

# Muestreo estratificado

## Muestreo aleatorio estratificado

Estratificación de la población

```
U1 <- c("Erik", "Sharon")
N1 <- length(U1)
U2 <- c("Yves", "Ken", "Leslie")
N2 <- length(U2)

U <- union(U1,U2)
N <- N1+N2
U

## [1] "Erik"    "Sharon"  "Yves"    "Ken"     "Leslie"
```

Selección de muestra aleatoria por estrato

```
library(TeachingSampling)

sam1 <- sample(N1, 1, replace=FALSE)
U1[sam1]

## [1] "Erik"

sam2 <- S.SI(N2,2)
U2[sam2]

## [1] "Ken"      "Leslie"
sam <- union(U1[sam1],U2[sam2])
sam

## [1] "Erik"    "Ken"     "Leslie"
```

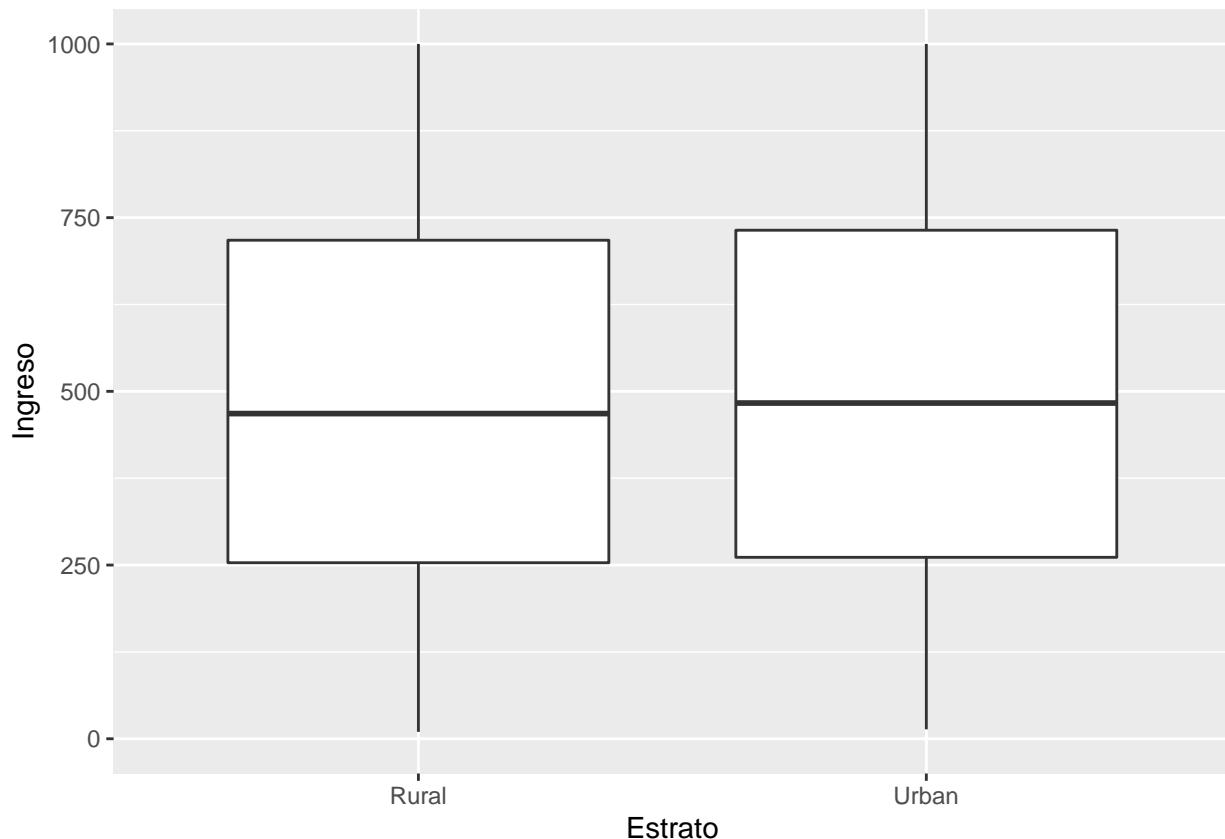
## Selección de muestra proporcional al tamaño del estrato

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

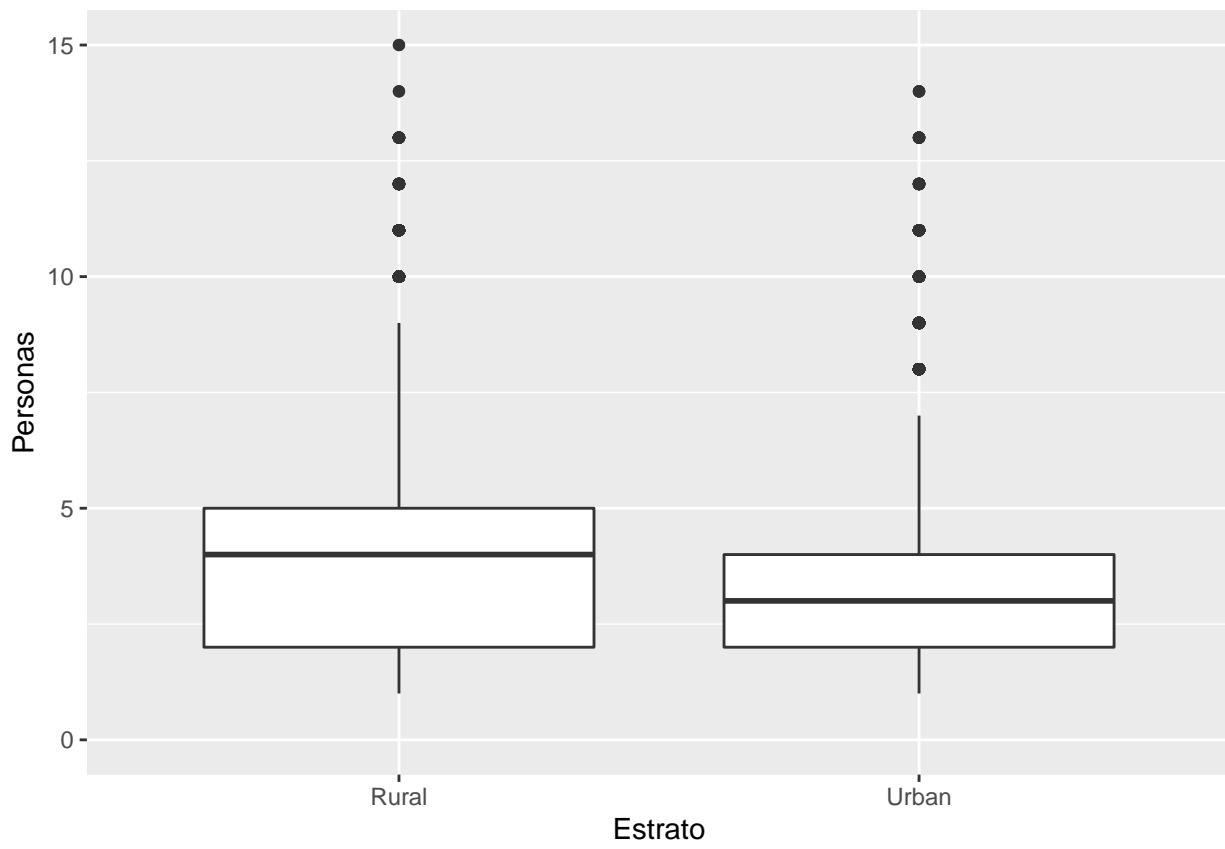
Observemos los boxplot respectivos para cada una de las variables

```
library(ggplot2)
attach(Hogares)

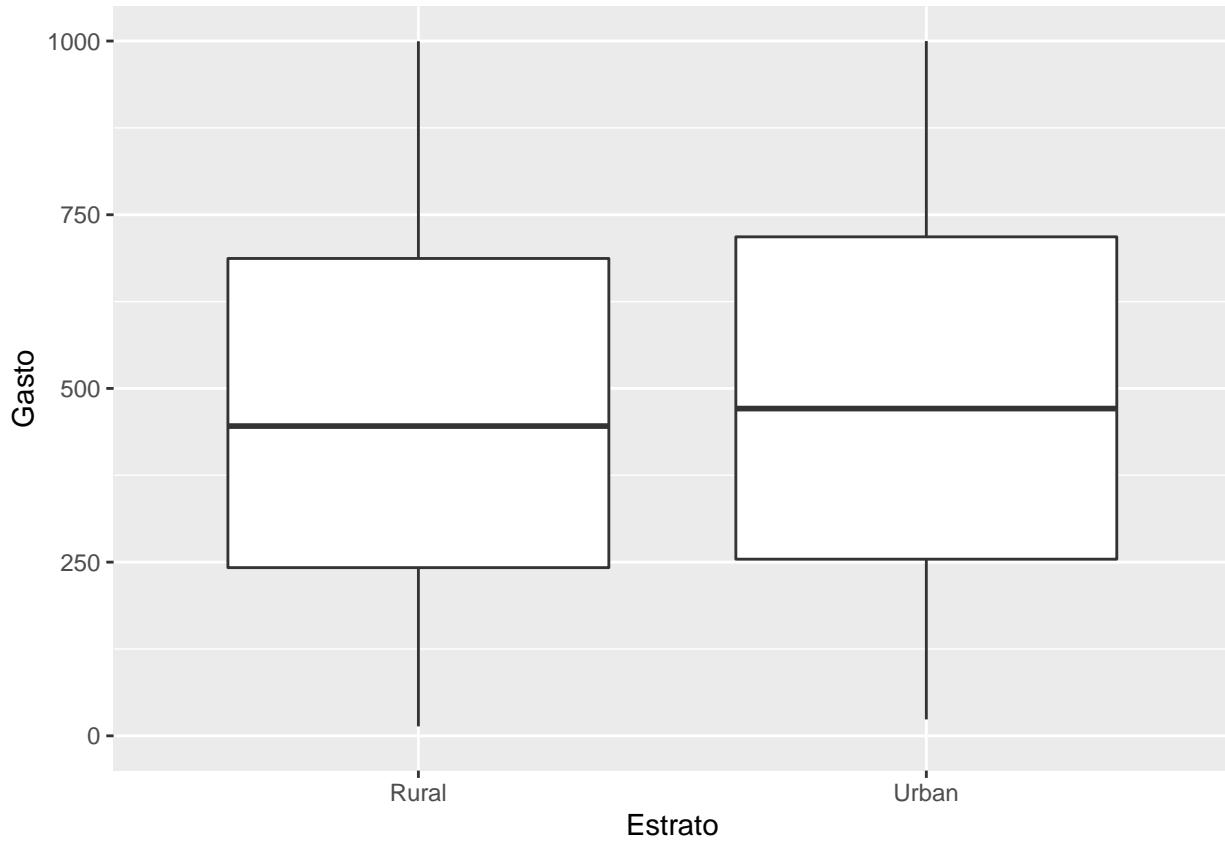
q1 = qplot(Estrato, Ingreso , data =Hogares, geom=c("boxplot"), ylim = c(0,1000))
q1
```



```
q2 = qplot(Estrato, Personas, data=Hogares, geom=c("boxplot"), ylim = c(0,15))
q2
```



```
q3 = qplot(Estrato, Gasto, data=Hogares, geom=c("boxplot"), ylim = c(0,1000))  
q3
```



Obtenemos tamaños poblacionales de cada estrato y el tamaño de muestra proporcional al tamaño del estrato

```
tamano.muestra = Hogares %>% group_by(Estrato) %>%
  summarise(Nh = n(),
            nh = round(2000*Nh/dim(Hogares)[1]))
```

```
tamano.muestra
```

```
## # A tibble: 2 x 3
##   Estrato     Nh     nh
##   <chr>    <int>  <dbl>
## 1 Rural     19174   929
## 2 Urban     22116   1071
```

```
sum(tamano.muestra$Nh)
```

```
## [1] 41290
```

```
sum(tamano.muestra$nh)
```

```
## [1] 2000
```

Obtenemos la muestra estratificada proporcional al tamaño

```
Nh <- tamano.muestra$Nh
nh <- tamano.muestra$nh
sam <- S.STSI(Estrato, Nh, nh)
muestra <- Hogares[sam,]
attach(muestra)
head(muestra)
```

```
## # A tibble: 6 x 6
```

```

##   HHID      Estrato Personas Ingreso Gasto Pobreza
##   <chr>     <chr>    <int>   <dbl> <dbl> <fct>
## 1 idHH00024 Rural          4    2757. 1429. NotPoor
## 2 idHH00086 Rural          4    1260   650. NotPoor
## 3 idHH00088 Rural          1    335.  199. NotPoor
## 4 idHH00091 Rural          3    455.  399. Relative
## 5 idHH00098 Rural          5   1321. 1528. Relative
## 6 idHH00112 Rural          4    557.  811. Relative

```

Con la función incorporada en la librería `TeachingSampling` es posible obtener las estimaciones correspondientes

```

estima <- data.frame(Ingreso, Gasto, Personas)
E.STSI(Estrato, Nh, nh, estima)

```

```

## , , N
##
##           Rural Urban Population
## Estimation 19174 22116     41290
## Standard Error 0     0         0
## CVE          0     0         0
## DEFF         NaN  NaN       NaN
##
## , , Ingreso
##
##           Rural      Urban Population
## Estimation 3.320787e+07 5.244166e+07 8.564953e+07
## Standard Error 1.235313e+06 1.602131e+06 2.023072e+06
## CVE          3.719939e+00 3.055073e+00 2.362035e+00
## DEFF         1.000000e+00 1.000000e+00 9.808009e-01
##
## , , Gasto
##
##           Rural      Urban Population
## Estimation 2.177444e+07 3.423700e+07 5.601143e+07
## Standard Error 6.817910e+05 1.029286e+06 1.234613e+06
## CVE          3.131153e+00 3.006356e+00 2.204216e+00
## DEFF         1.000000e+00 1.000000e+00 9.785838e-01
##
## , , Personas
##
##           Rural      Urban Population
## Estimation 74632.060280 75991.484594 1.506235e+05
## Standard Error 1174.802378 1119.424156 1.622736e+03
## CVE          1.574126    1.473092 1.077345e+00
## DEFF         1.000000    1.000000 9.847214e-01

```

Tamaño poblacional

```
E.STSI(Estrato, Nh, nh, estima)[, , "N"]
```

```

##           Rural Urban Population
## Estimation 19174 22116     41290
## Standard Error 0     0         0
## CVE          0     0         0
## DEFF         NaN  NaN       NaN

```

Estimación del total para el ingreso

```
E.STSI(Estrato, Nh, nh, estima)[, , "Ingreso"]
```

```
##                      Rural      Urban Population
## Estimation      3.320787e+07 5.244166e+07 8.564953e+07
## Standard Error 1.235313e+06 1.602131e+06 2.023072e+06
## CVE            3.719939e+00 3.055073e+00 2.362035e+00
## DEFF           1.000000e+00 1.000000e+00 9.808009e-01
```

Estimación total para el gasto

```
E.STSI(Estrato, Nh, nh, estima)[, , "Gasto"]
```

```
##                      Rural      Urban Population
## Estimation      2.177444e+07 3.423700e+07 5.601143e+07
## Standard Error 6.817910e+05 1.029286e+06 1.234613e+06
## CVE            3.131153e+00 3.006356e+00 2.204216e+00
## DEFF           1.000000e+00 1.000000e+00 9.785838e-01
```

Estimación de los totales poblacionales

```
E.STSI(Estrato, Nh, nh, estima)[, 3, ]
```

```
##          N     Ingreso      Gasto    Personas
## Estimation 41290 8.564953e+07 5.601143e+07 1.506235e+05
## Standard Error 0 2.023072e+06 1.234613e+06 1.622736e+03
## CVE          0 2.362035e+00 2.204216e+00 1.077345e+00
## DEFF         NaN 9.808009e-01 9.785838e-01 9.847214e-01
```

En R es posible encontrar el estimador del total para cada nivel según el dominio

```
Dominios <- Domains(Pobreza)
dom.NP <- Dominios[,3]*estima
dom.PR <- Dominios[,2]*estima
dom.PE <- Dominios[,1]*estima
E.STSI(Estrato, Nh, nh, Dominios)
```

```
## , , N
##
##                      Rural Urban Population
## Estimation      19174 22116      41290
## Standard Error   0     0        0
## CVE            0     0        0
## DEFF           NaN   NaN      NaN
##
## , , NotPoor
##
##                      Rural      Urban Population
## Estimation      12796.426265 15693.893557 2.849032e+04
## Standard Error  289.276849   299.387309 4.163098e+02
## CVE            2.260607    1.907668 1.461232e+00
## DEFF           1.000000    1.000000 9.984088e-01
##
## , , Extreme
##
##                      Rural      Urban Population
## Estimation      1651.15178 1053.14286 2704.2946332
## Standard Error  172.24135 140.45318 222.2480131
```

```

## CVE           10.43159   13.33657   8.2183358
## DEFF          1.00000   1.00000   0.9943541
##
## , , Relative
##
##                   Rural      Urban   Population
## Estimation     4726.421959 5368.963585 10095.385545
## Standard Error 264.609453 282.776619  387.273519
## CVE            5.598515   5.266875   3.836144
## DEFF           1.000000   1.000000   1.000479

```

Estimación del tamaño para No pobre

```
E.STSI(Estrato, Nh, nh, Dominios)[, , "NotPoor"]
```

```

##                   Rural      Urban   Population
## Estimation     12796.426265 15693.893557 2.849032e+04
## Standard Error 289.276849   299.387309 4.163098e+02
## CVE            2.260607    1.907668 1.461232e+00
## DEFF           1.000000   1.000000 9.984088e-01

```

Estimación del tamaño para Pobreza relativa

```
E.STSI(Estrato, Nh, nh, Dominios)[, , "Relative"]
```

```

##                   Rural      Urban   Population
## Estimation     4726.421959 5368.963585 10095.385545
## Standard Error 264.609453 282.776619  387.273519
## CVE            5.598515   5.266875   3.836144
## DEFF           1.000000   1.000000   1.000479

```

Estimación del tamaño para Pobreza extrema

```
E.STSI(Estrato, Nh, nh, Dominios)[, , "Extreme"]
```

```

##                   Rural      Urban   Population
## Estimation     1651.15178 1053.14286 2704.2946332
## Standard Error 172.24135  140.45318 222.2480131
## CVE            10.43159   13.33657   8.2183358
## DEFF           1.000000   1.000000   0.9943541

```

Estimación del total para ingreso en Pobreza relativa

```
E.STSI(Estrato, Nh, nh, dom.PR)[, , "Ingreso"]
```

```

##                   Rural      Urban   Population
## Estimation     1.142081e+06 899680.73849 2.041762e+06
## Standard Error 1.686436e+05 197642.19537 2.598136e+05
## CVE            1.476634e+01   21.96804 1.272497e+01
## DEFF           1.000000e+00   1.000000 9.994727e-01

```

Estimación del total para gasto en Pobreza relativa

```
E.STSI(Estrato, Nh, nh, dom.PR)[, , "Gasto"]
```

```

##                   Rural      Urban   Population
## Estimation     1.508015e+06 1.085136e+06 2.593151e+06
## Standard Error 2.230503e+05 2.390256e+05 3.269322e+05
## CVE            1.479099e+01 2.202726e+01 1.260753e+01
## DEFF           1.000000e+00 1.000000e+00 9.988477e-01

```

Estimación del total para ingreso en No pobre

```
E.STSI(Estrato, Nh, nh, dom.NP) [, , "Ingreso"]
```

```
##                               Rural      Urban   Population
## Estimation      5.359605e+06 7.914616e+06 1.327422e+07
## Standard Error 4.538210e+05 6.063085e+05 7.573398e+05
## CVE            8.467433e+00 7.660619e+00 5.705342e+00
## DEFF           1.000000e+00 1.000000e+00 9.984368e-01
```

Estimación del total para gasto en No pobre

```
E.STSI(Estrato, Nh, nh, dom.NP) [, , "Gasto"]
```

```
##                               Rural      Urban   Population
## Estimation      4.816501e+06 7.246498e+06 1.206300e+07
## Standard Error 3.774265e+05 5.627838e+05 6.776255e+05
## CVE            7.836114e+00 7.766287e+00 5.617389e+00
## DEFF           1.000000e+00 1.000000e+00 9.980720e-01
```

Estimación del total para ingreso en Pobreza extrema

```
E.STSI(Estrato, Nh, nh, dom.PE) [, , "Ingreso"]
```

```
##                               Rural      Urban   Population
## Estimation      2.670618e+07 4.362736e+07 7.033354e+07
## Standard Error 1.288502e+06 1.690202e+06 2.125329e+06
## CVE            4.824735e+00 3.874179e+00 3.021785e+00
## DEFF           1.000000e+00 1.000000e+00 9.857678e-01
```

Estimación del total para gasto en Pobreza extrema

```
E.STSI(Estrato, Nh, nh, dom.PE) [, , "Gasto"]
```

```
##                               Rural      Urban   Population
## Estimation      1.544992e+07 2.590536e+07 4.135528e+07
## Standard Error 6.985111e+05 1.040888e+06 1.253541e+06
## CVE            4.521131e+00 4.018042e+00 3.031152e+00
## DEFF           1.000000e+00 1.000000e+00 9.837391e-01
```

## Muestreo estratificado proporcional al tamaño

obtención del tamaño de muestra

```
attach(Hogares)
tamano.muestra = Hogares %>%
  group_by(Estrato) %>%
  summarise(Nh = n(),
            mh = round(2000*Nh/dim(Hogares)[1]))
sum(tamano.muestra$Nh)

## [1] 41290
sum(tamano.muestra$mh)

## [1] 2000
```

Obtención muestra aleatoria

```

Nh <- tamano.muestra$Nh
mh <- tamano.muestra$mh
res <- S.STPPS(Estrato, Personas, mh)
sam <- res[,1]
pk <- res[,2]
muestra <- Hogares[sam,]

attach(muestra)
head(muestra)

## # A tibble: 6 x 6
##   HHID      Estrato Personas Ingreso Gasto Pobreza
##   <chr>     <chr>    <int>   <dbl> <dbl> <fct>
## 1 idHH36128 Rural        3     325.  498. Relative
## 2 idHH03558 Rural       11    8974. 3553. NotPoor
## 3 idHH15434 Rural        4    2192. 2002. NotPoor
## 4 idHH10768 Rural        3     504   568. Relative
## 5 idHH31740 Rural        4    4265. 1610 NotPoor
## 6 idHH01516 Rural        7    7998. 3387. NotPoor

```

Estimación del tamaño poblacional y total de las variables.

```

estima <- data.frame(Ingresa, Gasto)
E.STPPS(estima, pk, mh, Estrato)

```

```

## , , N
##
##           Rural      Urban Population
## Estimation 19219.460073 22186.232034 41405.692107
## Standard Error 383.886421 413.244913 564.039132
## CVE          1.997384 1.862619 1.362226
## DEFF          Inf        Inf        Inf
##
## , , Ingresa
##
##           Rural      Urban Population
## Estimation 3.051280e+07 5.447628e+07 8.498908e+07
## Standard Error 6.746140e+05 1.260649e+06 1.429804e+06
## CVE          2.210921e+00 2.314125e+00 1.682339e+00
## DEFF          2.209439e-01 1.901460e-01 1.906962e-01
##
## , , Gasto
##
##           Rural      Urban Population
## Estimation 2.033579e+07 3.501759e+07 5.535338e+07
## Standard Error 3.896711e+05 7.543089e+05 8.490145e+05
## CVE          1.916184e+00 2.154086e+00 1.533808e+00
## DEFF          2.288913e-01 2.028287e-01 2.016458e-01

```

Estimación del tamaño poblacional

```

E.STPPS(estima, pk, mh, Estrato)[, , "N"]

```

```

##           Rural      Urban Population
## Estimation 19219.460073 22186.232034 41405.692107
## Standard Error 383.886421 413.244913 564.039132

```

```
## CVE           1.997384    1.862619    1.362226  
## DEFF          Inf         Inf         Inf
```

Estimación del total para Ingreso

```
E.STPPS(estima, pk, mh, Estrato)[, , "Ingreso"]
```

```
##                   Rural      Urban   Population  
## Estimation     3.051280e+07 5.447628e+07 8.498908e+07  
## Standard Error 6.746140e+05 1.260649e+06 1.429804e+06  
## CVE            2.210921e+00 2.314125e+00 1.682339e+00  
## DEFF           2.209439e-01 1.901460e-01 1.906962e-01
```

Estimación del total para Gasto

```
E.STPPS(estima, pk, mh, Estrato)[, , "Gasto"]
```

```
##                   Rural      Urban   Population  
## Estimation     2.033579e+07 3.501759e+07 5.535338e+07  
## Standard Error 3.896711e+05 7.543089e+05 8.490145e+05  
## CVE            1.916184e+00 2.154086e+00 1.533808e+00  
## DEFF           2.288913e-01 2.028287e-01 2.016458e-01
```

Estimación para el tamaño poblacional y el total de las variables por dominio.

```
Dominios <- Domains(Pobreza)  
dom.NP <- Dominios[,3] * estima  
dom.PR <- Dominios[,2] * estima  
dom.PE <- Dominios[,1] * estima
```

Estimación de los tamaños en los dominios

```
E.STPPS(Dominios, pk, mh, Estrato)[, 3 , ]
```

```
##                   N    NotPoor    Extreme  Relative  
## Estimation     41405.692107 28698.714970 3015.0140254 9691.963112  
## Standard Error 564.039132   698.846162   240.2160617 407.612117  
## CVE            1.362226    2.435113    7.9673282  4.205671  
## DEFF           Inf         2.575652    0.8546585  1.018686
```

Estimación del tamaño para No pobre

```
E.STPPS(Dominios, pk, mh, Estrato)[, , "NotPoor"]
```

```
##                   Rural      Urban   Population  
## Estimation     13127.396995 15571.317975 28698.714970  
## Standard Error 476.795791   510.932218   698.846162  
## CVE            3.632067    3.281239    2.435113  
## DEFF           2.552957    2.595510    2.575652
```

Estimación del tamaño Pobreza relativa

```
E.STPPS(Dominios, pk, mh, Estrato)[, , "Relative"]
```

```
##                   Rural      Urban   Population  
## Estimation     4460.9325187 5231.030593 9691.963112  
## Standard Error 273.5966435 302.146512 407.612117  
## CVE            6.1331715   5.776042   4.205671  
## DEFF           0.9898744   1.042618   1.018686
```

Estimacion dominios

```
E.STPPS(dom.PR, pk, mh, Estrato)
```

```
## , , N
##
##          Rural      Urban Population
## Estimation 19219.460073 22186.232034 41405.692107
## Standard Error 383.886421 413.244913 564.039132
## CVE         1.997384 1.862619 1.362226
## DEFF        Inf       Inf       Inf
##
## , , Ingreso
##
##          Rural      Urban Population
## Estimation 9.464098e+05 1.385457e+06 2.331867e+06
## Standard Error 1.035028e+05 1.609978e+05 1.913978e+05
## CVE         1.093636e+01 1.162055e+01 8.207919e+00
## DEFF        2.606988e-01 1.898170e-01 2.061668e-01
##
## , , Gasto
##
##          Rural      Urban Population
## Estimation 1.323364e+06 1.795269e+06 3.118633e+06
## Standard Error 1.447001e+05 2.146213e+05 2.588444e+05
## CVE         1.093427e+01 1.195482e+01 8.299931e+00
## DEFF        2.918707e-01 2.194131e-01 2.379510e-01
```

Estimación del total para Ingreso en Pobreza relativa

```
E.STPPS(dom.PR, pk, mh, Estrato)[, , "Ingreso"]
```

```
##          Rural      Urban Population
## Estimation 9.464098e+05 1.385457e+06 2.331867e+06
## Standard Error 1.035028e+05 1.609978e+05 1.913978e+05
## CVE         1.093636e+01 1.162055e+01 8.207919e+00
## DEFF        2.606988e-01 1.898170e-01 2.061668e-01
```

Estimación del total para Gasto en Pobreza relativa

```
E.STPPS(dom.PR, pk, mh, Estrato)[, , "Gasto"]
```

```
##          Rural      Urban Population
## Estimation 1.323364e+06 1.795269e+06 3.118633e+06
## Standard Error 1.447001e+05 2.146213e+05 2.588444e+05
## CVE         1.093427e+01 1.195482e+01 8.299931e+00
## DEFF        2.918707e-01 2.194131e-01 2.379510e-01
```

Estimación del total para Ingreso en No pobre

```
E.STPPS(dom.NP, pk, mh, Estrato)[, , "Ingreso"]
```

```
##          Rural      Urban Population
## Estimation 5.361610e+06 8.797469e+06 1.415908e+07
## Standard Error 3.207666e+05 5.131874e+05 6.051880e+05
## CVE         5.982654e+00 5.833353e+00 4.274204e+00
## DEFF        2.445110e-01 1.048347e-01 1.245015e-01
```

Estimación del total para Gasto en Pobreza extrema

```
E.STPPS(dom.PE, pk, mh, Estrato)[, , "Gasto"]  
  
## Rural Urban Population  
## Estimation 1.421483e+07 2.604687e+07 4.026169e+07  
## Standard Error 4.942848e+05 8.975466e+05 1.024650e+06  
## CVE 3.477248e+00 3.445891e+00 2.544975e+00  
## DEFF 3.479625e-01 3.254445e-01 3.244420e-01
```