

Análisis de encuestas de hogares con R

Módulo 2: Análisis de variables categóricas

CEPAL - Unidad de Estadísticas Sociales

- 1 Motivación
- 2 Lectura y procesamientos de encuestas con R
- 3 Estimación del tamaño.
- 4 Estimación de la proporción.
- 5 Tablas cruzadas.

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Motivación

Motivación

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Esta sección esta enfocada en los procedimientos para análisis simples univariados, bivariados y multivariados seleccionados para respuestas de encuestas categóricas, centrándose en la adaptación de técnicas de inferencia estadística a datos de encuestas de muestras complejas

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Lectura y procesamientos de encuestas con R

Lectura de la base

Iniciemos con la lectura de la encuesta.

```
encuesta <- readRDS("../Data/encuesta.rds") %>%  
  transmute(estrato = as.character(`_estrato`),  
            upm = as.character(`_upm`),  
            fep = `_fep`,  
            Zone = as_factor(areageo2),  
            Sex = as_factor(sexo),  
            Age = edad,  
            lp, li, # Linea de pobreza e indigencia  
            Income = ingcorte,  
            anoest, # años de estudio  
            empleo = conduct3)
```

El paso siguiente es realizar declaración del objeto tipo diseño.

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Definir diseño de la muestra con srvyr

Análisis de encuestas de hogares con R

Motivación

Lectura y procesamiento de encuestas con R

Estimación del tamaño.

Estimación de la proporción.

Tablas cruzadas.

Para el desarrollo de la presentación se define el diseño muestral con la función `as_survey_design`.

```
options(survey.lonely.psu = "adjust")
library(srvyr)
diseno <- encuesta %>% # Base de datos.
  as_survey_design(
    strata = estrato, # Id de los estratos.
    ids = upm, # Id para las observaciones.
    weights = fep, # Factores de expansión.
    nest = T # Valida el anidado dentro
              # del estrato
  )
```

Creación de nuevas variables

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Durante los análisis de encuesta surge la necesidad de crear nuevas variables a partir de las existentes, aquí mostramos la definición de algunas de ellas.

```
disenio <- disenio %>% mutate(  
  empleo = as_factor(empleo),  
  pobreza = ifelse(Income < lp, 1, 0),  
  desempleo = ifelse(empleo == "Desocupado", 1, 0)  
)
```


Dividiendo la muestra en Sub-grupos

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

En ocasiones se desea realizar estimaciones por sub-grupos de la población, en este caso se extraer 4 sub-grupos de la encuesta.

```
sub_Urbano <- diseno %>% filter(Zone == "Urbana")
sub_Rural  <- diseno %>% filter(Zone == "Rural")
sub_Mujer  <- diseno %>% filter(Sex == "Mujer")
sub_Hombre <- diseno %>% filter(Sex == "Hombre")
```

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

**Estimación del
tamaño.**

Estimación de
la proporción.

Tablas
cruzadas.

Estimación del tamaño.

Estimación del tamaño.

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

El primer parámetro estimado serán los tamaños de la población y subpoblaciones.

Estimación de tamaño

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
(tamano_zona <- diseno %>% group_by(Zone) %>%  
  summarise(  
    n = unweighted(n()), # Observaciones en la muestra  
    Nd = survey_total(vartype = c("se", "ci"))))
```

Zone	n	Nd	Nd_se	Nd_low	Nd_upp
Urbana	51737	8594098	79028	8439022	8749173
Rural	19641	1853727	28460	1797880	1909574

En la tabla n denota el número de observaciones en la muestra por Zona y Nd denota la estimación del total de observaciones en la población.

Estimación de tamaño

Análisis de encuestas de hogares con R

Motivación

Lectura y procesamiento de encuestas con R

Estimación del tamaño.

Estimación de la proporción.

Tablas cruzadas.

Empleando una sintaxis similar es posible estimar el número de personas en condición de pobreza extrema, pobreza y no pobres.

```
(tamano_pobreza <- diseno %>% group_by(pobreza) %>%  
  summarise(  
    Nd = survey_total(vartype = c("se","ci"))))
```

pobreza	Nd	Nd_se	Nd_low	Nd_upp
0	8166609	81442	8006798	8326420
1	2281215	71910	2140107	2422323

Estimación de tamaño

Otra variable de interés es conocer el estado de ocupación de la personas.

```
(tamano_ocupacion <- diseno %>%  
  group_by(empleo) %>%  
  summarise(  
    Nd = survey_total(vartype = c("se","ci"))))
```

empleo	Nd	Nd_se	Nd_low	Nd_upp
NA	1778302	38325.4	1703097	1853507
Ocupado	4397676	44381.1	4310588	4484764
Desocupado	274029	10832.9	252772	295286
Inactivo	3994586	53146.3	3890298	4098874
NS/NR	3231	864.5	1534	4927

Estimación de tamaño

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Utilizando la función `group_by` es posible obtener resultados por más de un nivel de agregación.

```
(tamano_ocupacion_pobreza <- diseno %>%  
  group_by(Sex, Zone) %>%  
  cascade(  
    Nd = survey_total(vartype = c("se","ci")),  
    .fill = "Total") %>%  
  data.frame()  
)
```

Estimación de tamaño

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Sex	Zone	Nd	Nd_se	Nd_low	Nd_upp
Hombre	Urbana	4142094	49550	4044862	4239325
Hombre	Rural	940580	18609	904064	977095
Hombre	Total	5082673	52929	4978811	5186535
Mujer	Urbana	4452004	50413	4353079	4550929
Mujer	Rural	913147	19457	874967	951328
Mujer	Total	5365151	54038	5259114	5471188
Total	Total	10447824	83997	10283000	10612649

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

**Estimación de
la proporción.**

Tablas
cruzadas.

Estimación de la proporción.

Estimación de la proporción.

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Otro resultado de interés es la estimación de las proporciones, dado que estas entregan una mayor información sobre las tendencias en la población, siendo de mucha importancia en la toma de decisiones.

Estimación de proporción de urbano y rural

Análisis de encuestas de hogares con R

Motivación

Lectura y procesamiento de encuestas con R

Estimación del tamaño.

Estimación de la proporción.

Tablas cruzadas.

El procedimiento estándar para el cálculo de proporciones es crear una *variable dummy* y sobre ella realizar las operaciones. Sin embargo, la librería *srvy* nos simplifica el cálculo, mediante la sintaxis.

```
(prop_zona <- diseno %>% group_by(Zone) %>%  
  summarise(  
    prop = survey_mean(vartype = c("se", "ci"),  
      proportion = TRUE )))
```

Zone	prop	prop_se	prop_low	prop_upp
Urbana	0.8226	0.0026	0.8174	0.8276
Rural	0.1774	0.0026	0.1724	0.1826

Note que, se utilizó la función `survey_mean` para la estimación.

Estimación de proporción de urbano y rural

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

La función idónea para realizar la estimación de las proporciones es `survey_prop` y la sintaxis es como sigue:

```
(prop_zona2 <- diseno %>% group_by(Zone) %>%  
  summarise(  
    prop = survey_prop(vartype = c("se", "ci") )))
```

Zone	prop	prop_se	prop_low	prop_upp
Urbana	0.8226	0.0026	0.8174	0.8277
Rural	0.1774	0.0026	0.1723	0.1826

Proporción de hombres y mujeres en la zona urbana

Análisis de encuestas de hogares con R

Si el interés es obtener la estimación para una subpoblación, procedemos así:

Motivación

Lectura y procesamiento de encuestas con R

Estimación del tamaño.

Estimación de la proporción.

Tablas cruzadas.

```
(prop_sexoU <- sub_Urbano %>% group_by(Sex) %>%  
  summarise(  
    prop = survey_prop(vartype = c("se","ci"))))
```

Sex	prop	prop_se	prop_low	prop_upp
Hombre	0.482	0.0036	0.475	0.489
Mujer	0.518	0.0036	0.511	0.525

Proporción de hombres y mujeres en la zona urbana

Análisis de encuestas de hogares con R

Motivación

Lectura y procesamiento de encuestas con R

Estimación del tamaño.

Estimación de la proporción.

Tablas cruzadas.

Si el interés es obtener la estimación para una subpoblación, procedemos así:

```
(prop_sexoU <- sub_Urbano %>% group_by(Sex) %>%  
  summarise(  
    prop = survey_prop(vartype = c("se","ci"))))
```

Sex	prop	prop_se	prop_low	prop_upp
Hombre	0.482	0.0036	0.475	0.489
Mujer	0.518	0.0036	0.511	0.525

Ejercicio

¿Cómo estimar el Proporción de hombres dado que están en zona rural?

Proporción de hombres y mujeres en la zona rural

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
(prop_sexor <- sub_Rural %>% group_by(Sex) %>%  
  summarise(  
    n = unweighted(n()),  
    prop = survey_prop(vartype = c("se", "ci"))))
```

Sex	n	prop	prop_se	prop_low	prop_upp
Hombre	10003	0.5074	0.0068	0.4940	0.5208
Mujer	9638	0.4926	0.0068	0.4792	0.5060

Proporción de hombres y mujeres en la zona rural

Análisis de encuestas de hogares con R

Motivación

Lectura y procesamiento de encuestas con R

Estimación del tamaño.

Estimación de la proporción.

Tablas cruzadas.

```
(prop_sexoR <- sub_Rural %>% group_by(Sex) %>%  
  summarise(  
    n = unweighted(n()),  
    prop = survey_prop(vartype = c("se", "ci"))))
```

Sex	n	prop	prop_se	prop_low	prop_upp
Hombre	10003	0.5074	0.0068	0.4940	0.5208
Mujer	9638	0.4926	0.0068	0.4792	0.5060

Ejercicio

¿Cómo estimar el Proporción de hombres en la zona rural dado que es hombre?

Proporción de hombres en la zona urbana y rural

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
(prop_ZonaH <- sub_Hombre %>% group_by(Zone) %>%  
  summarise(  
    prop = survey_prop(vartype = c("se", "ci"))))
```

Zone	prop	prop_se	prop_low	prop_upp
Urbana	0.8149	0.0035	0.8081	0.8218
Rural	0.1851	0.0035	0.1782	0.1919

Proporción de hombres en la zona urbana y rural

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
(prop_ZonaH <- sub_Hombre %>% group_by(Zone) %>%  
  summarise(  
    prop = survey_prop(vartype = c("se", "ci"))))
```

Zone	prop	prop_se	prop_low	prop_upp
Urbana	0.8149	0.0035	0.8081	0.8218
Rural	0.1851	0.0035	0.1782	0.1919

Ejercicio

¿Cómo estimar el Proporción de mujeres en la zona rural dado que es mujer?

Proporción de mujeres en la zona urbana y rural

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
(prop_ZonaM <- sub_Mujer %>% group_by(Zone) %>%  
  summarise(  
    prop = survey_prop(vartype = c("se","ci"))))
```

Zone	prop	prop_se	prop_low	prop_upp
Urbana	0.8298	0.0034	0.8231	0.8365
Rural	0.1702	0.0034	0.1635	0.1769

Proporción de hombres en la zona urbana y rural

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Con el uso de la función `group_by` es posible estimar un mayor número de niveles de agregación al combinar dos o más variables.

```
(prop_ZonaH_Pobreza <- sub_Hombre %>%  
  group_by(Zone, pobreza) %>%  
  summarise(  
    prop = survey_prop(vartype = c("se", "ci")))%>%  
  data.frame())
```

Proporción de hombres en la zona urbana y rural

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Zone	pobreza	prop	prop_se	prop_low	prop_upp
Urbana	0	0.7925	0.0073	0.7782	0.8069
Urbana	1	0.2075	0.0073	0.1931	0.2218
Rural	0	0.8037	0.0119	0.7804	0.8270
Rural	1	0.1963	0.0119	0.1730	0.2196

Proporción de mujeres en la zona urbana y rural

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
(prop_ZonaM_Pobreza <- sub_Mujer %>%  
  group_by(Zone, pobreza) %>%  
  summarise(  
    prop = survey_prop(vartype = c("se", "ci"))) %>%  
  data.frame())
```

Zone	pobreza	prop	prop_se	prop_low	prop_upp
Urbana	0	0.7706	0.0078	0.7553	0.7858
Urbana	1	0.2294	0.0078	0.2142	0.2447
Rural	0	0.7636	0.0139	0.7362	0.7910
Rural	1	0.2364	0.0139	0.2090	0.2638

Proporción de hombres en la zona y empleado

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
(prop_ZonaH_Ocupacion <- sub_Hombre %>%  
  group_by(Zone, desempleo) %>%  
  summarise(  
    prop = survey_prop(vartype = c("se", "ci")))%>%  
  data.frame())
```

Zone	desempleo	prop	prop_se	prop_low	prop_upp
Urbana	0	0.9772	0.0014	0.9744	0.9799
Urbana	1	0.0228	0.0014	0.0201	0.0256
Rural	0	0.9859	0.0018	0.9823	0.9894
Rural	1	0.0141	0.0018	0.0106	0.0177

Proporción de mujeres en la zona urbana y rural

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
(prop_ZonaM_Ocupacion <- sub_Mujer %>%  
  group_by(Zone, desempleo) %>%  
  summarise(  
    prop = survey_prop(  
      vartype = c("se", "ci")) %>%  
    data.frame())
```

Zone	desempleo	prop	prop_se	prop_low	prop_upp
Urbana	0	0.9682	0.0017	0.9649	0.9715
Urbana	1	0.0318	0.0017	0.0285	0.0351
Rural	0	0.9730	0.0029	0.9673	0.9788
Rural	1	0.0270	0.0029	0.0212	0.0327

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

**Tablas
cruzadas.**

Tablas cruzadas.

Tabla Zona Vs Sexo

Análisis de encuestas de hogares con R

Motivación

Lectura y procesamiento de encuestas con R

Estimación del tamaño.

Estimación de la proporción.

Tablas cruzadas.

Haciendo uso de la función `group_by` organizada en forma de `data.frame`.

```
(  
  prop_sexo_zona <- diseno %>%  
    group_by(pobreza, Sex) %>%  
    summarise(  
      prop = survey_prop(vartype = c("se", "ci"))) %>%  
      data.frame()  
)
```

Esta forma de organizar la información es recomendable cuando el realizar el análisis sobre las estimaciones puntuales.

Tabla Zona Vs Sexo

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

pobreza	Sex	prop	prop_se	prop_low	prop_upp
0	Hombre	0.4945	0.0035	0.4877	0.5014
0	Mujer	0.5055	0.0035	0.4986	0.5123
1	Hombre	0.4576	0.0056	0.4466	0.4687
1	Mujer	0.5424	0.0056	0.5313	0.5534

Tablas de doble entrada.

Una alternativa es utilizar la función `svyby` con la siguiente sintaxis.

```
tab_Sex_Pobr <- svyby(~Sex, ~pobreza, diseno, svymean)
tab_Sex_Pobr %>% select(-se.SexMujer, -se.SexHombre)
```

	pobreza	SexHombre	SexMujer
0	0	0.4945	0.5055
1	1	0.4576	0.5424

```
tab_Sex_Pobr %>% select(-SexMujer, -SexHombre)
```

	pobreza	se.SexHombre	se.SexMujer
0	0	0.0035	0.0035
1	1	0.0056	0.0056

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Tablas de doble entrada.

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Para la estimación de los intervalos de confianza utilizar la función `confint`.

```
confint(tab_Sex_Pobr) %>% as.data.frame()
```

	2.5 %	97.5 %
0:SexHombre	0.4877	0.5014
1:SexHombre	0.4466	0.4686
0:SexMujer	0.4986	0.5123
1:SexMujer	0.5314	0.5534

Prueba de independencia.

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Para realizar la prueba de independencia χ^2 puede ejecuta la siguiente linea de código.

```
svychisq(~Sex + pobreza, diseno, statistic="F")
```

```
##
```

```
## Pearson's X^2: Rao & Scott adjustment
```

```
##
```

```
## data: NextMethod()
```

```
## F = 36, ndf = 1, ddf = 1027, p-value = 3e-09
```

Más adelante se profundiza en la metodología de esta prueba.

Tablas de doble entrada.

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
(tab_Sex_Ocupa <- svyby(~Sex, ~empleo,  
  diseno %>% filter(empleo %in% c("Ocupado", "Desocupa  
    "Inactivo")), svymea
```

	empleo	SexHombre	SexMujer	se.SexHombre	se.SexMujer
Ocupado	Ocupado	0.6010	0.3990	0.0047	0.0047
Desocupado	Desocupado	0.3938	0.6062	0.0168	0.0168
Inactivo	Inactivo	0.3549	0.6451	0.0055	0.0055

Tablas de doble entrada

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
confint(tab_Sex_Ocupa) %>% as.data.frame()
```

	2.5 %	97.5 %
Ocupado:SexHombre	0.5919	0.6102
Desocupado:SexHombre	0.3609	0.4266
Inactivo:SexHombre	0.3441	0.3658
Ocupado:SexMujer	0.3898	0.4081
Desocupado:SexMujer	0.5734	0.6391
Inactivo:SexMujer	0.6342	0.6559

Tablas de doble entrada.

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Dado que la variable *pobreza* es de tipo numérica, es necesario convertirla en factor.

```
tab_region_pobreza <-  
  svyby(~as.factor(pobreza), ~desempleo, diseno, svymean)  
tab_region_pobreza %>%  
  select(-"se.as.factor(pobreza)0",  
        -"se.as.factor(pobreza)1")
```

	desempleo	as.factor(pobreza)0	as.factor(pobreza)1
0	0	0.7844	0.2156
1	1	0.6784	0.3216

Tablas de doble entrada.

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
tab_region_pobreza %>%  
  select("se.as.factor(pobreza)0",  
         "se.as.factor(pobreza)1")
```

	se.as.factor(pobreza)0	se.as.factor(pobreza)1
0	0.0062	0.0062
1	0.0189	0.0189

Prueba de independencia.

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Una vez más la prueba de independencia es:

```
svychisq(~desempleo + pobreza,  
          design = diseno,  statistic="F")
```

```
##
```

```
## Pearson's X^2: Rao & Scott adjustment
```

```
##
```

```
## data:  NextMethod()
```

```
## F = 48, ndf = 1, ddf = 1027, p-value = 1e-11
```

Razón de obbs

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
(tab_Sex <- svyby(~pobreza, ~Sex, diseno,  
  svymean, vartype = c("se", "ci")))
```

	Sex	pobreza	se	ci_l	ci_u
Hombre	Hombre	0.2054	0.0063	0.1930	0.2178
Mujer	Mujer	0.2306	0.0069	0.2172	0.2441

```
svycontrast(tab_Sex, quote(`Mujer`/`Hombre`))
```

```
##          nlcon    SE  
## contrast  1.12 0.05
```

Razón de obbs

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

```
tab_Sex_Pobr <-  
  svymean(~interaction (Sex, pobreza), diseno,  
           se=T, na.rm=T, ci=T, keep.vars=T)  
tab_Sex_Pobr %>% as.data.frame()
```

	mean	SE
interaction(Sex, pobreza)Hombre.0	0.3866	0.0041
interaction(Sex, pobreza)Mujer.0	0.3951	0.0043
interaction(Sex, pobreza)Hombre.1	0.0999	0.0031
interaction(Sex, pobreza)Mujer.1	0.1184	0.0036

Razón de obbs

Suponga que se desea calcular la siguiente razón de obbs.

$$\frac{\frac{P(\text{Sex}=\text{Female}|\text{pobreza}=0)}{P(\text{Sex}=\text{Female}|\text{pobreza}=1)}}{\frac{P(\text{Sex}=\text{Male}|\text{pobreza}=1)}{P(\text{Sex}=\text{Male}|\text{pobreza}=0)}}$$

La forma de cálculo en sería:

```
svycontrast(tab_Sex_Pobr,  
  quote(`interaction(Sex, pobreza)Mujer.0`/  
    `interaction(Sex, pobreza)Mujer.1`)/  
  (`interaction(Sex, pobreza)Hombre.0`/  
    `interaction(Sex, pobreza)Hombre.1`) ))
```

```
##          nlcon    SE  
## contrast 0.862 0.02
```

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Razón de obbs

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Ahora, se desea realizar la siguiente razón:

$$\frac{\frac{P(\text{Sex}=\text{Male}|\text{pobreza}=1)}{P(\text{Sex}=\text{Female}|\text{pobreza}=1)}}{\frac{P(\text{Sex}=\text{Male}|\text{pobreza}=0)}{P(\text{Sex}=\text{Female}|\text{pobreza}=0)}}$$

```
svycontrast(tab_Sex_Pobr,  
             quote(`interaction(Sex, pobreza)Hombre.1`/  
                   `interaction(Sex, pobreza)Mujer.1`)/  
             (`interaction(Sex, pobreza)Hombre.0`/  
               `interaction(Sex, pobreza)Mujer.0`)))
```

```
##          nlcon    SE  
## contrast 0.862 0.02
```

Contrastes

Análisis de encuestas de hogares con R

Motivación

Lectura y procesamiento de encuestas con R

Estimación del tamaño.

Estimación de la proporción.

Tablas cruzadas.

El interés ahora es realizar un contraste de proporciones, por ejemplo: $\hat{p}_F - \hat{p}_M$

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

	Sex	pobreza	se	ci_l	ci_u
Hombre	Hombre	0.2054	0.0063	0.1930	0.2178
Mujer	Mujer	0.2306	0.0069	0.2172	0.2441

Contrastes de la diferencia de proporciones en R

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Para realizar la diferencia de proporciones se hace uso de la función `svycontrast`.

```
svycontrast(tab_sex_pobreza,  
             list(diff_Sex = c(1, -1))) %>%  
             data.frame()
```

	contrast	diff_Sex
diff_Sex	-0.0252	0.0043

Contrastes de la diferencia de proporciones

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Diferencia en desempleo por sexo.

```
(tab_sex_desempleo <- svyby(  
  ~desempleo, ~Sex,  
  diseno %>% filter(!is.na(desempleo)) ,  
  svymean, na.rm=T, covmat = TRUE,  
  vartype = c("se", "ci"))
```

	Sex	desempleo	se	ci_l	ci_u
Hombre	Hombre	0.0212	0.0012	0.0189	0.0236
Mujer	Mujer	0.0310	0.0015	0.0281	0.0339

Contrastes de la diferencia de proporciones en R

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Siguiendo el ejemplo anterior se tiene que:

```
svycontrast(tab_sex_desempleo,  
             list(diff_Sex = c(-1, 1))) %>%  
             data.frame()
```

	contrast	diff_Sex
diff_Sex	0.0097	0.0018

¡Gracias!

Análisis de
encuestas de
hogares con R

Motivación

Lectura y
procesamien-
tos de
encuestas con
R

Estimación del
tamaño.

Estimación de
la proporción.

Tablas
cruzadas.

Email: andres.gutierrez@cepal.org