Trabajo I

Matias Bajac

2023-04-10

```
library("sampling")
library("tidyverse")
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr   0.3.5
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.5.0
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library("here")
```

```
## here() starts at /Users/matiasbajac/Desktop/Muestreo-I-
```

```
datos=load(here("Datos","RB (1).RData"))
rm(datos)
datos = rio_branco
rm(rio_branco)
```

```
var_names <- names(datos)[grepl("^NBI_", names(datos))][-13]


for (var_name in var_names) {
 datos[[var_name]] <- gsub("[89]", "0", datos[[var_name]])
}

for (var_name in var_names) {
 datos[[var_name]] <- as.numeric(datos[[var_name]]) ## pasamos las variables a numericas
}
datos_hogares=datos %>% mutate(NBI= NBI_EDUCACIÓN + NBI_HAC + NBI_MAT+NBI_COC +NBI_VIV +NBI_AGUA+NBI_SA

  datos_hogares2=datos_hogares %>% mutate( NBI = if_else(NBI>3,1,0))
```

El total poblacional de NBI a nivel hogares es :

```
t = sum(datos_hogares2$NBI)
```

$$Nos\ basaremos\ en\ el\ estmiador\ Horvitz\ thompson\ para\ estimar\ el\ total\ poblacional\ de\ la\ variable\ NBI$$

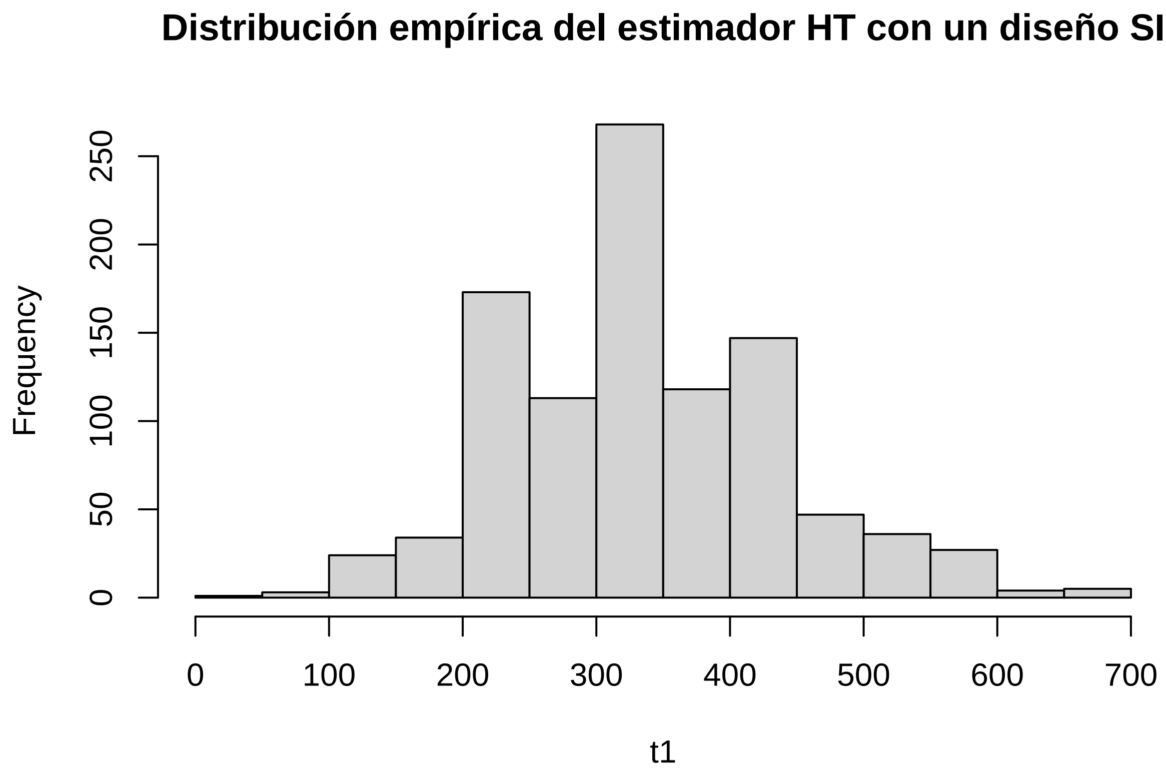
$$en\ el\ Diseño\ Simple\ la\ probabilidad\ de\ inclusion\ de\ primer\ orden\ es\ \pi_k = n/N$$

$$del\ estimador\ H-T\ sabemos\ que\ t_\pi = \sum_s y_k/\pi_k$$

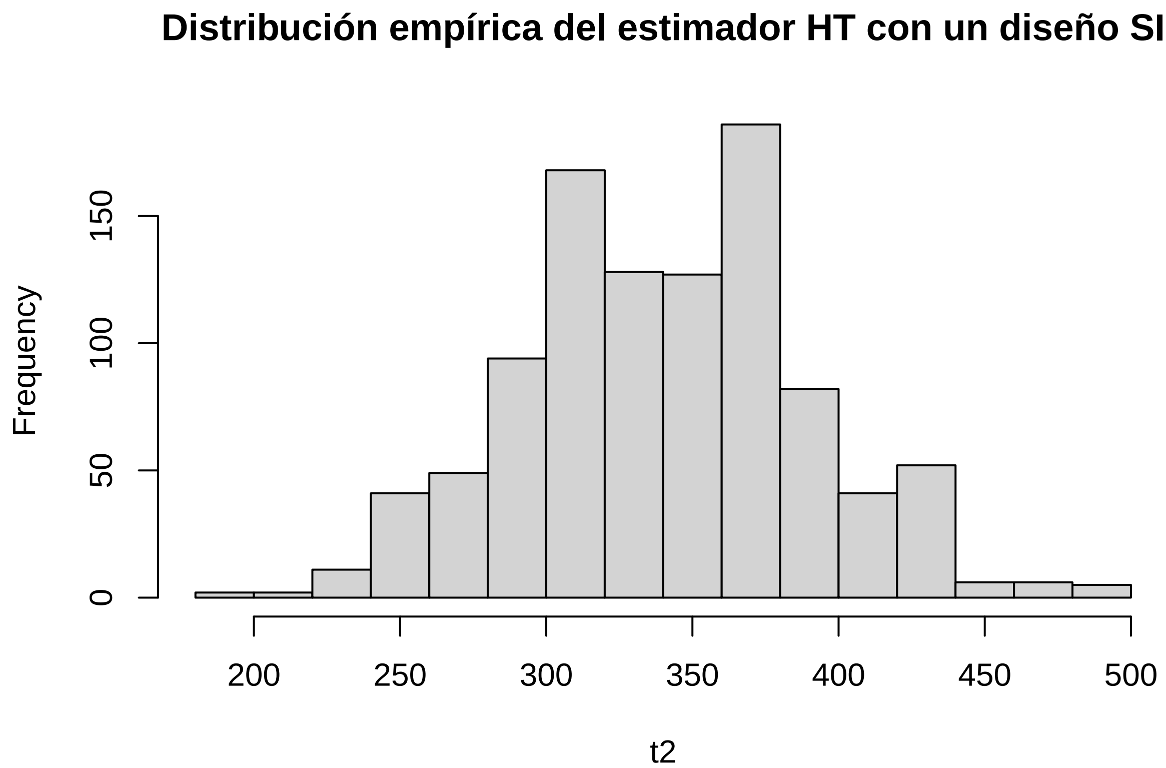
$$por\ lo\ tanto\ t_\pi = N * \bar{y_s}$$

Creamos una funcion en el cual calcula la estimacion del total para cada muestra.

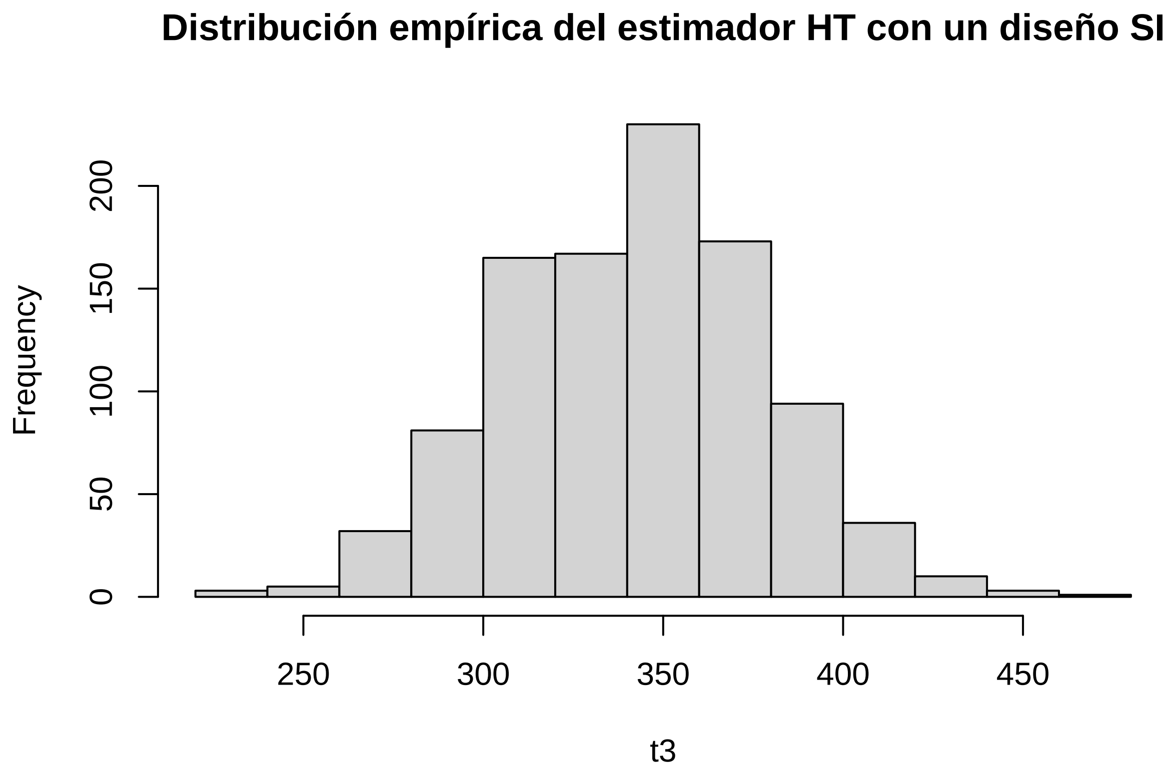
```r
N=nrow(datos_hogares2)

R=1000

 set.seed(1234)

t_si1=numeric()
totales_simple = function(n){
for(i in 1:R){
s=srswor(n,N)
m=getdata(datos_hogares2,s)
t_si1[i]=N*mean(m$NBI)

}
  return(t_si1)
}
t1 = totales_simple(150)
hist(t1,main="Distribución empírica del estimador HT con un diseño SI")
```




```r
t2 = totales_simple(600)
hist(t2,main="Distribución empírica del estimador HT con un diseño SI")
```




```r
t3 = totales_simple(1000)
hist(t3,main="Distribución empírica del estimador HT con un diseño SI")
```



Observamos que en los 3 casos se cumple que el estimador $\hat{t}$ es insesgado

```
set.seed(1234)

esperanza_total1=mean(t1)
esperanza_total2 = mean(t2)
esperanza_total3 = mean(t3)
```

Calculamos la varianza del estimador t para cada numero de muestra como paso previo para luego calcular el efecto diseño.

```
set.seed(1234)
varianza_tota_1_si=var(t1)
varianza_total_2_si = var(t2)
varianza_total_3_si= var(t3)
```

Calculamos con la varianza teorica y la comparamos con la simulada

```
n=150
N2= nrow(datos)

v_si1=N^2*(1-n/N2)*var(datos_hogares2$NBI)/n
n2=600
V_si2 = N^2*(1-n2/N2)*var(datos_hogares2$NBI)/n2
n3=1000
v_si3= N^2*(1-n3/N2)*var(datos_hogares2$NBI)/n3
```

Estimaremos el total de hogares que tienen XO