



FACULTAD DE CIENCIAS SOCIALES - PUCP

Curso: POL 278 - Estadística para el análisis político 1

Semestre 2021-2

Docente: Marylia Cruz

****Sesión 6: RAnálisis de variables categóricas***

El objetivo de esta sesión es conocer los comandos básicos de la prueba Chi-cuadrado y medidas de asociación.

La base de datos es del LATINOBARÓMETRO :

<https://www.latinobarometro.org/latContents.jsp>

Abrir la base de datos

```
library(rio)
data <- import("Latinobarometro_2018_Esp_Spss_v20190303.sav")
```

SELECCIONAR DATA

```
data=subset(data, IDENPA==68 | IDENPA==170 | IDENPA==604 | IDENPA==218)
table(data$IDENPA)
```

```
##
##      68   170   218   604
## 1200 1200 1200 1200
```

```
library(car)
```

```
## Loading required package: carData
```

```
data$pais<- recode(data$IDENPA, "68='Bolivia'; 170='Colombia';604='Peru';218='Ecuador'")
table(data$pais)
```

```
##
## Bolivia Colombia Ecuador Peru
##      1200      1200      1200      1200
```

Descripción de la base

Sexo

```
data$SEXO=as.factor(data$SEXO)
levels(data$SEXO)=c("Hombre", "Mujer")
```

Autoubicación en escala de pobreza-riqueza: Personal

```
data$pobr_riq=as.numeric(data$P10STC.A)
```

En términos generales ¿Diría usted que (PAÍS) está gobernado por unos cuantos grupos poderosos en su propio beneficio, o que está gobernado para el bien de todo el pueblo? (ESPERE RESPUESTA Y MARQUE UNA SOLA)

```
data$pode_pueblo=as.factor(data$P14ST)
levels(data$pode_pueblo)=c("Poderosos", "Pueblo")
```

¿Y en los próximos doce meses, cree que su situación económica y la de su familia será mucho mejor, un poco mejor, igual, un poco peor, o mucho peor que la que tiene hoy? (ORDINAL) 5 Mucho peor 1 Mucho mejor

```
data$situ_eco=as.factor(data$P9STGBSC)
levels(data$situ_eco)=c("Mucho mejor", "Un poco mejor", "Igual", "Un poco peor", "Mucho peor")
```

¿Diría Ud. que este país... progresado 1
estancado 2 esta retroceso 3

```
library(car)
data$pais_pro=as.factor(data$P2STC)
data$pais_pro=factor(data$pais_pro,
                     levels = levels(data$pais_pro),
                     labels = c("Progreso", "Estancado", "Retrocedio"),
                     ordered = T)
table(data$pais_pro)
```

```
##
##   Progreso  Estancado Retrocedio
##      1336      2375      948
```

En general, ¿Diría Ud. que está muy satisfecho, más bien satisfecho, no muy satisfecho o nada satisfecho con el funcionamiento de la democracia en (PAÍS)?

```
library(car)
table(data$P13STGBS.A)
```

```
##
##    1    2    3    4
##  271  899 2383 1015
```

```
data$satis_democracia=as.factor(data$P13STGBS.A)
data$satis_democracia <- recode(data$P13STGBS.A, "1:2=1; 3:4=2")
table(data$satis_democracia)
```

```
##
##      1      2
## 1170 3398
```

```
data$satis_democracia=as.factor(data$satis_democracia)
levels(data$satis_democracia)=c("Satisfecho", "No Satisfecho")
```

¿Cuán justa cree Ud. que es la distribución del ingreso en (PAÍS)?

```
str(data$P23ST)
```

```
##  num [1:4800] 1 3 3 3 4 2 2 NA 3 3 ...
```

```
data$justo=as.factor(data$P23ST)
levels(data$justo)=c("Muy justa", "Justa", "Injusta", "Muy injusta")
```

DESCRIPCIÓN DE VARIABLES

```
library(jmv)
descriptives(data, vars=vars(justo, pode_pueblo, pobr_riq, situ_eco, pais_pro, satis_democracia), freq = TRUE)
```

```
##
## DESCRIPTIVES
##
## Descriptives
##
##      justo      pode_pueblo      pobr_riq      situ_eco      pais_pro      satis_democracia
##
## N      4588      4623      4752      4562      4659      4568
## Missing 212      177      48      238      141      232
## Mean
## Median
## Minimum
## Maximum
##
##
## FREQUENCIES
##
## Frequencies of justo
##
##      Levels      Counts      % of Total      Cumulative %
##
##      Muy justa      81      1.76548      1.76548
##      Justa      916      19.96513      21.73060
```

```
##      Injusta      2500      54.48997      76.22058
##      Muy injusta    1091      23.77942     100.00000
##
##
##
## Frequencies of pode_pueblo
##
##      Levels      Counts      % of Total      Cumulative %
##
##      Poderosos    3693      79.88319      79.88319
##      Pueblo       930      20.11681     100.00000
##
##
##
## Frequencies of situ_eco
##
##      Levels      Counts      % of Total      Cumulative %
##
##      Mucho mejor    663      14.53310      14.53310
##      Un poco mejor  1647     36.10259     50.63569
##      Igual          1662     36.43139     87.06708
##      Un poco peor   433      9.49145     96.55853
##      Mucho peor     157      3.44147     100.00000
##
##
##
## Frequencies of pais_pro
##
##      Levels      Counts      % of Total      Cumulative %
##
##      Progreso      1336     28.67568     28.67568
##      Estancado      2375     50.97660     79.65229
##      Retrocedio     948     20.34771     100.00000
##
##
##
## Frequencies of satis_democracia
##
##      Levels      Counts      % of Total      Cumulative %
##
##      Satisfecho     1170     25.61296     25.61296
##      No Satisfecho  3398     74.38704     100.00000
##
```

TABLA DE CONTINGENCIA

UNIDAD DE ANÁLISIS : Personas entrevistadas

```
table(data$pais,data$satis_democracia)
```

```
##
##      Satisfecho No Satisfecho
##      Bolivia      314      816
```

```
## Colombia      304      841
## Ecuador       426      740
## Peru          126     1001
```

```
addmargins(table(data$pais,data$satis_democracia))
```

```
##
##      Satisfecho No Satisfecho Sum
## Bolivia      314      816 1130
## Colombia     304      841 1145
## Ecuador      426      740 1166
## Peru         126     1001 1127
## Sum         1170     3398 4568
```

Distribución conjunta

Porcentaje en relación al total general

Bolivia&Satisfechos= $314/4568=0.068$

```
prop.table(table(data$pais,data$satis_democracia))
```

```
##
##      Satisfecho No Satisfecho
## Bolivia 0.06873905 0.17863398
## Colombia 0.06654991 0.18410683
## Ecuador 0.09325744 0.16199650
## Peru    0.02758319 0.21913310
```

Distribución Marginal

Porcentaje en relación al total de las filas

```
prop.table(table(data$pais,data$satis_democracia),1)
```

```
##
##      Satisfecho No Satisfecho
## Bolivia 0.2778761 0.7221239
## Colombia 0.2655022 0.7344978
## Ecuador 0.3653516 0.6346484
## Peru    0.1118012 0.8881988
```

Porcentaje en relación al total de las columnas

```
prop.table(table(data$pais,data$satis_democracia),2)
```

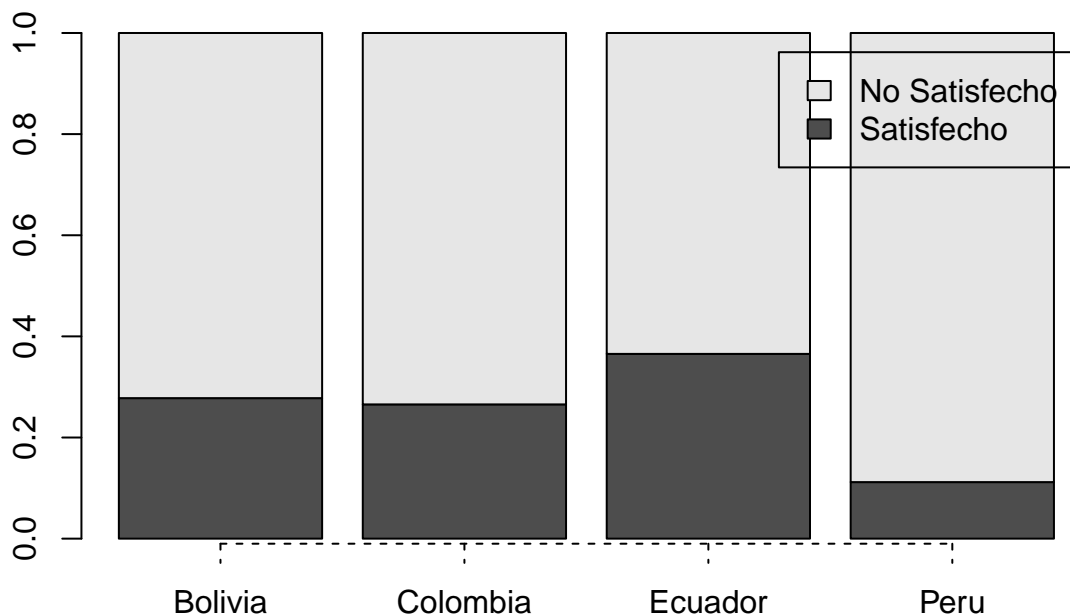
```
##
##      Satisfecho No Satisfecho
## Bolivia 0.2683761 0.2401413
## Colombia 0.2598291 0.2474985
## Ecuador 0.3641026 0.2177752
## Peru    0.1076923 0.2945851
```

```
prop.table(table(data$satis_democracia,data$pais),2)
```

```
##
##          Bolivia Colombia Ecuador Peru
## Satisfecho  0.2778761 0.2655022 0.3653516 0.1118012
## No Satisfecho 0.7221239 0.7344978 0.6346484 0.8881988
```

GRÁFICO

```
cuadro=prop.table(table(data$satis_democracia,data$pais),2)
barplot(cuadro, legend = rownames(cuadro),axis.lty = 2)
```



FRECUENCIAS ESPERADAS

```
table(data$satis_democracia,data$pais)
```

```
##
##          Bolivia Colombia Ecuador Peru
## Satisfecho    314      304    426  126
## No Satisfecho  816      841    740 1001
```

```
tablaFesperadas <- chisq.test(data$satis_democracia,data$pais)$expected
tablaFesperadas
```

```
##          data$pais
## data$satis_democracia Bolivia Colombia Ecuador Peru
## Satisfecho    289.4264 293.2684 298.6471 288.6581
## No Satisfecho  840.5736 851.7316 867.3529 838.3419
```

CHI CUADRADO

grados de libertad= (filas-1)(columnas-1) *grados de libertad= (2-1)(4-1)* grados de libertad=3

```
qchisq(0.95,3) # Valor de la tabla
```

```
## [1] 7.814728
```

```
chisq.test(data$pais,data$satis_democracia)
```

```
##
## Pearson's Chi-squared test
##
## data: data$pais and data$satis_democracia
## X-squared = 199.56, df = 3, p-value < 2.2e-16
```

H0: NO HAY ASOCIACIÓN ENTRE LA SATISFACCIÓN DE LA DEMOCRACIA Y EL PAÍS

P-VALOR= 0.000000000000000022

Rechazo la hipótesis nula. Si es posible una relación entre la satisfacción de la democracia y el país.

MEDIDAS DE ASOCIACIÓN PARA ALMENOS UNA VARIABLE NOMINAL

```
cuadro=table(data$satis_democracia,data$pais)
#install.packages("vcd")
library(vcd)
```

```
## Loading required package: grid
```

```
assocstats(cuadro)
```

```
##
##              X^2 df P(> X^2)
## Likelihood Ratio 216.99  3      0
## Pearson          199.56  3      0
##
## Phi-Coefficient   : NA
## Contingency Coeff.: 0.205
## Cramer's V       : 0.209
```

```
#install.packages("DescTools")
library(DescTools)
```

```
##
## Attaching package: 'DescTools'

## The following object is masked from 'package:car':
##
## Recode
```

```
Phi(cuadro,conf.level=0.95)
```

```
## [1] 0.2090112
```

```
UncertCoef(cuadro,conf.level=0.95)
```

```
##          uc      lwr.ci      upr.ci  
## 0.02429518 0.01824672 0.03034365
```

```
Lambda(cuadro, conf.level=0.95)
```

```
##      lambda      lwr.ci      upr.ci  
## 0.05708661 0.03969189 0.07448134
```

MEDIDAS DE ASOCIACIÓN PARA RELACIÓN ENTRE DOS VARIABLES ORDINALES

```
table(data$justo,data$pais_pro)
```

```
##  
##          Progreso Estancado Retrocedio  
## Muy justa      32      28      21  
## Justa          439     353     104  
## Injusta        615    1347     478  
## Muy injusta    196     561     308
```

```
tabla=prop.table(table(data$justo,data$pais_pro),1)*100  
addmargins(tabla)
```

```
##  
##          Progreso Estancado Retrocedio      Sum  
## Muy justa  39.50617  34.56790  25.92593 100.00000  
## Justa      48.99554  39.39732  11.60714 100.00000  
## Injusta    25.20492  55.20492  19.59016 100.00000  
## Muy injusta 18.40376  52.67606  28.92019 100.00000  
## Sum        132.11038 181.84620  86.04342 400.00000
```

```
cuadro2=table(data$justo,data$pais_pro)  
chisq.test(cuadro2)
```

```
##  
## Pearson's Chi-squared test  
##  
## data:  cuadro2  
## X-squared = 293.26, df = 6, p-value < 2.2e-16
```



```
# Medidas de asociación para variables ordinales
```

```
GoodmanKruskalGamma(cuadro2, conf.level=0.95)
```

```
##      gamma    lwr.ci    upr.ci  
## 0.3208106 0.2796476 0.3619736
```

```
KendallTauB(cuadro2, conf.level=0.95)
```

```
##      tau_b    lwr.ci    upr.ci  
## 0.2005291 0.1740476 0.2270107
```

```
StuartTauC(cuadro2, conf.level=0.95)
```

```
##      tauc    lwr.ci    upr.ci  
## 0.1839186 0.1593666 0.2084706
```

```
#Medidas asimetricas
```

```
cuadro2=table(data$pais_pro,data$justo)  
cuadro2
```

```
##  
##           Muy justa Justa Injusta Muy injusta  
## Progreso           32  439    615         196  
## Estancado           28  353   1347         561  
## Retrocedio          21  104    478         308
```

```
SomersDelta(cuadro2, direction = "row") #fila = variable dependiente #row reemplazar por col si es que
```

```
## [1] 0.2020387
```

DIRECTA: A medida que se cree que la distribución de ingresos es muy injusta, se considera que el país retrocedio. A medida que se cree que la distribución de ingresos es muy justa, se considera que el país progreso

INTENSIDAD: BAJA 0 a 0.3 es relación baja.