

Parcial 2 muestreo

Jhonatan Smith Garcia - 1049705595

30/7/2021

R Markdown

Ejercicio 1:

- Se pide numero promedio de horas a la semana estimadas para mu.
- Se pide Numero total A, en horas de dicha region que tienen mas de 2 niños.

Datos:

```
Nh = c(98,65,152) #Poblacion en cada estrato
N = sum(Nh)
nh = c(21,16,13) # Muestras tomadas para cada estrato
n = sum(nh)
yh = c(24,32,27) # Promedio en cada estrato
ph = c(0.8095238,0.5625,0.6153846 ) # Proporcion c/d estrato si ven tv
#en cada estrato (A,B y C) respectivamente
sh = c(9,7,12)
s2h = sh^2
```

Ahora, para el inciso a se pide estimar promedio de horas que cada hogar dedica a ver tv. Para ello, se tiene que:

```

$$\bar{y}_{est} = \sum_{h=1}^H (Nh/N) * \bar{y}_h$$

y_bar_est = sum((Nh/N)*yh)
y_bar_est
```

```
## [1] 27.09841
```

Por tanto, se tiene que la media de horas por semana que ve cada hogar en la region de estudio es de 27.09841

Ahora, para calcular la varianza del estimador esta dada por la siguiente expresion

varianza estratificada de la media “Y barra estratificado”:

$$Var[\bar{y}_{est}] = \frac{1}{N^2} * Var[\bar{y}_h]$$

Donde;

```
Var[yh_bar] = (1 - nh/Nh) * sigma^2/nh
var_yh_bar= (1-nh/Nh)*s2h/nh #Varianza de Y-Barra para cada estrato
var_y_est = 1/N^2*sum(Nh^2*var_yh_bar) # Varianza media estratificada
var_y_est
```

```
## [1] 2.750247
```

La varianza estratificada de \bar{y} es 2.750247

Y por definicion se tiene que; *con una confianza de aproximadamente 95%*

```
ee = sqrt(var_y_est )
B1 = 2*ee
B1
```

```
## [1] 3.316774
```

El limite del error estimado es $B1 = 3.316774$

b) *Proporcion:*

Para el calculo de dicha proporcion (proporcion de hogares de la region >2 niños por hogar)

se tiene que:

$$\hat{P}_{est} = \sum_{h=1}^H \frac{N_h}{N} * p_h$$

ademas;

$$\hat{A}_{est} = N * \hat{P}_{est}$$

```
p_est = sum((Nh/N)*ph)
p_est
```

```
## [1] 0.6648708
```

Dicha proporcion es del 0.6648708 (aproximadamente 67%)

Ahora, para calcular un B (LEE) se necesita la varianza del estimador:

```
qh= 1-ph
#Recuerde que varianza proporcion esta dada por:
var_ph = (1 - nh/Nh)*ph*qh/(nh-1) # <- es la varianza para cada ph.

var_p_est = 1/N^2*sum(Nh^2*var_ph)
var_p_est
```

```
## [1] 0.005312749
```

Por tanto, la varianza de $\hat{p}_{est} = 0.005312749$

Ahora:

```
ee = sqrt(var_p_est) # Error estimado

B = ee*2
B # Limite del error estimado
```

```
## [1] 0.1457772
```

Y este es el limite del error estimado con una confianza de aproximadamente 95%

c) Tamaño de la muestra

La formula tamaño de la muestra para la media con LEE B es:

$$n = \frac{\sum \frac{N_h^2 * \sigma_h^2}{w_h}}{N^2 * D + \sum_{h=1}^H N_h * \sigma_h^2}$$

Donde se tiene que $w_h = \frac{n_h}{n}$ y ademas

$$D = \frac{1}{N^2} * \sum N_h^2 * \frac{\sigma_h^2}{n * w_h} - \frac{1}{N^2} \sum N_h^2 * \frac{\sigma^2}{N_h}$$

Recuerde que:

$$D = Var[\bar{y}_{est}] = \frac{B^2}{Z^2}$$

$$B = Z * \sqrt{Var[\bar{y}_{est}]} = Z * E.E(\bar{y}_{est})$$

Entonces, con esto en mente, se tiene que:

Se usa afijacion optima de Neyman pues no se habla de costos de muestreo, asi que se asumen costos iguales.

```
B= 3 #Limite del error
Z = 2 # Debido a que se pide de "aproximadamente 95%"

wh_optima_costo_igual = (Nh*sh)/(sum(Nh*sh)) #Neyman
wh = wh_optima_costo_igual #Aqui va la afijacion de la muestra

D = B^2/(Z^2) # D es igual a la varianza de ybar estratificado

n_optimo = sum(((Nh^2*s2h)/wh))/(N^2*D+sum(Nh*s2h))
ceiling(n_optimo) # Redondeo por encima

## [1] 39

asignacion = wh*n_optimo
ceiling(asignacion)
```

```
## [1] 11 6 23
```

Con esto, el tamaño optimo para n será de 41 muestras, tomando una afijacion proporcional para cada estrato.

Ademas, segun la afijacion de la muestra, ha de tomarse asi:

11 primer estrato.

6 del segundo.

23 del tercero.

Ejercicio 2:

Estimar valor promedio pagado por almuerzo en dicha cafeteria, a travez de Muestreo Sistemático-Estándar y establezca el LEE= con un aprox del 95%

Datos:

```
N = 820
n = 40
k = 20
y_bar = 4800
sh = 1045
```

Ademas, no hay indicion acerca del orden de los elementos de la poblacion, por tanto el muestreo Sistemático Lineal Estandar será equivalente a usar Muestreo Aleatorio Simple (MAS):

$\bar{y}_{sis} = \frac{\sum (y_i)_{i=1}^n}{n}$ Y claramente se tiene que la media muestral se estima similarmente que un muestreo aleatorio simple:

Como \bar{y} es el estimador de μ , según enunciado es 4800

Calculemos su varianza:

```
s2h = sh^2  
var_y_bar_mas = (s2h/n)*(N-n)/(N)
```

Finalmente, el Límite del error estimado B2 estará dado por

```
B2 = 2*sqrt(var_y_bar_mas)  
B2
```

```
## [1] 322.2973
```

Con una confianza de *aproximadamente* el 95%