

Examen 1 muestreo

Jhonatan Smith G

24/3/2021

```
datos =read.csv("datos.csv",sep=";")
```

- A) Segun el ejercicio, se nos solicita calcular el tamaño de una muestra tal que: Se permita estimar el gasto promedio en Salario minimo por hogar El limite del error sea de 0.6 SMLV. Con una confianza dell 99%.

Se sabe que el promedio poblacional de una poblacion será μ . Ahora, con esta informacion podemos aplicar la formula del tamaño muestral dada por:

- $n = (N * S^2) / ((N - 1) * D + S^2)$ con N el tamaño de la poblacion y $D = \frac{B^2}{Z_{\frac{\alpha}{2}}}$

Se conoce el valor de N pero se desconoce el de S^2 (La varianza muestral).

Calculando la varianza muestral del ingreso de la muestra piloto (dada por aparte de la base de datos) con la funcion de R var(x) entonces:

Los hogares seleccionados en la muestra pilotos fueron los hogares identificados con los codigos 3,65,24,105,136,67,9,95 y 69 respectivamente. En ese orden, si vamos a la base de datos y nos fijamos en los datos de gastos tendremos que, en ese mismo orden, los gastos fueron de:

1.05,0.88,0.59,1.00,1.01,0.74,0.85,1.18 y 1.14 respectivamente.

```
piloto= c(1.05,0.88,0.59,1.00,1.01,0.74,0.85,1.18,1.14 )
var(piloto)
```

```
## [1] 0.03654444
```

$S^2 = 0.03654444$

Ahora, teniendo calculado la varianza muestral, calculemos el valor de D .

Para ello, tenemos que: $D = \frac{B^2}{Z_{\frac{\alpha}{2}}}$ y de aqui, necesitamos una aproximacion a la normal al 99%. Este valor es:

```
qnorm(0.995)
```

```
## [1] 2.575829
```

$Z_{\frac{\alpha}{2}} = 2.575829$

finalmente, $D = 0.001397608$

N = al tamaño total de la poblacion estudiada, es de 200. Entonces reemplazando en la formula dada para calcular el tamaño de n tenemos que: $n = (200 * 0.03654444) / (0.03654444 + 1990.001397608) = 23.22727$

Esto implica que el tamaño de la muestra para tener un error del 0.06 con un nivel de confianza al 99% debe de ser de 24 hogares.

B)

Un estimador de μ (media poblacional) para estimar el promedio de ingresos de los salarios, es la media muestral.

Sea X la variable aleatoria que mide el ingreso de los encuestados en la muestra piloto, entonces \bar{X} es un estimador insesgado de μ

entonces, el valor de $\mu = 1.011818$

$$1.011818 + \sqrt{((0.02583409167)/(33200))2.575829} = 1.077675 \quad 1.011818 - \sqrt{((0.02583409167)/(33200))2.575829} = 0.9459613$$

Entonces el IC pedido es (1.077675, 0.9459613)

- C) Con la muestra dada en (b) realice la estimación del gasto promedio por hogar para los hogares Sin Niños de la población en estudio junto con su intervalo de confianza del 99 % .

```
nohijos = c(0.93,0.88,0.79,0.74,1,0.85,0.75)
mean(nohijos)
```

```
## [1] 0.8485714
```

```
var(nohijos)
```

```
## [1] 0.009247619
```

Nuevamente, en este caso tenemos un calculo de un IC al 99% de confianza, sin embargo, esta ya es el calculo de una subpoblacion ya que la poblacion objetivos es, de los hogares que NO tienen hijos, calcular el ingreso promedio. De estos 33 hogares, 7 no tienen hijos. La media muestral se calcula sobre esos 7 hogares.

Ahora, podemos calcular un IC con la formula anterior, teniendo en cuenta el analisis subpoblacional realizado.

$$\text{Un IC pedido es } (0.7654692, 0.9316736) \quad 0.8485714 - \sqrt{((0.00924761926)/(337))2.575829} = 0.7654692 \\ 0.8485714 + \sqrt{((0.00924761926)/(337))2.575829} = 0.9316736$$

- D) Este ejercicio se realiza igual que el anterior, debido a problemas de conexion, no tuve tiempo de realizarlo. Sin embargo, bastaria con sacar la subpoblacion pedida de la muestra dada (en este caso de 27, pues 7 no tienen hijos y 26 si) y realizar los mismos calculos del IC