**Universidad del Valle**
**Programa Académico de Estadística**
**Muestreo Estadístico**
Cali, febrero de 2019

# Comparando Estimadores

Kevin Steven García[a], Cesar Andres Saavedra[b], Bryan Martinez[c]

Suponga que se tiene como población el conjunto $U = \{u_1, u_2, u_3, u_4, u_5\}$ y que al medir en cada unidad la característica X; se tiene el conjunto:

$$P(X) = \{x_1 = 6, x_2 = 4, x_3 = 3, x_4 = 8, x_5 = 9\}$$

a. Si se escogen muestras de tamaño 3 sin reemplazo conforme la distribución de las muestras conformando el siguiente cuadro:

| Muestra | A | B | C | D | E |
|---|---|---|---|---|---|
| 1 | 6 | 4 | 3 | | |
| 2 | 6 | 4 | | 8 | |
| 3 | 6 | 4 | | | 9 |
| 4 | 6 | | 3 | 8 | |
| 5 | 6 | | 3 | | 9 |
| 6 | 6 | | | 8 | 9 |
| 7 | | 4 | 3 | 8 | |
| 8 | | 4 | 3 | | 9 |
| 9 | | 4 | | 8 | 9 |
| 10 | | | 3 | 8 | 9 |

b. Una vez confeccionado el cuadro anterior, encuentre las siguientes distribuciones de probabilidad de los estimadores i)Media, Moda, Mediana y Media armónica. Complete las siguientes tablas:

| Distribución de la Mediana | | Distribución de la Media | | Distribución de la Moda | | Distribución de la Media Armónica | |
|---|---|---|---|---|---|---|---|
| Valor | $n_i$ | Valor | $n_i$ | Valor | $n_i$ | Valor | $n_i$ |
| 4 | 3 | 4.333333 | 1 | | | 4 | 1 |
| 6 | 4 | 5 | 1 | | | 4.235294 | 1 |
| 8 | 3 | 5.333333 | 1 | | | 4.32 | 1 |
| | | 5.666666 | 1 | | | 4.8 | 1 |
| | | 6 | 2 | | | 4.909090 | 1 |
| | | 6.333333 | 1 | | | 5.268292 | 1 |
| | | 6.666666 | 1 | | | 5.538461 | 1 |
| | | 7 | 1 | | | 5.684210 | 1 |
| | | 7.666666 | 1 | | | 6.171428 | 1 |
| | | | | | | 7.448275 | 1 |

La moda no existe en ninguna de las muestras, ya que en la población no hay valores repetidos y aparte de ello, el muestreo es sin reemplazo, por lo tanto, no hay posibilidades de que salga un mismo valor más de una vez en alguna muestra.

[a]Código: 1533173. E-mail: kevin.chica@correounivalle.edu.co
[b]Código:. E-mail:
[c]Código:. E-mail:

c. Encuentre la Mediana, la Moda, la Media y la Media Armónica poblacionales (parámetros).

Los parámetros poblacionales son:

| $M_e$ | $\mu$ | $M$ | $M_H$ |
|---|---|---|---|
| 6 | 6 | | 5.070423 |

No existe moda, ya que en la población no hay valores que se repitan(con mayor frecuencia que los demás).

d. Encuentre las esperanzas matemáticas para cada estimador, comente.
La esperanza matemática se define como:

$$E(X) = \sum_{i=1}^{n} x_i \cdot P(X = x_i)$$

Las esperanzas matemáticas son:

| $E[\hat{M}_e]$ | $E[\hat{\mu}]$ | $E[\hat{M}]$ | $E[\hat{M}_H]$ |
|---|---|---|---|
| 6 | 6 | | 5.237505 |

Se puede observar que la esperanza de los estimadores de la mediana y de la media son iguales al valor del parámetro poblacional, por lo cuál son estimadores insesgados para la mediana y la media poblacional. Por el contrario, la esperanza de la media armónica no es igual al valor del parámetro poblacional, entonces este estimador es sesgado para la media armónica poblacional.

e. Encuentre el valor del sesgo en cada caso; comente.

El sesgo de un estimador se define como:

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta$$

En este caso el sesgo para cada estimador es:

| $B(\hat{M}_e)$ | $B(\hat{\mu})$ | $B(\hat{M})$ | $B(\hat{M}_H)$ |
|---|---|---|---|
| 0 | 0 | | 0.167082 |

Se puede observar lo que se mencionó anteriormente, los dos primeros estimadores(para la mediana y la media) son insesgados, y el estimador de la media armónica es sesgado pero podríamos decir que presenta un sesgo pequeño.

f. Encuentre la varianza de cada estimador.

La varianza para cada estimador se define como:

$$V(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$$

donde $\theta$ es la esperanza de $\hat{\theta}$, $E[\hat{\theta}]$

Las varianzas para cada estimador son:

| $V(\hat{M}_e)$ | $V(\hat{\mu})$ | $V(\hat{M})$ | $V(\hat{M}_H)$ |
|---|---|---|---|
| 2.4 | 0.866666 | | 0.972771 |

g. Encuentre el Error Cuadrático Medio para cada estimador. Verifique la siguiente igualdad:

$$ECM(\hat{\theta}) = Var(\hat{\theta}) + [B(\hat{\theta})]^2$$

El error cuadrático medio se define como:

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

donde $\theta$ es el parámetro poblacional

El error cuadrático medio para cada estimador es:

| $ECM(\hat{M}_e)$ | $ECM(\hat{\mu})$ | $ECM(\hat{M})$ | $ECM(\hat{M}_H)$ |
|---|---|---|---|
| 2.4 | 0.866666 | | 1.000687 |

La igualdad $ECM(\hat{\theta}) = Var(\hat{\theta}) + [B(\hat{\theta})]^2$ se verifica en todos los casos.

h. Se dice que el sesgo del estimador $\hat{\phi}$ es despreciable cuando: $\left|\frac{B(\hat{\phi})}{\sqrt{V(\hat{\phi})}}\right| < \frac{1}{10}$. ¿Es el sesgo despreciable en cada estimador? Comente.

Calculando $\left|\frac{B(\hat{\phi})}{\sqrt{V(\hat{\phi})}}\right|$ para cada estimador, tenemos los valores en la siguiente tabla:

| $M_e$ | $\mu$ | $M$ | $M_H$ |
|---|---|---|---|
| 0 | 0 | | 0.169404 |

Dado que los estimadores para $M_e$ y para $\mu$ son insesgados, claramente su sesgo es despreciable ya que este es 0. Pero en el estimador para $M_H$, tenemos que $\left|\frac{B(\hat{\phi})}{\sqrt{V(\hat{\phi})}}\right| = 0.169404 > \frac{1}{10}$ por lo tanto, el sesgo de este estimador no es despreciable, es decir, el sesgo del estimador para $M_H$ representa una parte considerable de la desviación estándar de la estimación.

i. ¿Cuál de todos es el mejor estimador y porque?

Dado que estamos comparando dos estimadores sesgados y uno insesgado, se realiza la comparación por medio del error cuadrático medio. En este caso, tenemos que el mejor estimador es $\hat{\mu}$ ya que tiene el menor ECM que es 0.866666. Entre los dos estimadores restantes, es mejor el estimador de la media armónica que el de la mediana, ya que su ECM respectivamente es de 1.000687 y 2.4. Aquí podemos evidenciar que no necesariamente los mejores estimadores son los insesgados, ya que estos pueden tener una varianza alta como es el caso del estimador $\hat{M}_e$, en cambio, el estimador de la media armonica $\hat{M}_H$ tiene un sesgo, pero su varianza es mucho más pequeña, por lo cuál el error cuadrático medio también es menor.

j. ¿Qué tanto representa el Sesgo de cada estimador respecto del ECM. Comente.

Para saber el peso que tiene el sesgo de cada estimador respecto del ECM, calculamos la fracción $\frac{B(\hat{\theta})}{ECM(\hat{\theta})}$, este resultado para cada estimador se observa en la siguiente tabla:

| $M_e$ | $\mu$ | $M$ | $M_H$ |
|---|---|---|---|
| 0 | 0 | | 0.166967 |

Dado que los primeros dos estimadores son insesgados, lógicamente el sesgo de estos no aporta a su error cuadrático medio, el error cuadrático medio queda totalmente determinado por la varianza que ellos tengan. En el estimador de la media armónica, el sesgo representa el 16.7 % aproximadamente del error cuadrático medio, el resto (83.3 %) queda determinado por la varianza que este estimador tenga.