modelo econometrico

Ana Munoz

2024-02-15

Pre procesamiento de datos

1. Lectura de datos y formato panel

```r
df <- read.csv('df_17_21.csv') %>%
  dplyr::select(year, country, total_earnings, # orden de datos panel
         -total_players, -iso, #no aplica el modelo
         gdp_gr, -GNI_cap, -pbicap, -poder_adq, # las engloba pbippa
         pbippa, # variables macroeconomicas
         gastoedu, # edu
         CPI, # corrupcion
         desempleo,# work
         desnutricion, # pobreza
         internet, elect_acc, movil, # tech access
         age_game, # edades
         poblacion, pop_growth, rural_per# people
        ) %>%
  arrange(country, decreasing = FALSE)

head(df, 3)
```

```
##         year country total_earnings   gdp_gr   pbippa gastoedu CPI desempleo
## 1 2017-01-01 Albania        2868.16 3.898112 12771.05 3.611720  38     13.62
## 2 2018-01-01 Albania        1346.55 4.276312 13498.25 3.152945  36     12.30
## 3 2019-01-01 Albania       37459.64 2.523541 14407.37 3.916240  35     11.47
##   desnutricion internet elect_acc     movil age_game poblacion pop_growth
## 1          4.2 62.40000     99.89 125.92053   30.705   2873457   -0.09197
## 2          4.1 65.40000    100.00  94.36447   30.486   2866376   -0.24673
## 3          4.1 68.55039    100.00  91.51643   30.185   2854191   -0.42601
##   rural_per
## 1    40.617
## 2    39.681
## 3    38.771
```

2. Valores faltantes

- Numero de Valores faltantes por variable

```r
sapply(df, function(x) sum(is.na(x)))
```

```
##            year         country total_earnings          gdp_gr          pbippa 
##               0               0              0               0               0 
##        gastoedu             CPI      desempleo    desnutricion        internet 
##              52               0              0               0               2 
##       elect_acc           movil       age_game       poblacion      pop_growth 
##               0               0              0               0               0 
##       rural_per 
##               0
```

- corriegiendo los NAs

```r
# Internet: 2 faltantes -> 2018 cambodia y trinidad y tobago

### Cambodia, hueco en 2018, reemplazdo por el promedio

df[df$country=='Cambodia', 'internet'][2] <-
          (df[df$country=='Cambodia', 'internet'][1] +
             df[df$country=='Cambodia', 'internet'][3])/2

### trinidad y tobago, reemplazdo por el promedio
df[df$country=='Trinidad and Tobago', 'internet'][2] <-
          (df[df$country=='Trinidad and Tobago', 'internet'][1]+
              df[df$country=='Trinidad and Tobago', 'internet'][3])/2

# gasto en educacion

### para continuar con el analisis sin NAs, elimnare esta varible temporal
df <- df%>%dplyr::select(-gastoedu)

# verificamos NAs, ahora no tengo NAS
sapply(df, function(x) sum(is.na(x)))
```

```
##            year         country total_earnings          gdp_gr          pbippa 
##               0               0              0               0               0 
##             CPI       desempleo   desnutricion        internet       elect_acc 
##               0               0              0               0               0 
##           movil        age_game      poblacion      pop_growth       rural_per 
##               0               0              0               0               0
```

3. Normalizacion con logaritmo

- valores con varianzas muy grandes

```r
summary(df)
```

```
##      year             country          total_earnings        gdp_gr        
##  Length:450         Length:450         Min.   :     55   Min.   :-18.8544  
##  Class :character   Class :character   1st Qu.:  45536   1st Qu.: -0.2595
```

```
##  Mode  :character   Mode  :character   Median :  273947   Median :  1.9775
##                                        Mean   : 1955884   Mean   :  1.5685
##                                        3rd Qu.: 1389779   3rd Qu.:  4.4551
##                                        Max.   :50676418   Max.   : 18.7329
##      pbippa           CPI          desempleo       desnutricion
##  Min.   :  3973   Min.   :18.00   Min.   : 0.116   Min.   : 2.500
##  1st Qu.: 13371   1st Qu.:35.00   1st Qu.: 3.947   1st Qu.: 2.500
##  Median : 25632   Median :44.00   Median : 5.535   Median : 2.500
##  Mean   : 30442   Mean   :49.84   Mean   : 6.926   Mean   : 4.462
##  3rd Qu.: 44172   3rd Qu.:67.00   3rd Qu.: 8.700   3rd Qu.: 5.075
##  Max.   :131631   Max.   :89.00   Max.   :28.770   Max.   :19.400
##     internet         elect_acc          movil           age_game
##  Min.   : 13.78   Min.   : 80.70   Min.   : 51.55   Min.   :19.37
##  1st Qu.: 64.31   1st Qu.: 99.79   1st Qu.:106.18   1st Qu.:24.52
##  Median : 78.80   Median :100.00   Median :122.07   Median :28.58
##  Mean   : 74.38   Mean   : 98.90   Mean   :121.02   Mean   :28.80
##  3rd Qu.: 88.61   3rd Qu.:100.00   3rd Qu.:132.07   3rd Qu.:32.85
##  Max.   :100.00   Max.   :100.00   Max.   :219.70   Max.   :48.57
##    poblacion           pop_growth        rural_per
##  Min.   :3.434e+05   Min.   :-4.2566   Min.   : 0.00
##  1st Qu.:5.356e+06   1st Qu.: 0.1053   1st Qu.:18.19
##  Median :1.193e+07   Median : 0.7568   Median :29.79
##  Mean   :6.979e+07   Mean   : 0.7107   Mean   :31.18
##  3rd Qu.:4.996e+07   3rd Qu.: 1.3826   3rd Qu.:42.31
##  Max.   :1.412e+09   Max.   : 4.5561   Max.   :81.62
```

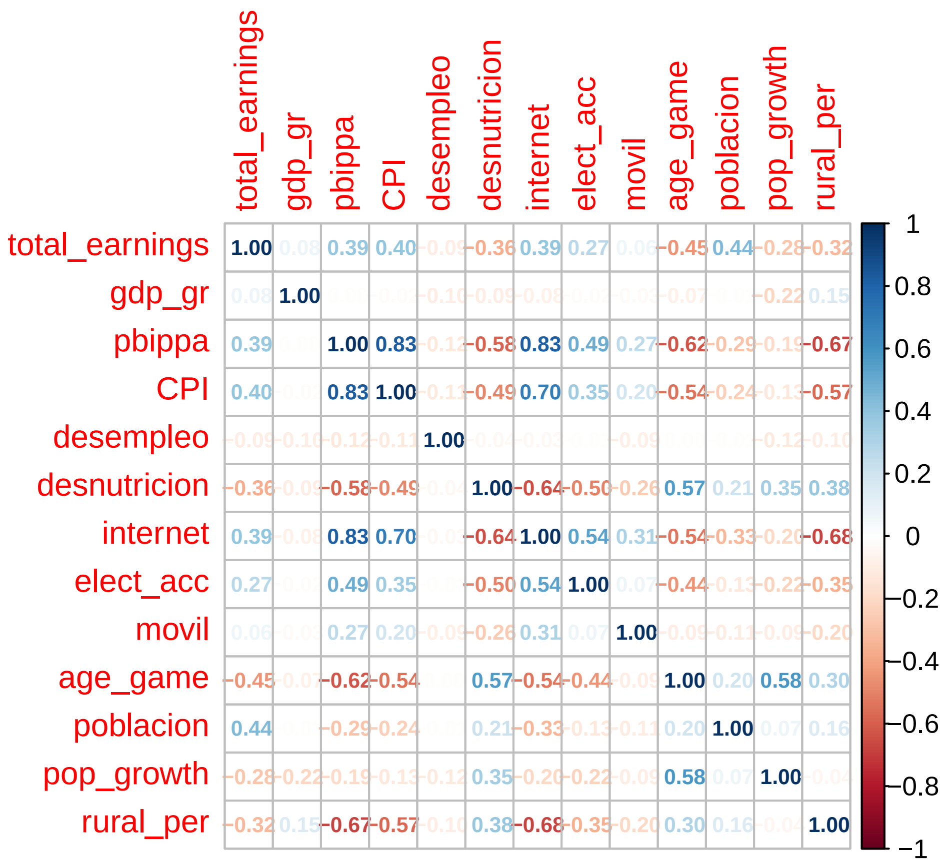
- aplico normalizacion logaritmica en algunas variables

```r
df_standar <- df %>%
  mutate(across(c("total_earnings", "pbippa", "poblacion"), ~log(.) %>% as.vector))
```

4. Correlacion y eliminacion de variables

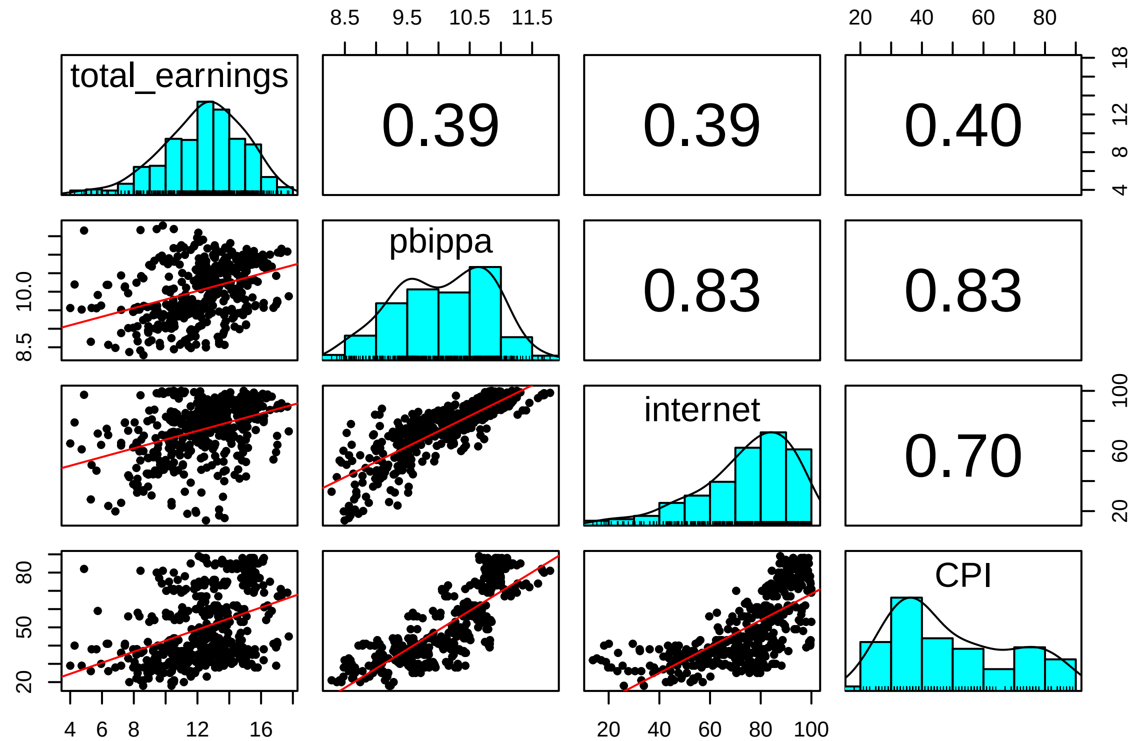
- verificacion de correlaciones

```r
matriz_corr <- cor(df_standar[3:15])
corrplot(matriz_corr, method = 'number', number.cex = 0.7)
```



- Veamos mas a detalle las correlaciones encontradas, pbippa, cpi e acceso internet Se debe crear otra variable para eliminar la correlacion? estas correlaciones no implican causalidad, pero como debo tratarlas?

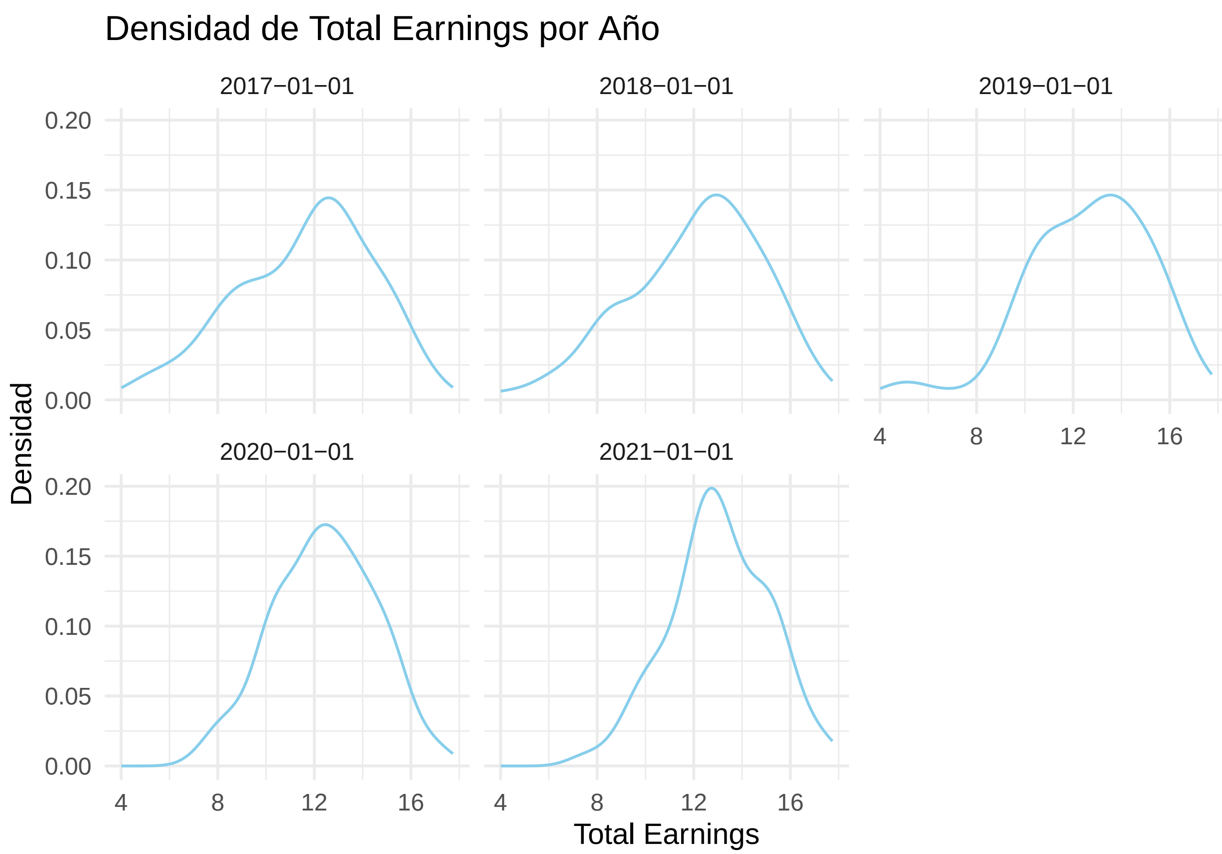
```r
pairs.panels(df_standar%>%dplyr::select(total_earnings,pbippa, internet,CPI),
             smooth = FALSE,      # Si TRUE, dibuja ajuste suavizados de tipo loess
             scale = FALSE,       # Si TRUE, escala la fuente al grado de correlación
             density = TRUE,      # Si TRUE, añade histogramas y curvas de densidad
             ellipses = FALSE,    # Si TRUE, dibuja elipses
             method = "pearson",  # Método de correlación (también "spearman" o "kendall")
             lm = TRUE,           # Si TRUE, dibuja un ajuste lineal en lugar de un ajuste LOESS
             cor = TRUE,          # Si TRUE, agrega correlaciones
             jiggle = FALSE,      # Si TRUE, se añade ruido a los dato
)
```



5. Distribucion de los datos

- Al parecer, cada year las ganacias se estan concentrando como una distribucionlognormal

```
ggplot(df_standar, aes(x = total_earnings)) +
  geom_density(color = "skyblue", linewidth = 0.5) +
  facet_wrap(~ year, ncol = 3) +
  labs(title = "Densidad de Total Earnings por Año",
       x = "Total Earnings",
       y = "Densidad") +
  theme_minimal()
```



* conociendo las distribuciones de los ingrsos por ano

```
# df_year <- df_standar%>%filter(year=='2020-01-01')
# descdist(df_year$total_earnings)
```

- un test de shapiro para comprobar y vemos que solo el 2019 no tendria una distribucion normal

```
years <- c('2017-01-01','2018-01-01','2019-01-01','2020-01-01','2021-01-01')
resultados <- data.frame(year = character(0), p_value = numeric(0))
for (i in years) {
  df_year <- df_standar %>% filter(year == i)
  p_value <- shapiro.test(df_year$total_earnings)$p.value
  resultados <- rbind(resultados, data.frame(year = i, p_value = p_value))
}
print(resultados)
```

```
##         year     p_value
## 1 2017-01-01 0.054053368
## 2 2018-01-01 0.087156308
## 3 2019-01-01 0.007368974
## 4 2020-01-01 0.741725375
## 5 2021-01-01 0.924970527
```

Implementando modelos

0. Preparando los datos

- tenemos datos panel con la siguente forma 90 paises 5 anios y estas cols

```r
# df_panel <- df_standar # por si sera necesario quitar algunas var
dim(table(df_standar$country,df_standar$year))
```

```
## [1] 90  5
```

```r
colnames(df_standar)
```

```
##  [1] "year"           "country"        "total_earnings" "gdp_gr"
##  [5] "pbippa"         "CPI"            "desempleo"      "desnutricion"
##  [9] "internet"       "elect_acc"      "movil"          "age_game"
## [13] "poblacion"      "pop_growth"     "rural_per"
```

- definimos las variables para el modelo

```r
attach(df_standar)
Y <- cbind(total_earnings)
X <- cbind(gdp_gr, pbippa, CPI, desempleo, desnutricion, internet,
           elect_acc, movil, age_game, poblacion, pop_growth, rural_per)

df_panel <- pdata.frame(df_standar, index=c('country','year'))

head(df_panel,3)
```

```
##                          year country total_earnings   gdp_gr   pbippa CPI
## Albania-2017-01-01 2017-01-01 Albania       7.961426 3.898112 9.454936  38
## Albania-2018-01-01 2018-01-01 Albania       7.205301 4.276312 9.510315  36
## Albania-2019-01-01 2019-01-01 Albania      10.531019 2.523541 9.575495  35
##                    desempleo desnutricion internet elect_acc     movil age_game
## Albania-2017-01-01     13.62          4.2 62.40000     99.89 125.92053   30.705
## Albania-2018-01-01     12.30          4.1 65.40000    100.00  94.36447   30.486
## Albania-2019-01-01     11.47          4.1 68.55039    100.00  91.51643   30.185
##                    poblacion pop_growth rural_per
## Albania-2017-01-01  14.87103   -0.09197    40.617
## Albania-2018-01-01  14.86856   -0.24673    39.681
## Albania-2019-01-01  14.86430   -0.42601    38.771
```

1. Pooled OLS estimator

```r
pooling <- plm(Y ~ X, data=df_panel, model= "pooling")
summary(pooling)
```

```
## Pooling Model
##
## Call:
## plm(formula = Y ~ X, data = df_panel, model = "pooling")
##
## Balanced Panel: n = 90, T = 5, N = 450
##
## Residuals:
##     Min.  1st Qu.   Median  3rd Qu.     Max.
## -5.99382 -0.78891  0.14527  1.01084  3.30005
##
## Coefficients:
##                  Estimate Std. Error t-value  Pr(>|t|)
## (Intercept)    6.3293279  3.7726240  1.6777 0.0941209 .
## Xgdp_gr        0.0411723  0.0160128  2.5712 0.0104640 *
## Xpbippa       -0.7340748  0.2418631 -3.0351 0.0025483 **
## XCPI           0.0286549  0.0067842  4.2238 2.925e-05 ***
## Xdesempleo    -0.0580979  0.0169899 -3.4195 0.0006861 ***
## Xdesnutricion -0.0474253  0.0282815 -1.6769 0.0942772 .
## Xinternet      0.0514245  0.0081935  6.2762 8.366e-10 ***
## Xelect_acc    -0.0267875  0.0292633 -0.9154 0.3604887
## Xmovil        -0.0058761  0.0033805 -1.7382 0.0828712 .
## Xage_game     -0.1347383  0.0237750 -5.6672 2.638e-08 ***
## Xpoblacion     1.0050873  0.0470406 21.3664 < 2.2e-16 ***
## Xpop_growth   -0.2438635  0.1007484 -2.4205 0.0159051 *
## Xrural_per    -0.0188562  0.0062061 -3.0383 0.0025218 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    2857.9
## Residual Sum of Squares: 986.47
## R-Squared:      0.65483
## Adj. R-Squared: 0.64535
## F-statistic: 69.0864 on 12 and 437 DF, p-value: < 2.22e-16
```

2. Between estimator

```r
between <- plm(Y ~ X, data=df_panel, model= "between")
summary(between)
```

```
## Oneway (individual) effect Between Model
##
## Call:
## plm(formula = Y ~ X, data = df_panel, model = "between")
##
## Balanced Panel: n = 90, T = 5, N = 450
## Observations used in estimation: 90
##
## Residuals:
##     Min.  1st Qu.   Median  3rd Qu.     Max.
## -3.44775 -0.73095  0.13601  0.87842  2.66608
##
```

```
## Coefficients:
##                 Estimate Std. Error t-value Pr(>|t|)
## (Intercept)    8.2316693  7.8331475  1.0509  0.29660
## Xgdp_gr        0.1665052  0.0965397  1.7247  0.08859 .
## Xpbippa       -0.6587200  0.5003725 -1.3165  0.19193
## XCPI           0.0290623  0.0137310  2.1165  0.03753 *
## Xdesempleo    -0.0570910  0.0353055 -1.6171  0.10996
## Xdesnutricion -0.0423711  0.0674028 -0.6286  0.53146
## Xinternet      0.0439754  0.0214472  2.0504  0.04373 *
## Xelect_acc    -0.0490616  0.0620170 -0.7911  0.43132
## Xmovil        -0.0050015  0.0072482 -0.6900  0.49225
## Xage_game     -0.1263979  0.0506282 -2.4966  0.01468 *
## Xpoblacion     0.9937806  0.0944904 10.5173  < 2e-16 ***
## Xpop_growth   -0.2520795  0.2511692 -1.0036  0.31870
## Xrural_per    -0.0285515  0.0131811 -2.1661  0.03340 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    496.58
## Residual Sum of Squares: 134.88
## R-Squared:      0.72839
## Adj. R-Squared: 0.68606
## F-statistic: 17.2078 on 12 and 77 DF, p-value: < 2.22e-16
```

3. First differences estimator

```r
firstdiff <- plm(Y ~ X, data=df_panel, model= "fd")
summary(firstdiff)
```

```
## Oneway (individual) effect First-Difference Model
##
## Call:
## plm(formula = Y ~ X, data = df_panel, model = "fd")
##
## Balanced Panel: n = 90, T = 5, N = 450
## Observations used in estimation: 360
##
## Residuals:
##     Min.  1st Qu.   Median  3rd Qu.     Max.
## -3.30665 -0.52452 -0.12607  0.43031  5.35957
##
## Coefficients:
##                 Estimate Std. Error t-value Pr(>|t|)
## (Intercept)    0.0278211  0.1255328  0.2216  0.82474
## Xgdp_gr        0.0088964  0.0125810  0.7071  0.47996
## Xpbippa        3.0005425  1.7646791  1.7003  0.08996 .
## XCPI           0.0211291  0.0301956  0.6997  0.48456
## Xdesempleo    -0.0816519  0.0523844 -1.5587  0.11998
## Xdesnutricion  0.0896063  0.1164275  0.7696  0.44204
## Xinternet     -0.0107310  0.0153285 -0.7001  0.48436
## Xelect_acc     0.0473437  0.0590559  0.8017  0.42329
## Xmovil        -0.0022971  0.0082925 -0.2770  0.78194
```

```
## Xage_game     -0.1732732  0.1378263 -1.2572  0.20953
## Xpoblacion    11.8287551  6.0732335  1.9477  0.05226 .
## Xpop_growth   -0.0328385  0.1253554 -0.2620  0.79351
## Xrural_per    -0.3540517  0.2472457 -1.4320  0.15305
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    383.36
## Residual Sum of Squares: 350.71
## R-Squared:      0.085162
## Adj. R-Squared: 0.053525
## F-statistic: 2.69186 on 12 and 347 DF, p-value: 0.0017564
```

4. Fixed effects or within estimator

```r
fixed <- plm(Y ~ X, data=df_panel, model= "within")
summary(fixed)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = Y ~ X, data = df_panel, model = "within")
##
## Balanced Panel: n = 90, T = 5, N = 450
##
## Residuals:
##      Min.   1st Qu.    Median   3rd Qu.      Max.
## -3.181978 -0.397522  0.012079  0.432181  2.266035
##
## Coefficients:
##                 Estimate Std. Error t-value Pr(>|t|)
## Xgdp_gr        0.0194150  0.0111453  1.7420 0.082394 .
## Xpbippa        0.7292312  1.1028603  0.6612 0.508910
## XCPI           0.0258059  0.0258489  0.9983 0.318809
## Xdesempleo    -0.0764798  0.0439533 -1.7400 0.082739 .
## Xdesnutricion -0.0379720  0.0734253 -0.5172 0.605379
## Xinternet      0.0245815  0.0117958  2.0839 0.037896 *
## Xelect_acc     0.1183639  0.0499909  2.3677 0.018445 *
## Xmovil         0.0028160  0.0070564  0.3991 0.690085
## Xage_game     -0.1223952  0.0744455 -1.6441 0.101060
## Xpoblacion     9.8371455  3.3849460  2.9061 0.003893 **
## Xpop_growth   -0.0234210  0.1026724 -0.2281 0.819692
## Xrural_per    -0.3226088  0.1307078 -2.4682 0.014061 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    375.01
## Residual Sum of Squares: 248.4
## R-Squared:      0.3376
## Adj. R-Squared: 0.14536
## F-statistic: 14.7805 on 12 and 348 DF, p-value: < 2.22e-16
```

5. Random effects estimator

```r
random <- plm(Y ~ X, data=df_panel, model= "random")
summary(random)
```

```
## Oneway (individual) effect Random Effect Model
##    (Swamy-Arora's transformation)
##
## Call:
## plm(formula = Y ~ X, data = df_panel, model = "random")
##
## Balanced Panel: n = 90, T = 5, N = 450
##
## Effects:
##                  var std.dev share
## idiosyncratic 0.7138  0.8449 0.307
## individual    1.6089  1.2684 0.693
## theta: 0.7145
##
## Residuals:
##      Min.   1st Qu.    Median   3rd Qu.      Max.
## -4.189467 -0.405262  0.089268  0.489445  2.410630
##
## Coefficients:
##                  Estimate  Std. Error z-value  Pr(>|z|)
## (Intercept)   -6.55845143  5.82609617 -1.1257   0.26029
## Xgdp_gr        0.02769039  0.01012484  2.7349   0.00624 **
## Xpbippa       -0.57185116  0.41500383 -1.3779   0.16822
## XCPI           0.01887024  0.01175176  1.6057   0.10833
## Xdesempleo    -0.04397381  0.02648519 -1.6603   0.09685 .
## Xdesnutricion -0.00955893  0.04164162 -0.2296   0.81844
## Xinternet      0.05889296  0.00759178  7.7575 8.665e-15 ***
## Xelect_acc     0.07081591  0.03825329  1.8512   0.06414 .
## Xmovil         0.00062328  0.00496433  0.1256   0.90009
## Xage_game     -0.14950056  0.03775334 -3.9599 7.497e-05 ***
## Xpoblacion     1.03803861  0.09293369 11.1697 < 2.2e-16 ***
## Xpop_growth   -0.07877996  0.09527150 -0.8269   0.40829
## Xrural_per    -0.01042271  0.01156477 -0.9012   0.36746
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    577.37
## Residual Sum of Squares: 329.19
## R-Squared:      0.42984
## Adj. R-Squared: 0.41419
## Chisq: 329.454 on 12 DF, p-value: < 2.22e-16
```

Test

LM test for random effects versus OLS

- H0 = los efectos aleatorios no son significativos no soportan el modelo (p rechazo)
- H1 = los efectos aleatorios sí son significativos
- los efectos aleatorios son importantes y deben considerarse en el modelo.
- OLS no captura completamente la variabilidad debida a los efectos aleatorios

```r
plmtest(pooling)
```

```
##
##  Lagrange Multiplier Test - (Honda)
##
## data:  Y ~ X
## normal = 19.035, p-value < 2.2e-16
## alternative hypothesis: significant effects
```

LM test for fixed effects versus OLS

- H0 = los efectos fijos no son significativos no soportan el modelo (p rechazo)
- los efectos fijos son importantes y deben considerarse en el modelo.
- OLS no captura completamente la variabilidad debida a los efectos fijos

```r
pFtest(fixed, pooling)
```

```
##
##  F test for individual effects
##
## data:  Y ~ X
## F = 11.618, df1 = 89, df2 = 348, p-value < 2.2e-16
## alternative hypothesis: significant effects
```

Hausman test for fixed versus random effects model

- hipótesis alternativa es que uno de los modelos es inconsistente.
- p es 0.0001451 rechazo la hipótesis nula
- uno de los modelos es inconsistente.

```r
phtest(random, fixed)
```

```
##
##  Hausman Test
##
## data:  Y ~ X
## chisq = 38.151, df = 12, p-value = 0.0001451
## alternative hypothesis: one model is inconsistent
```

Cluster standar error

- Para los efectos fijos

```r
coeftest(fixed, vcovHC(fixed, type='HC0', cluster='group'))
```

```
##
## t test of coefficients:
##
##                 Estimate Std. Error t value Pr(>|t|)
## Xgdp_gr        0.0194150  0.0080084  2.4243  0.01585 *
## Xpbippa        0.7292312  1.3382435  0.5449  0.58616
## XCPI           0.0258059  0.0250824  1.0288  0.30427
## Xdesempleo    -0.0764798  0.0623953 -1.2257  0.22113
## Xdesnutricion -0.0379720  0.0737233 -0.5151  0.60684
## Xinternet      0.0245815  0.0144007  1.7070  0.08872 .
## Xelect_acc     0.1183639  0.0836781  1.4145  0.15811
## Xmovil         0.0028160  0.0128257  0.2196  0.82634
## Xage_game     -0.1223952  0.1220172 -1.0031  0.31651
## Xpoblacion     9.8371455  4.3836553  2.2441  0.02546 *
## Xpop_growth   -0.0234210  0.1095989 -0.2137  0.83091
## Xrural_per    -0.3226088  0.1953753 -1.6512  0.09959 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Modificaciones el modelo elegido

Cambios en el nuevo modelo

```r
X_2 <- cbind(gdp_gr, desempleo,
             # desnutricion,
             internet,
           elect_acc, age_game, poblacion, rural_per)


fixed_2 <- plm(Y ~ X_2, data=df_panel, model= "within")
summary(fixed_2)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = Y ~ X_2, data = df_panel, model = "within")
##
## Balanced Panel: n = 90, T = 5, N = 450
##
## Residuals:
##      Min.   1st Qu.    Median   3rd Qu.      Max.
## -3.179025 -0.381123  0.011576  0.437741  2.255980
##
```

```
## Coefficients:
##               Estimate Std. Error t-value Pr(>|t|)
## X_2gdp_gr     0.0227277  0.0095441  2.3813 0.017779 *
## X_2desempleo -0.0889895  0.0407190 -2.1855 0.029512 *
## X_2internet   0.0255665  0.0112405  2.2745 0.023536 *
## X_2elect_acc  0.1160595  0.0489885  2.3691 0.018368 *
## X_2age_game  -0.1469211  0.0678067 -2.1668 0.030921 *
## X_2poblacion  9.7128866  3.3375366  2.9102 0.003841 **
## X_2rural_per -0.3620123  0.1163247 -3.1121 0.002009 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    375.01
## Residual Sum of Squares: 249.79
## R-Squared:      0.33391
## Adj. R-Squared: 0.15276
## F-statistic: 25.2797 on 7 and 353 DF, p-value: < 2.22e-16
```

```r
coeftest(fixed_2, vcovHC(fixed_2, type='HC0', cluster='group'))
```

```
##
## t test of coefficients:
##
##               Estimate Std. Error t value Pr(>|t|)
## X_2gdp_gr     0.0227277  0.0059679  3.8083 0.000165 ***
## X_2desempleo -0.0889895  0.0560786 -1.5869 0.113437
## X_2internet   0.0255665  0.0154600  1.6537 0.099074 .
## X_2elect_acc  0.1160595  0.0855774  1.3562 0.175904
## X_2age_game  -0.1469211  0.1074638 -1.3672 0.172442
## X_2poblacion  9.7128866  4.6866439  2.0725 0.038948 *
## X_2rural_per -0.3620123  0.1796418 -2.0152 0.044642 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```