



INSTITUTO DE SOCIOLOGÍA
FACULTAD DE CIENCIAS SOCIALES

Guía N°1

Análisis de Datos Multinivel - SOL3051

Estudiante [Andreas Laffert](#)

Profesora Camila Ortiz

Ayudante Andres González

jueves 12, septiembre 2024

Enunciado 1

Estime un modelo con confianza política como variable dependiente que le permita calcular la correlación intraclase, usando como unidad de agrupamiento la variable “país”. Reporte sus resultados, intérpretelos y argumente si en este caso se justifica aplicar modelamiento multinivel como estrategia de análisis. (6 pts).

La correlación intraclase (ICC) para el modelo nulo (no mostrado) es igual a 0.109, lo cual indica la cantidad de varianza de confianza política que puede atribuirse a la estructura de agrupación en la población, en este caso, los países (ICC en base a Hox et al., 2017, p. 13). Esto significa que cerca de un 11% de la varianza total de la confianza política de los individuos se asocia a características específicas de los países, asentando como precedente que existen diferencias significativas entre estos respecto al nivel de confianza política.

$$ICC = \frac{\sigma_{\mu_0}^2}{\sigma_{\mu_0}^2 + \sigma_{\epsilon}^2} = \frac{0.1376}{0.1376 + 1.1207} = 0.1094$$

El uso de modelos multinivel en este análisis se justifica tanto por razones sustantivas como metodológicas. En términos sustantivos, el coeficiente de correlación intraclase (ICC) del 11% indica que una parte significativa de la variabilidad en la confianza política de los individuos está asociada a las diferencias entre países, lo que hace necesario modelar esta variabilidad a nivel contextual para una comprensión más completa de la confianza política. Desde una perspectiva metodológica, ignorar la estructura jerárquica de los datos (individuos anidados en países) y emplear estimadores convencionales, como OLS, puede introducir sesgos. Como señalan Finch et al. (2014), aplicar métodos estadísticos convencionales a datos jerárquicos viola el supuesto de independencia de los residuos, lo que resulta en errores estándar mal estimados y valores p más pequeños de lo apropiado, incrementando el riesgo de cometer un error de tipo I.

Enunciado 2

Estime tres modelos usando confianza política como variable dependiente: i) En primer lugar, un modelo con intercepto aleatorio que sólo incluye efectos fijos de nivel individual de las variables sexo, edad, educación, empleado, casado, raza e ideología. ii) Luego, un modelo con la misma estructura fija que el modelo anterior pero que agrega efectos aleatorios de la variable ideología. iii) Por último, un modelo que añade a lo anterior una interacción cruzada entre ideología y el indicador de Presidencia de Izquierda.

- Reporte sus resultados en una tabla en formato académico (2 puntos).
- En base al primer modelo, grafique los interceptos aleatorios. ¿Qué se puede concluir de estos resultados? (6 puntos).

En la Figura 1 se muestran los interceptos aleatorios por país para el Modelo 1 que incluye los predictores de nivel individual. Los resultados indican que la mayoría de los países, con

Table 1: Modelos multinivel para confianza política, ideología política y países con presidencia de izquierda

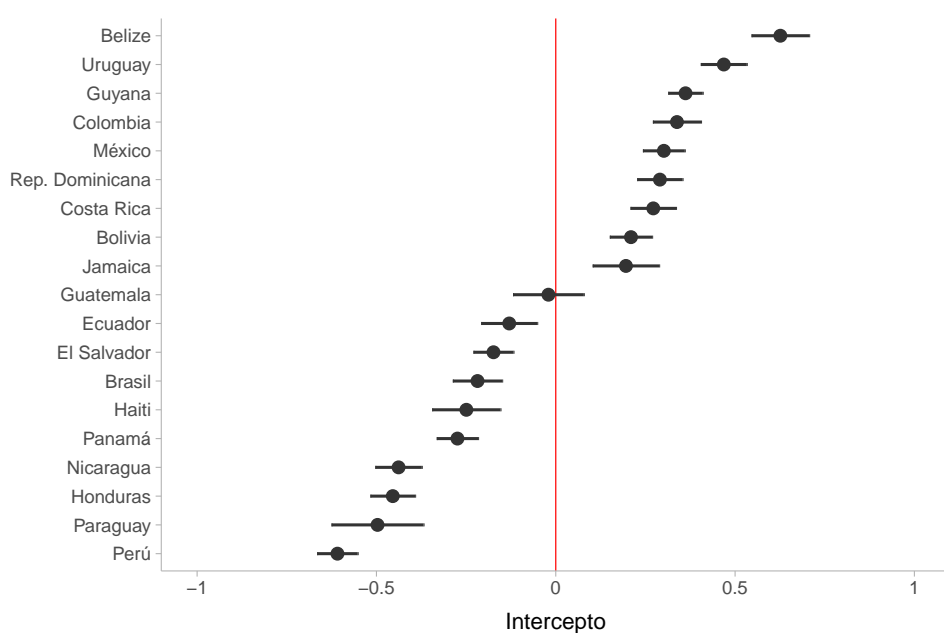
	Modelo 1	Modelo 2	Modelo 3
Intercepto	0.43*** (0.09)	0.34** (0.12)	0.46*** (0.13)
Mujer (Ref.= Hombre)	-0.03 (0.02)	-0.03 (0.02)	-0.03 (0.02)
Edad	0.00 (0.01)	-0.00 (0.01)	-0.00 (0.01)
Nivel educacional (en años)	-0.01*** (0.00)	-0.01** (0.00)	-0.01** (0.00)
Empleado (Ref.= Desempleado)	-0.06** (0.02)	-0.06*** (0.02)	-0.06*** (0.02)
Casado (Ref.= Otro)	0.05** (0.02)	0.05** (0.02)	0.05** (0.02)
Blanco (Ref.= Otro)	-0.01 (0.02)	-0.02 (0.02)	-0.02 (0.02)
Ideología política	-0.07*** (0.00)	-0.05*** (0.01)	-0.07*** (0.02)
Presidencia Izquierda (Ref. = Otra)			-0.42 (0.25)
Ideología política x Presidencia Izquierda (Ref. = Otra)			0.07* (0.03)
AIC	52879.32	52533.14	52538.58
BIC	52957.31	52626.73	52647.77
Log-likelihood	-26429.66	-26254.57	-26255.29
Num. obs	18012	18012	18012
Num. grupos: Países	19	19	19
Var: Países (Intercepto)	0.14	0.24	0.22
Var: Residual	1.09	1.07	1.07
Var: Países Ideología		0.00	0.00
Cov: Países (Intercepto), Ideología		-0.02	-0.01

Nota: Celdas contienen coeficientes de regresión con errores estándares entre paréntesis. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Fuente: Elaboración propia en base a LAPOP 2008.

la excepción de Jamaica, presentan niveles promedio de confianza política que se desvían significativamente del gran intercepto (γ_{00}). Esto se confirma en la parte aleatoria del Modelo 1 (Tabla 1), donde la varianza entre países en torno al intercepto ($\sigma_{\mu_0}^2$) es cercana al 14%. Así, países como Perú y Paraguay se desvían negativamente del promedio general, reflejando niveles más bajos de confianza política, mientras que países como Belize y Uruguay se desvían positivamente, indicando mayores niveles de confianza política en promedio.

Figure 1: Interceptos aleatorios por país



Fuente: Elaboración propia en base a LAPOP 2008

c) En base al primer modelo, interprete el efecto de ser casado y de la ideología (5 puntos).

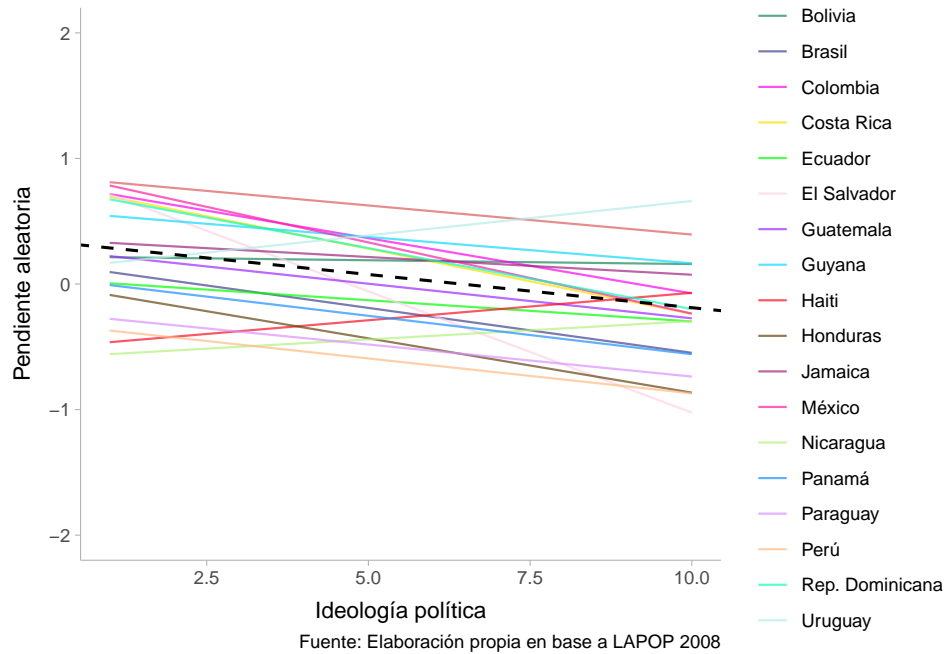
En la Tabla 1 se muestran los resultados de los modelos multinivel para la confianza política del enunciado 2. Los resultados del Modelo 1 indican que el efecto de estar casado es positivo y estadísticamente significativo ($\beta = 0.05$, $p < 0.01$). Específicamente, *ceteris paribus*, las personas casadas obtienen, en promedio, 0.05 puntos adicionales en la escala de confianza política en comparación con quienes no lo están, siendo este efecto significativo al 99% de confianza. Por otro lado, el efecto de la ideología política, medida en una escala de 1 a 10 (donde 1 representa posiciones más cercanas a la derecha y 10 a la izquierda), es negativo y estadísticamente significativo con un 99.9% de confianza. Esto implica que, por cada unidad que aumenta la escala de ideología, la confianza política disminuye en promedio 0.07 puntos, manteniendo constantes los demás predictores.

d) En base al segundo modelo, obtenga las pendientes aleatorias correspondientes al efecto de la ideología y genere manualmente un gráfico que permita visualizar la heterogeneidad de estos efectos. Para ello, mantenga constantes las variables continuas en su media y las variables dicotómicas en la moda muestral. Interprete sustantivamente estos resultados (6 puntos).

La Figura 2 presenta las pendientes aleatorias de la ideología política para cada país, junto con una línea punteada que representa la pendiente fija de esta variable en el Modelo 2. A partir de los resultados de este modelo, se observa que, aunque el efecto fijo de la ideología política *entre* países es negativo ($\beta = -0.05$, $p < 0.001$), al descomponer este efecto *dentro* de los países se evidencia una heterogeneidad en su dirección. Por ejemplo, en países como El Salvador y Honduras el efecto aleatorio de la ideología es negativo, mientras que en otros, como Uruguay y Haití, es positivo. Estos resultados sugieren que el efecto de la ideología (es decir, ser de

izquierda) sobre la confianza política varía entre países; en algunos se asocia con una menor confianza, mientras que en otros con una mayor, lo que releva las diferencias contextuales entre países (que serán objeto de modelamiento en análisis siguientes).

Figure 2: Pendientes aleatorias de ideología política por país



e) Calcule e interprete la varianza explicada por el modelo 2. (5 puntos).

Siguiendo a Snijders y Bosker (2012), dado que el Modelo 2 incluye efectos aleatorios de la variable ideología, se utiliza la estimación de la varianza explicada por el Modelo 1 para calcular la varianza explicada por el Modelo 2.

A nivel individual:

$$R_1^2 = \frac{\sigma_\epsilon^2 | \text{NULO} - \sigma_\epsilon^2 | \text{COMPLETO}}{\sigma_\epsilon^2 | \text{NULO}} = \frac{1.1207 - 1.0932}{1.1207} = 0.025$$

A nivel contextual:

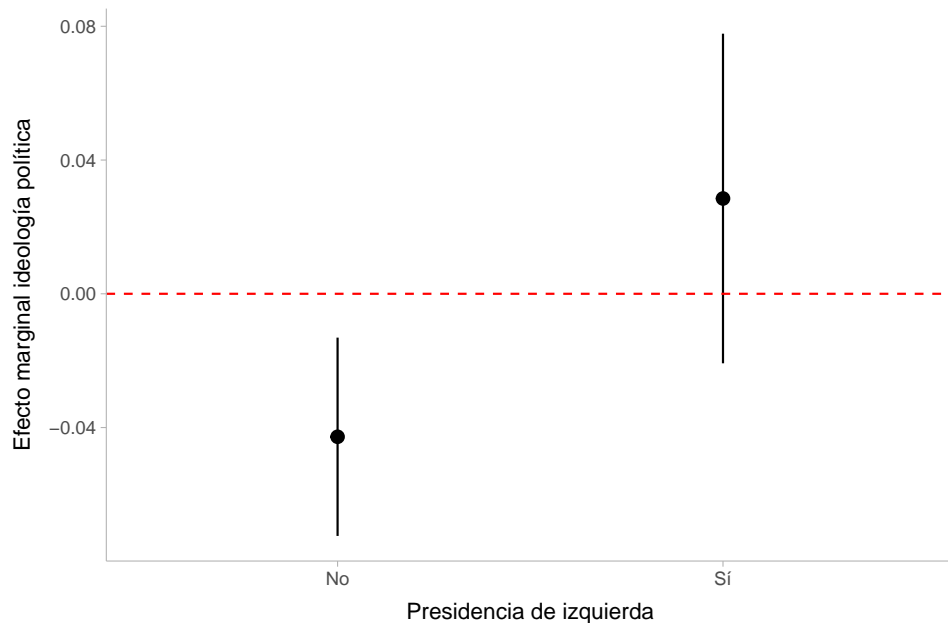
$$R_2^2 = \frac{\sigma_{\mu_0}^2 | \text{NULO} - \sigma_{\mu_0}^2 | \text{COMPLETO}}{\sigma_{\mu_0}^2 | \text{NULO}} = \frac{0.1376 - 0.1369}{0.1376} = 0.005$$

Los resultados indican que el R^2 a nivel individual (*within*) es de 0.025, mientras que el de nivel contextual (*between*) es de 0.005. De este modo, los predictores individuales del Modelo 1 explican un 2.5% de la varianza de la confianza política a nivel individual. Por su parte, este modelo solo logra explicar el 0.5% de la varianza a nivel contextual, lo que corresponde al 0.5% del 11% indicado por la correlación intraclase. Con todo, este no es un buen modelo para dar cuenta de la varianza de la confianza política.

- f) Tomando como referencia el tercer modelo, obtenga los efectos marginales de ideología moderado por “presidencia de izquierda”. Interprete sustantivamente estos efectos (6 puntos).

En la Figura 3 se muestran los efectos marginales de la ideología moderada por la presencia de presidencia de izquierda en el país, según el Modelo 3. Los resultados indican que, cuando el país cuenta con una presidencia de izquierda (leftpres = 1), el efecto marginal de la ideología es positivo, pero no es estadísticamente significativo al 95% de confianza (estimación = 0.00243, $p = 0.923$). En contraste, cuando el país no tiene una presidencia de izquierda (leftpres = 0), el efecto marginal de la ideología es negativo y estadísticamente significativo (estimación = -0.08163, $p < 0.001$).

Figure 3: Efecto marginal de ideología moderado por presidencia de izquierda



Los efectos marginales se pueden expresar con las siguientes ecuaciones:

1. Cuando el país tiene una presidencia de izquierda (leftpres = 1):

$$\left(\frac{\partial \text{Confianza política}}{\partial \text{Ideología}} \mid \text{Leftpres}=1 \right) = \beta_1 \text{Ideología} + \beta_3 \text{Ideología} \times \text{Leftpres}$$

Donde el efecto marginal estimado de la ideología es 0.00243 con un valor p de 0.923, el efecto marginal de la ideología es positivo pero no significativo al 95% de confianza.

2. Cuando el país no tiene una presidencia de izquierda (leftpres = 0):

$$\left(\frac{\partial \text{Confianza política}}{\partial \text{Ideología}} \mid \text{Leftpres}=0 \right) = \beta_1 \text{Ideología}$$

Aquí, el efecto marginal estimado de la ideología es -0.08163 con un valor $p < 0.001$, lo que indica un efecto negativo y estadísticamente significativo.

Con todo, esto sugiere que la influencia de la ideología sobre la confianza política varía dependiendo de la presencia de una presidencia de izquierda, siendo significativa solo en contextos sin dicha presidencia.

Enunciado 3

Contraste el modelo 1 estimado en la pregunta 2) con un modelo con la misma estructura fija pero que incorpora efectos aleatorios de la variable edad (modelo 4) y reporte sus resultados en una tabla. Formule la hipótesis apropiada para evaluar si se justifica la inclusión de este parámetro y realice el test estadístico correspondiente. Interprete sus resultados (6 puntos).

En la Tabla 2 se reportan los estadísticos del análisis de devianza (likelihood ratio test) para el contraste entre el Modelo 1 y 4. La prueba de razón de verosimilitud se utiliza para comparar modelos anidados y determinar cuál se ajusta mejor a los datos (Peugh, 2010).

La hipótesis nula para esta prueba es que el efecto aleatorio de la edad es igual a cero, mientras que la hipótesis alternativa sostiene que el efecto aleatorio es distinto de cero. Formalmente, las hipótesis son:

$$H_0 : \sigma_{edad}^2 = 0$$

$$H_A : \sigma_{edad}^2 \neq 0$$

Los resultados del análisis muestran que la prueba de razón de verosimilitud es estadísticamente significativa, $\chi^2(2) = 75.918$, $p < 0.001$. Esto indica que el Modelo 4, que incluye efectos aleatorios para la edad, se ajusta significativamente mejor a los datos que el Modelo 1, que solo considera interceptos aleatorios. Además, el Modelo 4 presenta valores menores de AIC y BIC en comparación con el Modelo 1, lo que refuerza la idea de un mejor ajuste (Hox et al., 2017). Con todo, los resultados sugieren que es posible rechazar la H_0 y entregan evidencia favorable a la H_A respecto a que el efecto aleatorio de la edad es diferente de cero y mejora el ajuste.

Table 2: Estadísticos de bondad de ajuste

Modelo	AIC	BIC	Deviance	X2	Df	p-value
Modelo 1	52822.33	52900.32	52802.33			
Modelo 4	52750.41	52844.00	52726.41	75.92	2	< 0.001 ***

Enunciado 4

Una de las hipótesis que se trabajan en el artículo de Morgan y Buice (2013) es que el peso de las mujeres en la fuerza laboral impacta de forma positiva el apoyo a que las mujeres ejerzan roles de liderazgo político. Evalúe esta hipótesis pero utilizando confianza política como variable dependiente. Para ello, estime un modelo en base al modelo 4 pero agregando un efecto fijo de la variable “Participación femenina en la fuerza laboral” (FLP). Reporte sus resultados en una tabla. Calcule un test para evaluar la significancia estadística de este efecto y señale a qué nivel de confianza sería estadísticamente significativo (7 puntos).

En la Tabla 3 se muestran los resultados de los modelos multinivel para la confianza política, incluyendo el Modelo 5, que incorpora un efecto fijo de la participación femenina en la fuerza laboral (FLP). Los resultados del Modelo 5 indican que, aunque el efecto de la participación laboral femenina es positivo ($\beta = 0.61$), no es estadísticamente significativo al 95% de confianza ($SE = 0.73$, $p > 0.05$), manteniendo los demás predictores constantes.

Para evaluar la significancia estadística de este coeficiente, se realiza un test de Wald. La hipótesis nula (H_0) es que el coeficiente de la participación laboral femenina es igual a cero:

$$H_0 : \beta_{FLP} = 0$$

El valor z se calcula como el cociente entre el coeficiente estimado y su error estándar:

$$z = \frac{\beta_{FLP}}{SE_{FLP}} = \frac{0.61}{0.73} = 0.83$$

Este valor z se compara con el valor crítico de una distribución normal estándar, que es 1.96 para un nivel de significancia $\alpha = 0.05$. Dado que el valor z estimado (0.8356) es menor que el valor crítico de 1.96 y el valor p asociado (0.4163) es mayor que 0.05, no se rechaza la hipótesis nula. Por lo tanto, con un 95% de confianza, no hay evidencia estadísticamente significativa para apoyar la hipótesis de que la participación laboral femenina tenga un efecto en la confianza política.

Para determinar el nivel de confianza en el que el coeficiente sería estadísticamente significativo, usamos el valor $z = 0.8356$, que corresponde a un nivel de significancia (α) de 0.2027 o 20.27%. El nivel de confianza vendría siendo $1 - \alpha$, lo que resuelve como $1 - 0.2027 = 0.7973$. Por tanto, el efecto de la participación laboral femenina sería estadísticamente significativo solo con un nivel de confianza del 79.73%, lo cual es inferior al nivel de confianza convencional del 95%.

Enunciado 5

En base al modelo 4, especifique un modelo que le permita evaluar si el efecto de la edad es moderado por el Índice de democracia a nivel país (Freedom House Index, fhouse).

Table 3: Modelos multinivel para confianza política, edad e índice de democracia del país

	Modelo 4	Modelo 5	Modelo 6
Intercepto	0.41*** (0.08)	0.08 (0.41)	0.61* (0.24)
Mujer (Ref.= Hombre)	-0.02 (0.02)	-0.02 (0.02)	-0.02 (0.02)
Edad	0.00 (0.01)	0.00 (0.01)	0.09* (0.04)
Nivel educacional (en años)	-0.01*** (0.00)	-0.01*** (0.00)	-0.01*** (0.00)
Empleado (Ref.= Desempleado)	-0.05* (0.02)	-0.05* (0.02)	-0.04* (0.02)
Casado (Ref.= Otro)	0.05** (0.02)	0.05** (0.02)	0.05** (0.02)
Blanco (Ref.= Otro)	-0.02 (0.02)	-0.02 (0.02)	-0.02 (0.02)
Ideología política	-0.07*** (0.00)	-0.07*** (0.00)	-0.07*** (0.00)
Participación laboral femenina		0.61 (0.73)	
Índice Freedom House			-0.08 (0.09)
Edad x Índice Freedom House			-0.04** (0.01)
AIC	52805.69	52805.82	52812.29
BIC	52899.27	52907.20	52921.47
Log-likelihood	-26390.84	-26389.91	-26392.14
Num. obs	18012	18012	18012
Num. grupos: Países	19	19	19
Var: Países (Intercepto)	0.10	0.11	0.10
Var: Países Edad	0.00	0.00	0.00
Cov: Países (Intercepto), Edad	0.00	0.00	0.00
Var: Residual	1.09	1.09	1.09

Nota: Celdas contienen coeficientes de regresión con errores estándares entre paréntesis. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Fuente: Elaboración propia en base a LAPOP 2008.

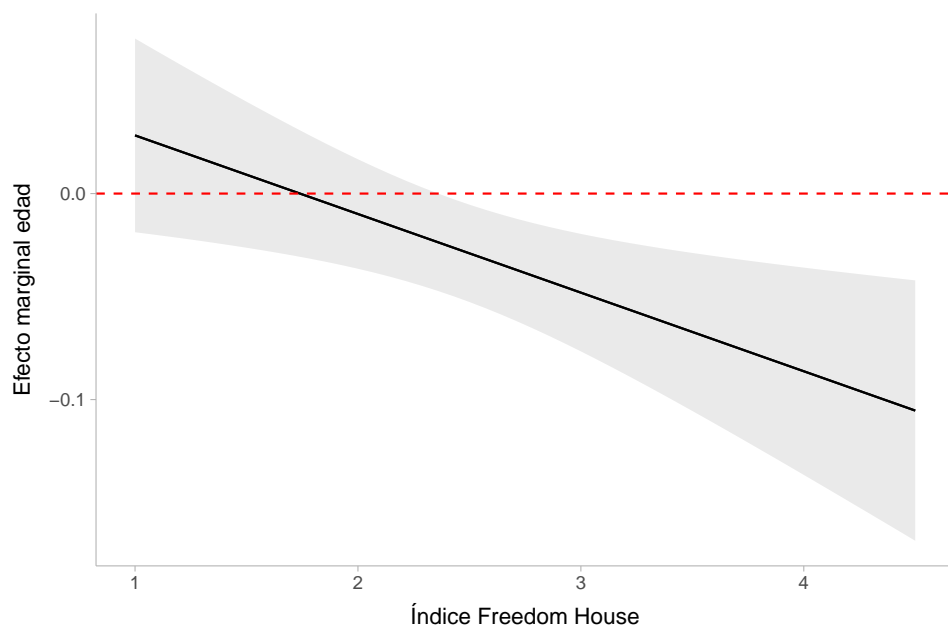
a) Reporte e interprete sustantivamente el intercepto en este modelo (4 puntos)

En la Tabla 3 se muestran los resultados de los modelos multinivel para la confianza política incorporando el Modelo 6 de este enunciado. De acuerdo con este modelo, el intercepto de este modelo tiene un efecto positivo y estadísticamente significativo ($\beta = 0.61$, $p < 0.05$). Este coeficiente representa la confianza política promedio para un individuo típico de referencia, es decir, cuando todas las demás variables predictoras (como sexo, edad, nivel educativo, empleo, casado, raza, ideología, y el índice de democracia) están en sus valores de referencia o cero. En este caso, los valores de referencia serían: hombre, desempleado, no casado, de otra raza (no blanco), con ideología 0, y en un país con el nivel más bajo de democracia (fhouse = 0).

Además, el efecto de interacción entre la edad y el Índice de democracia también afecta la

interpretación del intercepto, puesto que, sumado a lo anterior, el impacto de la edad depende del nivel de democracia de los países. Esto significa que, en países con mayores niveles de democracia, el impacto positivo de la edad sobre la confianza política será menor que en países con menores niveles de democracia, debido a la interacción negativa y estadísticamente significativa entre edad y fhouse ($\beta = -0.038$, $p < 0.01$).

Figure 4: Efecto marginal de edad moderado por índice de democracia del país



Fuente: Elaboración propia en base a LAPOP 2008

- b) Reporte sus resultados a través de un gráfico e interprete sus resultados. ¿Qué se puede concluir en relación al efecto moderador del grado de democracia a nivel país sobre el efecto de la edad? (7 puntos).

En la Figura 4 se presentan los efectos marginales de la edad moderados por el nivel de democracia del país, según el Modelo 6. Los resultados sugieren que el efecto marginal de la edad varía significativamente de acuerdo con el nivel del Índice Freedom House (fhouse). En la Tabla 4 se muestran los efectos marginales estimados para cada valor del Índice Freedom House.

Table 4: Efectos marginales de edad para valores de Índice Freedom House

Fhouse	Estimación	Error Estándar	Estadístico	p-value
1.0	0.0282434	0.0239885	1.1773724	
1.5	0.0091603	0.0180875	0.5064462	
2.0	-0.0099227	0.0135610	-0.7317095	
2.5	-0.0290058	0.0120586	-2.4053980	< 0.05*
3.0	-0.0480889	0.0145527	-3.3044549	< 0.001***
3.5	-0.0671719	0.0195695	-3.4324843	< 0.001***
4.5	-0.1053381	0.0322456	-3.2667441	< 0.01**

De este modo, en países con bajos niveles de democracia ($f_{\text{house}} = 1.0$), el efecto de la edad es positivo pero no estadísticamente significativo ($\beta = 0.028$, $p = 0.239$). Sin embargo, a medida que aumenta el nivel de democracia ($f_{\text{house}} = 2.5$ a 4.5), el efecto marginal de la edad se torna negativo y estadísticamente significativo. Por ejemplo, cuando f_{house} es 3.0 , el efecto es -0.048 ($p < 0.001$) y se vuelve más negativo a medida que el nivel de democracia aumenta, alcanzando -0.105 en f_{house} igual a 4.5 ($p < 0.01$).

Esto sugiere que en países con altos niveles de democracia, el impacto de la edad en la confianza política es negativo y más pronunciado, mientras que en países con bajos niveles de democracia, este efecto es insignificante. En definitiva, el grado de democracia modera significativamente el efecto de la edad sobre la confianza política, haciendo que este efecto sea más negativo en contextos más democráticos.

Referencias

- Finch, W. H., Bolin, J. E., & Kelley, K. (2014). *Multilevel Modeling Using R*.
- Hox, J., Moerbeek, M., & Schoot, R. van de. (2017). *Multilevel Analysis: Techniques and Applications, Third Edition* (3rd ed.). New York: Routledge. <https://doi.org/10.4324/9781315650982>
- Morgan, J., & Buice, M. (2013). Latin American Attitudes toward Women in Politics: The Influence of Elite Cues, Female Advancement, and Individual Characteristics. *American Political Science Review*, 107(4), 644–662. <https://doi.org/10.1017/S0003055413000385>
- Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, 48(1), 85–112. <https://doi.org/10.1016/j.jsp.2009.09.002>
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2. ed). Los Angeles, Calif.: SAGE.

Código de R

```
knitr::opts_chunk$set(echo = F,
                      warning = F,
                      error = F,
                      message = F)
if (! require("pacman")) install.packages("pacman")

pacman::p_load(tidyverse,
              sjmisc,
              sjPlot,
              lme4,
              easystats,
              influence.ME,
              broom.mixed,
              here,
              texreg,
              ggeffects,
              marginaleffects,
              naniar,
              ggdist,
              Polychrome,
              misty,
              kableExtra,
              sjlabelled)

options(scipen=999)
rm(list = ls())

miles <- function(x) {
  format(round(as.numeric(x), 0), big.mark = ".")
}

decimales <- function(x) {
  format(round(as.numeric(x), 2), decimal.mark = ",")
}

# set theme

theme_set(theme_ggdist())

options(knitr.kable.NA = "")
```

```

options(knitr.table.format="latex")

load(file = here("input/data/morgan2013.RData"))

names(morgan2013)
glimpse(morgan2013)

# seleccionar ----

db <- morgan2013 %>%
  dplyr::select(country, ID, trustgov, sex = female, age, educ, employed,
  sjlabelled::remove_all_labels() %>%
  janitor::clean_names() %>%
  as_tibble()

# filtrar: no -----

# recodificar y transformar ----

# trust
sjmisc::descr(db$in_trust)

# sexo
frq(db$sex)

db$sex <- car::recode(db$sex,
                      recodes= c("0='Hombre';1='Mujer'"),
                      levels = c("Hombre", "Mujer"),
                      as.factor = T)

# edad
sjmisc::descr(db$age)
frq(db$age)

db$age_f <- car::recode(db$age,
                      recodes = c("1='Tramo 1';
                                2='Tramo 2';
                                3='Tramo 3';
                                4='Tramo 4';
                                5='Tramo 5';
                                6='Tramo 6'"),

```

```

        levels = c("Tramo 1",
                    "Tramo 2",
                    "Tramo 3",
                    "Tramo 4",
                    "Tramo 5",
                    "Tramo 6"),
        as.factor = T
    )

# educ
sjmisc::descr(db$educ)

# employed
frq(db$employed)

db$employed <- car::recode(db$employed,
                           recodes= c("0='Desempleado';1='Empleado'"),
                           levels = c("Desempleado", "Empleado"),
                           as.factor = T)

# married
frq(db$married)

db$married <- car::recode(db$married,
                           recodes= c("0='No';1='Sí'"),
                           levels = c("No", "Sí"),
                           as.factor = T)

# race
frq(db$race)

db$race <- car::recode(db$race,
                       recodes= c("0='Otro';1='Blanco'"),
                       levels = c("Otro", "Blanco"),
                       as.factor = T)

# ideology
frq(db$ideology)

# left
frq(db$leftpres)

```

```

db$leftpres <- car::recode(db$leftpres,
                           recodes= c("0='No';1='Sí'"),
                           levels = c("No","Sí"),
                           as.factor = T)

# flp
sjmisc::descr(db$flp)

# fhouse
sjmisc::descr(db$fhouse)

# id
sjmisc::descr(db$id)

# country
frq(db$country)

# casos perdidos -----

colSums(is.na(db))

n_miss(db)

prop_miss(db)*100

miss_var_summary(db)

miss_var_table(db)

vis_miss(db) + theme(axis.text.x = element_text(angle=80))

db <- na.omit(db)

# Null model
model_0 <- lmer(in_trust ~ 1 + (1 | country),
                data = db, REML = T)

performance::icc(model_0, by_group = T)
## ICC Country = 0.11

```

```

sigma_e <- sigma(model_0)^2
tau_mu <- VarCorr(model_0)$country[[1]]
cic <- tau_mu / (tau_mu+sigma_e)

# Influence test
inf_m0 <- influence(model_0, group = "country")

# D cook
cooks.distance(inf_m0, parameters = 1, sort = T) # cut point is 4/19

n_country <- length(unique(db$country))

plot(inf_m0, which="cook",
      cutoff=(4/n_country), sort=TRUE,
      xlab="Distancia de Cook",
      ylab="País", width=60, height=40)

# no obs influyentes

# Modelo 1: Indicadores N1
model_1 <- lmer(in_trust ~ 1 + sex + age + educ + employed + married +
               race + ideology + (1 | country),
               data = db,
               REML = T)

# Modelo 2: Pendiente aleatoria ideology
model_2 <- lmer(in_trust ~ 1 + sex + age + educ + employed + married +
               race + ideology + (1 + ideology | country),
               data = db,
               REML = T)

# Modelo 3: Interaccion ideology y leftpres
model_3 <- lmer(in_trust ~ 1 + sex + age + educ + employed + married +
               race + ideology + leftpres + ideology*leftpres +
               (1 + ideology | country),
               data = db,
               REML = T)

# Modelo 4: Pendiente aleatoria edad
model_4 <- lmer(in_trust ~ 1 + sex + age + educ + employed + married +
               race + ideology + (1 + age | country),

```



```

        data = db,
        REML = T)

# Modelo 5: Pendiente aleatoria edad + flp
model_5 <- lmer(in_trust ~ 1 + sex + age + educ + employed + married +
               race + ideology + flp + (1 + age| country),
               data = db,
               REML = T)

# Modelo 6: Pendiente aleatoria edad + flp + interaccion fhouse
model_6 <- lmer(in_trust ~ 1 + sex + age + educ + employed + married +
               race + ideology + age*fhouse + (1 + age| country),
               data = db,
               REML = T)

cccoef <- list(
  "(Intercept)" = "Intercepto",
  sexMujer = "Mujer (Ref.= Hombre)",
  age = "Edad",
  educ = "Nivel educacional (en años)",
  employedEmpleado = "Empleado (Ref.= Desempleado)",
  marriedSí = "Casado (Ref.= Otro)",
  raceBlanco = "Blanco (Ref.= Otro)",
  ideology = "Ideología política",
  leftpresSí = "Presidencia Izquierda (Ref. = Otra)",
  "ideology:leftpresSí" = "Ideología política x Presidencia Izquierda (Re

texreg::texreg(list(model_1, model_2, model_3),
               custom.model.names = c("Modelo 1",
                                     "Modelo 2",
                                     "Modelo 3"),
               caption = paste("(\\#tab:table1)", "Modelos multinivel para
               stars = c(0.05, 0.01, 0.001),
               custom.coef.map = ccoef,
               custom.note = "\\item Nota: Celdas contienen coeficientes
               threeparttable = T,
               leading.zero = T,
               float.pos = "h!",
               use.packages = F,

```

```

booktabs = TRUE,
scalebox = 0.9,
custom.gof.names = c("AIC",
                      "BIC",
                      "Log-likelihood",
                      "Num. obs",
                      "Num. grupos: Países",
                      "Var: Países (Intercepto)",
                      "Var: Residual",
                      "Var: Países Ideología",
                      "Cov: Países (Intercepto), Ideología"
                      ))

sjPlot::plot_model(model_1,
                   type = "re",
                   vline.color = "red",
                   grid = F,
                   sort.est = "sort.all",
                   ci.lvl = .95,
                   colors = "grey20") +
labs(title = NULL,
     y = "Intercepto",
     caption = "Fuente: Elaboración propia en base a LAPOP 2008")+
theme_ggdist()

stats_m2 <- broom.mixed::tidy(model_2)

ideology_rang <- seq(1:10)
in_trust2 <- as.data.frame(sapply(ideology_rang, function(x) fixef(model_2)
                                                                    fixef(model_2) ["sexMujer"] * 0 +
                                                                    fixef(model_2) ["age"] * 2.74 +
                                                                    fixef(model_2) ["educ"] * 9.23 +
                                                                    fixef(model_2) ["employedEmpleado"] * 1
                                                                    fixef(model_2) ["marriedSí"] * 1 +
                                                                    fixef(model_2) ["raceBlanco"] * 0 +
                                                                    fixef(model_2) ["ideology"] * x +
                                                                    ranef(model_2)$country[, 1] +
                                                                    ranef(model_2)$country[, 2] * x))

in_trust2$country <- seq(1:19)
in_trust2 <- reshape(in_trust2,

```

```

        direction = "long", #formato long
        v.names = "pred",
        varying = list(names(in_trust2)[1:10]),
        idvar = c("country"),
        timevar = "ideology")

colores <- setNames(createPalette(19, c("#E16462", "#177b4b", "#0D0887")))

in_trust2 %>%
  mutate(country = factor(country,
                           levels = 1:19,
                           labels = levels(db$country))) %>%
  ggplot(aes(x = ideology, y = pred, group = country, color = country)) +
  geom_line(alpha = 0.7) +
  ylim(-2,2) +
  geom_abline(intercept = stats_m2$estimate[stats_m2$term == "(Intercept)",
        slope = stats_m2$estimate[stats_m2$term == "ideology"],
        color = "black",
        linetype = "dashed",
        linewidth = 0.7) +
  scale_color_manual(values = colores) +
  labs(y = "Pendiente aleatoria",
        x = "Ideología política",
        color = "País",
        caption = "Fuente: Elaboración propia en base a LAPOP 2008")

# manual

varcomp_0 <- as.data.frame(VarCorr(model_0))
tau00_0 <- varcomp_0[1,4]
sigma2_0 <- varcomp_0[2,4]

varcomp_1 <- as.data.frame(VarCorr(model_1))
tau00_1 <- varcomp_1[1,4]
sigma2_1 <- varcomp_1[2,4]

R2_1_L1 <- (sigma2_0-sigma2_1)/sigma2_0
R2_1_L1

```

```

R2_2_L1 <- (tau00_0-tau00_1)/tau00_0
R2_2_L1

# misty
db_n <- db %>%
  mutate(
    across(.cols = c(sex, age, employed, married, race),
           .fns = ~ as.numeric(.))
  )

model_1_n <- lmer(in_trust ~ 1 + sex + age + educ + employed + married +
  race + ideology + (1 | country),
  data = db_n,
  REML = T)

misty::multilevel.r2(model = model_1_n, print = "all")

plot_slopes(model_3,
  variables = "ideology",
  condition = "leftpres",
  conf_level = .95) +
  geom_hline(yintercept = 0,
    color = "red",
    linetype = "dashed") +
  labs(y = "Efecto marginal ideología política",
    x = "Presidencia de izquierda",
    caption = "Fuente: Elaboración propia en base a LAPOP 2008",
    title = NULL)

#performance::test_likelihoodratio(model_1, model_4)

res_fit1 <- anova(model_1, model_4)

fit_tab <- res_fit1[c(2,3,5,6,7,8)] %>% as.tibble(.)

fit_tab$mod <- c("Modelo 1", "Modelo 4")

fit_tab$p <- if_else(fit_tab$`Pr(>Chisq)` < 0.001, "< 0.001 ***", NA)

```

```

fit_tab$Chisq <- round(fit_tab$Chisq, digits = 2)

fit_tab <- fit_tab %>%
  select(mod, AIC, BIC, deviance, Chisq, Df, p)

colnames <- c("Modelo", "AIC", "BIC", "Deviance", "X2", "Df", "p-value")

fit_tab <- kableExtra::kable(fit_tab,
                             format = "latex",
                             col.names = colnames,
                             row.names = F,
                             booktabs = T,
                             caption = paste("\\#tab:table3)", "Estadísti
kableExtra::kable_styling(latex_options = "hold_position",
                           position = "center") %>%
kableExtra::kable_styling(bootstrap_options = c("striped", "hover", "co
column_spec(1, width = "2cm") %>%
row_spec(0, bold = T)

fit_tab

ccoef <- list(
  "(Intercept)" = "Intercepto",
  sexMujer = "Mujer (Ref.= Hombre)",
  age = "Edad",
  educ = "Nivel educacional (en años)",
  employedEmpleado = "Empleado (Ref.= Desempleado)",
  marriedSí = "Casado (Ref.= Otro)",
  raceBlanco = "Blanco (Ref.= Otro)",
  ideology = "Ideología política",
  flp = "Participación laboral femenina",
  fhouse = "Índice Freedom House",
  "age:fhouse" = "Edad x Índice Freedom House"
)

texreg::texreg(list(model_4, model_5, model_6),
                custom.model.names = c("Modelo 4",
                                       "Modelo 5",
                                       "Modelo 6"),

```

```

caption = paste("(\\#tab:table2)", "Modelos multinivel para
stars = c(0.05, 0.01, 0.001),
custom.coef.map = ccoef,
custom.note = "\\item Nota: Celdas contienen coeficientes
threeparttable = T,
leading.zero = T,
float.pos = "h!",
use.packages = F,
booktabs = TRUE,
scalebox = 0.9,
custom.gof.names = c("AIC",
                      "BIC",
                      "Log-likelihood",
                      "Num. obs",
                      "Num. grupos: Países",
                      "Var: Países (Intercepto)",
                      "Var: Países Edad",
                      "Cov: Países (Intercepto), Edad",
                      "Var: Residual"

                      ))

n2 <- length(unique(db$country))

z_fpl <- (fixef(model_5)[9])/(sqrt(diag(vcov(model_5))[9]))

pval_flp <- 2 * pt(-abs(z_fpl), df=n2-1-1, lower.tail = T)

model_parameters(model_5, ci_method = "wald", summary = T)

plot_slopes(model_6,
             variables = "age",
             condition = "fhouse",
             conf_level = .95) +
geom_hline(yintercept = 0.0,
           color = "red",
           linetype = "dashed") +
labs(y = "Efecto marginal edad",
     x = "Índice Freedom House",
     caption = "Fuente: Elaboración propia en base a LAPOP 2008",

```

```

    title = NULL)

slopes(model_6, variables = "age", by = "fhouse", newdata = "mean") %>%
  tidy() %>%
  select(4:8) %>%
  mutate(p = case_when(p.value < 0.05 & p.value > 0.01 ~ "< 0.05*",
                        p.value < 0.01 & p.value > 0.001 ~ "< 0.01**",
                        p.value < 0.001 ~ "< 0.001***",
                        TRUE ~ "")) %>%
  select(-p.value) %>%
  kableExtra::kable(format = "latex",
                    col.names = c("Fhouse", "Estimación", "Error Estándar", "p-value"),
                    row.names = F,
                    booktabs = T, align = 'c',
                    caption = paste("(\\#tab:table4)", "Efectos marginales"))
kableExtra::kable_styling(latex_options = "hold_position",
                          position = "center") %>%
kableExtra::kable_styling(bootstrap_options = c("striped", "hover", "condensed"),
                          column_spec(1, width = "2cm") %>%
  row_spec(0, bold = T)

```