

¿La tasa de Alfabetismo es significativa para explicar el PIB per  
cápita?

Rosana Cardona, Angelo Colombo, Camila Toro

2023-07-03



**udp** UNIVERSIDAD  
DIEGO PORTALES

## *Introducción*

Nuestro estudio tiene como objetivo analizar si la tasa de alfabetización puede explicar las diferencias en términos de PIB per cápita, que existen entre 203 países elegidos de manera aleatoria

Para llevar a cabo este proyecto, emplearemos un modelo econométrico, el cual nos permitirá determinar el impacto de diversos factores en el desarrollo económico. Este modelo busca evaluar de manera empírica la relación entre ambas variables y se basa en la teoría de crecimiento endógeno, el cual sugiere que la tasa de alfabetización desempeña un rol fundamental en el crecimiento económico a largo plazo, ya que se enfoca en que las habilidades y conocimiento adquiridos mediante la alfabetización, y educación en general, son determinantes para la productividad y crecimiento económico de un país, sin embargo, es esta variable estadísticamente significativa.

Dentro del marco de investigación sobre la relación entre la tasa de alfabetización y el PIB per cápita, existen diversos estudios que estiman similares modelos y poseen diversas conclusiones.

Por ejemplo, “Education and Economic Growth: Endogenous Growth Theory Test. The French Case” por Marielle Monteils. Este estudio utiliza un enfoque de crecimiento endógeno para analizar la relación entre educación y crecimiento económico en Francia, estimando variables tales como la tasa de alfabetización, gasto en educación y PIB per cápita. Este estudio determina que existe una relación positiva entre la educación y el crecimiento económico en Francia, sugiriendo que la inversión en educación es clave para impulsar el desarrollo económico del país.

Otro caso, es el del artículo “Income Inequality, Illiteracy Rate, and Life Expectancy in Brazil” por Erick Messias, el cual examina la relación entre la desigualdad de ingresos, tasa de analfabetismo y esperanza de vida en Brasil. Los resultados de este estudio indican que la tasa de alfabetización tiene un efecto positivo en la esperanza de vida y sugiere que la reducción de la tasa de analfabetismo puede mejorar la esperanza de vida en el país.

Estos artículos demuestran la importancia de analizar la relación entre la tasa de alfabetización y PIB per cápita, debido a la variedad conclusiones que se han obtenido según el enfoque de los investigadores, por lo que nuestro estudio plantea la hipótesis que la tasa de alfabetización es estadísticamente significativa para explicar el PIB per cápita, por lo que buscamos aportar nueva evidencia empírica y contribución al debate sobre lo esencial que es la alfabetización y educación para el desarrollo económico de un país.

## Datos

La base de datos se obtuvo de *Kaggle*, la cual es una recopilación de diversos datos de 227 países del mundo para el año 2006. Los datos con los que se formuló dicha base, provienen de la página “*The World Factbook*”.

La base de datos se compone de los siguientes datos:

- *Densidad Poblacional(den)*: Corresponde a la cantidad de habitantes que viven en por milla cuadrada por país. Es una variable cuantitativa.
- *Migración neta(mi)*: Corresponde a la cantidad de personas que migran hacia o desde un país dividido por la cantidad de habitantes. Se mide en porcentaje, si es positiva significa que el país recibió migrantes, si es negativa significa que el país perdió habitantes. Es una variable cuantitativa.
- *Mortalidad infantil(inf)*: Cuántos niños mueren por cada mil nacimientos. Es una variable cuantitativa.
- *PIB per cápita(pib)*: Producto interno bruto por persona, año 2006. (*PIB/población*). Es una variable cuantitativa.
- *Alfabetización(alf)*: Porcentaje de la población de un país, que sabe leer y escribir. Es una variable cuantitativa.
- *Teléfono(tel)*: Cantidad de teléfonos por cada 1000 habitantes. Es una variable cuantitativa.
- *Tasa de natalidad(nat)*: Cantidad de nacimientos por cada 1000 habitantes. Es una variable cuantitativa.
- *Tasa de mortalidad(mor)*: Cantidad de fallecimientos por cada 1000 habitantes. Es una variable cuantitativa.
- *Costa(cos)*: Es una variable binaria, si toma el valor de 1 el país tiene salida al mar, si toma el valor de 0, el país no tiene salida al mar.
- *Migración x Costa(mixcos)*: Esta variable corresponde a una interacción entre Migración y Costa.

La estadística descriptiva de cada variable anteriormente mencionadas es la siguiente:

	Media	Desviacion	Mediana	Minimo	Maximo
den	360.89	1716.13	74.00	1.80	16271.50
mi	0.01	5.01	0.00	-20.99	23.06
inf	37.36	36.33	23.51	2.29	191.19
pib	9727.59	10213.01	5600.00	500.00	55100.00
alf	82.87	19.72	92.30	17.60	100.00
tel	225.33	219.49	153.80	0.20	1035.60
nat	22.37	11.27	18.90	7.29	50.73
mor	9.40	5.17	8.10	2.29	29.74
cos	0.80	0.40	1.00	0.00	1.00
mixcos	-0.18	4.52	0.00	-20.99	18.75

# Resultados empíricos

## Modelo de regresión lineal simple

Inicialmente, se realizó un modelo de regresión lineal simple para evaluar si la variable explicativa de nuestra hipótesis es significativa.

Se planteó la siguiente regresión lineal simple, en la cual, el PIB per cápita es la variable a explicar y la variable explicativa es la alfabetización.

$$pib = \beta_0 + \beta_1 \times alf + \mu$$

Se trata de un modelo *nivel-nivel*, en el cual, la variable PIB per cápita está en *dólares* y la alfabetización está expresada como *porcentaje*.

Se espera que la variable alfabetización explique el PIB per cápita de manera significativa. Además, que ésta tenga un signo positivo, es decir, mientras aumente la tasa de alfabetización, el PIB per cápita aumente.

Al realizar la estimación se obtuvo el siguiente modelo:

Dependent variable:	
pib	
alf	267.440*** (31.278)
Constant	-12,434.340*** (2,663.984)
-----	
Observations	203
R2	0.267
Adjusted R2	0.263
Residual Std. Error	8,767.359 (df = 201)
F Statistic	73.108*** (df = 1; 201)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

**Modelo estimado:**  $pib\ gorro = -12,434.340 + 267.440 \times alf$

### Interpretación de parámetros

1.  $\beta_0$  **gorro**: -12,434.340, signo negativo

Se trata de un valor contraintuitivo, dado que el pib per cápita no debiese ser negativo. Cuando la alfabetización es cero, en promedio, el PIB per cápita es -12,434.340 dólares, con todo lo demás constante.

2.  $\beta_1$  **gorro**: 267.440, signo positivo

Este resultado es intuitivo. Cuando la alfabetización aumenta en una unidad porcentual, el PIB per cápita aumenta en promedio 267.440 dólares, con todo lo demás constante.

- R2 Ajustado

Se tiene un  $R^2$  Ajustado de 0.26, es decir que, aproximadamente un 26% de la variación de la variable PIB per cápita es explicada, en promedio, por el incremento de un por ciento de la alfabetización.

## Supuestos del modelo de regresión lineal simple

Los supuestos de linealidad y muestra aleatoria se asumen que se cumplen, debido a que los parámetros son lineales. Además, se tiene una muestra de 203 observaciones, en la cual se observa aleatoriedad.

- **Supuesto 3: Variación muestral**

Para verificar el supuesto de variación muestral se observa la variación de la variable alfabetización, la cual es de 388.95. Por lo tanto, el supuesto tres se cumple.

- **Supuesto 4: Media condicional cero**

Para el supuesto cuatro, se obtiene el valor de los residuos con *residuals* en *r studio*. Por otra parte, se realiza la correlación entre la variable explicativa *Alfabetización* y los errores obtenidos anteriormente, arrojando como resultado una correlación muy cercana a 0.

0.000000000000000000003690258

El supuesto, de *Gauss-Markov*, establece que la media condicional de los errores debe ser cero. Sin embargo, según estudios, en la práctica es muy difícil que la media condicional sea estrictamente cero en todos los casos. Pequeñas desviaciones de cero, como el valor obtenido, pueden ocurrir debido a errores de medición o errores de aproximación de los datos.

Dicho lo anterior, el supuesto cuatro de Gauss Markov se cumple. Al cumplirse los primeros 4 supuestos, el parámetro de la variable alfabetización es *insesgado* y *consistente*.

Por otra parte, es importante comprobar que el error del modelo y la variable alfabetización no están correlacionados, debido a que, si esto no ocurre, el parámetro estimado será sesgado e inconsistente, alejándose del verdadero valor del parámetro poblacional, obteniendo valores sobreestimados o subestimados de la variable explicativa.

- **Supuesto 5: Homocedasticidad**

Para evaluar el supuesto de homocedasticidad, se realiza el test de *breusch-pagan*.

Se tienen las siguientes hipótesis y se plantea un alfa del 5%

- *Hipotesis*

*Ho: Varianza de los residuos es constante. (homocedasticidad)*

*H1: No se cumple H0. (heterocedasticidad)*

studentized Breusch-Pagan test

```
data:  rgs
```

BP = 18.42, df = 1, p-value = 0.00001772

Según el bptest, se tiene un p valor de *0.00001772*, el cual es menor al alfa de 0.05. Por lo tanto, se rechaza  $H_0$  con un 95% de confianza. **Los errores no son homocedásticos.**

Para corregir la heterocedasticidad en la varianza de los errores, se utiliza coeftest. Se obtienen errores robustos con un valor de *13.760* para el parámetro de la variable Alfabetización.

Anteriormente, el modelo con errores heterocedásticos, nos brinda un error estándar del parámetro  $\beta_1$  del *31.278*. Al realizar el test t, nos muestra que la variable es significativa, con un valor de *8.550*. Sin embargo, no podemos quedarnos con este resultado, ya que el error estándar del parámetro de la variable alfabetización es heterocedástico, por lo tanto debemos utilizar errores robustos a la heterocedasticidad.

Finalmente, al realizar nuevamente la inferencia de la variable alfabetización con errores robustos a la heterocedasticidad, se observa que la variable independiente sigue siendo significativa a un nivel de confianza del *95%*.

### Modelo de regresión lineal multiple

Se planteó la siguiente regresión lineal multiple, en la cual, el PIB per cápita es la variable a explicar y las variables explicativas son *alfabetización, densidad poblacional, migración, mortalidad infantil, cantidad de teléfonos, natalidad, mortalidad, costa* y la variable interacción entre la variable continua *migración* y la variable ficticia *costa*.

$$pib = \beta_0 + \beta_1 \times alf + \beta_2 \times den + \beta_3 \times mi + \beta_4 \times inf + \beta_5 \times tel + \beta_6 \times nat + \beta_7 \times mor + \beta_8 \times cos + \beta_9 \times mixcos + \mu$$

Se espera un  $R^2$  ajustado mayor, es decir, que las variables agregadas expliquen en mayor porcentaje la variación del PIB per cápita.

Estimar este modelo es de gran utilidad, para saber con certeza si la alfabetización es realmente una variable significativa o no para predecir el PIB per cápita.

Dependent variable:	
	pib
alf	-12.732 (31.206)
den	-0.259 (0.216)
mi	577.523*** (173.546)
inf	-72.027*** (26.558)
tel	32.194*** (2.630)
nat	22.078 (73.327)

mor	199.259*
	(102.174)
cos	-708.141
	(992.329)
mixcos	-151.742
	(186.568)
Constant	4,478.767
	(3,996.105)

```

-----
Observations          203
R2                    0.774
Adjusted R2           0.763
Residual Std. Error   4,971.099 (df = 193)
F Statistic           73.291*** (df = 9; 193)
=====
Note:                  *p<0.1; **p<0.05; ***p<0.01

```

### Modelo estimado:

$$pib = 4,478.767 - 12.732 \times alf - 0.259 \times den + 577.523 \times mi - 72.027 \times inf + 32.194 \times tel + 22.078 \times nat + 199.259 \times mor - 708.141 \times cos - 151.742 \times mixcos$$

### Interpretación de parámetros

#### 1. $\beta_0$ gorro, constante

Magnitud: 4478.767, signo positivo.

En promedio, la variable PIB per cápita es de 4,478.7673 dolares, cuando todas las demás variables explicativas son cero.

#### 2. $\beta_1$ gorro, Alfabetización

Magnitud: 12.732, signo negativo

El signo del parámetro de la variable alfabetización es contraintuitivo, dado que se esperaría que a más porcentaje de alfabetización, el PIB per cápita aumentará.

El PIB per cápita disminuye en 12.732 dolares, cuando la tasa de alfabetización aumenta en 1 unidad porcentual, con todo lo demás constante.

#### 3. $\beta_2$ gorro, Densidad poblacional

Magnitud: 0.259, signo negativo.

El PIB per cápita disminuye en 0.259 dolares, cuando la densidad poblacional aumenta en 1 unidad, con todo lo demás constante.

#### 3. $\beta_3$ gorro, Migración

Magnitud: 577.523, signo positivo.

El PIB per cápita aumenta en *577.523* dolares cuando la tasa de migración neta aumenta en 1 unidad porcentual, con todas las demás variables explicativas constantes.

#### 4. $\beta_4$ gorro, Infantil

Magnitud: 72.026, signo negativo.

El PIB per cápita disminuye en 72.0266 dolares, cuando la tasa de mortalidad infantil aumenta en 1 unidad, con todo lo demás constante.

#### 5. $\beta_5$ gorro, Teléfono

Magnitud: 32.394, signo positivo.

El PIB per cápita aumenta en 32.3944 unidades, cuando la cantidad de teléfonos aumenta en 1 unidad, con todo lo demás constante.

#### 6. $\beta_6$ gorro, Natalidad

Magnitud: 22.078, signo positivo.

El PIB per cápita aumenta en 22.07 dolares, cuando la tasa de natalidad aumenta en una unidad, con todo lo demás constante.

#### 7. $\beta_7$ gorro, Mortalidad

Magnitud: 199.259, signo positivo.

El PIB per cápita aumenta en *199.259* unidades, cuando la tasa de mortalidad aumenta en 1 unidad, con todo lo demás constante.

#### 8. $\beta_8$ gorro, Costa

Magnitud: 708.140, signo negativo.

El PIB per cápita disminuye 708.1405 dolares cuando el país tiene acceso al mar, que cuando no lo tiene, con todo lo demás ceteris paribus.

#### 9. $\beta_9$ gorro, Migración por Costa

Magnitud: 151.742, signo negativo.

El PIB per cápita disminuye en *151.742* dolares, cuando la interacción entre migración y costa aumenta en 1 unidad, cuando todas las demás variables explicativas son cero.

#### • R<sup>2</sup> ajustado

El R<sup>2</sup> ajustado del modelo indica que aproximadamente un *77%* de la variabilidad del PIB per cápita *es explicado* por las variables explicativas del modelo.



## Supuestos para el modelo de regresión lineal múltiple

Los supuestos de linealidad y muestra aleatoria se asumen que se cumplen, debido a que los parámetros son lineales. Además se tiene una muestra de 203 observaciones, en la cual se observa aleatoriedad.

- **Supuesto 3: Multicolinealidad** Se realiza una matriz para observar la correlación entre las variables explicativas del modelo de regresión lineal múltiple.

	den	mi	inf	alf	tel	nat	mor	cos	mixcos
den	1.00	0.18	-0.14	0.10	0.28	-0.18	-0.06	0.08	0.20
mi	0.18	1.00	-0.03	-0.01	0.27	-0.08	0.03	-0.09	0.90
inf	-0.14	-0.03	1.00	-0.77	-0.69	0.86	0.67	-0.33	-0.07
alf	0.10	-0.01	-0.77	1.00	0.60	-0.79	-0.40	0.12	0.03
tel	0.28	0.27	-0.69	0.60	1.00	-0.73	-0.30	0.18	0.24
nat	-0.18	-0.08	0.86	-0.79	-0.73	1.00	0.43	-0.24	-0.10
mor	-0.06	0.03	0.67	-0.40	-0.30	0.43	1.00	-0.39	-0.01
cos	0.08	-0.09	-0.33	0.12	0.18	-0.24	-0.39	1.00	-0.02
mixcos	0.20	0.90	-0.07	0.03	0.24	-0.10	-0.01	-0.02	1.00

Se cumple el supuesto, dado que no existe colinealidad perfecta entre ninguna de las variables explicativas.

- **Supuesto 4: Media condicional cero**

El supuesto, de Gauss-Markov, establece que la media condicional de los errores debe ser cero. Sin embargo, según estudios, en la práctica es muy difícil que la media condicional sea estrictamente cero en todos los casos. Pequeñas desviaciones de cero, como el valor obtenido, pueden ocurrir debido a errores de medición o errores de aproximación de los datos. Las variables son exógenas.

```
[,1]
den  0.0000000000000000002272209
mi   0.0000000000000000003099854
inf  0.00000000000000000006885128
alf -0.00000000000000000007023660
tel  0.00000000000000000003306398
nat -0.000000000000000000010741082
mor  0.000000000000000000030206512
cos  0.0000000000000000000106259316
```

Dicho lo anterior, el supuesto cuatro de Gauss Markov se cumple. Al cumplirse los primeros cuatros supuestos, los parámetros estimados son *insesgados y consistentes*.

- **Supuesto 5: homocedasticidad.**

Para evaluar el supuesto de homocedasticidad, se realiza el test de *breusch-pagan*.

Se tienen las siguientes hipótesis y se plantea un alfa del 5%

- **Hipotesis**

*H<sub>0</sub>: Varianza de los residuos es constante. (homocedasticidad)*

*H<sub>1</sub>: No se cumple H<sub>0</sub>. (heterocedasticidad)*

#### studentized Breusch-Pagan test

```
data:  rgm
BP = 37.291, df = 9, p-value = 0.00002335
```

Al ser el valor p menor a la significancia de 5%, se rechaza la hipótesis nula de que los residuos son homocedásticos con un 95% de confianza. Por ende, se asume que los residuos son heterocedásticos.

Para arreglar este problema, que es fundamental para poder hacer inferencia correctamente y que nuestro modelo sea eficiente, se deben calcular errores robustos para las variables.

La heterocedasticidad en el modelo de regresión lineal múltiple se puede generar debido a:

1. **Valores atípicos (outliers):** Cuando existen valores atípicos en los datos puede generar heterocedasticidad. Estos valores pueden tener un impacto negativo en la variabilidad de los errores en comparación con el resto de los datos, lo que conduce a heterocedasticidad.
2. **Variables omitidas:** Si hay variables importantes que no se incluyen en el modelo de regresión, puede haber una relación entre esas variables omitidas y los errores del modelo. Esta relación puede llevar a heterocedasticidad.
3. **Error de medición:** Errores de medición no aleatorios en las variables independientes o dependientes pueden introducir heterocedasticidad. Si hay un error sistemático en la medición de una variable, esto puede afectar la varianza de los errores y generar heterocedasticidad.

Por otra parte, si la heterocedasticidad no se resuelve, la inferencia sería ineficiente, puesto que la varianza de los estimadores no es la mínima, generando estimaciones menos precisas. Por otra parte, la presencia de heterocedasticidad en los errores puede dificultar la interpretación de los resultados del modelo para determinar la verdadera relación entre las variables explicativas y la variable dependiente del modelo.

#### Corrección de heterocedasticidad

Los errores robustos nos permiten inferir con mayor precisión, dado que se aproximan a la distribución t y distribución f.

#### Errores robustos

Obtención de errores robustos utilizando corrección de *Huber-White*. Se utiliza la librería `coeftest` para generar errores robustos a la heterocedasticidad.

#### Tabla de modelos de regresión lineal

Dependent variable:			
	pib OLS		coefficient test
	(1)	(2)	(3)
alf	267.440*** (31.278)	-12.732 (31.206)	-12.732 (16.300)

den		-0.259 (0.216)	-0.259 (0.343)
mi		577.523*** (173.546)	577.523 (381.770)
inf		-72.027*** (26.558)	-72.027*** (24.643)
tel		32.194*** (2.630)	32.194*** (3.094)
nat		22.078 (73.327)	22.078 (54.274)
mor		199.259* (102.174)	199.259*** (76.001)
cos		-708.141 (992.329)	-708.141 (980.648)
mixcos		-151.742 (186.568)	-151.742 (386.456)
Constant	-12,434.340*** (2,663.984)	4,478.767 (3,996.105)	4,478.767* (2,468.818)

Observations	203	203
R2	0.267	0.774
Adjusted R2	0.263	0.763
Residual Std. Error	8,767.359 (df = 201)	4,971.099 (df = 193)
F Statistic	73.108*** (df = 1; 201)	73.291*** (df = 9; 193)

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

El modelo (1) corresponde a la regresión lineal simple, el modelo (2) a la regresión lineal múltiple y el modelo (3) a la regresión lineal múltiple con errores robustos a heterocedasticidad.

## Inferencia

### • Test de significancia individual

	rlineals	rlinealm	rlienalmrob
alf	90% 95% 99%	Ns	Ns
den	-	Ns	Ns
mi	-	90% 95% 99%	Ns
inf	-	90% 95% 99%	90% 95% 99%
tel	-	90% 95% 99%	90% 95% 99%
nat	-	Ns	Ns
mor	-	90% 95%	90% 95% 99%
cos	-	Ns	Ns
mixcos	-	Ns	Ns
constant	90% 95% 99%	Ns	90% 95%

- (-)= No se incluye esta variable en el modelo
- Ns= Variable no es significativa a los niveles de significancia de 10%,5% y 1%
- 90%= variable significativa al 90% de confianza
- 95%= variable significativa al 90% de confianza
- 99%= variable significativa al 99% de confianza

A nivel individual, las únicas variables significativas, en el modelo robusto a heterocedasticidad son la mortalidad infantil, teléfono y mortalidad a niveles de confianza del 90%, 95% y 99%. La constante es significativa al 90% y al 95%.

Si se hubiese hecho inferencia con el modelo de regresión lineal múltiple que no es robusto a heterocedasticidad, en el caso de la variable migración, se hubiese cometido un error de tipo I, dado que al realizar la inferencia, se rechaza que el parámetro de la variable es igual a 0, cuando no se debe rechazar.

Hubiese sucedido un error del tipo 2 al 90% y 95% de confianza con la constante, si no se hubiese rechazado la hipótesis nula de que la constante es igual 0, cuando sí se debería haber rechazado al 10% y al 5% de significancia.

Por último la variable alfabetización y la constante son significantes a todos los niveles en el modelo de regresión lineal simple. Al agregar variables, que estaban contenidas en el término de error, la alfabetización deja de ser significativa al 90%, al 95% y al 99%, mientras que la constante solo al 99%.

Nuestra variable de interés no es significativa a nivel individual, sin embargo podría serlo de forma conjunta, por ende para saber si la variable alfabetización explica el PIB per cápita se deben probar test de significancia conjunto.

## Inferencia modelo regresión multiple

- **Test conjunto con nuevo modelo.**

Debido a que se tienen variables no significativas en el modelo de regresión lineal multiple anteriormente planteado, se crea un nuevo modelo en el cual solo se incluyan las variables significativas para realizar un test conjunto.

$$pib = \beta_0 + \beta_1 \times inf + \beta_2 \times tel + \beta_3 \times mor + \mu$$

Dependent variable:	
pib	
inf	-38.969** (19.317)
tel	36.322*** (2.492)
mor	181.975* (103.295)
Constant	1,287.846

(1,101.700)

```
-----
Observations                203
R2                          0.726
Adjusted R2                 0.722
Residual Std. Error    5,383.611 (df = 199)
F Statistic             175.987*** (df = 3; 199)
=====
Note:          *p<0.1; **p<0.05; ***p<0.01
```

### 1. $\beta_1$ , *Infantil*

Magnitud: 38.969, signo negativo

El PIB per cápita disminuye en promedio 38.969 dólares, cuando la tasa de mortalidad infantil, aumenta en una unidad, ceteris paribus.

### 2. $\beta_2$ , *Teléfono*

Magnitud: 36.322, signo positivo

El PIB per cápita aumenta en promedio 36.322 dólares, cuando la cantidad de teléfonos aumenta en una unidad, con todo lo demás constante.

### 3. $\beta_3$ , *Mortalidad*

Magnitud: 181.975, signo positivo.

El PIB per cápita aumenta en promedio 181.975 dólares, cuando la tasa de mortalidad aumenta en una unidad, cuando las demás variables explicativas son cero.

#### • *R2 ajustado*

El R2 ajustado del modelo indica que aproximadamente un 72% de la variabilidad del PIB per cápita *es explicado* por las variables explicativas este modelo.

#### *Modelo no restringido:*

$$pib = \beta_0 \text{ gorro} + \beta_1 \text{ gorro} \times \text{alf} + \beta_2 \text{ gorro} \times \text{den} + \beta_3 \text{ gorro} \times \text{mi} + \beta_4 \text{ gorro} \times \text{inf} + \beta_5 \text{ gorro} \times \text{tel} + \beta_6 \text{ gorro} \times \text{nat} + \beta_7 \text{ gorro} \times \text{mor} + \beta_8 \text{ gorro} \times \text{cos} + \beta_9 \text{ gorro} \times \text{mixcos}$$

#### *Modelo restringido:*

$$pib = \beta_0 \text{ gorro} + \beta_1 \text{ gorro} \times \text{infantil} + \beta_2 \text{ gorro} \times \text{telefono} + \beta_3 \text{ gorro} \times \text{mortalidad}$$

#### • *Hipotesis*

$H_0: \beta_1 \text{ alf} = 0, \beta_2 \text{ den} = 0, \beta_3 \text{ mi} = 0, \beta_6 \text{ nat} = 0, \beta_8 \text{ cos} = 0, \beta_9 \text{ mixcos} = 0$

$H_0: \beta_1 \text{ alf} \neq 0, \beta_2 \text{ den} \neq 0, \beta_3 \text{ mi} \neq 0, \beta_6 \text{ nat} \neq 0, \beta_8 \text{ cos} \neq 0, \beta_9 \text{ mixcos} \neq 0$

Se restringen 6 parámetros

***F calculado***

6.789159

***F en tabla, alfa 5%***

2.145801

***Regla de rechazo***

```
cat(fc > ft)
```

TRUE

Se rechaza H0, con un 95% de confianza, por lo que las variables excluidas en conjunto son significativas.

- Test conjunto con modelo de regresión simple

***Modelo no restringido:***

$$pib = \beta_0 \text{ gorro} + \beta_1 \text{ gorro} \times \text{alf} + \beta_2 \text{ gorro} \times \text{den} + \beta_3 \text{ gorro} \times \text{mi} + \beta_4 \text{ gorro} \times \text{inf} + \beta_5 \text{ gorro} \times \text{tel} + \beta_6 \text{ gorro} \times \text{nat} + \beta_7 \text{ gorro} \times \text{mor} + \beta_8 \text{ gorro} \times \text{cos} + \beta_9 \text{ gorro} \times \text{mixcos}$$

***Modelo restringido:***

$$pib = \beta_0 \text{ gorro} + \beta_1 \text{ gorro} \times \text{alf}$$

- Hipotesis

$$H_0: \beta_2 \text{ denl} = 0, \beta_3 \text{ mi} = 0, \beta_4 \text{ inf} = 0, \beta_5 \text{ tel} = 0, \beta_6 \text{ nat} = 0, \beta_7 \text{ mor} = 0, \beta_8 \text{ cos} = 0, \beta_9 \text{ mixcos} = 0$$

$$H_1: \beta_2 \text{ denl} \neq 0, \beta_3 \text{ mi} \neq 0, \beta_4 \text{ inf} \neq 0, \beta_5 \text{ tel} \neq 0, \beta_6 \text{ nat} \neq 0, \beta_7 \text{ mor} \neq 0, \beta_8 \text{ cos} \neq 0, \beta_9 \text{ mixcos} \neq 0$$

***F calculado***

54.12113

***F en tabla, alfa 5%***

1.986628

***Regla de rechazo***

```
cat(fc2 > ft2)
```

TRUE

Se rechaza H0, con un 95% de confianza, por ende las variables son significativas conjuntamente para explicar la variable dependiente.

- **Test de significancia global**

Se evaluará si el modelo de regresión lineal multiple global explica la variabilidad del PIB per cápita.

- **Hipotesis**

$H_0$ :  $\beta_1 \text{ alf} = 0$ ,  $\beta_2 \text{ den} = 0$ ,  $\beta_3 \text{ mi} = 0$ ,  $\beta_4 \text{ inf} = 0$ ,  $\beta_5 \text{ tel} = 0$ ,  $\beta_6 \text{ nat} = 0$ ,  $\beta_7 \text{ mor} = 0$ ,  $\beta_8 \text{ cos} = 0$ ,  $\beta_9 \text{ mixcos} = 0$

$H_1$ :  $\beta_1 \text{ alf} \neq 0$ ,  $\beta_2 \text{ den} \neq 0$ ,  $\beta_3 \text{ mi} \neq 0$ ,  $\beta_4 \text{ inf} \neq 0$ ,  $\beta_5 \text{ tel} \neq 0$ ,  $\beta_6 \text{ nat} \neq 0$ ,  $\beta_7 \text{ mor} \neq 0$ ,  $\beta_8 \text{ cos} \neq 0$ ,  $\beta_9 \text{ mixcos} \neq 0$

Según el modelo analizado, se observa un F statistic de 73.291. Sin embargo, se realiza el calculo de manera manual.

***F calculado***

73.44248

***F en tabla, alfa 5%***

1.928647

***Regla de rechazo***

```
cat(fc3>ft3)
```

TRUE

Se rechaza  $H_0$  con un 95% de confianza, por ende, las variables explicativas en su conjunto son significativas.

## Conclusión

El estudio analizó si la tasa de alfabetización explica las diferencias en el PIB per cápita entre 203 países. Utilizando un modelo econométrico, se encontró que la tasa de alfabetización tiene un efecto estadísticamente significativo en el PIB per cápita.

Inicialmente, se realizó un modelo de regresión lineal simple, en el cual se encontró que un aumento del 1% en la tasa de alfabetización se asoció con un aumento promedio de 267.440 dólares en el PIB per cápita, manteniendo constantes las demás variables. Sin embargo, el valor del parámetro del intercepto ( $-12,434.340$  dólares) resultó contraintuitivo, ya que el PIB per cápita no debería ser negativo cuando la tasa de alfabetización es cero.

Se realizaron comparaciones con las investigaciones, que buscaron la relación entre la educación y el crecimiento económico en diferentes países. Estos estudios respaldan la idea de que la inversión en educación, incluida la alfabetización, es fundamental para impulsar el desarrollo económico.

Por otra parte, se planteó un modelo de regresión lineal múltiple, en el que se incluye variables adicionales como densidad poblacional, migración, mortalidad infantil, cantidad de teléfonos, natalidad, mortalidad y la presencia de costa. Este modelo mostró un  $R^2$  ajustado de 0.763, lo que significa que aproximadamente el 76.3% de la variación en el PIB per cápita puede ser explicada por estas variables. Aunque la tasa de alfabetización sigue siendo una variable significativa, su impacto disminuye cuando se consideran otras variables explicativas.

Finalmente, el estudio proporciona evidencia que respalda la importancia de la alfabetización en el desarrollo económico de un país. Aunque existen otros factores que influyen en el PIB per cápita, la tasa de alfabetización sigue desempeñando un papel significativo. Estos hallazgos son importantes para conocer la importancia de la educación en el desarrollo económico.

## Bibliografía

- Messias, E. (2003, septiembre). *Income Inequality, Illiteracy Rate, and Life Expectancy in Brazil*. American Journal of Public Health.
- Monteils, M. (2002). *Education and Economic Growth: Endogenous Growth Theory Test. The French Case*. GESIS - Leibniz Institute for the Social Sciences
- Wooldridge, J. (2009). *Introducción a la econometría, Un enfoque moderno* - Cengage Learning Editores, S.A. de C.V.