac{1}{\overline{T}} \quad \text{por lo tanto} \quad \boxed{\sigma^2 = \frac{1}{\hat{\lambda}^2}}$$

Resumen

- Estimador Puntual
- Estimadores Usuales
- Criterios
  - Insesgamiento
  - Varianza Mínima
  - ECM
  - Eficiencia Relativa

- Métodos de estimación puntual
  - Método de Momentos
  - Método de máxima verosimilitud
    - Propiedades