

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))][1:13]
```

```
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_SAL)
```

```
datos_hogares$X0 = ifelse(datos_hogares$X0==0,0,1)
```

El total poblacional de NBI a nivel hogares es :

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

```
N=nrow(datos_hogares)
n=150
set.seed(1234)
```

Nos basaremos en el estimador 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 $t_\pi = \sum_s y_k / \pi_k$

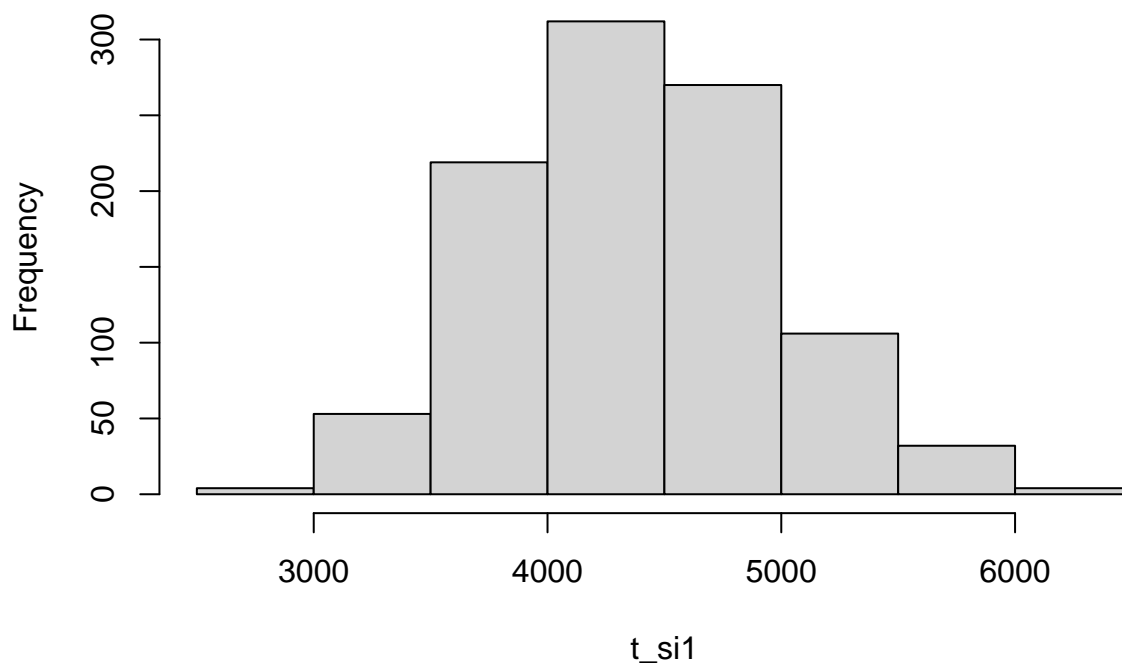
por lo tanto $t_\pi = N * \bar{y}_s$

En primer lugar, seleccionamos 1000 muestras con $n = 150$

```
R=1000
t_si1=numeric()
for(i in 1:R){
  s=srswor(n,N)
  m=getdata(datos_hogares,s)
  t_si1[i]=N*mean(m$NBI)
}

hist(t_si1,main="Distribución empírica del estimador HT con un diseño SI")
```

Distribución empírica del estimador HT con un diseño SI



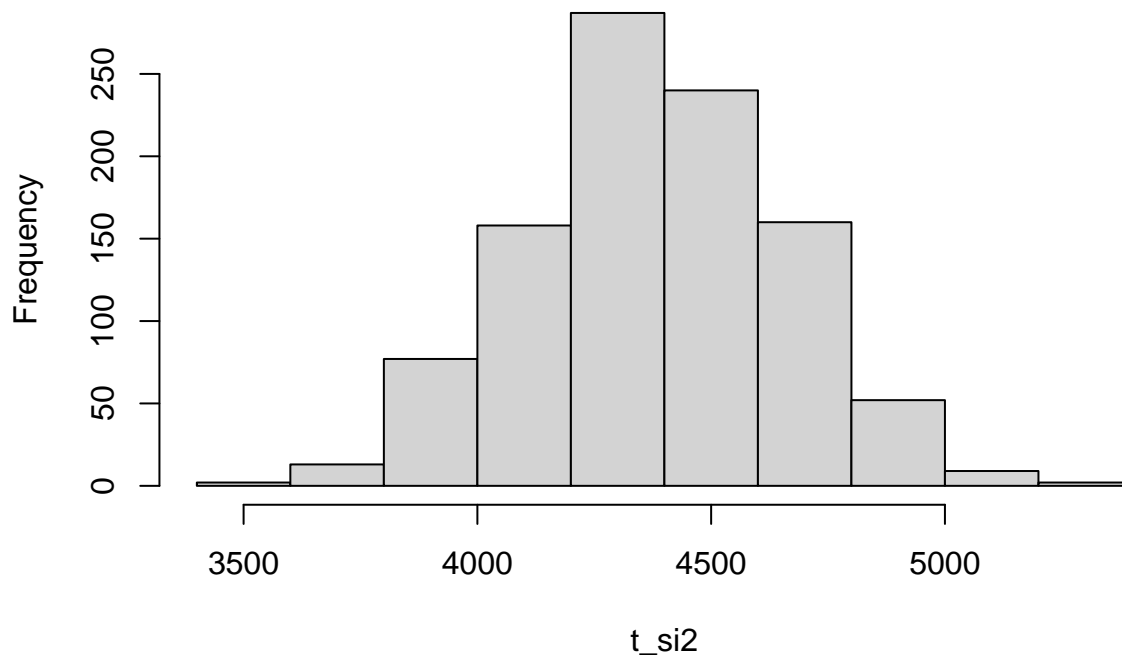
Seleccionamos 1000 muestras con $n = 600$

```
set.seed(1234)
n2=600

R=1000
t_si2=numeric()
for(i in 1:R){
  s=srswor(n2,N)
  m=getdata(datos_hogares,s)
  t_si2[i]=N*mean(m$NBI)
}

hist(t_si2,main="Distribución empírica del estimador HT con un diseño SI")
```

Distribución empírica del estimador HT con un diseño SI

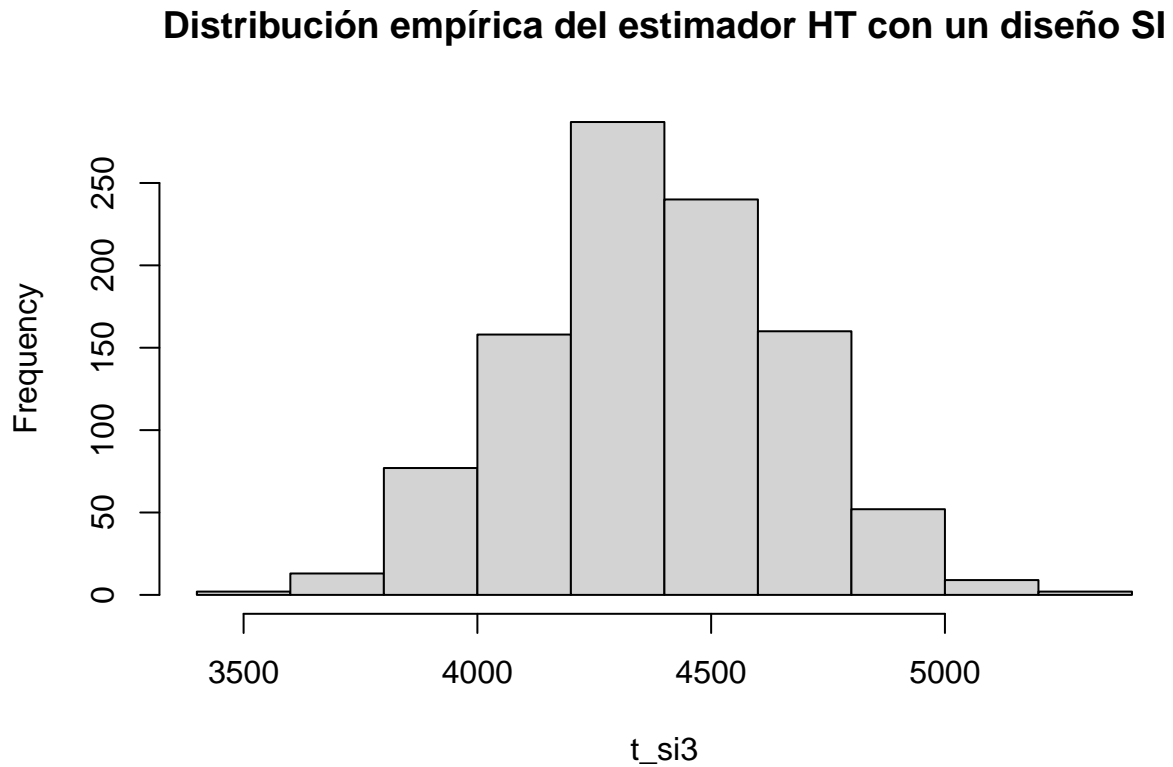


Seleccionamos 1000 muestras con $n = 1000$

```
set.seed(1234)
n3=1000

R=1000
t_si3=numeric()
for(i in 1:R){
  s=srswor(n2,N)
  m=getdata(datos_hogares,s)
  t_si3[i]=N*mean(m$NBI)
}
```

```
}
hist(t_si3,main="Distribución empírica del estimador HT con un diseño SI")
```



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

```
var_si1 = var(t_si1)
var_si2 = var(t_si2)
var_si3 = var(t_si3)
```

Estimaremos el total de hogares que tienen XO