

# 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.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

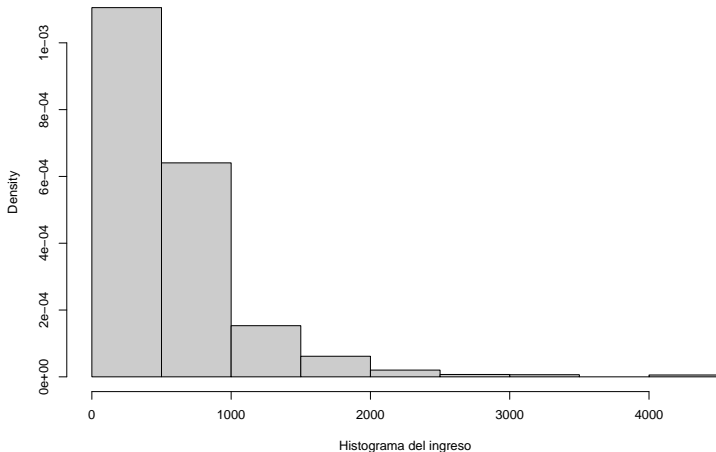
Andrés  
Gutiérrez,  
Ph.D.  
Stalyn  
Guerrero,  
M.Sc.

```
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

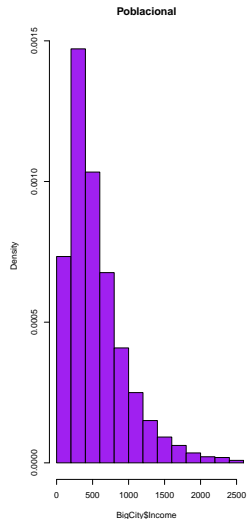
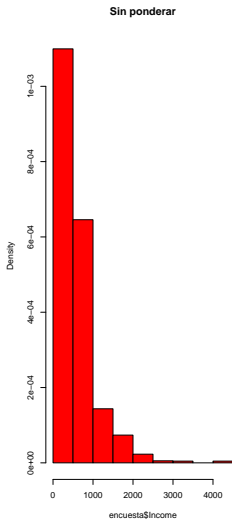
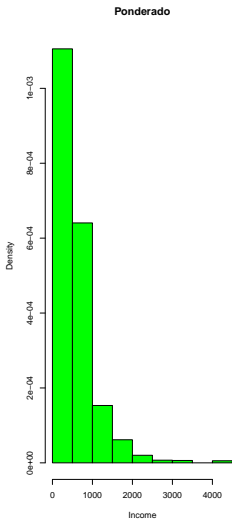
Andrés  
Gutiérrez,  
Ph.D.  
Stalyn  
Guerrero,  
M.Sc.

```
data("BigCity", package = "TeachingSampling")
par(mfrow = c(1,3))
svyhist(  ~ Income,
         diseno, 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.

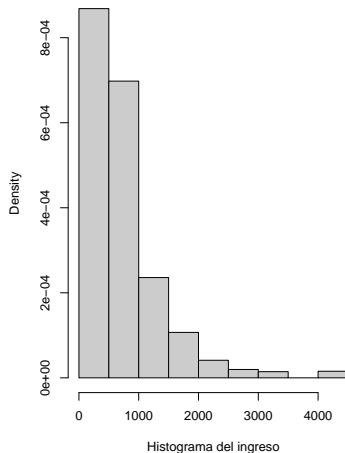
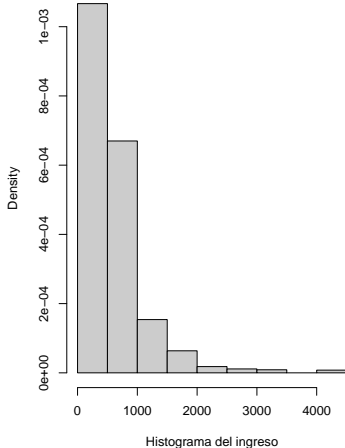
```
par(mfrow = c(1,2))
svyhist(
  ~ Income ,
  subset (sub_Mujer, Age >= 18),
  main = "",
  col = "grey80",
  xlab = "Histograma del ingreso"
)

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.

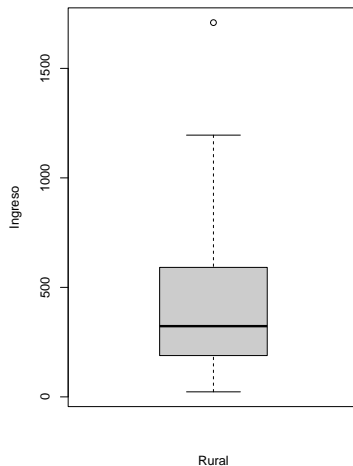
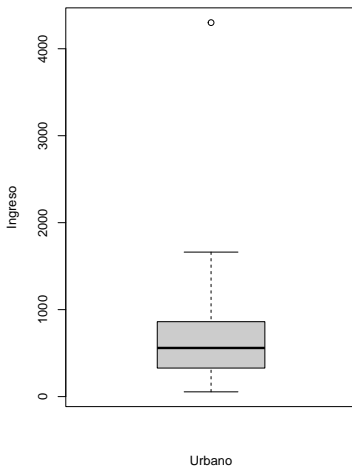
```
par(mfrow = c(1,2))  
svyboxplot(  
  Income ~1 ,  
  sub_Urbano,  
  col = "grey80",  
  ylab = "Ingreso",  
  xlab = "Urbano")
```

```
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

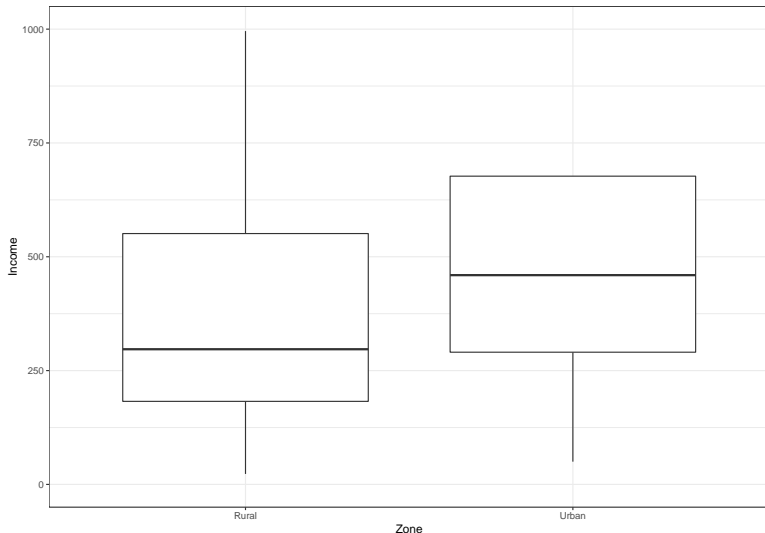
Andrés  
Gutiérrez,  
Ph.D.  
Stalyn  
Guerrero,  
M.Sc.

```
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	85793667	4778674	11

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

	2.5 %	97.5 %
Income	76427637	95159697



# 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	55677504	2604138	10.22

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

	2.5 %	97.5 %
Expenditure	50573486	60781522

# 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	44153820	2324452	39551172	48756467
Male	41639847	2870194	35956576	47323118
Total ingreso	85793667	4778674	76331414	95255920

# 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	570.9	28.48	8.821

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

	2.5 %	97.5 %
Income	515.1	626.8

# 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	370.5	13.29	6.016

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

	2.5 %	97.5 %
Expenditure	344.5	396.6

# 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	374.4	16.06	342.6	406.2
Female	367.0	12.34	342.6	391.5
El gasto medio	370.5	13.29	344.2	396.9

# 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	459.6	22.21	415.6	503.6
Rural	273.9	10.26	253.6	294.3
El gasto medio	370.5	13.29	344.2	396.9

# 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	469.8	26.96	416.4	523.2
Urban	Female	450.8	20.12	411.0	490.7
Urban	El gasto medio	459.6	22.21	415.6	503.6
Rural	Male	275.3	10.25	255.0	295.6
Rural	Female	272.7	11.61	249.7	295.7
Rural	El gasto medio	273.9	10.26	253.6	294.3
El gasto medio	El gasto medio	370.5	13.29	344.2	396.9

# 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	310.3	117.4	262.6	351.6
Urban	581.9	285.0	421.6	706.7



# 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, 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	294.9	111.6	249.6	334.1
Rural	Male	325.8	125.0	274.2	370.2
Urban	Female	568.4	286.5	400.7	696.8
Urban	Male	596.8	288.9	436.8	722.1

# 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
298.3	8.825	282.2	317.2

# Estimación de la mediana por sub-grupo

Análisis de  
encuestas de  
hogares con R

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

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

Zone	Mediana	Mediana_se	Mediana_low	Mediana_upp
Rural	240.7	11.00	214.2	258.3
Urban	380.7	19.84	337.1	416.3

# 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	299.9	10.499	282.2	323.8
Male	297.3	9.287	277.3	314.1

# 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
298.3	11.96	264.8	312.1

# 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	209.7	14.91	169.0	228.1
Male	192.5	10.41	163.5	204.7

# Estimación de quantile para el gasto por sub-grupo

Análisis de  
encuestas de  
hogares con R

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

```
diseño %>% 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	159.9	4.641	144.9	163.5
Urban	258.2	9.048	256.0	292.2

# 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.

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

Razon	Razon_se	Razon_low	Razon_upp
0.649	0.0232	0.6031	0.6949



# 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	0.0351	1.045	1.184

# 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	0.0352	0.9975	1.139

# 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
0.6583	0.0199	0.619	0.6976

# 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
0.6391	0.0288	0.5821	0.696

# 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	0.4133	0.0187

# 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()
```

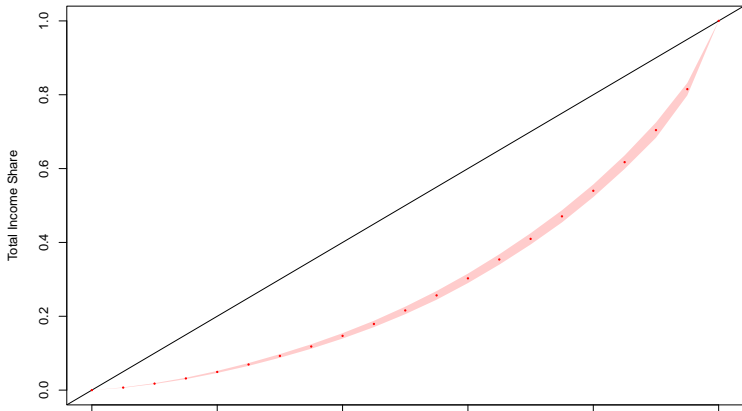
	gini	Expenditure
Expenditure	0.3509	0.0141

# 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.

```
svyttest(Income ~ Sex, disen0)
```

```
##
```

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

```
##
```

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

```
## t = 1.4, df = 118, p-value = 0.2
```

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

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

```
## -12.82 69.39
```

```
## sample estimates:
```

```
## difference in mean
```

```
## 28.28
```



# 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 = 1.6, df = 63, p-value = 0.1
```

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

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

```
## -12.32 101.74
```

```
## sample estimates:
```

```
## difference in mean
```

```
## 44.71
```

# 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 = 1.5, df = 118, p-value = 0.1
```

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

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

```
## -10.73 82.85
```

```
## sample estimates:
```

```
## difference in mean
```

```
## 36.06
```

# 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	552.4	55.36	443.9	660.9
Sur	Sur	625.8	62.41	503.5	748.1
Centro	Centro	650.8	61.47	530.3	771.3
Occidente	Occidente	517.0	46.22	426.4	607.6
Oriente	Oriente	541.8	71.66	401.3	682.2

```
# 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	3065	0	0	0	0
Sur	0	3894	0	0	0
Centro	0	0	3778	0	0
Occidente	0	0	0	2136	0
Oriente	0	0	0	0	5136

```
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	-73.41	83.42

# 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	1178.1	83.42
centro	650.8	61.47
Occidente_Oriente	1058.8	85.28

```
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	557.6	25.83	506.9	608.2
Male	Male	585.8	34.59	518.1	653.6

# 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	-28.28	20.76



# Contrastes no independiente

Análisis de  
encuestas de  
hogares con R

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

```
vcov(prom_sexo)
```

	Female	Male
Female	667.2	716.3
Male	716.3	1196.3

*# 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	14277323	1507575	11322530	17232100
Sur	Sur	16068151	1877989	12387359	19748543
Centro	Centro	16483319	2383556	11811634	21155004
Occidente	Occidente	16853540	1823807	13278944	20428136
Oriente	Oriente	22111335	2833460	16557856	27664814

# 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	46828792	3388357

# 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	2272782099289	0	0	0	0
Sur	0	3526843231468	0	0	0
Centro	0	0	5681340267222	0	0
Occidente	0	0	0	3326270307526	0
Oriente	0	0	0	0	8028493876790

```
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	463.8	28.87
6-15	6-15	511.6	34.88
16-30	16-30	607.3	37.42
31-45	31-45	573.4	26.95
46-60	46-60	763.1	58.97
Más de 60	Más de 60	466.6	31.21

# 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	564.3	25.4

# Contrastes no independiente

`vcov(prom_edad)`

	0-5	6-15	16-30	31-45	46-60	Más de 60
0-5	833.4	548.4	361.1	262.3	132.7	312.6
6-15	548.4	1216.6	739.7	528.1	565.5	120.1
16-30	361.1	739.7	1399.9	534.9	1564.6	412.5
31-45	262.3	528.1	534.9	726.2	642.3	161.5
46-60	132.7	565.5	1564.6	642.3	3477.7	416.6
Más de 60	312.6	120.1	412.5	161.5	416.6	973.9

```
(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	1.519	0.0458	1.429	1.60
Male	Male	1.565	0.0704	1.427	1.70



# 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	-0.0457	0.0416

# Contrastes no independiente

Análisis de  
encuestas de  
hogares con R

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

```
vcov(razon_sexo)
```

	Female	Male
Female	0.0021	0.0027
Male	0.0027	0.0050

```
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](mailto:andres.gutierrez@cepal.org)