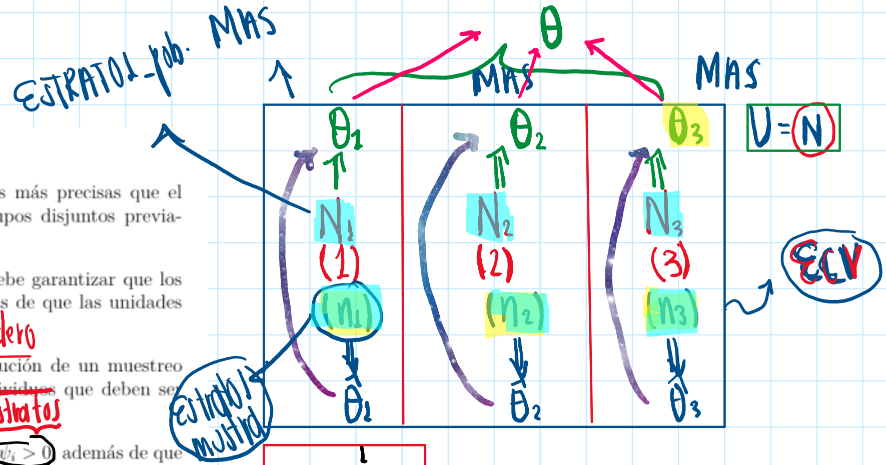
Segundo Taller

miércoles, 26 de febrero de 2025 2:02 p. m.



1. Determine el valor de verdad de las siguientes afirmaciones.

- (a) Un muestreo aleatorio simple conllevará estimaciones más precisas que el muestreo aleatorio estratificado incluso si existen subgrupos disjuntos previamente identificados. Falsa $\Rightarrow$ Más precisas
- (b) Para realizar un muestreo aleatorio estratificado se debe garantizar que los estratos son suficientemente heterogéneos entre sí, además de que las unidades intra- estrato son adecuadamente homogéneas. Verdadero
- (c) Se establece la restricción $N = \sum_{i=1}^{L} N_i$ para la ejecución de un muestreo aleatorio estratificado, donde L refiere al número de ~~individuos~~ Estratos que deben ser considerados. Falso
- (d) El factor de afijación muestral ψ_i cumple la restricción $\psi_i > 0$ además de que su estimación depende de la información conocida, permitiendo la determinación de las sub- muestras n_i. $0 \leq \psi_i \leq 1$

$N = \sum_{i=1}^{L} N_i$; $\boxed{L = \#\ \text{Estratos}}$; $\boxed{L=3}$

$n = \sum_{i=1}^{L} n_i \Rightarrow n_i = \psi_i \cdot n$ Proporción

$\psi_i = \frac{N_i}{N}$

Proporcional ; Neymann ; Costos

- (e) La estimación de un parámetro poblacional cualquiera $\hat{\theta}$ corresponde a una suma ponderada de los parámetros estimados por estrato $\hat{\theta}_i$ por el factor de afijación definido ψ_i. Falsa

$\boxed{\hat{\theta} = \sum_{i=1}^{L} m_i \cdot \hat{\theta}_i}$; $m_i = \frac{N_i}{N}$

Estimación global : $\mathrm{Var}[\hat{\theta}] = \mathrm{Var}\left[\sum_{i=1}^{L} m_i \cdot \hat{\theta}_i\right] = \sum_{i=1}^{L} m_i^2\, \mathrm{Var}[\hat{\theta}_i]$ MAS

$N = 81961$; $N_1 =$ $N_2 =$ $N_3 =$

$\boxed{N_i = [N_1, N_2, N_3] ; \quad N : 81961} \Rightarrow N_i/N : \boxed{\psi_i = [\psi_1, \psi_2, \psi_3]}$

$$n_i = \frac{N_i \sigma_i}{\sum_{k=1}^{L} N_k \sigma_k} n.$$

ψ_i = Factor afijación ; $\boxed{\sigma_i = \text{Desviación estándar}}$

```
> n_iNeymann
[1] 211.66275 381.12436  83.80088
```

n_1: Tamaño mínimo primer estrato ; n_2 ; n_3

$\boxed{\text{Estimar } \mu : (\text{Ingreso promedio poblacional} : N)}$: $\mu \rightarrow \hat{\mu} \leftarrow \hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$

- $\hat{\mu} = \sum_{i=1}^{3} \hat{\mu}_i \cdot m_i$ MAS : $\boxed{\text{Estimar\_parametro()}}$

```
> estimar_mu1
$Estimador
[1] 1319589

$Varianza
[1] 8320882272

$IC
[1] 1140133 1499045
```

$\hat{\mu}_1$; $\mathrm{Var}[\hat{\mu}_1]$; LI ; LS

$\alpha = 0,05$

Estimación estrato 1

Con una confianza del 95% se puede determinar que el ingreso promedio por hogar está comprendido entre LI y LS

LI LS

Con una confianza del 95% se puede determinar que el ingreso promedio por hogar está comprendido entre LI y LS.

```
> estimar_mu2
$Estimador
[1] 2589569

$Varianza
[1] 30548070450

$IC
[1] 2245718 2933420
```

$\hat{\mu}_2$

$\widehat{Var}[\hat{\mu}_2]$

Interpretarlo

```
> estimar_mu3
$Estimador
[1] 5188056

$Varianza
[1] 614797678569

$IC
[1] 3581920 6794193
```

$\hat{\mu}_3$

$\therefore \hat{\mu} = \sum_{i=1}^{3} m_i \cdot \hat{\mu}_i$

$\hat{\mu} = m_1 \cdot \hat{\mu}_1 + m_2 \cdot \hat{\mu}_2 + m_3 \cdot \hat{\mu}_3$

```
> mu_hat_global
[1] 2091992
```

$\Rightarrow \hat{\mu}$: El ingreso promedio por hogar para toda la población (estimación).

• Por estrato: $\hat{\mu}_i \pm t_{\alpha/2, n-1} \cdot \sqrt{\widehat{Var}[\hat{\mu}_i]}$; $\hat{\mu} \pm t_{\alpha/2, n-1} \cdot \sqrt{\widehat{Var}[\hat{\mu}]}$

$\Downarrow$

$m_i = \frac{N_i}{N}$

$\sum_{i=1}^{3} m_i \cdot \hat{\mu}_i \pm t_{\alpha/2, n-1} \cdot \sqrt{\sum_{i=1}^{3} m_i^2 \cdot \widehat{Var}[\hat{\mu}_i]}$

```
> IC_muglobal
[1] 1895520 2288465
```

LI LS

Con una confianza del 95% se puede determinar que el ingreso promedio por hogar para toda la población varía entre LI y LS

$IC: = \sum_{i=1}^{3} \frac{N_i}{N} \cdot \hat{\mu}_i \pm t_{\alpha/2, n-1} \sqrt{\sum_{i=1}^{3} \frac{N_i^2}{N^2} \cdot \widehat{Var}[\hat{\mu}_i]}$

```
# Para el cálculo de los intervalos de confianza
# varianzas estimadas de todos los estratos:
varianzaest_mu <- c(estimar_mu1$Varianza, estimar_mu2$Varianza, esti
# Cuando se complete la información anterior, correr la línea:
varianzaest_global <- sum(m_i^2) * varianzaest_mu)
```

Proporción de hogares que se perciben como seguros : Sec (0,1)

```
> estimar_p1
$Estimador
[1] 0.8615385

$Varianza
[1] 0.0003651336

$IC
[1] 0.8239461 0.8991308
```

LI LS

$\hat{P}_1$: Proporción estimada de personas en el estrato 1 que se sienten seguras.

Con una confianza del 95% se puede determinar que la proporción de personas que se sienten seguras en el estrato 1 varía entre LI y LS

$\rightsquigarrow \hat{p}_i \pm t_{\alpha/2, n-1} \cdot \sqrt{\widehat{Var}[\hat{p}_i]}$

```
> estimar_p2
$Estimador
[1] 0.8796296

$Varianza
[1] 0.0003250936

$IC
[1] 0.8441579 0.9151014
```

$\hat{p}_2$

```
> estimar_p3
$Estimador
[1] 0.862069

$Varianza
[1] 0.004211559

$IC
[1] 0.7291345 0.9950035
```

$\hat{p}_3$

$\Rightarrow \hat{p} = \sum_{i=1}^{3} m_i \cdot \hat{p}_i = m_1 \cdot \hat{p}_1 + m_2 \cdot \hat{p}_2 + m_3 \cdot \hat{p}_3$

```
> p_hat_global
[1] 0.8702043
```

$\Rightarrow \hat{p}$: La proporción de personas que se sienten seguras (población) es de 0,8702.

$\hat{p}$ ✓

$\Uparrow$

$\Rightarrow \sum_{i=1}^{3} \frac{N_i}{N} \cdot \hat{p}_i \pm t_{\alpha/2, n-1} \sqrt{\sum_{i=1}^{3} \frac{N_i^2}{N^2} \cdot \widehat{Var}[\hat{p}_i]}$

```
> IC_pglobal
[1] 0.8449179 0.8954907
```

LI LS

$\Rightarrow \hat{p}$.

∴ Con una confianza del 95%

$$\sum_{i=1} \frac{N_i}{N} \cdot p_i \pm t_{\alpha/2, n-L} \sqrt{\sum_{i=1} \frac{N_i^2}{N^2} \cdot Var[p_i]}$$

```
[1] 0.8449179 0.8954907
```

LI LS → p.

• Con una confianza del 95% se puede determinar que la proporción de personas en toda la población que se sienten seguras varía entre LI y LS

4. Determine el tamaño de muestra mínimo necesario para estimar la proporción de personas que poseen propiedad recreativa con un límite para el error de estimación del 2% y una confianza del 95%. **Para los p_i utilice las estimaciones intra- estrato y para el valor de la afijación ψ_i emplee la afijación proporcional.**

Cuando se impone un límite para el error de estimación δ con un nivel de confianza del 95%, la ecuación anterior toma la siguiente forma

Error estimación

$$n = \frac{\sum_{i=1}^{L} \frac{N_i^2 p_i (1-p_i)}{\psi_i}}{N^2 D + \sum_{i=1}^{L} N_i p_i (1 - p_i)} \qquad (**)$$

$\hat{p_i}$

con $D = \frac{\delta^2}{Z^2_{\alpha/2}}$ al estimar a p y $D = \frac{\delta^2}{N^2 Z^2_{\alpha/2}}$ al estimar a A.

A continuación se definen los criterios de afijación los cuáles determinan los valores de los ψ_i en las ecuaciones anteriores.

```
> n
[1] 1069.792
```

El tamaño mínimo necesario para estimar la proporción de personas que se sienten seguras (población) con una confianza del 95% y un límite error del 2%.

Redondeado

$\boxed{n_{redondeado} = n_{1\_redond} + n_{2\_redond} + n_{3\_redond}}$: n_sinredond : $n_i = \psi_i \cdot n_{sinredondear}$

$n_i : \lceil n_i \rceil = n_{i\_redond}$

(No hay restricción presupuestal) (costos)

```
> n_1
[1] 513
> n_2
[1] 512
> n_3
[1] 46
> n_redond <- n_1 + n_2 + n_3
> n_redond
[1] 1071
```

"Con una confianza del 95% y un límite de error de estimación del 2% se puede determinar que el tamaño mínimo para estimar la proporción de personas que se sienten seguras (Población global) es 1071 hogares".