Calculo del tamaño de muestra

```
library(dplyr)
library(TeachingSampling)
data("BigCity")
Hogares <- BigCity %>% group_by(HHID) %>%
          summarise(Estrato = unique(Zone),
                    Personas = n(),
                    Ingreso = sum(Income),
                    Gasto = sum(Expenditure),
                    Pobreza = unique(Poverty))
attach(Hogares)
```

Tamaño de muestra para estimar la media de la variable ingreso con un margen de error relativo menor al 5 %.

```
library(samplesize4surveys)
N <- nrow(Hogares)
mu <- mean(Ingreso)
sigma <- sd(Ingreso)
rme <- 0.05
ss4m(N = N, mu = mu, sigma = sigma,error = "rme", delta = rme)
```

```
## [1] 2678
```

Tamaño de muestra para estimar la media de la variable ingreso con un error relativo menor a 10 dólares.

```
me <- 10
rme <- me/mu
rme
```

```
## [1] 0.004697751
```

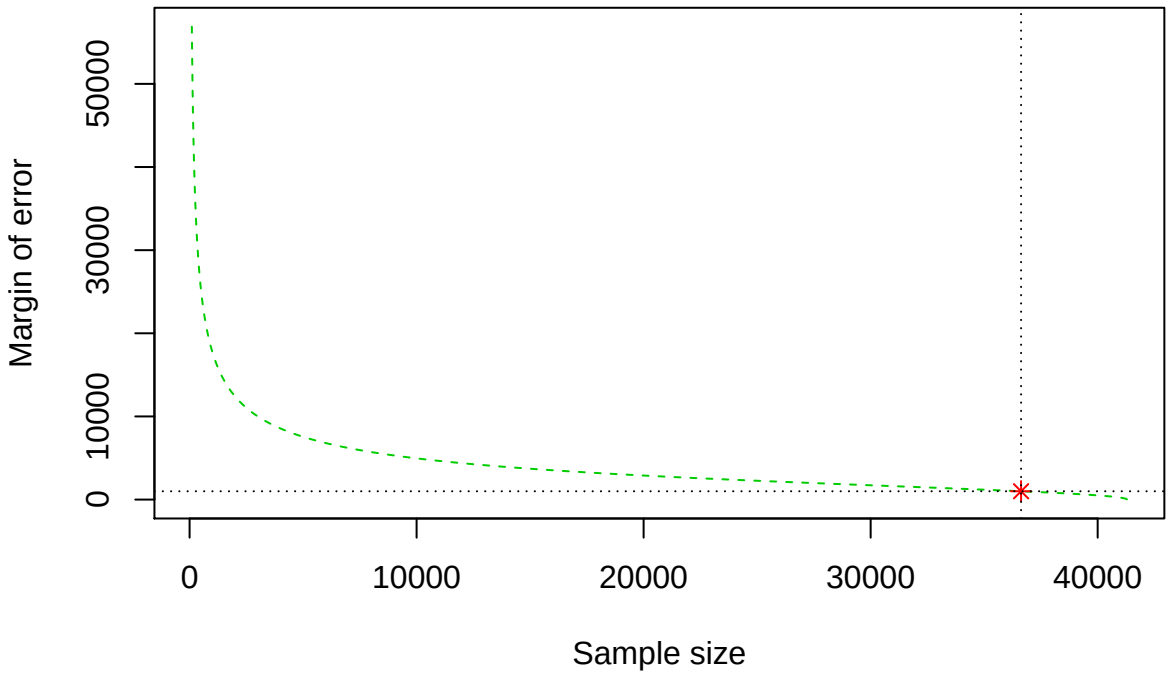
```
ss4m(N = N, mu = mu, sigma = sigma, error = "me", delta = 10)
```

```
## [1] 36627
```

```
ss4m(N = N, mu = mu, sigma = sigma, error = "rme", delta = rme)
```

```
## [1] 36627
```

```
ss4m(N = N, mu = mu, sigma = sigma, error = "me", delta = 10, plot = TRUE)
```




```
## [1] 36627
```

Tamaño de muestra para estimar la proporción de la variable pobreza == "pobreza relativa" con un margen de error menor al 2 %.

```
prop.table(table(Pobreza))
```

```
## Pobreza
##    NotPoor    Extreme   Relative
## 0.68743037 0.06989586 0.24267377
```

```
N <- nrow(Hogares)
P <- prop.table(table(Pobreza))[3]
me <- 0.02
ss4p(N = N, P = P, error = "me", delta = me)
```

```
## Relative
##     1693
```

Tamaño de muestra para estimar la proporción de la variable pobreza == "pobreza relativa" con un margen de error relativo menor al 2 %.

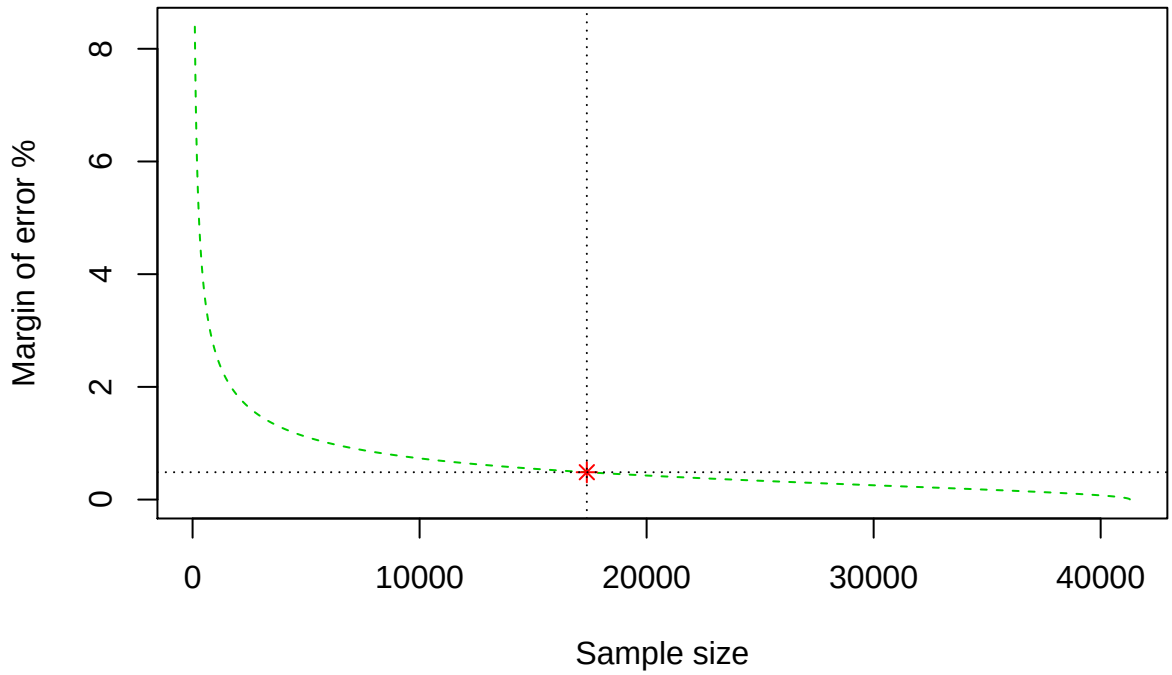
```
rme <- 0.02
me <- rme * P
me
```

```
##    Relative
## 0.004853475
```

```
ss4p(N = N, P = P, error = "me", delta = me)
```

```
## Relative
##    17366
```

```
ss4p(N = N, P = P, error = "me",delta = me, plot = TRUE)
```




```
## Relative
##    17366
```

Tamaño de muestra para estimar la media de la variable ingreso con un margen de error relativo menor al 5 %.

```
attach(Hogares)
N <- nrow(Hogares)
mu <- mean(Ingreso)
sigma <- sd(Ingreso)
deff <- 2.5
rme <- 0.05
```

```
ss4m(N = N, mu = mu, sigma = sigma,DEFF = deff, delta = rme, error = "rme")
```

```
## [1] 6100
```

tamaño de muestra para estimar la media de la variable ingreso con un error relativo menor a 10 dólares.

```
me <- 10
rme <- me/mu
rme
```

```
## [1] 0.004697751
```

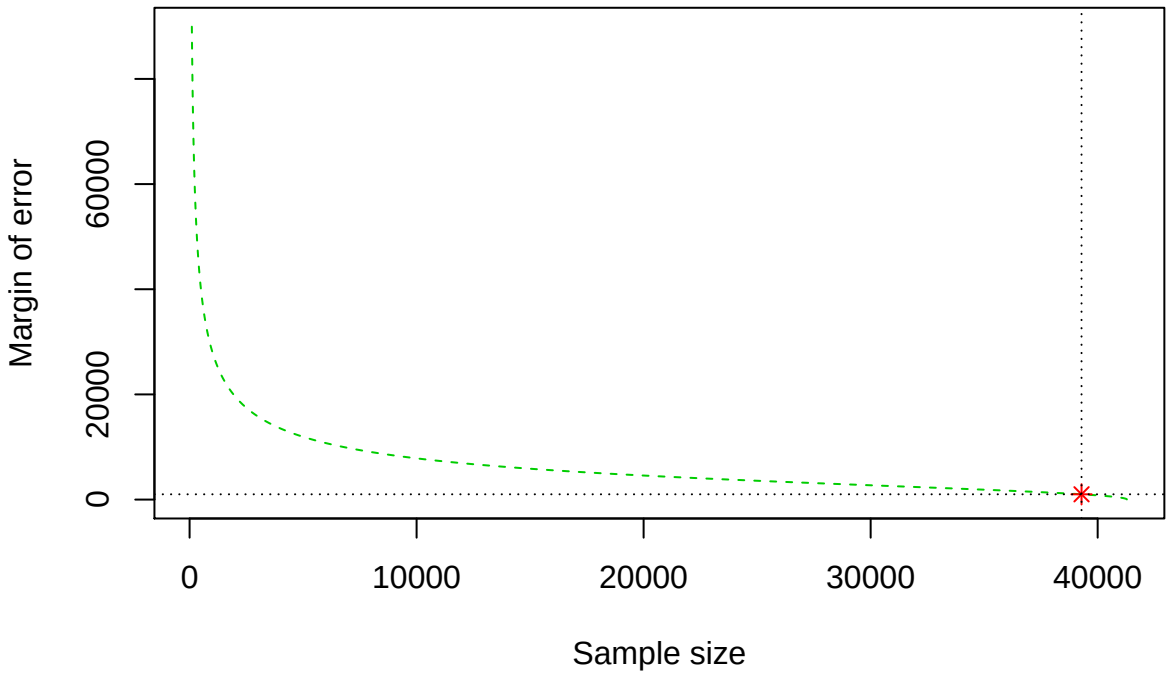
```
ss4m(N = N, mu = mu, sigma = sigma,DEFF = deff, delta = rme, error = "rme")
```

```
## [1] 39290
```

```
ss4m(N = N, mu = mu, sigma = sigma,DEFF = deff, delta = me, error = "me")
```

```
## [1] 39290
```

```
ss4m(N = N, mu = mu, sigma = sigma, DEFF = deff, delta = me, error = "me", plot = TRUE)
```




```
## [1] 39290
```

Tamaño de muestra para estimar la proporción de la variable pobreza == "pobreza relativa" con un margen de error menor al 2 %.

```
prop.table(table(Pobreza))
```

```
## Pobreza
##    NotPoor    Extreme   Relative
## 0.68743037 0.06989586 0.24267377
```

```
N <- nrow(Hogares)
P <- prop.table(table(Pobreza))[3]
deff <- 2.5
me <- 0.02
```

```
ss4p(N = N, P = P, DEFF = deff, delta = me, error = "me")
```

```
## Relative
##     3987
```

tamaño de muestra para estimar la proporción de la variable pobreza == "pobreza relativa" con un margen de error relativo menor al 2 %.

```
rme <- 0.02
me <- rme * P
me
```

```
##    Relative
## 0.004853475
```

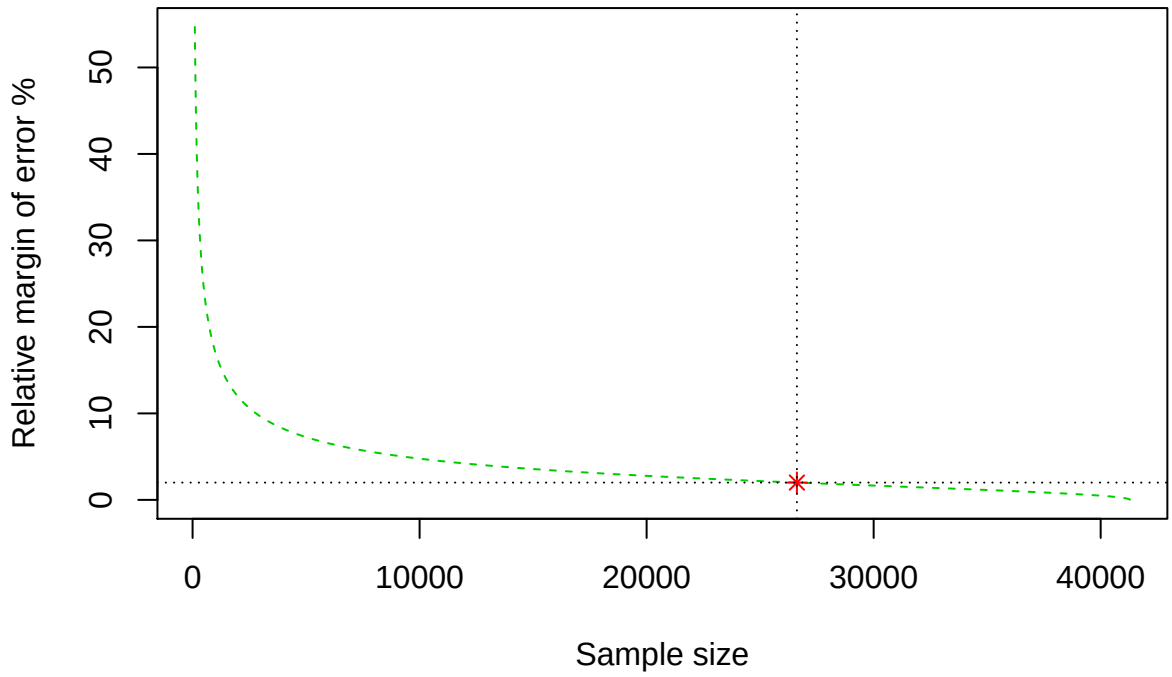
```
ss4p(N = N, P = P, DEFF = deff, delta = me, error = "me")
```

```
## Relative
##    26621
```

```
ss4p(N = N, P = P, DEFF = deff, delta = rme, error = "rme")
```

```
## Relative
##    26621
```

```
ss4p(N = N, P = P, DEFF = deff, delta = rme, error = "rme",plot = TRUE)
```




```
## Relative
##    26621
```

Calculo del tamaño para la proporcion

Para desocupados, ocupados e inactivos

```
library(TeachingSampling)
data(BigCity)
BigCity1 <- BigCity[!is.na(BigCity$Employment), ]
summary(BigCity1$Employment)
```

```
## Unemployed   Inactive   Employed
##       4630      44104      62188
```

```
BigCity1$Unemp <- Domains(BigCity1$Employment)[, 1]
BigCity1$Active <- Domains(BigCity1$Employment)[, 1] +
Domains(BigCity1$Employment)[, 3]
```

```
N <- nrow(BigCity)
M <- length(unique(BigCity$PSU))
r <- sum(BigCity1$Active)/N
b <- N/length(unique(BigCity$HHID))
rho <- ICC(BigCity1$Unemp, BigCity1$PSU)$ICC
P <- sum(BigCity1$Unemp)/sum(BigCity1$Active)
delta <- 0.07
```

```r
conf <- 0.95
m <- c(5:15)
```

```r
error <- P * delta
P - 1.96 * error
```

```
## [1] 0.05978575
```

```r
P + 1.96 * error
```

```
## [1] 0.07879966
```

```r
ss4HHSp(N, M, r, b, rho, P, delta, conf, m)
```

```
##    HouseholdsPerPSU PersonsPerPSU DEFF PSUinSample HouseholdsInSample
## 1                 5             8 1.21        1454               7269
## 2                 6            10 1.26        1256               7535
## 3                 7            11 1.31        1114               7800
## 4                 8            13 1.36        1008               8064
## 5                 9            15 1.41         925               8325
## 6                10            16 1.45         859               8585
## 7                11            18 1.50         804               8843
## 8                12            19 1.55         758               9100
## 9                13            21 1.60         720               9355
## 10               14            23 1.65         686               9609
## 11               15            24 1.70         657               9861
##    PersonsInSample
## 1            11763
## 2            12194
## 3            12623
## 4            13049
## 5            13472
## 6            13893
## 7            14311
## 8            14726
## 9            15139
## 10           15550
## 11           15958
```

Calculo para la media de los ingresos

```r
data(BigCity)
BigCity1 <- BigCity %>%
              group_by(HHID) %>%
              summarise(IncomeHH = sum(Income),
                        PSU = unique(PSU))
head(BigCity1)
```

```
## # A tibble: 6 x 3
##   HHID      IncomeHH PSU
##   <chr>        <dbl> <chr>
## 1 idHH00001    2775  PSU0001
## 2 idHH00002    1492. PSU0001
## 3 idHH00003    4280  PSU0001
```

```
## 4 idHH00004    2200  PSU0001
## 5 idHH00005    3119. PSU0001
## 6 idHH00006     675  PSU0001
```

```
summary(BigCity1$IncomeHH)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
##     10.0    596.6   1303.0   2128.7   2646.0 131680.0
```

```
mean(BigCity1$IncomeHH)
```

```
## [1] 2128.678
```

```
sd(BigCity1$IncomeHH)
```

```
## [1] 2905.597
```

```
N <- nrow(BigCity)
M <- length(unique(BigCity$PSU))
rho <- ICC(BigCity1$IncomeHH, BigCity1$PSU)$ICC
mu <- mean(BigCity1$IncomeHH)
sigma <- sd(BigCity1$IncomeHH)

delta <- 0.05
conf <- 0.95
m <- c(5:15)
ss4HHSm(N, M, rho, mu, sigma, delta, conf, m)
```

```
##    HouseholdsPerPSU DEFF PSUinSample HouseholdsInSample
## 1                 5 1.44         800               4000
## 2                 6 1.54         716               4295
## 3                 7 1.65         655               4588
## 4                 8 1.76         610               4880
## 5                 9 1.87         575               5171
## 6                10 1.98         546               5461
## 7                11 2.09         523               5750
## 8                12 2.20         503               6037
## 9                13 2.31         486               6324
## 10               14 2.41         472               6609
## 11               15 2.52         460               6893
```