



Facultad de ciencias

Modelo Lineales con R

“El precio de la tortilla en México”

Alumno:

Jesus Alfredo Aguilar Guadalupe

27 de mayo de 2024

Índice

1. Resumen

2. Introducción

3. Planteamiento del problema

4. Objetivos

5. Marco teórico

- 5.1. Modelos de regresión lineal simple
- 5.2. Estimación por mínimos cuadrados
- 5.3. Modelos lineales con factores
- 5.4. Criterio de Akaike
- 5.5. Razón de verosimilitud (log-likelihood)
- 5.6. Modelos lineales mixto en general
- 5.7. Índice de precios

6. Metodología

- 6.1. Obtención y descripción de los datos
- 6.2. Construcción de los modelos simples
- 6.3. Construcción del modelo mixto 1
- 6.4. Factores aleatorios
- 6.5. Obtención y descripción de los datos(precio maíz)
- 6.6. Construcción del modelo mixto 2

7. Resultados

- 7.1. Modelo lineal simple
- 7.2. Modelo lineal con tipo de cambio
- 7.3. modelo lineal mixto

8. Conclusión

9. Referencias bibliográficas

10. Anexo

1. Resumen

En el presente trabajo se enfoca en analizar los factores que influyen en el precio de la tortilla en México utilizando modelos de regresión lineal simple y mixtos en R. Las tortillas en México son un alimento básico en la dieta mexicana, y su precio es cuestión de interés público, ya que su precio afecta significativamente a la economía de los hogares, especialmente a los de bajos recursos. Se recolectaron datos históricos sobre el precio de la tortilla y el tipo de cambio desde el 2007 hasta el 2024. Se propusieron tanto modelos simples como modelos mixtos, para identificar y cuantificar las principales determinantes del precio de la tortilla, considerando variables como el año, la región y el tipo de cambio así también el índice del precio del maíz. Los resultados indican que tanto el año como el tipo de cambio tienen un efecto significativo en el precio de la tortilla, variaciones regionales importantes.

2. Introducción

Las tortillas son un alimento fundamental en la dieta mexicana, con una importancia cultural y económica significativa en el país. El precio de las tortillas es un tema de interés público, ya que puede afectar a millones de personas en México, especialmente a aquellas de bajos ingresos para quienes las tortillas son un alimento básico.

Existen factores que están involucrados en el aumento de este producto; unas las principales es el costo del maíz que también va subiendo por lo tanto, la harina con la que se hace la tortilla tiene un valor más elevado

Según la calculadora de inflación del INEGI, los precios en general en México han subido 55.45% en los últimos 10 años, en pocas palabras las tortillas han subido de precio más de lo normal.

En México en promedio se producen 20 millones de toneladas de maíz anuales, sin embargo todo ese producto se consume (ya que el maíz blanco es el mejor para producir tortilla). El resto se importa, provocando que el tipo de cambio afecte negativamente.

En el año 2015 en Torreón, Coahuila. El alza del dólar tuvo un impacto negativo en el precio del maíz lo que obligó a la industria de la masa a incrementar los precios de la tortilla. Los insumos fueron subiendo provocando que el maíz frente al deslice del peso con el dólar aumentara nuevamente. Algunas industrias de las masas incrementaron el precio de la tortilla de 12 a 16 pesos el kilogramo

Sinaloa es el principal productor de maíz en México. La secretaria de Agricultura y Desarrollo Rural señalo que el 22 por ciento de este producto (blanco y amarillo) proviene de tierra sinaloense.

A inicios de 2023, la inflación afectó fuertemente la economía mexicana. Productores, tortillerías e industriales han sido un golpe fuerte. El presidente del Comité Directivo de la Cámara Nacional de la Industria de Producción de Masa y Tortilla en Quintana Roo, Ruben

Montalvo Morales señaló que el incremento al kilo de la tortilla no es debido al precio que los productores tengan. De acuerdo al especialista el año pasado (2022) y antepasado(2021) se especuló mucho con el precio de la tortilla que los intermediarios que adquieren el maíz a bajo precio y lo vendieron mucho más caro

3. Planteamiento del problema

La situación del precio de la tortilla tiene implicaciones amplias en la estabilidad económica y social del país. Varios son los factores que contribuyen en la fluctuación del precio de la tortilla, siendo los más destacados el costo del maíz, variaciones en el tipo de cambio y las diferencias regionales.

4. Objetivos

- Realizar un análisis del precio de la tortilla en la región norte y sur.
- Identificar tendencias de aumento en el precio de la tortilla desde 2007 hasta el 2023.
- Analizar las fluctuaciones del tipo de cambio y su efecto en el precio de la tortilla.
- Construir modelos de regresión lineal simple y mixtos.

5. Marco teórico

El precio de la tortilla es un tema de gran relevancia en México debido a su importancia en la dieta diaria de la población. Es importante entender los factores que influyen. Este estudio se enfoca en el análisis del precio de la tortilla utilizando modelos de regresión lineal en R, con el objetivo de identificar y cuantificar las determinantes más importantes que influyen o afectan su precio

5.1 Modelos de regresión lineal simple

La regresión lineal es un método útil para predecir una respuesta cuantitativa Y partiendo de una sola variable predictora X, se asume que hay una relación lineal entre X e Y.

Matemáticamente, esta relación lineal se representa como:

$$Y = \beta_0 + \beta_1 X + \epsilon \quad (1)$$

Donde β_0 es un parámetro desconocido, y ϵ son variables aleatorias mutuamente independientes con media cero y varianza σ^2

5.2. Estimación por mínimos cuadrados

El objetivo es obtener unos estimadores insesgados con los que el modelo lineal se ajuste bien a los datos disponibles. El método de mínimos cuadrados se encarga de minimizar la suma de los residuos al cuadrado:

$$s = \sum_{i=0}^n (y_i - \mu_i)^2 \quad (2)$$

Los residuos es la diferencia entre cada valor de la variable observada y la predicha del modelo. Algunos residuos serán positivos y otros negativos. La recta que se ajuste mejor a los datos tendrá residuos pequeños.

5.3. Modelos lineales con factores

Los modelos donde los datos se dividen en distintos grupos cada uno con distinta media, son menos evidentes de presentar la forma:

$$\mu = x\beta \quad (3)$$

Considere el modelo $\mu_i = \beta_j$ si la observación i se encuentra en el grupo j .

5.4. Criterio de Akaike

El AIC busca un balance entre la capacidad predictiva de un modelo (varianza explicada) y la cantidad de parámetros que este debe considerar para lograr un mejor ajuste. AIC no solamente recompensa la bondad del ajuste, sino también incluye una penalidad, la función creciente del número de parámetros estimados.

5.5. Razón de verosimilitud (log-likelihood)

Es una prueba de hipótesis que ayuda a elegir el mejor modelo entre dos modelos anidados o jerarquizados. La verosimilitud, en términos simples, mide qué tan probable es que un conjunto de datos haya sido generado por un modelo particular.

5.6. Modelos lineales mixto en general

Los modelos lineales mixtos generalizados amplían el modelo lineal de modelo que:

- El objetivo está linealmente relacionado con los factores y covariables mediante una función de enlace especificada.
- El objetivo puede tener una distribución no normal
- Las observaciones se pueden correlacionar

El modelo lineal mixto es una extensión del modelo lineal general, en el que se supone que los factores y las variables tienen una relación lineal con la variable dependiente.

Factores

Los predictores categóricos deben seleccionarse como factores en el modelo. Cada nivel de un factor puede tener un efecto lineal diferente en el valor de la variable dependiente

Los factores de efectos fijos generalmente se consideran variables cuyos valores de interés se representan todos en el archivo de datos,

Los factores de efectos aleatorios son variables cuyos valores en el archivo de datos se pueden considerar una muestra aleatoria de una población mayor de valores.

Matemáticamente se representa como:

$$y = X\beta + Zb + \varepsilon \quad (4)$$

$$b \sim N(0, \Sigma_{\theta})$$

$$\varepsilon \sim N(0, \Sigma_{\phi})$$

donde X es una matriz de coeficientes fijos que describe la dependencia de y en relación a los efectos aleatorios. El principal parámetro a estimar es Σ_{θ}

Se tiene que

$$y \sim N(X\beta, Z\Sigma_{\theta}Z^T + \Sigma_{\phi}) \quad (5)$$

5.7. Índice de precios

Índice base:

- Un índice de precios establece un “año base” o “mes base” y asigna a ese punto un valor de 100.
- Todos los otros valores del índice son comparados con este punto base.

6. Metodología

Para llevar a cabo el análisis, se utilizaron datos históricos sobre el precio de la tortilla, el tipo de cambio. Se emplean modelos lineales en R para realizar los análisis estadísticos y construir los modelos de regresión lineal, múltiple o mixto.

Algunos pasos específicos que incluyen son:

- ❖ Obtencion y preparacion de datos
- ❖ Análisis exploratorio de datos
- ❖ Ajuste de modelos

❖ Evaluación de modelos

6.1. Obtención y descripción de los datos

Los datos de los precios de la tortilla fueron recabados de la página kaggle, contiene 270 mil datos del precio de la tortilla, de 53 ciudades de todo México desde 2007 hasta la fecha (2024), la información está dividida por estado, ciudad, año, día, tienda y precio por kilogramo para este análisis solo se tomará en cuenta los estados del norte y sur, año y precio por kilogramo.

Las regiones que se definieron para este análisis fueron norte y sur. Los estados del norte son: Baja California, Sonora, Chihuahua, Nuevo León y Tamaulipas; mientras que los del sur son: Chiapas, Oaxaca, Guerrero, Quintana Roo, Yucatán y Campeche

Se implementó un modelo lineal inicial solo con el año como variable para analizar cómo ha cambiado el precio de la tortilla a lo largo del tiempo. Mostrando así sus estadísticas y visualizaciones básicas

Los datos del tipo de cambio fueron recabados de Investing.com se recolectaron los datos desde el año 2007 hasta la fecha (2024) está organizado por fechas (día, mes y año), cierre, apertura, máximo y mínimo.

Se implementó la combinación de los datos del precio de la tortilla con los datos del tipo de cambio. Se ajustó un modelo lineal que incluye el tipo de cambio como variable predictiva.

6.2. Construcción del modelo simple

El modelo simple compara el precio de la tortilla por kilogramo en función de la región y el año matemáticamente:

$$y = \beta_0 + \beta_1 \text{Año} + \beta_2 \text{Región sur} + \varepsilon \quad (6)$$

- β_0 : Es el intercepto o el precio base.
- β_1 : Es el cambio en el precio por cada aumento de un año
- β_2 : Es la diferencia en el precio entre la región Norte(categoría de referencia) y la región sur.
- ε : Es el término de error.

El modelo que incluye el tipo de cambio, la región y el año matemáticamente es:

$$y = \beta_0 + \beta_1 \text{Año} + \beta_2 \text{Región sur} + \beta_3 \text{Tipo de cambio} + \varepsilon \quad (7)$$

- $\beta_0, \beta_1, \beta_2$: Mismas interpretaciones que el modelo anterior
- β_3 : Es el efecto adicional en el precio por cada unidad de aumento en el tipo de cambio.

6.3. Construcción del modelo mixto

En la construcción del modelo mixto se involucró el precio de la tortilla y el tipo de cambio, considerando las diferencias regionales (norte y sur).

Factores fijos:

Los factores fijos son aquellos que esperamos que tengan un efecto sistemático y constante sobre la respuesta:

- Año: para capturar las tendencias generales a lo largo del tiempo
- Tipo de cambio: para ver cómo las fluctuaciones en el tipo de cambio afectan los precios de la tortilla
- Región: norte y sur, para modelar diferencias estructurales en los precios entre estas áreas.

6.4. Factores aleatorios

Interceptos aleatorios por año dentro de cada región: Permite que el nivel base de los precios de la tortilla varíe cada año y por región, capturando variabilidad específica no explicada por las tendencias de tiempo lineales o los efectos del tipo de cambio

$$y_{ij} = \beta_0 + \beta_1\{Year\}_{ij} + \beta_2\{tipocambio\}_{ij} + \beta_3\{RegionSur\}_{ij} + \mu_{ji} + \epsilon_{ij} \quad (8)$$

- y_{ij} : Precio por kilogramo de la tortilla para el año i en la región j .
- β_0 : Intercepto, representa el precio base cuando todas las variables explicativas son cero.
- β_1 : Efecto de año sobre el precio de la tortilla
- β_2 : Efecto del tipo de cambio sobre el precio de la tortilla.
- β_3 : Diferencia de precios entre las regiones (sur comparado con el norte como referencia)

- μ_{jt} : Efectos aleatorios por combinación de año y región, que captura las variaciones específicas no explicadas por los efectos fijos.
- ϵ_{ij} : Error residual, asumido normalmente distribuido

El grupo “Year:Región” (intercepto): varianza y desviación estándar de los interceptos aleatorios para cada combinación de año y región, indicando cuánto varían los precios base entre diferentes combinaciones de año y región

6.5. Obtención y descripción de los datos(precio maíz)

Los datos fueron recabados del INEGI en el apartado de índices de precios de genéricos para producción total. Los datos del índice de precio de maíz está en base julio 2019 = 100

Si el índice en un mes posterior es mayor que 100, indica que los precios han aumentado comparación con junio de 2012 y viceversa.

Esto permite que los efectos de la inflación general estén implícitamente ajustados, facilita la comparación de los precios del maíz a lo largo del tiempo de una manera estandarizada.

En el gráfico se observa la tendencia ascendente en el precio del maíz. aunque hay fluctuaciones, el precio generalmente ha aumentado



Figura 1. Gráfico evolución precio del maíz

La tabla (figura 1.1) muestra un resumen estadístico de las columnas “fecha” y “V2”(índices de precio maíz).

fecha	V2
Min. :1981-01-01	Min. : 0.05412
1st Qu.:1991-10-24	1st Qu.: 14.97156
Median :2002-08-16	Median : 52.10171
Mean :2002-08-16	Mean : 53.31859
3rd Qu.:2013-06-08	3rd Qu.: 85.80959
Max. :2024-04-01	Max. :166.90556

Figura 1.1. Resumen estadístico

- El primer cuartil es de 114.9756, lo que significa que el 25% de los precios del maíz están por debajo de este valor.
- El tercer cuartil es de 85.8959, el 75% de los precios están por debajo de este valor, implica que el 25% restante está por encima.

6.6. Construcción del modelo mixto 2

El siguiente modelo está diseñado para analizar cómo varía el precio de las tortillas en función del año, del precio del maíz y de la región geográfica, tomando en cuenta las diferencias fijas entre regiones como variaciones aleatorias dentro de ellas.

$$y_{ij} = \beta_0 + \beta_1\{Year\}_{ij} + \beta_2\{V2\}_{ij} + \beta_3\{RegionSur\}_{ij} + \mu_{ji} + \epsilon_{ij} \quad (9)$$

- y_{ij} : Precio por kilogramo de la tortilla para el año i en la región j .
- β_0 : Interceptor, representa el precio base cuando todas las variables explicativas son cero.
- β_1 : Efecto de año sobre el precio de la tortilla
- β_2 : Efecto del precio del maíz sobre el precio de la tortilla.

- β_3 : Diferencia de precios entre las regiones (sur comparado con el norte como referencia)
- μ_{jt} : Efectos aleatorios por combinación de año y región, que captura las variaciones específicas no explicadas por los efectos fijos.
- ϵ_{ij} : Error residual, asumido normalmente distribuido

7. Resultados

El coeficiente Year es positivo, indicando que hay una tendencia creciente en el precio por kilogramo, específicamente cada año, el precio aumenta en promedio 0.76 unidades monetarias. El residual standard error es de 2.979 esto nos dice cuánto varían los precios reales del modelo predicho (en promedio), R-squared es de 0.5274 indica que aproximadamente el 52.74% de la variabilidad en el precio por kilogramo se explica por las variables incluidas en el modelo este valor es ligeramente más alto que la del modelo simple (0.521) , esto sugiere que el modelo con el tipo de cambio explica una mayor proporción de la variabilidad.

El modelo con tipo de cambio tiene un valor F más alto, lo que indica una mayor significancia estadística.

```

Call:
lm(formula = Price.per.kilogram ~ Region + Year, data = datos)

Residuals:
    Min       1Q   Median       3Q      Max
-6.9278 -2.5113 -0.3075  2.1301 12.5225

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.357e+03  4.267e+00 -318.00  <2e-16 ***
RegionSur    -9.501e-01  2.239e-02  -42.43  <2e-16 ***
Year         6.797e-01  2.117e-03  321.04  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.236 on 96740 degrees of freedom
Multiple R-squared:  0.521,    Adjusted R-squared:  0.521
F-statistic: 5.261e+04 on 2 and 96740 DF,  p-value: < 2.2e-16

Call:
lm(formula = Price.per.kilogram ~ Region + Year + tipocambio,
    data = combined_data)

Residuals:
    Min       1Q   Median       3Q      Max
-8.1000 -2.2475 -0.7216  2.1490 13.5963

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.736e+03  4.785e+00 -362.84  < 2e-16 ***
RegionSur    4.995e-01  1.054e-01   4.74 2.13e-06 ***
Year         8.704e-01  2.398e-03  363.00  < 2e-16 ***
tipocambio   -3.862e-01  3.342e-03 -115.57  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.979 on 270122 degrees of freedom
(5672 observations deleted due to missingness)
Multiple R-squared:  0.5274,    Adjusted R-squared:  0.5273
F-statistic: 1.005e+05 on 3 and 270122 DF,  p-value: < 2.2e-16

```

Figura 2. Resumen estadístico del modelo simple y modelo simple con tipo de cambio.

El gráfico de la figura 3 muestra la evolución del precio de la tortilla por año para cada región (norte y sur). Las líneas azules y rojas representan las regiones norte y sur respectivamente. Las variaciones anuales en el precio se observan claramente y parece que hay cierta tendencia de aumento a lo largo del tiempo.

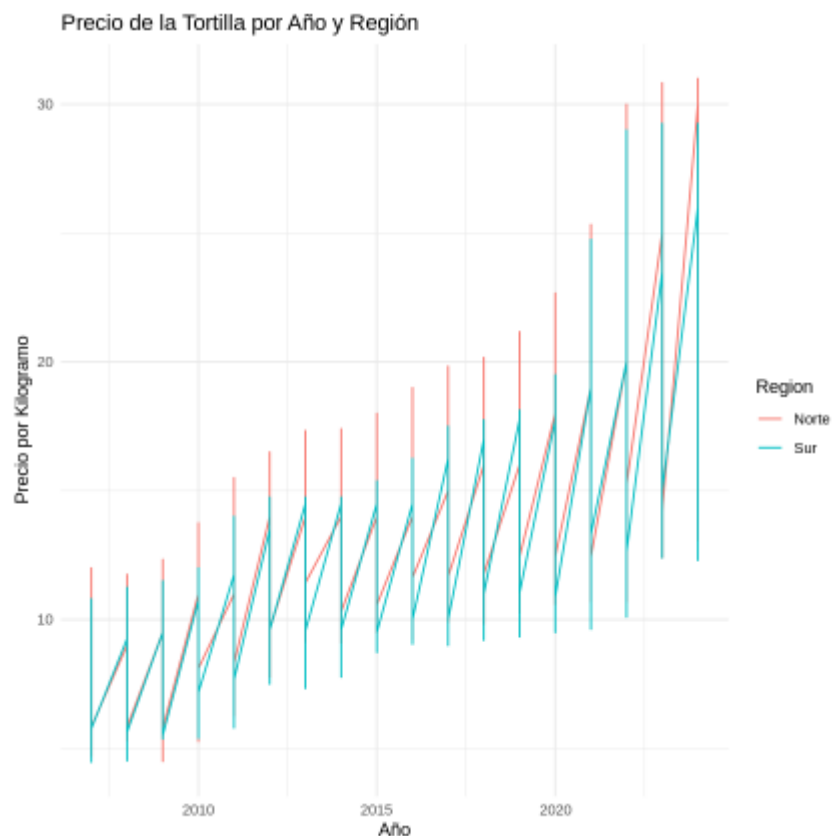


Figura 3. Gráfico precio de la tortilla por año y región (norte y sur).

En el gráfico q-q del modelo simple se nota que los puntos siguen bastante de cerca la línea verde, lo que indica que los residuos en el centro de la distribución son aproximadamente normales. Pero hay desviaciones en los extremos del gráfico, lo que indica que hay colas más pesadas que lo que se esperaría bajo una distribución normal.

En el caso del modelo simple con el tipo de cambio el gráfico muestra un patrón similar pero con una ligera mejora en el ajuste a la línea azul. Los puntos de adhieren bastante bien a la línea, sugiriendo así una normalidad en esa parte de la distribución de residuos.

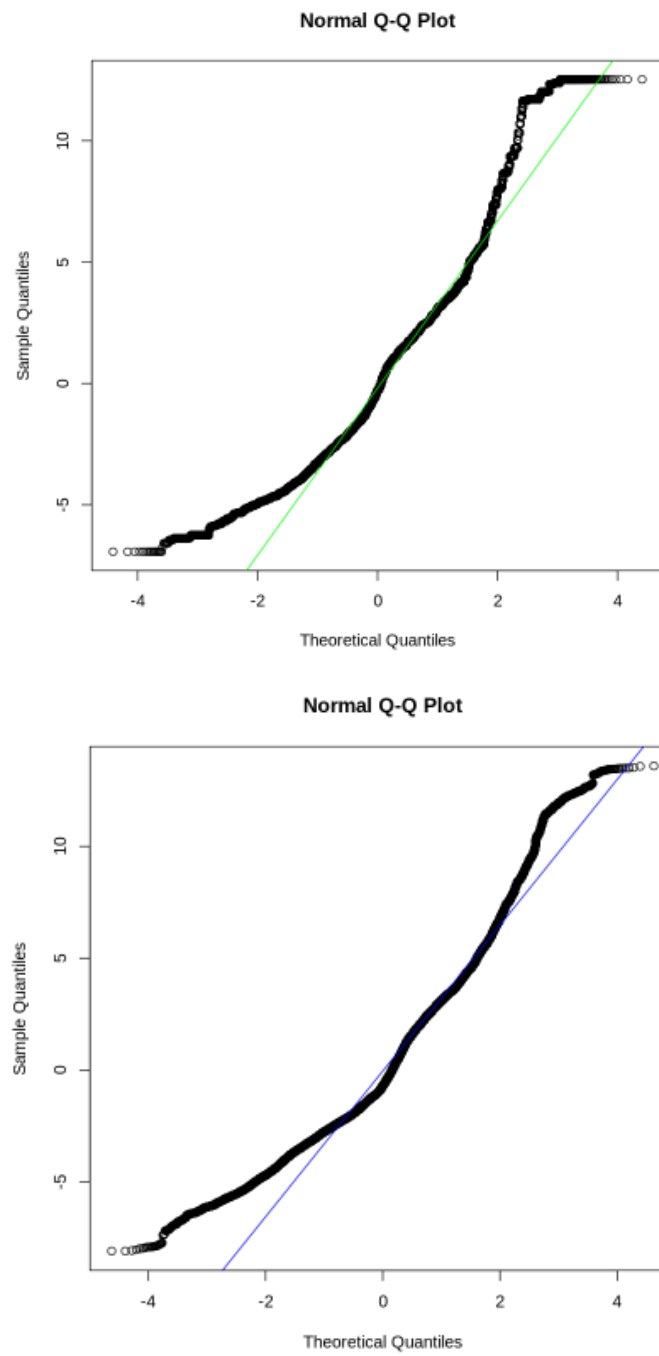
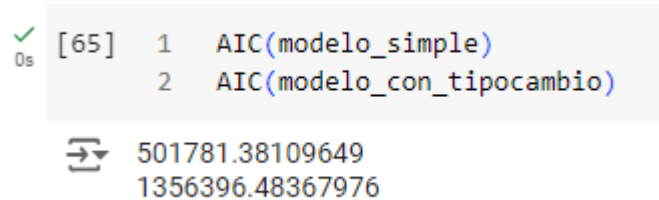


Figura 4. Gráficos Q-Q del modelo simple y modelo simple con tipo de cambio.

El modelo simple es más preferible para la toma de decisión ya que tiene el AIC más bajo, sugiere que el modelo es más eficaz en términos de equilibrio entre la complejidad del modelo y el ajuste al conjunto de datos.

Entonces aunque añadamos el tipo de cambio puede mejorar el ajuste del modelo, esto no justifica el costo adicional de complejidad,



A screenshot of an R console window. On the left, there is a green checkmark icon and the text '0s'. The main area shows the output of the `AIC()` function for two models. The first line is '[65] 1 AIC(modelo_simple)' and the second line is '2 AIC(modelo_con_tipocambio)'. Below these, there are two numerical values: '501781.38109649' and '1356396.48367976'. A small icon of a double-headed arrow is positioned to the left of these numbers.

Model	AIC Value
1 AIC(modelo_simple)	501781.38109649
2 AIC(modelo_con_tipocambio)	1356396.48367976

Figura 4.1. Criterio de Akaike.

El gráfico de cajas que se muestra a continuación compara los precios de la tortilla en las regiones norte y sur de México.

Donde tenemos un error estándar de los residuos de 3.1198, que mide la cantidad típica que los valores observados difieren de la línea de regresión, el R-squared es 0.4994, indica que aproximadamente el 49% de la variabilidad en los precios de la tortilla está explicada por el modelo. El Adjusted R-squared es similar a R-squared lo que indica que el modelo es adecuado en términos de número de predictores y el estadístico F es 2.937e-04 con un p-value muy pequeño, lo que indica que el modelo es estadísticamente significativo. Esto indica que hay una diferencia en las tendencias de los precios entre el norte y el sur.

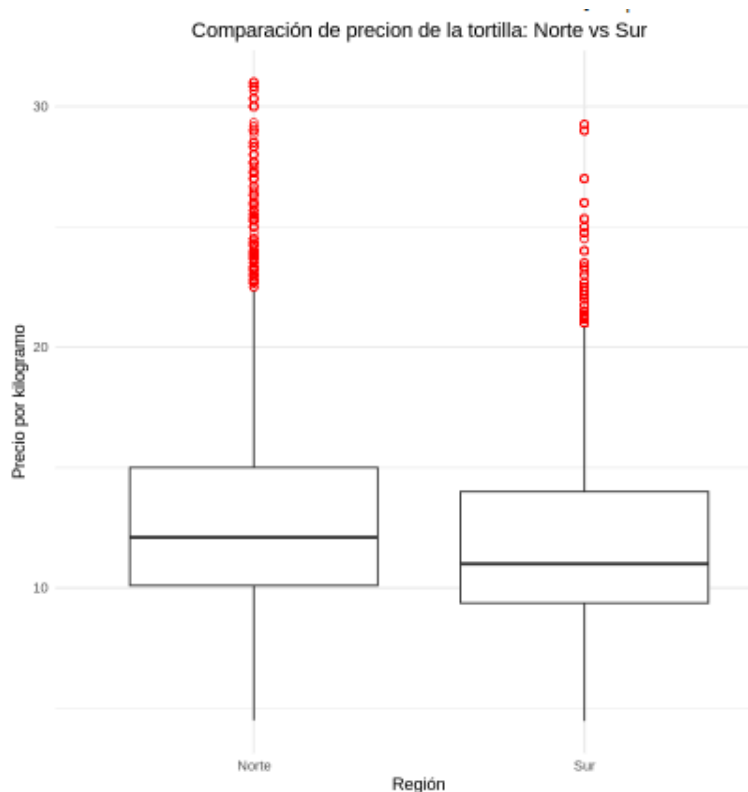


Figura 5. Gráfico precio de la tortilla por año y región (norte y sur)

Cada año adicional está asociado con un aumento aproximado de 0.699 en el precio por kilogramo de tortilla. Este efecto es estadísticamente significativo. El tipo de cambio tiene un pequeño efecto negativo en el precio de la tortilla, lo que significa que un aumento en el tipo de cambio está asociado con una disminución en el precio.

Comparando la región norte (la categoría de referencia), estar en la región sur está asociado con una disminución promedio en el precio por kilogramo de tortilla de aproximadamente 0.990. Este coeficiente es estadísticamente significativo (“3.034”)

El gráfico tiene una dispersión bastante uniforme de los residuos alrededor del cero, con algunos valores extremos. Generalmente esto es bueno, para los valores extremos pueden ser valores atípicos.


```

Linear mixed model fit by REML ['lmerMod']
Formula: Price.per.kilogram ~ Year + tipocambio + Region + (1 | Year:Region)
Data: combined_data

REML criterion at convergence: 490164

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.7083 -0.7209 -0.1947  0.7262  3.8016

Random effects:
 Groups      Name      Variance Std.Dev.
Year:Region (Intercept) 0.955    0.9773
Residual          9.670    3.1097
Number of obs: 95940, groups: Year:Region, 36

Fixed effects:
              Estimate Std. Error t value
(Intercept) -1.397e+03  6.519e+01 -21.424
Year         6.996e-01  3.237e-02  21.615
tipocambio  -1.213e-02  1.310e-02  -0.926
RegionSur   -9.909e-01  3.266e-01  -3.034

Correlation of Fixed Effects:
              (Intr) Year   tipcmb
Year          -1.000
tipocambio    0.226 -0.229
RegionSur    -0.003  0.001  0.000

```

Figura 6.Resumen estadístico modelo mixto 1.

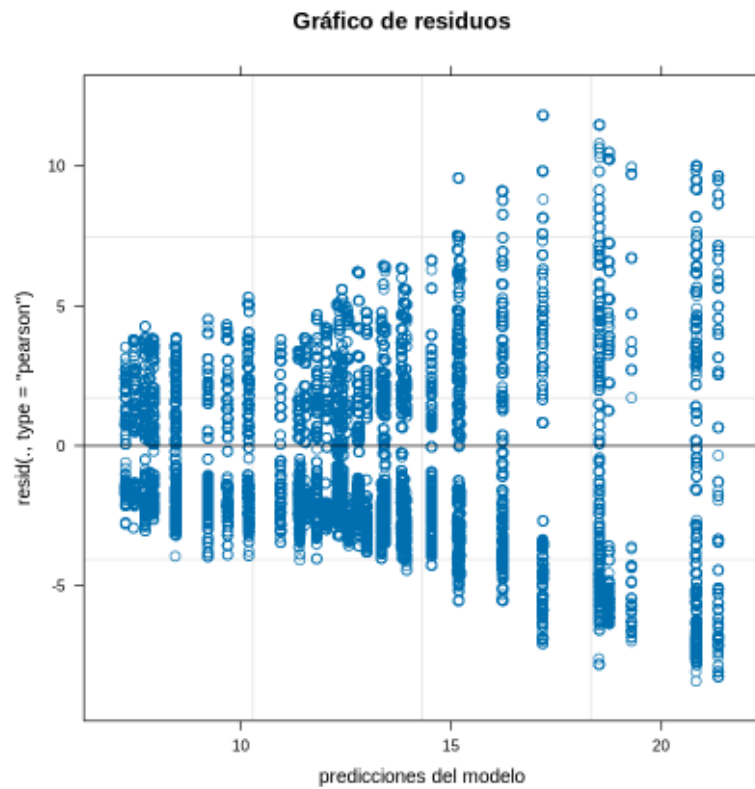


Figura 7. Gráfico de residuos del modelo mixto 1.

La línea roja representa la distribución esperada si los residuos fueran perfectamente normales. Los puntos parece que la mayoría siguen la línea roja, lo que indicaría que los residuos se ajustan bien a una distribución normal. Aunque hay algunas desviaciones en los extremos.

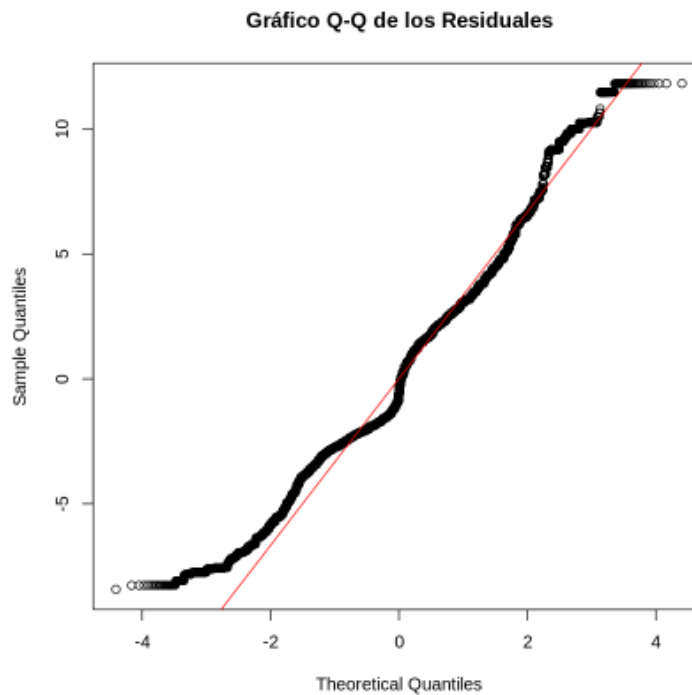


Figura 8. Gráfico Q-Q del modelo mixto 1.

Hay un aumento anual promedio en el precio por kilogramo de las tortillas de aproximadamente 0.576 unidades monetarias. Este efecto es estadísticamente significativo, lo que sugiere que el precio de las tortillas ha aumentado a lo largo de los años.

El efecto del precio del maíz(efecto fijo) nos dice que por cada unidad que aumenta el precio del maíz, el precio por kilogramo de las tortillas aumenta en 0.0249 unidades monetarias, esto indica que el precio del maíz tiene un impacto positivo en el precio de las tortillas.

comparando la región norte (la categoría de referencia), el precio por kilogramo de las tortillas en la región sur es en promedio 0.263 unidades monetarias más bajo.

Los efectos fijos “Year” y “V2” están altamente correlacionados, lo que podría indicar multicolinealidad entre estas variables. Esto puede afectar la precisión de las estimaciones de los coeficientes.

```

➡ Linear mixed model fit by REML ['lmerMod']
Formula: Price.per.kilogram ~ Year + V2 + Region + (1 | Year:Region)
Data: combined_data

REML criterion at convergence: 13814.1

Scaled residuals:
    Min       1Q   Median       3Q      Max
-2.3927 -0.7088 -0.1520  0.7111  3.3940

Random effects:
Groups      Name      Variance Std.Dev.
Year:Region (Intercept) 0.526   0.7253
Residual                10.168   3.1887
Number of obs: 2666, groups: Year:Region, 36

Fixed effects:
              Estimate Std. Error t value
(Intercept) -1.150e+03  9.776e+01 -11.763
Year         5.759e-01  4.886e-02  11.789
V2           2.480e-02  8.739e-03   2.838
RegionSur    -9.623e-01  2.803e-01  -3.433

Correlation of Fixed Effects:
          (Intr) Year    V2
Year      -1.000
V2         0.826 -0.829
RegionSur -0.018  0.017 -0.012

```

Figura 9. Resumen estadístico del modelo mixto 2.

El gráfico de residuos muestra la dispersión de los errores del modelo frente a los valores ajustados. El verde representa los datos que pertenecen a la región norte y el rojo los datos que pertenecen a la región sur.

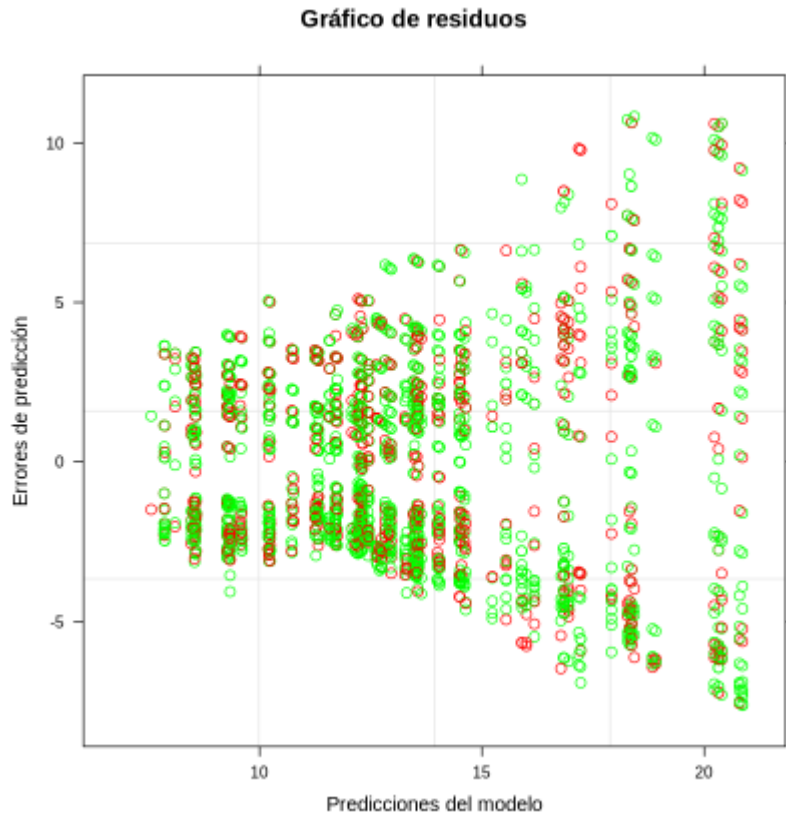


Figura 10. Gráfico de Residuos del modelo mixto 2.

Los puntos siguen de cerca la línea azul en gran parte del gráfico, sugiriendo que los residuos son aproximadamente normales. Esto es bueno por la suposición de la normalidad de los residuos. En los extremos hay ligeras desviaciones, lo que podría ser colas ligeramente más pesadas.

Los extremos son cruciales para identificar problemas de colas pesadas o ligeras. Una ligera desviación en los extremos (como la de la gráfica) puede ser aceptable.

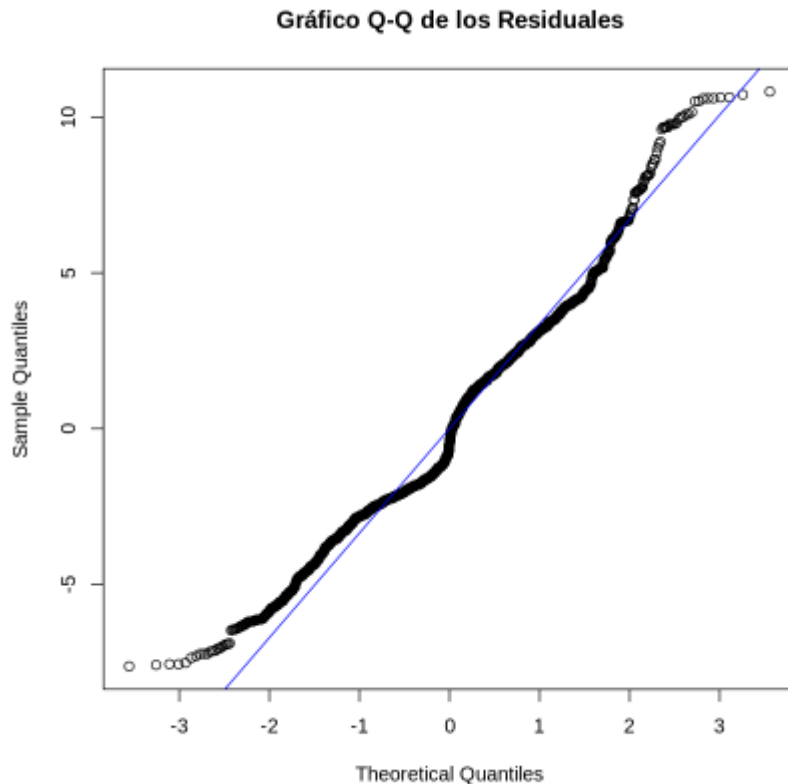


Figura 11. Gráfico Q-Q de los residuales del modelo mixto 2

De acuerdo a la R^2 para los dos modelos mixtos tenemos dos medidas R_m^2 (marginal) y R_c^2 (conditional). Para el modelo mixto(1) R_m^2 indica que el 52.45% de la variabilidad en el precio por kilogramo de la tortilla es explicada por los efectos fijos del modelo, mientras que R_c^2 incluye tanto efectos fijos como los aleatorios. El valor 0.5672013 sugiere que cuando se considera tanto los efectos fijos como los aleatorios, el modelo explica aproximadamente el 56.72% de la variabilidad en los precios.

Para el segundo modelo mixto R_m^2 aumentó ligeramente a 0.5403084, lo que sugiere una mejora en la cantidad de varianza en el precio por kilogramo de la tortilla explicada únicamente por las variables fijas (en comparación con el primer modelo). Esto nos dice que las variables fijas tienen una relación más fuerte con la respuesta en este modelo. R_c^2 ha aumentado a 0.5626557, indica que considerando tanto las variables fijas como los efectos aleatorios, este modelo puede explicar un poco más de la variabilidad en los precios que el modelo anterior.

```

      R2m      R2c
[1,] 0.5244575 0.5672013
Warning message in Year:Region:
"numerical expression has 2666 elements: only the first used"
      R2m      R2c
[1,] 0.5400304 0.5626557

```

Figura 12. Cálculo de R_m^2 y R_c^2

De acuerdo con la información de Akaike (AIC) el mejor ajuste es el modelo mixto 2 ya que el valor es significativamente menor que el primero, esto indica que el modelo mixto 2 es mucho más eficiente en términos de la relación entre el ajuste del modelo.

```

✓ [5] 1 AIC(modelo_mixto)
0s 2 AIC(modelo_mixto2)
3
⇌ 490175.959276914
13826.1356842058

```

Figura 13. Criterio de Akaike.

El P-value de la prueba, es de 0.0008414, este valor indica la probabilidad de obtener una estadística de prueba al menos tan extrema como la observada bajo la hipótesis nula de que ambos modelos se ajustan igualmente bien a los datos (es decir incluir “Región” no mejora el modelo).

El P-valor es extremadamente bajo lo que sugiere que hay una diferencia significativa en el ajuste entre el modelo reducido y completo.

Basado en la prueba Likelihood y la significancia estadística observada, es más recomendable utilizar el modelo completo que incluye la variable “región” para un mejor ajuste para un mejor ajuste y comprensión de los factores que afectan el precio por kilogramo en el conjunto de datos analizado.

```

⇒ refitting model(s) with ML (instead of REML)

Data: combined_data
Models:
modelo_reducido: Price.per.kilogram ~ Year + V2 + (1 | Year:Region)
modelo_completo: Price.per.kilogram ~ Year + V2 + Region + (1 | Year:Region)
               npar   AIC   BIC logLik deviance Chisq Df Pr(>Chisq)
modelo_reducido    5 13819 13849 -6904.6    13809
modelo_completo    6 13810 13845 -6899.1    13798 11.148  1 0.0008414 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Figura 14. Prueba Likelihood para modelo completo y reducido.

8. Conclusión

El análisis realizado en este trabajo muestra que el precio de la tortilla en México ha experimentado un aumento significativo a lo largo de los años, influenciado por factores como el tipo de cambio y las diferencias regionales. El modelo mixto proporciona un mejor ajuste y una mejor comprensión de las variaciones en el precio de la tortilla en comparación con los modelos simples. Se observa que el tipo de cambio tiene un efecto significativo, aunque menor, en el precio de la tortilla, mientras que las diferencias regionales son también un factor importante. Estos hallazgos pueden ser útiles para la formulación de políticas públicas y estrategias comerciales que busquen estabilizar y controlar el precio de la tortilla.

9. Referencias

- [1] CNN Español. (2022). Así ha aumentado el precio de la tortilla en todo México en los últimos 10 años. Recuperado de:
<https://cnnespanol.cnn.com/2022/08/15/asi-ha-aumentado-el-precio-de-la-tortilla-en-todo-mexico-en-los-ultimos-10-anos-orix/>
- [2] Vanguardia. (2023). Alza del dólar provoca incremento en el precio de la tortilla. Recuperado de:
<https://vanguardia.com.mx/coahuila/2630632-alza-del-dolar-provoca-incremento-en-el-precio-de-la-tortilla-HMVG2630632>
- [3] RPubS. (s.f). Regresion lineal simple. Recuperado de:
https://rpubs.com/Cristina_Gil/Regresion_Lineal_Simple
- [4] Radio Fórmula. (2023). Jaque a la tortilla; precio del maiz pone en riesgo su futuro, advierten campesinos y tortilleros. Recuperado de:
<https://www.radioformula.com.mx/economia/2023/4/26/jaque-la-tortilla-precio-del-maiz-pone-en-riesgo-su-futuro-advierten-campesinos-tortilleros-759612.html>.
- [5]. Kaggle. (s. f.). *Tortilla prices in Mexico*. Recuperado de:
<https://www.kaggle.com/datasets/richave/tortilla-prices-in-mexico>.
- [6]. INEGI. (s.f). Índices de precios de genéricos para producción total. Recuperado de:
<https://inegi.org.mx/app/indicesdeprecios/Estructura.aspx?idEstructura=1120015000300030&ST=%C3%8Dndices%20de%20precios%20de%20gen%C3%A9ricos%20para%20producci%C3%B3n%20total>

10. Anexo

<https://github.com/jesusaguilar12/Modelos-lineales-R.git>