

ANÁLISIS BIVARIADO SOBRE LIBERTADES EN ITALIA

Encuesta Social Europea, 2020/2022 (Ronda 10)

Técnicas de Investigación en Ciencia Política I. UBU
Práctica 2.

Tomàs Ferrandis Moscardó

2024-05-19

Índice

1 INTRODUCCIÓN	3
1.1 DESCRIPCIÓN DE LA PRÁCTICA	3
1.2 LIBRERÍAS R	3
1.3 GESTIÓN DE FICHEROS DE DATOS	3
2 ACTIVIDAD 1. CONTRASTE DE MEDIAS PARA DATOS RELACIONADOS.	5
2.1 DESCRIPCIÓN DE LA ACTIVIDAD 1	5
2.2 PREPARACIÓN DE LA BASE DE DATOS	5
2.3 HIPÓTESIS	6
2.4 PRUEBA DE T-STUDENT	6
2.5 CONCLUSIÓN SOBRE T-STUDENT	7
3 ACTIVIDAD 2. ASOCIACIÓN DE VARIABLES.	8
3.1 DESCRIPCIÓN DE LA ACTIVIDAD 2	8
3.2 PREPARACIÓN DE LA BASE DE DATOS	8
3.3 HIPÓTESIS	9
3.4 LA TABLA DE CONTINGENCIA (TABLA CRUZADA)	9
3.5 CONCLUSIONES SOBRE LA TABLA DE CONTINGENCIA	10
3.6 CÁLCULO DEL COEFICIENTE CHI CUADRADO	12
Comprobación con el Coeficiente Gamma	13
3.7 CONCLUSIÓN SOBRE CHI CUADRADO	13
4 ACTIVIDAD 3. RESPECTO A LA LIBERTAD DE LOS MEDIOS DE COMUNICACIÓN	15
4.1 DESCRIPCIÓN DE LA ACTIVIDAD 3	15
4.2 PREPARACIÓN DE LA BASE DE DATOS	15
4.3 HIPÓTESIS	15
4.4 COEFICIENTE TAU-B DE KENDALL	16

4.5 CONCLUSIÓN SOBRE TAU-B DE KENDALL	16
4.6 COEFICIENTE DELTA (D) DE SOMERS.	17
4.7 CONCLUSIÓN SOBRE DELTA (D) DE SOMERS	17
5 VALORACIÓN CRÍTICA	18
ANEXO 1. ANÁLISIS UNIVARIADO DE ALGUNAS VARIABLES	19
Actividad 1.	19
Actividad 2.	19
Actividad 3.	20
ANEXO 2. EJECUCIÓN DEL Rmd Y REPOSITORIO DE FICHEROS	22

1 INTRODUCCIÓN

1.1 DESCRIPCIÓN DE LA PRÁCTICA

Este trabajo versará sobre las libertades en Italia a partir de diversos **análisis bivariados** sobre datos de la Encuesta Social Europea de los años 2020 y 2022 (Ronda 10). Se escogerá en cada caso el test más adecuado y, finalmente, se hará una valoración crítica a partir de los datos analizado. Las variables que se analizarán en las siguientes actividades corresponden al cuestionario de la encuesta.

La práctica está relacionada con los temas 5 y 6 de la asignatura Técnicas de la Investigación en Ciencia Política I del Grado Universitario de Ciencia Política y Gestión Pública de la Universidad de Burgos, curso 2023-2024.

1.2 LIBRERÍAS R

Un primer paso será asegurar que en nuestra instalación de R están disponibles algunas funciones necesarias para abrir el fichero de datos externo (tipo .sav), para algún cálculo estadístico o formateo de datos en Rmd.

```
if(!require(haven)){install.packages("haven")} # para abrir fichero .sav
library(haven)

if(!require(knitr)){install.packages("knitr")} # Act. 2. Tabla Rd
library(knitr)

if(!require(DescTools)){install.packages("DescTools")} #actividad 3
library(DescTools)

if (!require(epiDisplay)){install.packages("epiDisplay")}
library(epiDisplay)

if (!require(DescTools)){install.packages("DescTools")}
library(DescTools)
```

1.3 GESTIÓN DE FICHEROS DE DATOS

Tanto el fichero de los datos de la encuesta (*datosprac2.sav*) como cualquier fichero de resultados se ubicarán en una subcarpeta DATOS del directorio donde se esté este mismo fichero

Rmd.

```
# El fichero .sav debe estar en una subcarpeta DATOS
nombreFichero<-"datosprac2.sav"
directorioTrabajo<-getwd()
rutaFichero<-paste(directorioTrabajo,"DATOS",nombreFichero,sep="/")
```

A partir del fichero de tipo sav (**datosprac2.sav**) de la Encuesta Social Europea, se creará el *data frame* de R. Esta será la **base de datos** inicial con todos los datos de la encuesta.

```
df_ronda10<-read_sav(rutaFichero)
```

2 ACTIVIDAD 1. CONTRASTE DE MEDIAS PARA DATOS RELACIONADOS.

2.1 DESCRIPCIÓN DE LA ACTIVIDAD 1

En esta primera actividad se realizará un contraste de hipótesis para ver si dos variables presentan diferencias estadísticamente significativas entre si.

Las variables son:

- **fairelcc:** Las elecciones son libres y limpias en Italia (Pregunta B3 del cuestionario).
- **cttresac:** Los tribunales en Italia tratan a todo el mundo por igual (Pregunta B18 del cuestionario).

Ambas son Variables cualitativas ordinarias y admiten un valor que va desde 0 (“No es válida en absoluto”) a 10 (“Es completamente válido”) además de los valores considerados como no válidos a efectos estadísticos 77, 88 y 99 (“Rechaza responder”, “NS No sabe” y “NC No contesta” respectivamente).

2.2 PREPARACIÓN DE LA BASE DE DATOS

Recodificación de valores no válidos

Se comprueba si hay valores que necesitan recodificarse. Serían los valores externos al intervalo [0-10] que según el codebook de la encuesta podrían ser 77, 88 ó 99 ya explicados.

```
unique(df_ronda10$fairelcc)
```

```
## [1] 7 6 2 5 8 1 0 3 9 4 NA 10
```

```
unique(df_ronda10$cttresac)
```

```
## [1] 1 0 2 8 3 7 6 5 9 4 NA 10
```

Se comprueba que ya deben estar recodificados o almenos no aparecen. No obstante, en caso de que no fuese así se procedería de la siguiente manera.

En primer lugar, se duplicarán los campos para hacer las modificación sobre estos.

```
df_ronda10$rec_cttresac <- df_ronda10$cttresac
df_ronda10$rec_fairelecc <-df_ronda10$fairelcc
```

Después, se recodifican los valores 77, 88 y 99 como NA

```
df_ronda10$rec_fairelecc[df_ronda10$rec_fairelecc >=77 &
                        df_ronda10$rec_fairelecc <=99]<-NA

df_ronda10$rec_cttresac[df_ronda10$rec_cttresac >=77 &
                        df_ronda10$rec_cttresac <=99]<-NA
```

2.3 HIPÓTESIS

- Se plantea como **hipótesis nula (H0) que ambas medias son iguales**.
- La hipótesis alternativa (H1) sería, por lo tanto, que las medias son distintas.

2.4 PRUEBA DE T-STUDENT

Aplicaremos la función que implementa el método T-Student sobre los campos duplicados en nuestro caso.

- *paired* = *T*, indica a la función que operará con 2 muestras emparejadas.
- *alternative* = *'two.sided'*, indica que la hipótesis alternativa es que las medias no son iguales.
- Se establece un **nivel de confianza** del 95%.

```
t.test(df_ronda10$rec_fairelecc,df_ronda10$rec_cttresac,
       alternative = 'two.sided',
       conf.level = .95,
       paired = T)
```

```
##
```

```
## Paired t-test
```

```
##
```

```
## data: df_ronda10$rec_fairelecc and df_ronda10$rec_cttresac
```

```
## t = 22.183, df = 2503, p-value < 2.2e-16
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##  1.110378 1.325724
## sample estimates:
## mean difference
##          1.218051
```

2.5 CONCLUSIÓN SOBRE T-STUDENT

A la vista del resultado de t_test , se comprueba que hay una diferencia significativa entre las medias de aproximadamente 1,22. El intervalo de confianza, con un nivel de confianza del 95%, es de 1.110378 a 1.325724. La probabilidad de errar por rechazar la hipótesis nula es prácticamente nula $p=2.2 \times 10^{-16}$; muy inferior al 0,05 que se establece habitualmente en ciencia política como límite para considerarlo un valor estadísticamente significativo.

Por lo que **se debe rechazar la hipótesis nula que plantea la igualdad de las medias: las medias son distintas.**

Esto se traduce en que existe una percepción distinta en los encuestados entre la libertad y limpieza del proceso electoral en Italia y el trato igualitario de su sistema judicial.

Dirección de la diferencia en la función de R

El valor positivo del estadístico *t-valor*, de los intervalos y de la *diferencia de medias* nos indica el orden de la diferencia. Viendo el orden de las variables en la función R, la primera variable “fairelecc” tiene una media en promedio mayor que la segunda “cttresac”, por lo que la percepción de los encuestados sobre la limpieza de las elecciones es más positiva que la que tienen sobre la igualdad de trato del sistema judicial.

Si se invierte el orden de las variables en la función de R dará los mismos valores absolutos de *t*, de *intervalos* y de *diferencia de media* pero con signo negativo (cambiaría la dirección de la diferencia). El valor *p-value* seguirá igualmente positivo o absoluto puesto que indica la significancia estadística.

3 ACTIVIDAD 2. ASOCIACIÓN DE VARIABLES.

3.1 DESCRIPCIÓN DE LA ACTIVIDAD 2

En esta actividad se procederá a comprobar la asociación entre dos variables cualitativas ordinales.

Las variables a analizar serán:

- **rec_hmsfmlsh**, *Me daría vergüenza que un familiar cercano fuese gay o lesbiana*, representa la **variable independiente** y, por lo tanto, sus categorías corresponderán a filas en la tabla de contingencia (Pregunta A48 del cuestionario).
- **rec_freehms**, *Los gays y lesbianas deberían tener libertad para vivir como quieran*, representa la **variable dependiente**. Sus categorías corresponderán a columnas en la tabla (Pregunta A47 del cuestionario).

3.2 PREPARACIÓN DE LA BASE DE DATOS

Duplicación de campos

Se duplicarán las columnas del dataframe a manipular puesto que vamos a manipularlos.

```
df_ronda10$rec_hmsfmlsh<-df_ronda10$hmsfmlsh  
df_ronda10$rec_freehms<-df_ronda10$freehms
```

Recodificación. Etiquetas de valores

Se añaden las etiquetas correspondientes a los valores en los campos tipo *factor* en los nuevos campos duplicados.

```
# Recodificación para que se muestre los valores de las etiquetas  
# De la variable independiente: "Me daría vergüenza que un familiar..."  
df_ronda10$rec_hmsfmlsh <- factor(df_ronda10$rec_hmsfmlsh,  
                                  levels = c(1,2,3,4,5),  
                                  labels = c("Muy de acuerdo", "De acuerdo",  
                                              "Ni de acuerdo ni en desacuerdo",  
                                              "En desacuerdo", "Muy en desacuerdo"))
```



```
# De la variable dependiente: "Los gays y lesbianas deberían tener
# libertad para..."

df_ronda10$rec_freehms <- factor(df_ronda10$rec_freehms,
                                levels = c(1,2,3,4,5),
                                labels = c("Muy de acuerdo", "De acuerdo",
                                           "Ni de acuerdo ni en desacuerdo",
                                           "En desacuerdo", "Muy en desacuerdo"))
```

3.3 HIPÓTESIS

- Hipótesis nula: **No hay asociación estadísticamente significativa entre las variables.**
- Hipótesis alternativa: Hay asociación estadísticamente significativa entre las variables

3.4 LA TABLA DE CONTINGENCIA (TABLA CRUZADA)

Se obtienen los valores de frecuencia absolutos con la función *xtabs*.

```
freqtabla <- xtabs(~rec_hmsfmlsh + rec_freehms, data = df_ronda10)
```

Pero lo que interesa es tener en cada celda el % de casos respecto a la fila. Teniendo en cuenta que cada fila representa una categoría de la variable independiente. Se usará la función *prop.table*.

```
# El Parámetro 1 de prop.table indica % por fila
freqtabla_100xFila <- prop.table((freqtabla), 1) * 100
```

Se ve que la función no aporta los marginales (fila de totales). Para ello R dispone de otra función, *margin.table* :

```
Totales <- margin.table((prop.table(freqtabla) * 100), 2)
```

Se combinan las dos matrices (*freqtabla_100xFila* y *Totales*) para elaborar la **Tabla de contingencia o Tabla cruzada** mediante la función *rbind*

```
tablaContingencia <- round(rbind(freqtabla_100xFila, Totales), 2)
kable(tablaContingencia,
caption="Opinión sobre la libertad de gays y lesbianas (columns)
```

```
según sentimiento de vergüenza por un familiar gay o lesbiana (filas)",
align=c("c","c","c","c","c"))
```

Tabla 1: Opinión sobre la libertad de gays y lesbianas (columnas) según sentimiento de vergüenza por un familiar gay o lesbiana (filas)

	De				
	Muy de acuerdo	acuer- do	Ni de acuerdo ni en desacuerdo	En des- acuerdo	Muy en desacuerdo
Muy de acuerdo	22.97	10.81	8.11	24.32	33.78
De acuerdo	4.07	36.65	25.79	26.24	7.24
Ni de acuerdo ni en desacuerdo	6.92	40.71	46.05	4.94	1.38
En desacuerdo	17.39	67.34	12.26	2.79	0.22
Muy en desacuerdo	66.87	28.80	1.69	0.60	2.05
Totales	30.54	45.02	16.61	5.18	2.65

3.5 CONCLUSIONES SOBRE LA TABLA DE CONTINGENCIA

En una primera observación sobre los **valores marginales** o totales se debe destacar que la mayoría de encuestados está “De acuerdo con las libertades de los gays y lesbianas” (45.02%) siendo la segunda opción “Muy de acuerdo” (30,54%), además la respuesta con menor porcentaje es la de “Muy en desacuerdo” (2,65%).

Al observar los valores de las diferentes **categorías de la variable independiente** o filas se tiene un primer bloque con las dos categorías de encuestados que reconocen que sentirían algún tipo de vergüenza por tener un familiar homosexual. Aquí se ve una diferencia notable entre ambas categorías:

- Por un parte los que responden estar “Muy de acuerdo” con este sentimiento de vergüenza, su posicionamiento mayoritario es contrario a la libertad, decantándose un 33,78% por “Muy en desacuerdo” y un 24,32% por “En desacuerdo”. No obstante aparece un dato muy interesante: un importante 22,97% que se manifiesta “Muy de acuerdo” con la libertad.

- Por otra parte, los que se muestran “De acuerdo” con la afirmación de vergüenza, su posición mayoritaria respecto a la libertad es estar “De acuerdo” (36,65%), un dato también interesante. Seguidamente optan por “En desacuerdo” y “Ni de acuerdo ni en desacuerdo” con cifras similares (26,24% y 25,79% respectivamente).

Un nivel intermedio lo configuraría la categoría más “neutral”: los que no están “Ni de acuerdo ni en desacuerdo” con sentir vergüenza y que también se decantan mayoritariamente por la misma posición neutra, ocupando el centro de la tabla, con un 46,05%. Valores muy alejados del 4,94 o 1,38 de oposición a las libertades. Aunque es importante señalar que su segunda opción es posicionarse “De acuerdo” con las libertades de gays y lesbianas (40,71%), se ve una relación “Ni de acuerdo ni en desacuerdo- Ni desacuerdo ni desacuerdo”

En un tercer bloque estarían las dos categorías que manifiestan rechazo a la idea de sentir vergüenza por tener un familiar homosexual. Aquí también se ven unas relaciones con cierta lógica.

- Los encuestados que están “En desacuerdo”, son los que más se posicionan mayoritariamente también como “De acuerdo” con las libertades (67,34%).
- Con cifras similares, en el extremo, los que están “Muy en desacuerdo” con la afirmación sobre avergonzarse se declaran “Muy de acuerdo con la libertad” (66,84%).

En resumen, entre los que no sentirían vergüenza por tener un familiar homosexual se ve una coincidencia “De acuerdo-desacuerdo” y “Muy desacuerdo-muy de acuerdo” en lo que respeta a las libertades de este colectivo.

Cuando se examina columna a columna toda la tabla, se observa que los valores de cada columna varían respecto a su correspondiente valor marginal. De hecho no se observa ninguna coincidencia lo que nos lleva a concluir una dependencia entre las variables y, consecuentemente, a rechazar la hipótesis nula.

3.6 CÁLCULO DEL COEFICIENTE CHI CUADRADO

Este coeficiente se muestra junto a las tablas cruzadas y nos indica si existe asociación entre la variable independiente (filas) y la dependiente (columnas). Solo arroja un valor igual o mayor que cero.

- Si vale 0 significa que no hay asociación.
- Si es >0 , existe asociación, cuanto mayor sea, más evidencia hay.
- Al dar siempre valores positivos (absolutos), no indica ninguna dirección de la asociación.

En R existe la función *chisq* para el cálculo de *Chi cuadrado* con dos parametrizaciones posibles.

- Pasando como argumentos directamente las variables a la función *chisq.test*

```
chi <- chisq.test( df_ronda10$rec_hmsfmlsh, df_ronda10$rec_freehms)

## Warning in chisq.test(df_ronda10$rec_hmsfmlsh, df_ronda10$rec_freehms):
## L'aproximació Chi-quadrat pot ser incorrecta

chi

##
##  Pearson's Chi-squared test
##
## data:  df_ronda10$rec_hmsfmlsh and df_ronda10$rec_freehms
## X-squared = 1723.8, df = 16, p-value < 2.2e-16
```

- Pasando como argumento la tabla en *chisq.test*

```
chi <- chisq.test(freqtabla)

## Warning in chisq.test(freqtabla): L'aproximació Chi-quadrat pot ser incorrecta

chi

##
##  Pearson's Chi-squared test
##
## data:  freqtabla
## X-squared = 1723.8, df = 16, p-value < 2.2e-16
```

Comprobación con el Coeficiente Gamma

Al tener valores de frecuencias menores de 5 en una proporción superior al 20% (8 de 25), Chi cuadrado advierte mediante un WARNING de un posible error.

Se puede optar por usar otro coeficiente como el **coeficiente Gamma** válido para variables ordinales. Los posibles resultados de Gamma son:

- Entre -1 y 0: asociación negativa.
- 0: No existe asociación.
- Entre 0 y 1: asociación positiva.

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

```
##      gamma      lwr.ci      upr.ci  
## -0.7150679 -0.7472630 -0.6828728
```

El coeficiente gamma de Goodman y Kruskal es aproximadamente -0.715.

El intervalo de confianza del 95% para el coeficiente gamma tiene un límite inferior

Este resultado de -0,747 debe interpretarse como asociación negativa importante (-0,747 está más cerca del -1 que del 0) entre las dos variables de la tabla de contingencia. El signo negativo indica la dirección de la asociación. En caso negativo, cuando una variable aumenta muy probablemente la otra disminuya y viceversa.

Los límites del intervalo de confianza no incluyen el cero ($0 > -0.6828728 > -0.7472630$) lo que redunda en la conclusión de rechazar la hipótesis nula.

3.7 CONCLUSIÓN SOBRE CHI CUADRADO

El coeficiente indica la magnitud de la asociación. El valor de 1723,8 es muy superior a 0, lo que indica una asociación fuerte. Hay una evidencia para el rechazo de la H_0 (independencia de las variables) y además, el *p-valor* es muy inferior (2.2×10^{-16}) a *alpha* (0,001) por lo que hay un nivel de confianza superior al 99,9%.

No obstante, al advertir R de la posibilidad de error que se produce por la cantidad de valores inferiores a 5, se ha optado por usar otro coeficiente para asegurar el resultado.

Se debe rechazar la hipótesis nula: las variables están asociadas. tal y como se ha podido

observar en la tabla de contingencia al comparar los valores en cada columna con su marginal correspondiente.

De las tres observaciones llegamos a la siguiente conclusión: la posición a favor o en contra de una igualdad en libertades de gays y lesbianas sería una variable dependiente del sentimiento de rechazo que causaría tener un familiar cercano homosexual.

4 ACTIVIDAD 3. RESPECTO A LA LIBERTAD DE LOS MEDIOS DE COMUNICACIÓN

4.1 DESCRIPCIÓN DE LA ACTIVIDAD 3

En esta actividad se pretende analizar si existe una dependencia entre la autoubicación ideológica de los encuestados y su percepción sobre el nivel de crítica que ejercen los medios de comunicación respecto al gobierno italiano.

Para ello se analizarán las variables:

- **lrscale**, *¿Dónde se colocaría usted en la escala? Escala de 0 a 10 (0 representa la ubicación más a la izquierda y 10 la ubicación más a la derecha)*. Esta es la variable independiente (Pregunta A39 del cuestionario)
- **medcrgvc**, *Los medios de comunicación en Italia pueden criticar al gobierno. Escala de 0 a 10 (0 indica No es válida en absoluto y 10, Es completamente válida)*. Esta es la variable dependiente (Pregunta B15 del cuestionario)

4.2 PREPARACIÓN DE LA BASE DE DATOS

Se comprueba que no hay valores que recodificar como NA

```
unique(df_ronda10$lrscale)
```

```
## [1] NA 5 8 6 4 3 10 1 2 7 9 0
```

```
unique(df_ronda10$medcrgvc)
```

```
## [1] 1 3 6 5 8 10 2 0 9 7 4 NA
```

Si fuese necesario recodificar por cualquier motivo (valores no válidos, añadir etiquetas...), convendría hacerlo sobre una duplicación de estos campos de forma análoga a lo realizado en los apartados 2.2 y 3.2 de la actividad anteriores.

4.3 HIPÓTESIS

- H0, hipótesis nula: **No hay asociación estadísticamente significativa entre las variables.**

- H1, hipótesis alternativa: Hay asociación estadísticamente significativa entre las variables.

4.4 COEFICIENTE TAU-B DE KENDALL

Se trata del test más adecuado para analizar dos variables cualitativas ordinales que tienen el mismo número de categorías como es este caso. Este coeficiente nos indica si hay asociación entre dos variables y la dirección de esta. Sus valores oscilan entre -1 y 1, siendo el 0 el valor resultante cuando no existe asociación. El signo nos indica la dirección de la asociación: valor negativo si una variable aumenta mientras la otra disminuye y valor positivo si ambas aumentan o disminuyen a la vez.

```
taub <- KendallTauB(df_ronda10$lrscale, df_ronda10$medcrgvc,
                    conf.level = 0.95)

taub
```

```
##          tau_b          lwr.ci          upr.ci
## -0.009130106 -0.043788752  0.025528540
```

El coeficiente *tau_b* (-0.009130106) es negativo y cercano a 0. Los valores *lwr.ci* y *upr.ci* (-0.043788752 0.025528540 respectivamente) son los límites inferior y superior del intervalo de confianza del 95%.

4.5 CONCLUSIÓN SOBRE TAU-B DE KENDALL

Sobre el coeficiente: El coeficiente, al ser negativo, indicaría que cuando aumenta una variable la otra disminuye (dirección de la asociación) pero al tratarse de un valor tan cercano a 0 (-0.009130106) nos sugiere que no hay asociación estadísticamente relevante entre las variables.

Sobre el intervalo de confianza: Los valores de *lwr.ci* y *upr.ci* indican que, con un 95% de confianza, el verdadero valor del coeficiente de Kendall se encuentra dentro de él. Este intervalo incluye al 0 (-0.043788752 < 0 < 0.025528540) hecho que refuerza la interpretación de que no hay una asociación estadísticamente significativa entre las variables.

Por todo esto:

NO se puede rechazar la hipótesis nula y consecuentemente no hay asociación estadística

entre las variables

4.6 COEFICIENTE DELTA (D) DE SOMERS.

Se trata de otro test de asociación para variables cualitativas ordinales que también nos informa de la magnitud y la dirección de la asociación mediante un coeficiente. Sus valores y la interpretación de estos es idéntica al Tau-B de Kendall: intervalo de -1 a 1, valor 0 para indicar la inexistencia de asociación y el signo para la dirección.

```
delta <- SomersDelta(df_ronda10$lrscale, df_ronda10$medcrgvc,  
                    conf.level = 0.95)  
  
delta
```

```
##          somers          lwr.ci          upr.ci  
## -0.009052743 -0.043415457  0.025309971
```

El coeficiente de *Somers* (-0,009052743) cercano a 0 y negativo.

Los valores *lwr.ci* y *upr.ci* (-0.043415457 y 0.025309971) son los límites inferior y superior del intervalo de confianza del 95%.

4.7 CONCLUSIÓN SOBRE DELTA (D) DE SOMERS

Sobre le coeficiente: Aunque el coeficiente, al ser negativo indica que cuando aumenta una variable la otra disminuye (dirección de la asociación), al tratarse de un valor tan cercano a 0 (-0.009052743) este sugiere que no hay asociación estadísticamente relevante entre las variables.

Sobre el intervalo de confianza: Los valores de *lwr.ci* y *upr.ci* indican que, con un 95% de confianza, el verdadero valor del coeficiente de Sommers se encuentra dentro de él. Este intervalo incluye al 0 (-0.043415457 < 0 < 0.025309971) hecho que refuerza la interpretación de que no hay una asociación estadísticamente significativa entre las variables.

Por todo esto, **NO se puede rechazar la hipótesis nula por lo tanto NO hay asociación estadística entre las variables** Lo que viene a decir que la opinión sobre la capacidad de criticar al gobierno por parte de la prensa es independiente de la autoubicación ideológica de las personas encuestadas.

Se comprueba que ambos coeficientes nos llevan a la misma conclusión.

5 VALORACIÓN CRÍTICA

Las libertades analizadas, operacionalizadas como variables ordinales, pueden estudiarse individualmente como cuantitativas como se estudió en la práctica anterior (Ver Anexo1). Así, se puede ver que los italianos encuestados “aprueban” su sistema electoral con un 5,8 de media y un 6 de mediana mientras que suspenderían a sus tribunales por trato desigual con una media de 4,52 y apenas una mediana de 5. Considerando que se compara la confianza en un proceso de máxima participación y reglamentación con la práctica diaria de funcionarios en una democracia consolidada, esta poca confianza en la limpieza y libertad de las elecciones es un hecho preocupante.

Los encuestados mayoritariamente rechazan la idea de sentir vergüenza por tener un familiar homosexual (media de 3,8 y mediana de 4 en un intervalo de 1 a 5). Un sentimiento relacionado, como se ha visto, con su posicionamiento favorable a las libertades de este colectivo (una media de 2,05 y una mediana de 2 en un intervalo de 1 a 5). Ambas medidas distan un punto, aproximadamente, del extremo más inclusivo. Sería interesante estudiar las rondas siguientes para ver la evolución en una época de polarización y megaidentities.

Con bastante independencia de la ideología, los encuestados aprueban la capacidad crítica de sus medios de comunicación respecto al gobierno. En la distribución de valores, tenemos que destacan los valores de 6 a 7 (entre 0 y 10) siendo los únicos con porcentajes superiores al 15%. Se puede entender que es una valoración bastante positiva tratándose de un modelo de prensa latino y, además ser transversal a los sectores ideológicos.

ANEXO 1. ANÁLISIS UNIVARIADO DE ALGUNAS VARIABLES

Análisis univariado en los que se apoyan aspectos de la valoración crítica.

Actividad 1.

Pese a no ser variables cuantitativas, tal como se ha estudiado, se analizan individualmente los valores de tendencia central de dos variables ordinales con valores de 0 a 10:

0: No es válido en absoluto. 10: Es totalmente válido.

Variable *fairelecc*. Sobre elecciones libres y limpias en Italia.

```
summary(df_ronda10$rec_fairelecc)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
##	0.000	4.000	6.000	5.739	8.000	10.000	99

Variable *cttresac*. Los tribunales en Italia tratan a todo el mundo por igual

```
summary(df_ronda10$rec_cttresac)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
##	0.000	3.000	5.000	4.522	6.000	10.000	73

Actividad 2.

Pese a no ser variables cuantitativas, tal como se ha estudiado, se analizan individualmente los valores de tendencia central dos variables ordinales con valores de 1 a 5:

1: Muy de acuerdo. 2: De acuerdo. 3: Ni de acuerdo ni en desacuerdo. 4: En desacuerdo. 5: Muy en desacuerdo. 10: Es totalmente válido.

Variable *freehms*. Los gays y las lesbianas deberían tener libertad para vivir como quieran

```
summary(df_ronda10$freehms)
```

##	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
##	1.00	1.00	2.00	2.05	2.00	5.00	65

Variable *hmsfmlsh*: Me daría vergüenza que un familiar cercano fuese gay o lesbiana

```
summary(df_ronda10$hmsfmlsh)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   NA's  
##    1.000   3.000   4.000   3.857   5.000   5.000    91
```

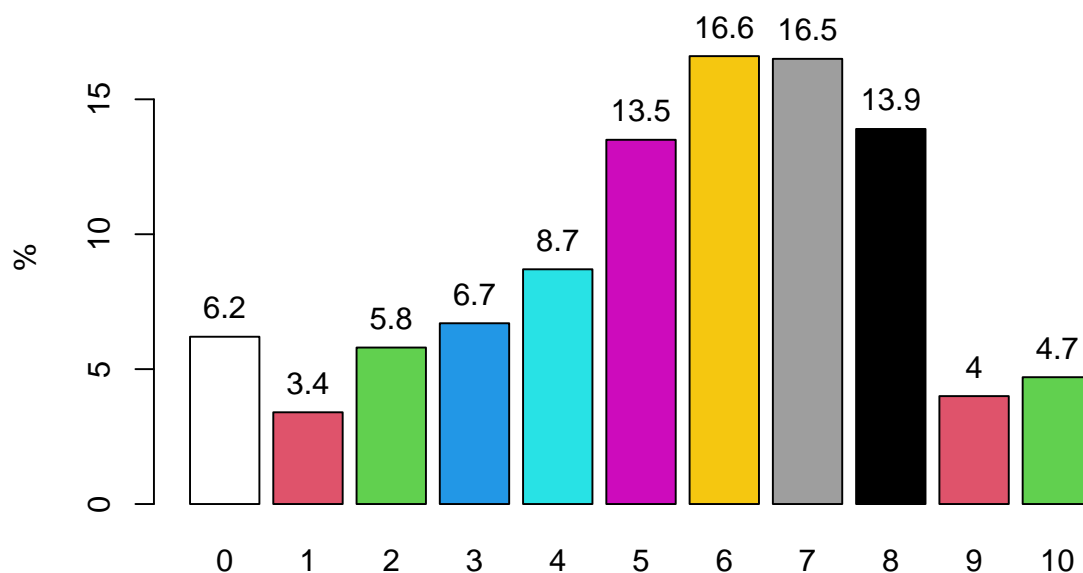
Actividad 3.

Se analiza la distribución de valores de una variable. Variable *medcrgvc*: Los medios de comunicación en Italia pueden criticar al gobierno.

0: No es válido en absoluto. 10: Es totalmente válido.

```
tablaGrafico <- tab1(df_ronda10$medcrgvc, cum.percent = TRUE,  
                      bar.values = "percent", missing = FALSE,  
                      xlab="% probabilidad", decimal=1,  
                      main="Gráfico 1. Capacidad de criticar")
```

Gráfico 1. Capacidad de criticar



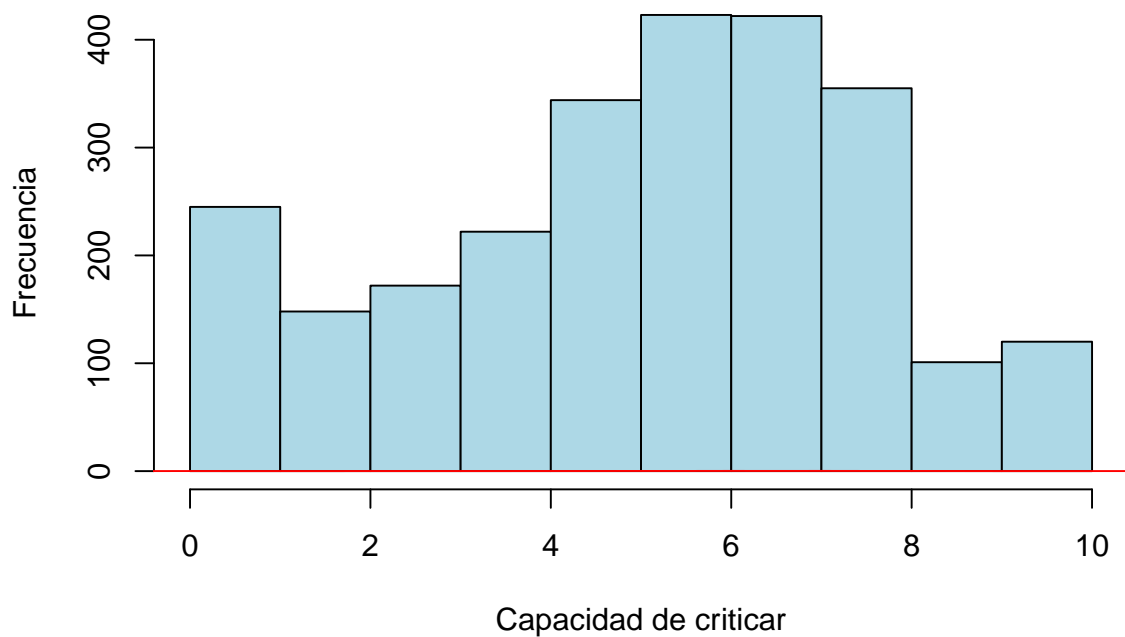
```
# Histograma
```

```
df_ronda10SinNA <- df_ronda10[complete.cases(
    df_ronda10$medcrgvc), ]

hist(df_ronda10$medcrgvc, col = "lightblue",
      xlab = "Capacidad de criticar",
      ylab = "Frecuencia",
      main = "Gráfico 2. Histograma de la Capacidad de criticar")

lines(density(df_ronda10SinNA$medcrgvc), col = "red", lwd = 1)
```







Gráfico 2. Histograma de la Capacidad de criticar



ANEXO 2. EJECUCIÓN DEL Rmd Y REPOSITORIO DE FICHEROS

Para la ejecución del código R, está disponible el fichero ejecutable Rmd y el fichero de datos *sav* en el repositorio GitHub. En el mismo repositorio se encuentra este PDF y un fichero HTML renderizados a partir del Rmd.

El fichero **datosprac2.sav** debe ubicarse en una subcarpeta del directorio de trabajo con nombre *DATA*.

	Enlace a GitHub	
	Fichero de datos tipo sav	datosprac2.sav
	Fichero Rmd	libertadesItalia2022.Rmd
	Fichero HTML	libertadesItalia2022.html
	Fichero PDF	libertadesItalia.pdf
	Repositorio tofermos	