

Análisis de encuestas de hogares con R

Módulo 1: Análisis de variables continuas

Andrés Gutiérrez, Ph.D.
Stalyn Guerrero, M.Sc.

CEPAL - Unidad de Estadísticas Sociales

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

Lectura de la base

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
encuesta <- readRDS("../Data/encuesta_pba.rds")
```

Definir diseño de la muestra con srvyr

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
options(survey.lonely.psu="adjust")  
library(srvyr)  
  
diseno <- encuesta %>%  
  as_survey_design(  
    strata = Stratum,  
    ids = PSU,  
    weights = wk,  
    nest = T  
  )
```

Histograma ponderado para la variable ingreso

Análisis de
encuestas de
hogares con R

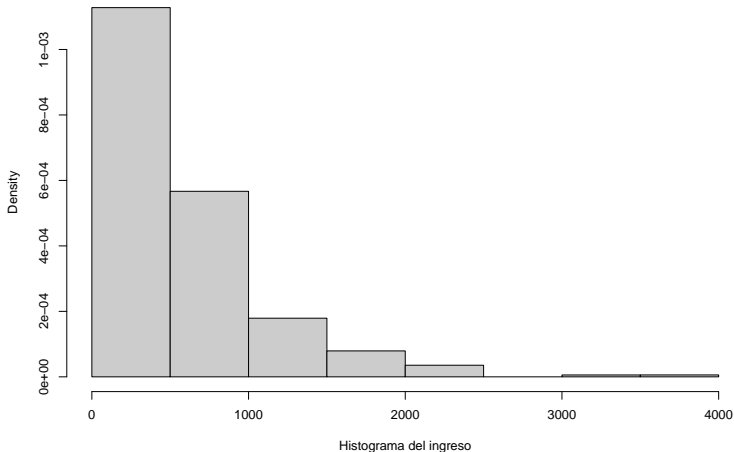
Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svyhist(  
  ~ Income ,  
  diseno,  
  main = "",  
  col = "grey80",  
  xlab = "Histograma del ingreso"  
)
```

Histograma ponderado para la variable ingreso

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.



Comparación de histogramas

Análisis de
encuestas de
hogares con R

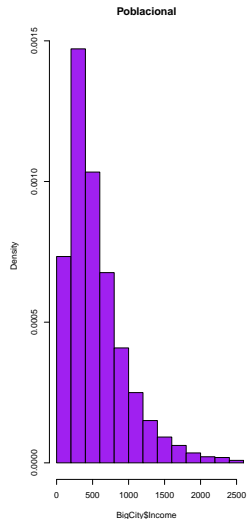
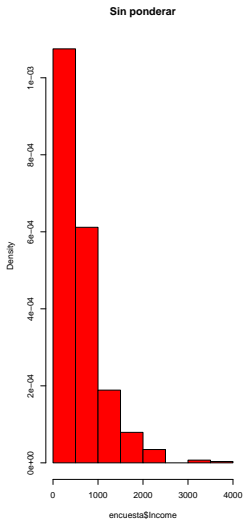
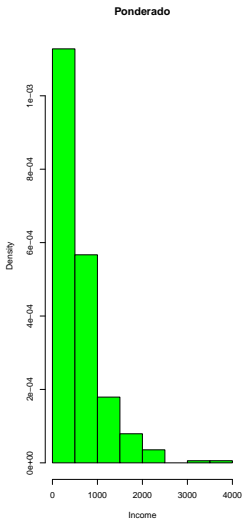
Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
data("BigCity", package = "TeachingSampling")
par(mfrow = c(1,3))
svyhist(      ~ Income,
  diseneno, main = "Ponderado",
  col = "green"
)
hist( encuesta$Income,
  main = "Sin ponderar",
  col = "red", prob = TRUE
)
hist( BigCity$Income,
  main = "Poblacional",
  col = "purple", prob = TRUE,
  xlim = c(0, 2500), breaks = 200
)
```

Comparación de histogramas

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.



Sub-grupos

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

Extraer sub-grupos de la encuesta.

```
sub_Urbano <- diseno %>% filter(Zone == "Urban")
sub_Rural  <- diseno %>% filter(Zone == "Rural")
sub_Mujer  <- diseno %>% filter(Sex == "Female")
sub_Hombre <- diseno %>% filter(Sex == "Male")
```

Histograma ponderado en sub-grupos

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

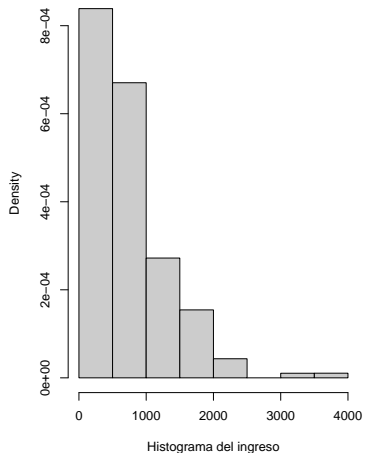
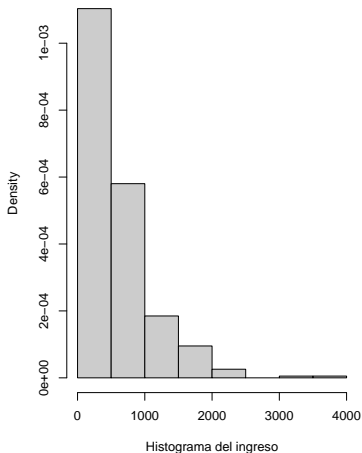
```
par(mfrow = c(1,2))
svyhist(
  ~ Income ,
  subset (sub_Mujer, Age >= 18),
  main = "",
  col = "grey80",
  xlab = "Histograma del ingreso"
)

svyhist(
  ~ Income ,
  subset (sub_Urbano, Age >= 18),
  main = "",
  col = "grey80",
  xlab = "Histograma del ingreso"
)
```

Histograma ponderado en sub-grupos

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.



Boxplot ponderado del ingreso por sub-grupos

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

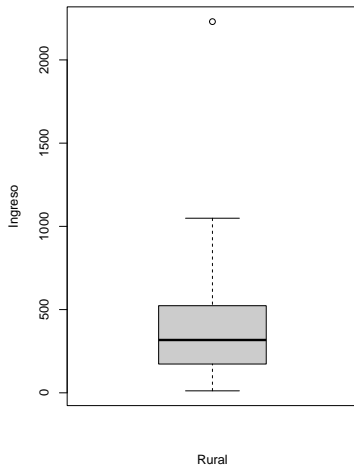
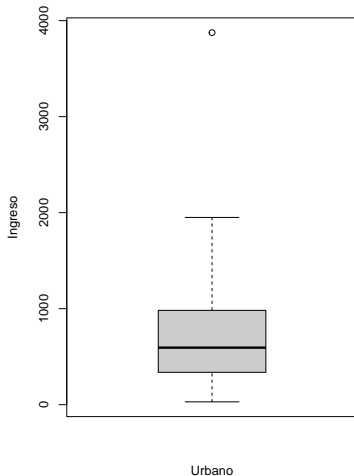
```
par(mfrow = c(1,2))  
svyboxplot(  
  Income ~1 ,  
  sub_Urbano,  
  col = "grey80",  
  ylab = "Ingreso",  
  xlab = "Urbano")
```

```
svyboxplot(  
  Income ~ 1 ,  
  sub_Rural,  
  col = "grey80",  
  ylab = "Ingreso",  
  xlab = "Rural"  
)
```

Boxplot ponderado del ingreso por sub-grupos

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.



Boxplot ponderado del ingreso por sub-grupos alternativa

Análisis de
encuestas de
hogares con R

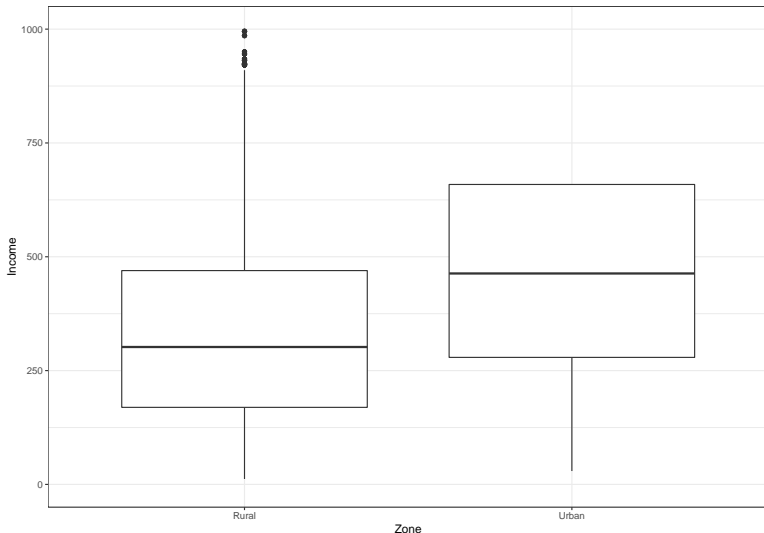
Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
library(ggplot2)
ggplot(data = encuesta,
       aes(y = Income, x = Zone,
           weight = wk)) +
  geom_boxplot() +
  ylim(c(0, 1000)) +
  theme_bw()
```

Boxplot ponderado del ingreso por sub-grupos alternativa

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.



Estimación de totales e intervalos de confianza del ingreso

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svytotal (~Income, diseno, deff=T) %>%  
  data.frame()
```

	total	Income	deff
Income	88683749	4892945	14.2

```
confint(svytotal (~Income, diseno, deff=T))
```

	2.5 %	97.5 %
Income	79093753	98273745

Estimación de totales e intervalos de confianza del gasto

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svytotal (~Expenditure, diseno, deff=T) %>%  
  data.frame()
```

	total	Expenditure	deff
Expenditure	52315200	2252648	14.22

```
confint(svytotal (~Expenditure, diseno, deff=T))
```

	2.5 %	97.5 %
Expenditure	47900091	56730309

Estimación de totales por sub-grupos

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
diseno %>% group_by(Sex) %>%  
  cascade(  
    Total = survey_total(Income, level = 0.95,  
                          vartype = c("se", "ci") ),  
    .fill = "Total ingreso")
```

Sex	Total	Total_se	Total_low	Total_upp
Female	46918242	2700595	41570793	52265690
Male	41765508	2484707	36845540	46685476
Total ingreso	88683749	4892945	78995230	98372269

Estimación de la media e intervalo de confianza del ingreso

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svymean(~Income, diseno, deff=T) %>%  
  data.frame()
```

	mean	Income	deff
Income	590.2	26.32	9.276

```
confint(svymean (~Income, diseno, deff=T))
```

	2.5 %	97.5 %
Income	538.6	641.8

Estimación de la media e intervalo de confianza del gasto

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svymean (~Expenditure, diseno, deff=T) %>%  
  data.frame()
```

	mean	Expenditure	deff
Expenditure	348.2	10.87	7.473

```
confint(svymean (~Expenditure, diseno, deff=T))
```

	2.5 %	97.5 %
Expenditure	326.9	369.4

Estimación de la media por sub-grupos

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
diseño %>% group_by(Sex) %>%  
  cascade(  
    Media = survey_mean(  
      Expenditure, level = 0.95,  
      vartype = c("se", "ci")),  
    .fill = "El gasto medio" ) %>%  
  arrange(desc(Sex))
```

Sex	Media	Media_se	Media_low	Media_upp
Male	343.2	12.15	319.1	367.2
Female	352.6	10.96	330.9	374.3
El gasto medio	348.2	10.87	326.6	369.7

Estimación de la media por sub-grupos

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
diseño %>% group_by(Zone) %>%  
  cascade(  
    Media = survey_mean(  
      Expenditure, level = 0.95,  
      vartype = c("se", "ci")),  
    .fill = "El gasto medio")%>%  
  arrange(desc(Zone))
```

Zone	Media	Media_se	Media_low	Media_upp
Urban	441.4	17.014	407.7	475.1
Rural	247.1	9.677	227.9	266.2
El gasto medio	348.2	10.866	326.6	369.7

Estimación de medias por sub-grupos

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
diseno %>% group_by(Zone, Sex) %>%  
  cascade(  
    Media = survey_mean(  
      Expenditure, level = 0.95,  
      vartype = c("se", "ci")),  
    .fill = "El gasto medio") %>%  
  arrange(desc(Zone), desc(Sex)) %>%  
  data.frame()
```

Zone	Sex	Media	Media_se	Media_low	Media_upp
Urban	Male	442.4	19.675	403.4	481.3
Urban	Female	440.6	16.750	407.4	473.7
Urban	El gasto medio	441.4	17.014	407.7	475.1
Rural	Male	240.2	9.966	220.4	259.9
Rural	Female	253.5	10.171	233.4	273.7
Rural	El gasto medio	247.1	9.677	227.9	266.2
El gasto medio	El gasto medio	348.2	10.866	326.6	369.7

Estimación de la desviación estándar de los ingresos por sub-grupo

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
(tab_sd <- diseno %>% group_by(Zone) %>%  
  summarise(Sd = sqrt(  
    survey_var(  
      Income,  
      level = 0.95,  
      vartype = c("se", "ci"),  
    )  
  )))
```

Zone	Sd	Sd_se	Sd_low	Sd_upp
Rural	372.1	215.2	216.2	479.7
Urban	570.8	248.1	451.6	669.1

Estimación de la desviación estándar de los ingresos por sub-grupo

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
(tab_sd <- diseno %>% group_by(Zone, Sex) %>%  
  summarise(Sd = sqrt(  
    survey_var(  
      Income,  
      level = 0.95,  
      vartype = c("se", "ci"),  
    )  
  ))) %>% data.frame()
```

Zone	Sex	Sd	Sd_se	Sd_low	Sd_upp
Rural	Female	375.6	222.8	206.8	489.3
Rural	Male	368.4	207.2	225.3	469.8
Urban	Female	551.8	241.8	434.5	648.3
Urban	Male	592.3	264.7	460.5	699.7

Estimación de la mediana para gastos

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
diseño %>% summarise(Mediana =  
  survey_median(  
    Expenditure,  
    level = 0.95,  
    vartype = c("se", "ci"),  
  ))
```

Mediana	Mediana_se	Mediana_low	Mediana_upp
289.3	10.9	264.7	307.9

Estimación de la mediana por sub-grupo

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
diseño %>% group_by(Zone) %>%  
  summarise(Mediana =  
    survey_median(  
      Expenditure,  
      level = 0.95,  
      vartype = c("se", "ci"),  
    ))
```

Zone	Mediana	Mediana_se	Mediana_low	Mediana_upp
Rural	203.8	16.55	179.3	245.7
Urban	385.7	20.76	348.6	431.6

Estimación de la mediana por sub-grupo

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
disenio %>% group_by(Sex) %>%  
  summarise(Mediana =  
    survey_median(  
      Expenditure,  
      level = 0.95,  
      vartype = c("se", "ci"),  
    ))
```

Sex	Mediana	Mediana_se	Mediana_low	Mediana_upp
Female	298.1	13.43	271.7	324.9
Male	274.3	12.37	252.0	301.0

Estimación de quantile para el gasto

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
disenio %>%  
  summarise(  
    Q = survey_quantile(  
      Expenditure,  
      quantiles = 0.5,  
      level = 0.95,  
      vartype = c("se", "ci"),  
      interval_type = "score"  
    ))
```

Q_q50	Q_q50_se	Q_q50_low	Q_q50_upp
289.3	14.69	240.3	298.5

Estimación de quantile para el gasto por sub-grupo

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
disenio %>% group_by(Sex) %>%  
  summarise(  
    Q = survey_quantile(  
      Expenditure,  
      quantiles = 0.25,  
      level = 0.95,  
      vartype = c("se", "ci"),  
      interval_type = "score"  
    )  
  )
```

Sex	Q_q25	Q_q25_se	Q_q25_low	Q_q25_upp
Female	181	8.545	146.7	180.5
Male	166	7.517	145.7	175.4

Estimación de quantile para el gasto por sub-grupo

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
disenio %>% group_by(Zone) %>%  
  summarise(  
    Q = survey_quantile(  
      Expenditure,  
      quantiles = 0.25,  
      level = 0.95,  
      vartype = c("se", "ci"),  
      interval_type = "score"  
    )  
  )
```

Zone	Q_q25	Q_q25_se	Q_q25_low	Q_q25_upp
Rural	140.8	7.118	113.8	142.3
Urban	238.7	13.896	222.6	278.2

Estimación de la razón entre el gasto y el ingreso

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
diseño %>% summarise(  
  Razon = survey_ratio(  
    numerator = Expenditure,  
    denominator = Income,  
    level = 0.95,  
    vartype = c("se", "ci")  
  ))
```

Razon	Razon_se	Razon_low	Razon_upp
0.5899	0.0194	0.5515	0.6284

Estimación de la razón entre hombres y mujeres

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
disenio %>% summarise(  
  Razon = survey_ratio(  
    numerator = (Sex == "Female"),  
    denominator = (Sex == "Male"),  
    level = 0.95,  
    vartype = c("se", "ci")  
  ))
```

Razon	Razon_se	Razon_low	Razon_upp
1.114	0.0308	1.053	1.175

Estimación de la razón entre hombres y mujeres en la zona rural

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
sub_Rural %>% summarise(  
  Razon = survey_ratio(  
    numerator = (Sex == "Female"),  
    denominator = (Sex == "Male"),  
    level = 0.95,  
    vartype = c("se", "ci")  
  ))
```

Razon	Razon_se	Razon_low	Razon_upp
1.068	0.0449	0.978	1.158

Estimación de la razón del gastos y los ingreso entre las mujeres

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
sub_Mujer %>% summarise(  
  Razon = survey_ratio(  
    numerator = Expenditure,  
    denominator = Income,  
    level = 0.95,  
    vartype = c("se", "ci")  
  ))
```

Razon	Razon_se	Razon_low	Razon_upp
0.5952	0.0206	0.5543	0.636

Estimación de la razón del gasto y los ingresos entre los hombres

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
sub_Hombre %>% summarise(  
  Razon = survey_ratio(  
    numerator = Expenditure,  
    denominator = Income,  
    level = 0.95,  
    vartype = c("se", "ci")  
  ))
```

Razon	Razon_se	Razon_low	Razon_upp
0.584	0.0195	0.5453	0.6227

Estimación del índice de GINI

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

La estimación del índice de GINI se realiza haciendo uso de la librería convey

```
library(convey)
diseno_gini <- convey_prep(diseno)
svygini( ~Income, design = diseno_gini) %>%
  data.frame()
```

	gini	Income
Income	0.4254	0.0132

Estimación del índice de GINI

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svygini( ~Expenditure, design = diseno_gini) %>%  
  data.frame()
```

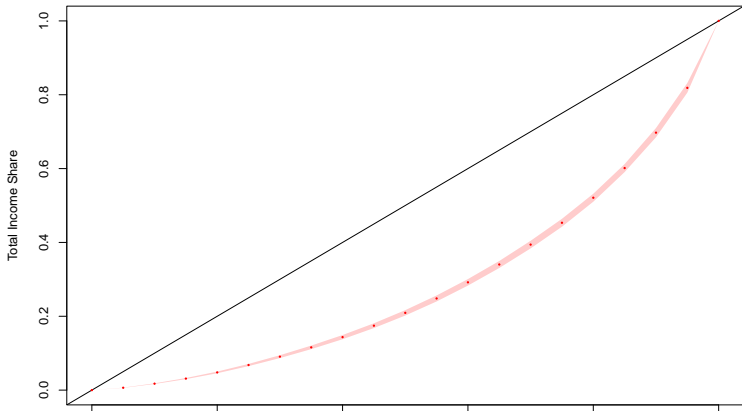
	gini	Expenditure
Expenditure	0.3538	0.0096

Estimación del índice de GINI por sub-grupo

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svylorenz( ~Income, diseno_gini,  
           seq(0,1,.05), alpha = .01 )
```



Pruebas de diferencia medias de los ingresos entre hombres y mujeres

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svytttest(Income ~ Sex, diseneno)
```

```
##
```

```
## Design-based t-test
```

```
##
```

```
## data: Income ~ Sex
```

```
## t = -0.28, df = 118, p-value = 0.8
```

```
## alternative hypothesis: true difference in mean is
```

```
## 95 percent confidence interval:
```

```
## -39.00 29.28
```

```
## sample estimates:
```

```
## difference in mean
```

```
## -4.859
```


Pruebas de diferencia medias de los ingresos entre hombres y mujeres en la zona urbana

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svyttest(Income ~ Sex, sub_Urbano)
```

```
##
```

```
## Design-based t-test
```

```
##
```

```
## data: Income ~ Sex
```

```
## t = 0.42, df = 63, p-value = 0.7
```

```
## alternative hypothesis: true difference in mean is
```

```
## 95 percent confidence interval:
```

```
## -44.12 67.76
```

```
## sample estimates:
```

```
## difference in mean
```

```
## 11.82
```

Pruebas de diferencia medias de los ingresos entre hombres y mujeres mayores a 18 años

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svyttest(Income ~ Sex, diseno %>% filter(Age > 18))
```

```
##
```

```
## Design-based t-test
```

```
##
```

```
## data: Income ~ Sex
```

```
## t = -0.41, df = 118, p-value = 0.7
```

```
## alternative hypothesis: true difference in mean is
```

```
## 95 percent confidence interval:
```

```
## -44.70 29.31
```

```
## sample estimates:
```

```
## difference in mean
```

```
## -7.694
```

Contrastes

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
(prom_region <- svyby(~Income, ~Region, diseno,  
                      svymean, na.rm=T, covmat = TRUE  
                      vartype = c("se", "ci")))
```

	Region	Income	se	ci_l	ci_u
Norte	Norte	487.3	59.48	370.7	603.8
Sur	Sur	656.3	47.51	563.2	749.5
Centro	Centro	671.3	37.56	597.6	744.9
Occidente	Occidente	531.5	57.13	419.5	643.5
Oriente	Oriente	605.8	67.84	472.8	738.8

```
# Paso 1: diferencia de estimaciones (Norte - Sur)  
461.9124 - 592.2575
```

```
## [1] -130.3
```

contrastes

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
# Paso 2: error estándar de la diferencia  
vcov(prom_region)
```

	Norte	Sur	Centro	Occidente	Oriente
Norte	3537	0	0	0	0
Sur	0	2257	0	0	0
Centro	0	0	1410	0	0
Occidente	0	0	0	3264	0
Oriente	0	0	0	0	4603

```
sqrt(759.8 + 3755 - 2*0)
```

```
## [1] 67.19
```

contrastes

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svycontrast(prom_region,  
             list(diff_NS = c(1, -1, 0, 0, 0))) %>%  
data.frame()
```

	contrast	diff_NS
diff_NS	-169.1	76.12

Contrastes

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svycontrast(prom_region, list(  
  Norte_sur = c(1, 1, 0, 0, 0),  
  centro = c(0, 0, 1, 0, 0),  
  Occidente_Oriente = c(0, 0, 0, 1, 1)  
)) %>% data.frame()
```

	contrast	SE
Norte_sur	1143.6	76.12
centro	671.3	37.56
Occidente_Oriente	1137.3	88.70

```
sqrt(759.8 + 3755 - 2*0) ; sqrt(2594);
```

```
## [1] 67.19
```

```
## [1] 50.93
```

```
sqrt(2591 + 1022 - 2*0)
```

```
## [1] 60.11
```

Contrastes no independiente

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
(prom_sexo <-  
  svyby(~Income, ~Sex, diseno,  
    svymean, na.rm=T, covmat = TRUE,  
    vartype = c("se", "ci")))
```

	Sex	Income	se	ci_l	ci_u
Female	Female	592.5	27.83	537.9	647.0
Male	Male	587.6	27.53	533.7	641.6

Contrastes no independiente

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svycontrast(prom_sexo,  
             list(diff_Sexo = c(1, -1))) %>%  
data.frame()
```

	contrast	diff_Sexo
diff_Sexo	4.859	17.24

Contrastes no independiente

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
vcov(prom_sexo)
```

	Female	Male
Female	774.7	617.8
Male	617.8	758.1

Note que el error estandar de la diferencia es igual a
`sqrt(512.3 + 591.6 - 2*308.8)`

```
## [1] 22.05
```

Contrastes no independiente

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
(sum_region <- svyby( ~ Income, ~ Region,  
                      diseno, svytotal, na.rm = T,  
                      covmat = TRUE,  
                      vartype = c("se", "ci")))
```

	Region	Income	se	ci_l	ci_u
Norte	Norte	12484837	1841598	8875370	16094104
Sur	Sur	18394966	1639406	15181790	21608742
Centro	Centro	16519417	1597354	13388660	19650174
Occidente	Occidente	16712469	2222473	12356502	21068436
Oriente	Oriente	24572061	3220380	18260232	30883890

Contrastes no independiente

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svycontrast(sum_region,  
             list(  
               Agregado_NCS = c(1, 1, 1, 0, 0)  
             )) %>% data.frame()
```

	contrast	Agregado_NCS
Agregado_NCS	47399219	2937801

Contrastes

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
require(kableExtra)
```

Note que el error estandar de la diferencia es igual

```
vcov(sum_region) %>% data.frame() %>%
```

```
  kable(digits = 10,
```

```
        format.args = list(scientific = FALSE))
```

	Norte	Sur	Centro	Occidente	Oriente
Norte	3391484124672	0	0	0	0
Sur	0	2687650650573	0	0	0
Centro	0	0	2551540395586	0	0
Occidente	0	0	0	4939385647690	0
Oriente	0	0	0	0	10370849850694

```
sqrt(2805154074898 + 3839259031856 + 7559032807016 )
```

```
## [1] 3768746
```

Contrastes no independiente

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
(prom_edad <- svyby(~Income, ~CatAge, diseno,  
                    svymean, na.rm=T, covmat = TRUE))
```

	CatAge	Income	se
0-5	0-5	570.6	38.15
6-15	6-15	593.5	42.88
16-30	16-30	592.2	29.35
31-45	31-45	616.6	35.41
46-60	46-60	595.3	36.16
Más de 60	Más de 60	555.3	49.50

Contrastes no independiente

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svycontrast(prom_edad,  
  list(  
    agregado_edad = c(1/6, 1/6, 1/6, 1/6, 1/6, 1/6)  
  )) %>% data.frame()
```

	contrast	agregado_edad
agregado_edad	587.3	27.08

Contrastes no independiente

`vcov(prom_edad)`

	0-5	6-15	16-30	31-45	46-60	Más de 60
0-5	1455.3	575.9	605.4	701.3	500.8	617.9
6-15	575.9	1838.6	618.1	602.0	768.7	637.0
16-30	605.4	618.1	861.3	305.8	727.1	320.7
31-45	701.3	602.0	305.8	1253.9	245.7	1162.9
46-60	500.8	768.7	727.1	245.7	1307.7	226.8
Más de 60	617.9	637.0	320.7	1162.9	226.8	2449.8

```
(1/6)*sqrt(4888 + 646 + 1611 + 2197 + 1181 + 5553
            + 2*22 + 2*(-1289) + 2*864 + 2*(-1387) + 2*189
            2*454 + 2*441 + 2*347 + 2*856 +
            2*290 + 2*820 + 2*1337+
            2*103 + 2*488 +
            2*268)
```

```
## [1] 25.65
```

Contrastes no independiente

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
(razon_sexo <- svyby(~Income, ~Sex,  
                    denominator = ~Expenditure,  
                    diseno, svyratio,  
                    na.rm=T, covmat = TRUE,  
                    vartype = c("se", "ci")))
```

	Sex	Income/Expenditure	se.Income/Expenditure	ci_l	ci_u
Female	Female	1.680	0.0583	1.566	1.79
Male	Male	1.712	0.0573	1.600	1.82

Contrastes no independiente

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
svycontrast(razon_sexo,  
             list(  
               diff_sexo = c(1, -1)  
             )) %>% data.frame()
```

	contrast	diff_sexo
diff_sexo	-0.032	0.0306

Contrastes no independiente

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

```
vcov(razon_sexo)
```

	Female	Male
Female	0.0034	0.0029
Male	0.0029	0.0033

```
sqrt(0.0013 + 0.0020 - 2*0.0011)
```

```
## [1] 0.03317
```

¡Gracias!

Análisis de
encuestas de
hogares con R

Andrés
Gutiérrez,
Ph.D.
Stalyn
Guerrero,
M.Sc.

Email: andres.gutierrez@cepal.org