SOC3070 Análisis de Datos Categóricos

Tarea corta 5

Ponderación: 6% de la nota final del curso. Entrega: Desde el momento de entrega, los estudiantes tienen 1 semana exacta de plazo para completar esta tarea.



Figure 1: NO

Problema:

En esta tarea usarás los datos de una encuesta realizada por FLACSO/Chile en Abril y Mayo de 1988 sobre intención de voto en el plebiscito de 1989, junto con otras variables socio-demográficas.

```
datos_chile <- datos_chile %>% mutate(vote = case_when(vote=="Y" ~ 0, vote=="N" ~ 1))
datos_chile %>% glimpse()
```

Rows: 2,700

En particular, trabajarás con el siguiente modelo de regresión logística que estima la probabilidad de votar NO en función de los ingresos, género (M=Hombre, F=Mujer) y el apoyo al status-quo (valores más altos indican mayor apoyo al régimen de Pinochet).

```
mymodel <- glm(vote ~ income + sex*statusquo, family=binomial, data = datos_chile )
summary(mymodel)</pre>
```

```
##
## Call:
  glm(formula = vote ~ income + sex * statusquo, family = binomial,
       data = datos_chile)
##
## Deviance Residuals:
##
      Min
                10
                     Median
                                   30
                                           Max
## -2.8454 -0.1965
                     0.1516
                               0.2754
                                        3.2390
##
## Coefficients:
##
                    Estimate Std. Error z value Pr(>|z|)
## (Intercept)
                  -8.026e-01 1.761e-01 -4.556 5.21e-06 ***
## income
                             2.416e-06
                                          2.588 0.00966 **
                   6.251e-06
## sexM
                   6.506e-01
                             2.067e-01
                                          3.147
                                                0.00165 **
## statusquo
                  -3.252e+00 2.153e-01 -15.101
                                                < 2e-16 ***
## sexM:statusquo 5.507e-02 2.900e-01
                                          0.190
                                                0.84940
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##
      Null deviance: 2368.68 on 1708 degrees of freedom
## Residual deviance: 719.63 on 1704 degrees of freedom
     (991 observations deleted due to missingness)
## AIC: 729.63
##
## Number of Fisher Scoring iterations: 6
```

1. Calcula el efecto marginal sobre la probabilida de votar por el NO de la variable "statusquo" para hombres y mujeres (por separado), fijando ingresos y apoyo al status-quo a sus respectivos valores medianos (en la muestra completa).

Nota: recuerda que dado la interacción sex*statusquo, el modelo contiene 2 efectos de statusquo.

Los efectos marginales correctos son los siguientes:

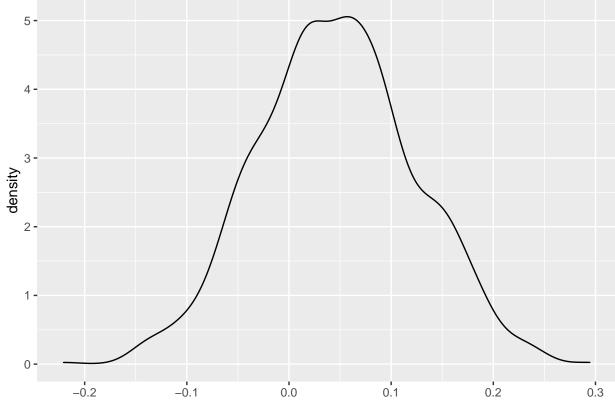
2 M 15000 -0.0456 -0.798

2. Usa el método de Bootstrap para crear un intervalo de confianza al 92% para la diferencia entre los efectos marginales de hombres y mujeres reportados en la pregunta 1.

El intervalo de confianza correcto es el siguiente:

4% 96% ## -0.08169914 0.18154377

La "bootstrap distribution" de la diferencia entre los efectos marginales de hombres y mujeres se ve así:



Dif. ME apoyo status-quo entre mujer y hombre mediano sobre propabilidad votar NO