

# Complex Surveys

Ignacio Sepulveda

2020

```
## Warning: package 'survey' was built under R version 4.0.3

## Loading required package: grid

## Loading required package: Matrix

## Loading required package: survival

##
## Attaching package: 'survey'

## The following object is masked from 'package:graphics':
##
##      dotchart
```

## Chapter II: SIMPLE AND STRATIFIED SAMPLING.

### ANALYZING SIMPLE RANDOM SAMPLES.

Datos y primeras estimaciones.

```
data(api)
srs_design = svydesign(id=~1,fpc=~fpc,data=apisrs)
srs_design
```

```
## Independent Sampling design
## svydesign(id = ~1, fpc = ~fpc, data = apisrs)
```

```
svytotal(~enroll,srs_design)
```

```
##           total      SE
## enroll 3621074 169520
```

```
svymeans(~enroll,srs_design)
```

```
##           mean      SE
## enroll 584.61 27.368
```

Pero y los pesos?

```
nofpc=svydesign(id=~1,weights=~pw,data=apisrs)
nofpc
```

```
## Independent Sampling design (with replacement)
## svydesign(id = ~1, weights = ~pw, data = apisrs)
```

```
svytotal(~enroll,nofpc)
```

```
##          total      SE
## enroll 3621074 172325
```

```
svymean(~enroll,nofpc)
```

```
##          mean      SE
## enroll 584.61 27.821
```

Vemos que no hay cambios ni en la media ni en el total pero si los hay en la desviacion estandar  
Pueden ser factores igual,

```
svytotal(~stype,srs_design)
```

```
##          total      SE
## stypeE 4397.74 196.00
## stypeH  774.25 142.85
## stypeM 1022.01 160.33
```

Y hacer contrastes,

```
means=svymean(~api00+api99,srs_design)
means
```

```
##          mean      SE
## api00 656.58 9.2497
## api99 624.68 9.5003
```

```
svycontrast(means,c(api00=1,api99=-1))
```

```
##          contrast      SE
## contrast      31.9 2.0905
```

```
#update para agregar variables.
srs_design=update(srs_design, apidiff=api00-api99)
srs_design=update(srs_design, apipct=apidiff/api99)
svymean(~apidiff+apipct,srs_design)
```

```
##          mean      SE
## apidiff 31.900000 2.0905
## apipct  0.056087 0.0041
```

## STRATIFIED SAMPLING

```
strat_design = svydesign(id=~1,strata=~stype,fpc=~fpc,data=apistrat)
strat_design
```

```
## Stratified Independent Sampling design
## svydesign(id = ~1, strata = ~stype, fpc = ~fpc, data = apistrat)
```

```
svytotal(~enroll,strat_design)
```

```
##          total      SE
## enroll 3687178 114642
```

```
svymean(~enroll,strat_design)
```

```
##          mean      SE
## enroll 595.28 18.509
```

```
svytotal(~stype,strat_design)
```

```
##          total SE
## stypeE  4421  0
## stypeH   755  0
## stypeM  1018  0
```

## REPLICATE WEIGHTS

```
chis_adult=read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/Adult.dta")
```

```
## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('sragef') for 'srage_p' are missing
```

```
## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('povgwd') for 'povgwd_p' are missing
```

```
## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw0' are missing
```

```
## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw1' are missing
```

```
## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw2' are missing
```

```
## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw3' are missing
```

```

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw4' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw5' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw6' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw7' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw8' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw9' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw10' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw11' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw12' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw13' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw14' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw15' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw16' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw17' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw18' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw19' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw20' are missing

```

```
## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw37' are missing
```





```
## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw72' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw73' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw74' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw75' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw76' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw77' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw78' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw79' are missing

## Warning in read.dta("C:/Users/nasho/Documents/GitHub/Notes_COMPLEX_SURVEYS/
## Adult.dta"): value labels ('rakedf') for 'rakedw80' are missing
```

```
#especificando los weights
chis = svrepdesign(variables=chis_adult[,1:418],
                  repweights=chis_adult[,420:499],
                  weights=chis_adult[,419],
                  combined.weights=TRUE,
                  type='other',scale=1,rscales=1)
chis
```

```
## Call: svrepdesign.default(variables = chis_adult[, 1:418], repweights = chis_adult[,
##      420:499], weights = chis_adult[, 419], combined.weights = TRUE,
##      type = "other", scale = 1, rscales = 1)
## with 80 replicates.
```

```
#Creando nuestros weights

boot_design = as.svrepdesign(strat_design,type='bootstrap',replicates=100)
jk_design=as.svrepdesign(strat_design)
svymean(~enroll,boot_design)
```

```
##           mean      SE
## enroll 595.28 21.731
```



```
svymean(~enroll,jk_design)
```

```
##           mean      SE  
## enroll 595.28 18.509
```