3.3. Propiedades de los estimadores

Para estimar un parámetro, puede existir en ocasiones más de un estimador, por lo que es necesario utilizar aquellos que tengan las propiedades que se explican a continuación.

Propiedades deseables de los estimadores

Insesgado
• La primera propiedad de un estimador es que estime lo que se quiere estimar; por ejemplo, si se realizara una estimación con muchas muestras aleatorias, el valor esperado del estimador es el parámetro poblacional de interés. Cuando esto ocurre, el estimador es *insesgado*.

Un estimador es insesgado si satisface la siguiente condición:

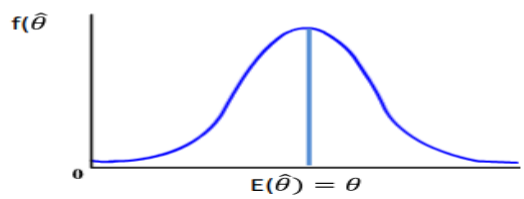
$$E[\hat{\theta}] = \theta$$

Si esto se cumple, entonces:

$$ECM[\hat{\theta}] = Var\,[\hat{\theta}]$$

En la figura 4 se ilustra esta propiedad.

Figura 4. Distribución de un estimador insesgado



La figura anterior ilustra la distribución de un estimador insesgado cuyo valor esperado es el parámetro. Es importante mencionar que la distribución acampanada de la figura solamente es con fines ilustrativos, ya que un estimador no necesariamente tiene esta distribución de probabilidades.

Con menos variabilidad
• La siguiente característica que se busca en un estimador es que sus estimaciones varíen lo menos posible del parámetro poblacional. Un estimador así es más eficiente o con menos variabilidad.

Sean $\hat{\boldsymbol{\theta}}_1$ y $\hat{\boldsymbol{\theta}}_2$ dos estimadores del parámetro θ:

$$\text{Si } Var\,[\hat{\theta}_1] < Var\,[\hat{\theta}_2]$$

Entonces, $\hat{\boldsymbol{\theta}}_1$ es más eficiente que $\hat{\boldsymbol{\theta}}_2$

La figura 5 ilustra esta característica.

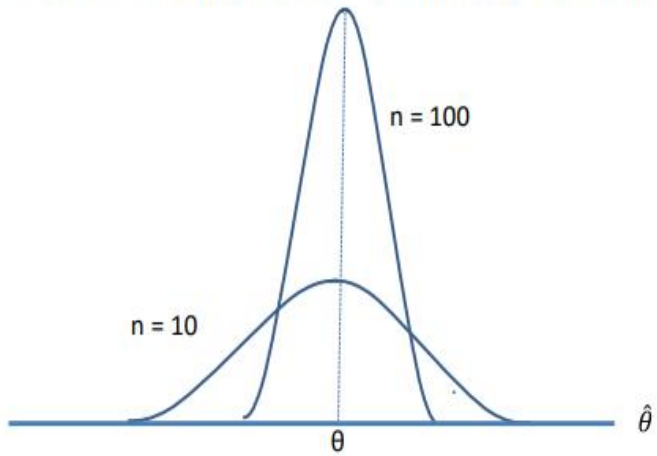
Figura 5. Eficiencia de dos estimadores insesgados

La figura 5 ilustra la distribución de dos estimadores $\hat{\boldsymbol{\theta}}_1$ y $\hat{\boldsymbol{\theta}}_2$ del parámetro poblacional θ. Aunque ambos estimadores son insesgados, el primero da mejores estimaciones, en tanto es más probable que arroje un valor más cercano al parámetro real respecto del segundo. Por tanto, $\hat{\boldsymbol{\theta}}_1$ es más eficiente que $\hat{\boldsymbol{\theta}}_2$.

Consistente
• La última propiedad esperada en un estimador es que, a medida que utilice mayor información de la población, su estimación sea cada vez más cercana al parámetro poblacional. Cuando esto ocurre, el estimador es *consistente*.

La figura 6 ilustra el comportamiento de un estimador consistente.

Figura 6. Comportamiento de un estimador consistente



La figura anterior ilustra el comportamiento de un estimador consistente. Conforme aumenta el tamaño de muestra, la variabilidad del estimador disminuye: las estimaciones son cada vez más cercanas al valor real del parámetro.

3.4. Estimación de una media con muestras grandes

El teorema del límite central garantiza que, conforme aumenta el tamaño de la muestra, la distribución del promedio muestral se acerca a una distribución normal cuya media es el promedio poblacional, y la varianza es la varianza poblacional entre el tamaño de la muestra. Como regla general:

se considera que con un tamaño de muestra al menos de 30 elementos la distribución del promedio muestral sigue una distribución normal.

Teniendo presente esta regla, en muestras grandes (al menos de 30 elementos) se empleará una distribución normal para realizar una estimación por intervalo de la media.

La tabla siguiente muestra el parámetro media, su estimador, el tamaño de la muestra, la fórmula para el intervalo de confianza de la media, la fórmula para calcular el estimador de la media y la fórmula del estimador de la media muestral.

Tabla 3. Elementos para realizar la estimación puntual y por intervalo de la media (promedio) con muestras grandes

Parámetro población	Estimador	Tamaño de la muestra	Fórmula		
			Intervalo de confianza	Estimador puntual	Desviación estándar del estimador
Promedio μ	$\bar{x}$	n>30	$IC = \bar{x} \pm z\frac{s}{\sqrt{n}}$	$\bar{x} = \frac{\sum x_i}{n}$	$s_{\bar{x}} = \frac{s}{\sqrt{n}}$

En la tabla anterior, columna 5, se muestra el estimador puntual de la media poblacional, que es el promedio muestral. En la columna 4, se presenta cómo calcular el intervalo de confianza. En este caso, cuando se conoce la varianza poblacional, se utiliza en los cálculos; de no ser así, se estima este valor empleando la varianza muestral. El valor *Z* representa el nivel de confianza buscado; en este planteamiento, z es el cuantil de una distribución normal estándar (Z) que parte la curva en dos áreas, una con valor $1-\frac{\alpha}{2}$ y otra de $\frac{\alpha}{2}$, siendo α un valor entre 0 y 1.

En la columna 6, se muestra la desviación del estimador, acorde con el teorema del límite central.

Para calcular el valor de z, se puede recurrir a tablas o algún paquete. A continuación, se plantea cómo calcularlo en MS-Excel.

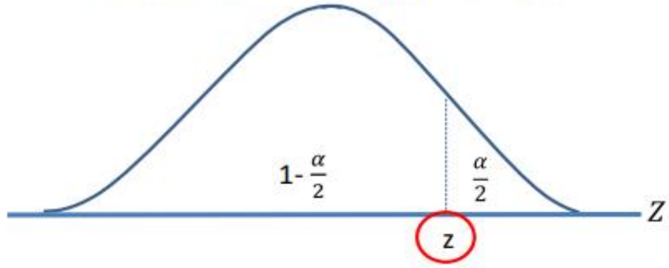
En Excel se emplea la función

DISTR.NORM.ESTAND.INV(probabilidad)

Para calcular el cuantil z donde se acumula una probabilidad de $1-\frac{\alpha}{2}$ $(0<\alpha<1)$

En la figura 7, se ilustra el valor que calcula la fórmula.

Figura 7.Valor calculado con la fórmula de Excel DISTR.NORM.ESTAND.INV(probabilidad)



En la figura anterior, en el valor encerrado en el círculo rojo se encuentra el punto que estima la fórmula. La información que se introduce en la fórmula es el área acumulada del lado izquierdo de z.

Supóngase que se desea realizar una estimación con un nivel de confianza del 95%, entonces:

$1 - \alpha = 0.95$ $\alpha = 1\text{-}0.95$ $\alpha = 0.05$	Por tanto:	$\frac{\alpha}{2} = \frac{0.05}{2} = 0.025$

Entonces, el valor *z* es el siguiente:

DISTR.NORM.ESTAND.INV(1-0.025) = 1.96

En la tabla 4 se muestran los valores de *z* para los niveles de confianza más usados.

Tabla 4. Valores de z obtenidos para los niveles de confianza más usados empleando Excel

Nivel de confianza	α 1-nivel de confianza	Función en MS-Excel *DISTR.NORM.ESTAND.INV(1-$\frac{\alpha}{2}$)*	z
90%	10%	DISTR.NORM.ESTAND.INV(1-0.10/2)	1.64
95%	5%	DISTR.NORM.ESTAND.INV(1-0.05/2)	1.96
99%	9%	DISTR.NORM.ESTAND.INV(1-0.01/2)	2.58

Ejemplos de estimación de una media con muestras grandes

Primer ejemplo

El director financiero de una agencia de publicidad desea conocer el gasto promedio de la organización, pues está preocupado por el nivel de gasto registrado recientemente. Por tal motivo realiza una auditoría a 30 facturas elegidas al azar.

La información de las erogaciones seleccionadas se muestra a continuación.

Monto de las facturas auditadas

Monto de facturas combinadas

Factura	Gasto en miles	Factura	Gasto en miles
1	99	16	96
2	15	17	79
3	59	18	71
4	14	19	56
5	72	20	51
6	59	21	72
7	68	22	25
8	22	23	71
9	40	24	52
10	79	25	99
11	97	26	70
12	82	27	82
13	93	28	47
14	76	29	35
15	48	30	93

Con la información de esta muestra, procede lo siguiente:

- Estimar el gasto promedio de la organización con una estimación puntual.
- Estimar un intervalo de confianza con un nivel de confianza del 99%.
- Interpretar los resultados.

Respuestas

Para solucionar este problema, se sugiere realizar lo siguiente.

1. Determinar el parámetro a estimar Promedio μ	**2.** Determinar el estimador del Parámetro $\bar{x}$
3. Calcular el estimador puntual a través de la fórmula correspondiente: $\bar{x} = \frac{\sum x_i}{n}$ $\bar{x} = \frac{99 + 15 + \cdots + 35 + 93}{30}$ $\bar{x} = \frac{1,922}{30} = 64.06$	**4.** Determinar la fórmula para realizar el cálculo de la estimación por intervalo del estimador: $IC = \bar{x} \pm Z\frac{s}{\sqrt{n}}$
5. Establecer el nivel de confianza para calcular z, a través del nivel de confianza (1-α) Nivel de confianza 99%, es decir, 0.99 Determinar el valor de α, donde α = 1 – nivel de confianza α = 1 – 0.99 α = 0.01 Calcular el valor de z, utilizando la función en Excel DISTR. NORM. ESTAND. INV (1- α /2) DISTR.NORM. ESTAND. INV (1-0.01/2) Z = 2.5758 = 2.58	**6.** Determinar la fórmula para calcular la desviación estándar del estimador $s_{\bar{x}} = \frac{s}{\sqrt{n}}$ Calcular la desviación estándar *s* y definir el valor de *n* n = 30 $s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$ $s = \sqrt{\frac{(99-64.06)^2 + (15-64.06)^2 + \cdots (35-64.06)^2 + (93-64.06)^2}{30-1}}$ $s = \sqrt{\frac{18,339.86}{29}}$ $s = \sqrt{632.409}$

5. Establecer el nivel de confianza para calcular *z*, a través del nivel de confianza (1-α)

Nivel de confianza 99%, es decir, 0.99

Determinar el valor de α, donde

α = 1 – nivel de confianza

α = 1 – 0.99

α = 0.01

Calcular el valor de **z**, utilizando la función en Excel

DISTR. NORM. ESTAND. INV (1- α /2)

DISTR.NORM. ESTAND. INV (1-0.01/2)

Z = 2.5758 = 2.58

6. Determinar la fórmula para calcular la desviación estándar del estimador

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

Calcular la desviación estándar *s* y definir el valor de *n*

n = 30

$$s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$$

$$s = \sqrt{\frac{(99-64.06)^2 + (15-64.06)^2 + \cdots (35-64.06)^2 + (93-64.06)^2}{30-1}}$$

$$s = \sqrt{\frac{18,339.86}{29}}$$

$$s = \sqrt{632.409}$$

$$s = 25.14$$

7. Calcular la desviación del estimador a través de la fórmula correspondiente:

$$s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

$$s_{\bar{x}} = \frac{25.14}{\sqrt{30}}$$

$$s_{\bar{x}} = \frac{25.14}{5.477}$$

$$s_{\bar{x}} = 4.59$$

8. Sustituir los valores en la fórmula general y calcular el límite inferior (LI) y límite superior (LS) del intervalo de confianza (IC):

$$IC = \bar{x} \pm Z\frac{s}{\sqrt{n}}$$

$$IC = 64.06 \pm 2.58 \cdot 4.59$$

$$IC = 64.06 \pm 11.845$$

$$LI = 64.06 - 11.845$$

$$LI = 52.22$$

$$LS = 64.06 + 11.845$$

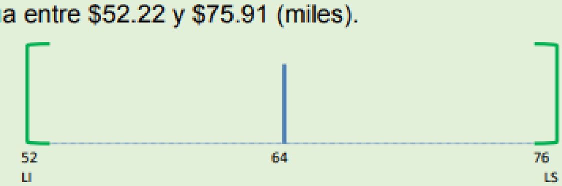
$$LS = 75.91$$

9. Construir el intervalo de confianza:

$$IC = (LI, LS)$$

$$IC = (52.22, 75.91)$$

Conforme a la estimación puntual, el promedio de gasto de la organización es de $64.06 (miles). De acuerdo con la estimación por intervalo, el gasto promedio de la organización a un nivel de confianza del 99% se sitúa entre $52.22 y $75.91 (miles).



Segundo ejemplo

Una farmacéutica cuenta con 500 representantes médicos. Con la intención de diseñar un plan de incentivos, se quiere conocer el promedio de visitas que realizan los representantes, para lo cual se analizó una muestra de 35 representantes médicos elegidos al azar.

En la siguiente tabla, se muestran las visitas realizada en un día por 35 representantes seleccionados.

a. Estimar el promedio de visitas que realizan los representantes médicos, con una estimación puntual.

b. Estimar un intervalo de confianza con un nivel de confianza del 95%.

c. Interpretar los resultados.

Representante	Número de visitas realizadas	Representante	Número de visitas realizadas
1	8	19	5
2	4	20	5
3	7	21	8
4	8	22	7
5	6	23	7
6	6	24	7
7	5	25	5
8	8	26	6
9	6	27	8
10	6	28	7
11	7	29	5
12	7	30	5
13	6	31	7
14	4	32	4
15	5	33	7
16	7	34	7
17	8	35	6
18	6		

Respuestas

1. Determinar el parámetro a estimar: Promedio μ	**2.** Determinar el estimador del parámetro $\bar{x}$
3. Calcular el estimador puntual a través de la fórmula correspondiente: $\bar{x} = \frac{\sum x_i}{n}$ $\bar{x} = \frac{8+4+\cdots+7+6}{35}$ $\bar{x} = \frac{220}{35} = 6.28$	**4.** Determinar la fórmula para realizar el cálculo de la estimación puntual del estimador: $IC = \bar{x} \pm Z\frac{s}{\sqrt{n}}$

5. Establecer el nivel de confianza para calcular *z*: 95%, es decir, 0.95 Determinar el valor de α: α = 1 – 0.95 α = 0.05 Calcular *z* con la función de Excel: DISTR. NORM. ESTAND. INV (1- α /2) DISTR. NORM. ESTAND. INV (1- 0.05/2) Z = 1.959 = 1.96	**6.** Determinar la fórmula para calcular la desviación estándar del estimador: $s_{\bar{x}} = \frac{s}{\sqrt{n}}$ Calcular la desviación estándar *s* y definir el valor de *n*: n = 35 $s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$ $s = \sqrt{\frac{(8-6.28)^2 + (4-6.28)^2 + \cdots (7-6.28)^2 + (6-6.28)}{35-1}}$ $s = \sqrt{\frac{51.14}{34}}$ $s = \sqrt{1.504}$ $s = 1.226$

7. Calcular la desviación del estimador a través de la fórmula correspondiente: $s_{\bar{x}} = \frac{s}{\sqrt{n}}$ $s_{\bar{x}} = \frac{1.226}{\sqrt{35}}$ $s_{\bar{x}} = \frac{1.226}{5.916}$ $s_{\bar{x}} = 0.2073$	**8.** Sustituir los valores en la fórmula general y calcular el límite inferior (LI) y límite superior (LS) del intervalo de confianza (IC): $IC = \bar{x} \pm Z\frac{s}{\sqrt{n}}$ $IC = 6.28 \pm 1.96 \cdot 0.207$ $IC = 6.28 \pm 0.4063$ $LI = 6.28 - 0.4063$ $LI = 5.87$ $LS = 6.28 + 0.4063$ $LS = 6.69$

9. Construir el intervalo de confianza:

$$IC = (LI, LS)$$

$$IC = (5.87, 6.69)$$

Con base en la estimación puntual, el promedio de visitas diarias efectuadas por un representante médico es de 6.

Conforme a la estimación por intervalo, el promedio de visitas que realiza un representante médico al día con un nivel de confianza del 95% se sitúa entre 6 y 7.