UNIVERSIDAD NACIONAL AGRARIA LA MOLINA

FACULTAD DE ECONOMÍA Y PLANIFICACIÓN

DEPARTAMENTO ACADÊMICO DE ESTADÍSTICA E INFORMÁTICA

Diseño de series temporales interrumpidas

- **Curso:** Diseños Experimentales 2
- **Profesor:** Beatriz Del Carmen Lidia Montaño Miranda

- **Integrantes:**

Apellidos y Nombres	Código
Blacido Hernandez, Pamela Brescia	20220758
Chavez Paucar, Carla Andrea	20220760
Untiveros cruzado, Alonso	20220783

2024-II

I. RESUMEN:

El diseño de series temporales interrumpidas (ITS) es una metodología frecuentemente utilizada para evaluar el impacto de intervenciones a gran escala, especialmente en salud pública. Este enfoque permite observar y medir los cambios antes y después de la implementación de una intervención. Sin embargo, cuando los datos presentan estacionalidad y autocorrelación, los métodos tradicionales como la regresión segmentada pueden no ser suficientes. En estos casos, los modelos Autorregresivos Integrados de Media Móvil (ARIMA) ofrecen una solución más robusta.

Este informe, basado en el artículo "Análisis de series temporales interrumpidas utilizando modelos ARIMA", describe los fundamentos de estos modelos y su aplicación en la evaluación de políticas de salud. Se utiliza como estudio de caso la política implementada en Australia para restringir la prescripción de quetiapina de 25 mg, lo que permite observar tanto los efectos inmediatos como las tendencias a largo plazo en la dispensación de este medicamento. A través de la aplicación de ARIMA, se demuestran reducciones significativas en la dispensación, que no habrían sido identificadas con otros enfoques. Los resultados destacan la utilidad de los modelos ARIMA para estudios donde la regresión segmentada no es suficiente, aportando flexibilidad y precisión en contextos con alta variabilidad temporal.

II. INTRODUCCION:

El diseño de series temporales interrumpidas es una herramienta clave para evaluar intervenciones a gran escala en el ámbito de la salud pública, donde los ensayos controlados aleatorizados no siempre son factibles o éticos. Este enfoque permite analizar los efectos de una intervención al observar patrones de datos antes y después de su implementación. A pesar de que el método de regresión segmentada es comúnmente utilizado, presenta limitaciones cuando los datos contienen autocorrelaciones o estacionalidades. En tales casos, los modelos Autorregresivos Integrados de Media Móvil (ARIMA) ofrecen una alternativa robusta, ya que son capaces de capturar estas complejidades.

El presente informe se basa en el artículo "Análisis de series temporales interrumpidas utilizando modelos ARIMA", el cual explora cómo los modelos ARIMA pueden mejorar la precisión en la evaluación de políticas de salud. Se utiliza un estudio de caso sobre la política del gobierno australiano para limitar la prescripción de quetiapina de 25 mg, demostrando cómo este enfoque permite modelar tanto cambios inmediatos como tendencias a largo plazo en la dispensación de medicamentos. Los resultados destacan cómo los modelos ARIMA pueden ser utilizados para estimar con precisión los cambios inmediatos y las tendencias a largo plazo resultantes de intervenciones en políticas de salud.

- ❖ **OBJETIVOS**
    - ➢ **Objetivo general:**

      Evaluar el impacto de las intervenciones de salud a gran escala utilizando el modelo ARIMA en análisis de series temporales interrumpidas, con el fin de mejorar la precisión en la medición de los efectos a corto y largo plazo.

    - ➢ **Objetivos específicos:**

      1. Aplicar el modelo ARIMA para abordar las limitaciones de la regresión segmentada en la presencia de estacionalidad y autocorrelación en los datos.

      2. Modelar y analizar los efectos inmediatos y sostenidos de una intervención de salud utilizando datos de series temporales antes y después de su implementación.

      3. Comparar los resultados obtenidos con los modelos ARIMA frente a enfoques más simples, destacando su capacidad para captar variaciones complejas en los datos.

      4. Proporcionar una guía práctica para la implementación de modelos ARIMA en estudios que involucren grandes intervenciones de salud pública.

Este estudio es importante ya que proporciona evidencia cuantitativa sobre la efectividad de políticas de salud pública a través de análisis robustos que controlan autocorrelación y estacionalidad. Además, el uso de modelos ARIMA añade precisión y confiabilidad en la evaluación de intervenciones cuando no se pueden realizar ensayos controlados aleatorios.

IV. REVISIÓN DE LA LITERATURA

Una serie temporal es una secuencia de puntos de datos en puntos espaciados de manera uniforme en el tiempo y ordenados cronológicamente. El uso del análisis de series temporales interrumpidas (ITS) ha sido ampliamente explorado como un método robusto para evaluar el impacto de intervenciones a nivel de políticas públicas, especialmente cuando no es factible realizar ensayos controlados aleatorizados (RCTs). Tradicionalmente, las regresiones segmentadas han sido una técnica común para el análisis ITS, que modela los datos pre y post intervención. Sin embargo, este enfoque tiene limitaciones importantes cuando se enfrenta a problemas de estacionalidad y autocorrelación en los datos, que son comunes en escenarios de salud pública.

Propiedades de series temporales interrumpidas
Las series temporales suelen presentar tres características:

- **No Estacionariedad:**
  Las principales fuentes de no estacionariedad son la varianza cambiante a lo largo del tiempo, llamada heteroscedasticidad, que puede corregirse mediante una transformación logarítmica, y la tendencia creciente o decreciente, que se aborda tomando la primera diferencia entre observaciones consecutivas. En ocasiones, puede ser necesario realizar una segunda diferenciación, aunque más allá de esto es poco común.

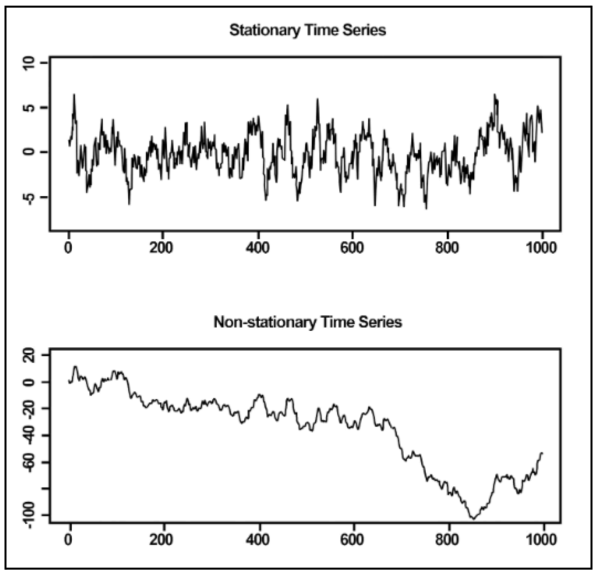

Figura 1. superior: serie de tiempo estacional, imagen inferior: serie de tiempo no estacional

- **Autocorrelación:**
  La autocorrelación, o correlación entre las observaciones de una serie temporal y sus valores en puntos anteriores, es común en las series temporales y rompe con los supuestos del análisis de regresión estándar. Para eliminar la autocorrelación, generalmente es suficiente diferenciar los datos. Las funciones de autocorrelación (ACF) ayudan a identificar la autocorrelación y la estacionariedad. La ACF gráfica la correlación entre cada observación y los valores en rezagos anteriores. Si la serie es estacionaria, la autocorrelación decae rápidamente; si no lo es, la ACF decae lentamente.

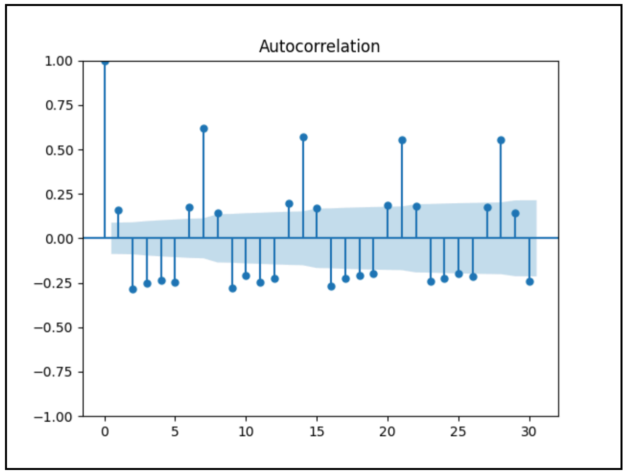

Figura 2. Autocorrelación en un diseño se series de tiempo

- **Estacionalidad:**
  Se refiere a fluctuaciones periódicas en intervalos regulares, como meses o días. En series temporales de datos de salud, la estacionalidad es común, influenciada por factores como los patrones climáticos o los ciclos administrativos.Para manejar la estacionalidad en modelos ARIMA, se aplica la diferencia estacional, que resta a cada observación el valor correspondiente en el mismo punto del ciclo anterior. Con datos estacionales es importante contar con suficientes puntos de tiempo para capturar adecuadamente estas fluctuaciones.

Según el artículo, *Schaffer et al. (2021)* destacan que los modelos autorregresivos integrados de media móvil (ARIMA) son una alternativa superior en estos casos, ya que permiten controlar tanto la autocorrelación como la estacionalidad, proporcionando un ajuste más preciso a los datos observacionales. Los estudios previos que han evaluado intervenciones a través de ITS con modelos ARIMA han mostrado resultados más confiables que con las regresiones segmentadas simples, lo que hace que este enfoque sea cada vez más utilizado en el ámbito de la salud pública para evaluar políticas como la reducción de prescripciones inadecuadas de medicamentos.

Componentes de los modelos ARIMA

Los modelos ARIMA se utilizan para predecir una variable dependiente Y_t basada en sus valores pasados y un término de error ε_t. Estos modelos asumen que los errores están distribuidos normalmente, lo que los hace adecuados para datos continuos, como tasas o medias, y recuentos grandes no limitados por cero. Sin embargo, no son apropiados para datos de recuento pequeños que siguen una distribución de Poisson.

Para estos casos, en los últimos años se han desarrollado enfoques que utilizan modelos lineales generalizados para modelar datos de recuento correlacionados en serie.

1. **Modelo autorregresivo (AR)**: El modelo Y_t se predice a partir de uno o varios valores rezagados de Y_t.

$$\boxed{Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_k Y_{t-k} + \varepsilon_t}$$

Donde:

- ❖ C : constante
- ❖ ϕ: magnitud de la autocorrelación
- ❖ k: número de rezagos
- ❖ ε_t: error

2. **Modelo de media móvil (MA):** Y_t se predice mediante uno o varios valores rezagados del error ε_t. Esto no debe confundirse con el suavizado de media móvil.

$$\boxed{Y_t = c + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}}$$

Donde:

- ❖ θ : valor de autocorrelación de los errores
- ❖ q : número de rezagos

3. **Modelo estacional:** Y_t se predice mediante valores rezagados de Y_t en un intervalo regular s $(la\ estacion)$.

$$\boxed{Y_t = c + \Phi Y_{t-s} + \varepsilon_t}$$

Donde:

- ❖ Φ: valor de la autocorrelación

- s : estacionalidad(52 para semanal, 12 para mensual, 4 para trimestral)

4. **<u>Integración o diferenciación:</u>** en un modelo ARIMA, la serie temporal que se modela debe ser estacionaria para obtener predicciones significativas. La estacionariedad se induce mediante la diferenciación, que se refiere al cálculo de la diferencia entre observaciones adyacentes.

$$\boxed{Y_t'' = Y_t - Y_{t-1}}$$

Un modelo ARIMA es una combinación de un modelo AR, un modelo MA y diferenciación (integración). Si $\phi = 0$ y $\theta = 0$ y $\Phi = 0$, entonces la serie temporal es expresado como $Y_t = c + \varepsilon_t$ donde c es una constante.

La notación básica para describir un modelo ARIMA no estacional es (p , d , q), donde p , d y q son números enteros positivos:

- p = el orden de la parte AR del modelo
- d = el grado de diferenciación no estacional
- q = el orden de la parte MA del modelo.

Por ejemplo:

- Modelo de ruido blanco (estacionario): ARIMA (0, 0, 0)
- Modelo AR con p rezagos: ARIMA(p , 0, 0)
- Modelo MA con q rezagos: ARIMA (0, 0, q)

Si hay estacionalidad, el modelo ARIMA se expresa como:

$$(p , d , q) \times (P , D , Q)$$

Aquí, D es el grado de diferenciación estacional, y P y Q son los términos AR y MA para el componente estacional.

Evaluación de intervenciones utilizando ARIMA

El objetivo del análisis ITS(series temporales interrumpidas) para evaluar intervenciones es estimar el impacto de la implementación de la intervención en un resultado determinado, o en otras palabras, el "efecto de la intervención". Si bien existe una amplia variedad de impactos que se pueden observar, aquí nos centraremos en tres tipos principales: cambio de nivel, pulso y rampa. Si utilizamos T_0 para representar el momento de inicio de la intervención, estos se resumen de la siguiente manera:

- **Cambio de nivel (escalonada) :** cambio repentino y sostenido en el que la serie temporal se desplaza hacia arriba o hacia abajo en un valor determinado inmediatamente después de la intervención. La variable de cambio de nivel toma el valor 0 antes del inicio de la intervención y 1 después.

$$\boxed{S_a = \begin{Bmatrix} 0, si\ t < T_0 \\ 1, si\ t > T_0 \end{Bmatrix}}$$

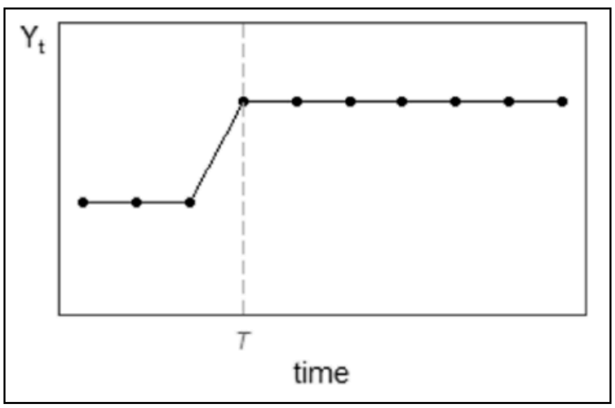

Figura 3. Función escalonada

- **Pulso:** cambio repentino y temporal que se observa durante uno o más puntos de tiempo inmediatamente después de la intervención y luego vuelve al nivel inicial. La variable pulso toma el valor 1 en la fecha de la intervención y 0 en caso contrario.

$$\boxed{P_t = \begin{Bmatrix} 0, si\ t \neq T_0 \\ 1, si\ t = T_0 \end{Bmatrix}}$$

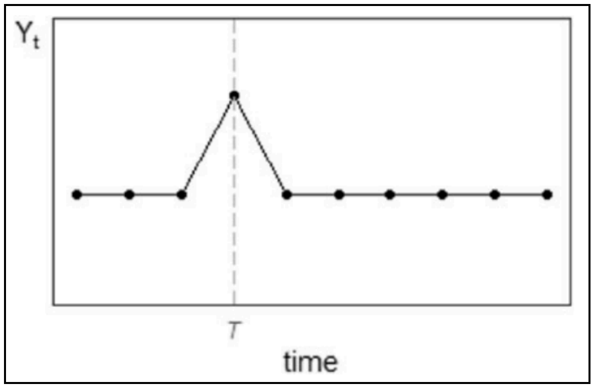

Figura 4. Función de pulso

- **Rampa:** Cambio de pendiente que se produce inmediatamente después de la intervención. La variable rampa toma el valor 0 antes del inicio de la intervención y aumenta en 1 después de la fecha de la intervención.

$$R_t = \begin{cases} 0, si\ t < T_0 \\ t - T_0 + 1\ , si\ t \geq T_0 \end{cases}$$

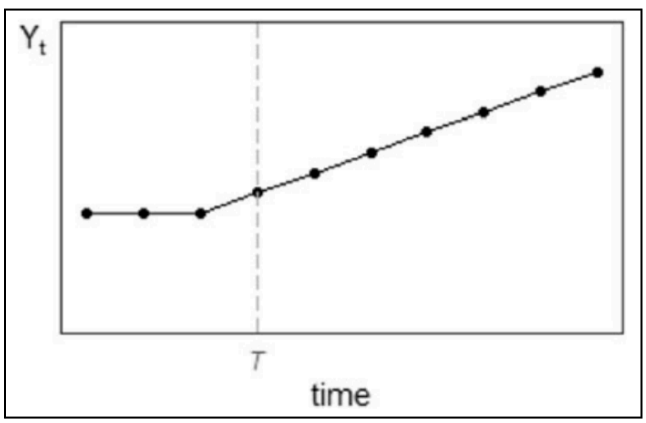

Figura 5. Función de rampa

El impacto de una intervención debe ser hipotetizado de antemano y su forma depende de factores como si es temporal o continua y del resultado evaluado. En un estudio de 2015, un evento temporal de medios negativos sobre estatinas causó un aumento temporal en la discontinuación (pulso) y una reducción sostenida en la dispensación (cambio de nivel). Las intervenciones permanentes, como restricciones en la prescripción o empaquetado genérico, tienden a tener efectos a largo plazo, ya sean inmediatos o graduales (pendiente). Algunas intervenciones pueden requerir una combinación de ambos tipos de impacto.

En el análisis de series temporales interrumpidas (ITS), ARIMA pronostica Y_t en ausencia de la intervención (el "contrafactual") y determina cómo lo observado se desvía de este pronóstico. A diferencia de la regresión segmentada, no es necesario incluir variables ficticias de tiempo o estacionales en el modelo ARIMA, ya que ARIMA puede eliminar tendencias y estacionalidad mediante diferenciación. Si la tendencia se elimina a través de la diferenciación, entonces las tendencias pre y post-intervención no pueden ser estimadas por el modelo. Sin embargo, si se desea estimar la pendiente pre y/o post-intervención, esto puede lograrse incluyendo el tiempo como una covariable e incorporando términos AR y MA para abordar la autocorrelación (por ejemplo, modelos ARMA).

<u>Funciones de transferencia</u>

Una ventaja de los modelos ARIMA es la capacidad de ir más allá de las formas básicas de impacto de la intervención y modelar impactos más complejos a través de "funciones de transferencia". Las funciones de transferencia describen la relación entre la intervención y la serie de resultados Y_t. Modifican la relación entre las entradas anteriores (cambio de paso, pulso, rampa) y la serie temporal para modelar relaciones más complejas, como cambios de nivel graduales o un pulso que decae gradualmente con el tiempo, y también pueden incorporar efectos retardados. La forma general de una función de transferencia es $\frac{\omega(B)}{\delta(B)}$ o:

$$\boxed{Y_t = \mu + \frac{\omega_0+\omega_1 B+\omega_2 B^2+\ldots+\omega_h B^h}{1-\delta_1 B-\delta_2 B^2-\ldots-\delta_r B^r} X_t + \varepsilon_t}$$

Donde:

- B: operador de retroceso (es decir, $B^p Y_t = Y_{t-p}$)
- ω_0: valor inicial del impacto de la intervención en el momento de la intervención (T)
- δ: tasa de decaimiento.
- X_t: variable de intervención (cambio de paso, pulso o rampa)

- h: cuándo ocurre el efecto
- r: patrón de decaimiento

Las estadísticas de ajuste del modelo (como AIC y BIC) pueden ayudar a determinar la forma más apropiada para la función de transferencia, así como el momento del evento (es decir, si el impacto se retrasó y, de ser así, cuánto).

El **AIC (Criterio de Información de Akaike**) y el **BIC** (**Criterio de Información Bayesiano)** son métricas utilizadas para evaluar la calidad de un modelo estadístico, especialmente en el contexto de modelos de series temporales, como ARIMA. Estas métricas ayudan a seleccionar el modelo más adecuado entre un conjunto de modelos candidatos, teniendo en cuenta tanto la bondad del ajuste como la complejidad del modelo.

V. METODOLOGIA

Para la identificación y ajuste del modelo ARIMA, se utilizó el enfoque clásico de Box-Jenkins, que implica los siguientes pasos: identificación del modelo, estimación de parámetros y verificación del modelo. Aunque hoy en día existen algoritmos automatizados en paquetes estadísticos, como R, que simplifican este proceso mediante la minimización de criterios de información (AIC, BIC), el proceso manual sigue siendo útil para la comprensión detallada del ajuste del modelo.

1. **Graficar los datos para comprender los patrones:** El primer paso consistió en graficar la serie temporal para identificar patrones, tendencias, efectos estacionales y posibles valores atípicos. Si hay valores atípicos, la forma de abordarlos dependerá de su causa e influencia en el modelo y las recomendaciones son las mismas para ARIMA que para otros modelos de regresión. Por ejemplo, si los investigadores saben que estos valores extremos se deben a factores externos, como otras intervenciones o una clasificación errónea conocida, estos deben modelarse explícitamente en los datos.

2. **Transformación de datos:** Para estabilizar la varianza de la serie temporal, si esta mostraba variabilidad cambiante a lo largo del tiempo, se aplicó una transformación logarítmica. Esto ayudó a manejar la heterocedasticidad y a preparar los datos para el ajuste del modelo.

3. **Selección de modelos:** Se emplearon gráficos de autocorrelación (ACF) y autocorrelación parcial (PACF) para identificar los posibles parámetros p (orden autorregresivo) y q (orden de media móvil) del modelo. Además, los algoritmos automatizados fueron utilizados para sugerir valores iniciales de estos parámetros, comparando modelos basados en los criterios de información (AIC y BIC).
    a) Determinar el orden de diferenciación para inducir la estacionariedad
    b) Graficar la ACF/PACF de los datos de estacionariedad para determinar los posibles órdenes AR/MA
    c) Estimar el modelo y utilizar criterios de información para encontrar el mejor modelo

Tabla 1: *Consejos para seleccionar los términos autorregresivos (p) y de promedio móvil (q) más apropiados a partir de la autocorrelación y la autocorrelación parcial.*

Tipo de modelo	**Características de ACF y PACF**	
	ACF	**PACF**
ARIMA(p,d,0)	Se desvía o es sinusoidal	Corta el retraso *p*
ARIMA(0,d,q)	Corta el retraso *q*	Se desvía o es sinusoidal
ARIMA(p,d,q)	Se desvía o es sinusoidal	Se desvía o es sinusoidal

4. **Compruebe si los residuos del modelo elegido son ruido blanco:** Esto se puede hacer observando los gráficos de residuos y probando formalmente la presencia de autocorrelación utilizando la prueba de **Ljung-Box** para ruido blanco.

Si la autocorrelación sigue presente en los residuos o su modelo no se ajusta bien, elija diferentes órdenes AR y/o MA. Si los datos no se han transformado previamente, una transformación puede ayudar con residuos que no se distribuyen normalmente. En general, determinar los términos AR y MA es un proceso iterativo que implica prueba y error. Es importante destacar que puede que no haya un modelo "correcto". El objetivo es seleccionar el modelo más parsimonioso (es decir, p / P y q / Q más pequeños) que tenga un buen ajuste y controle adecuadamente la autocorrelación y la estacionalidad. Una vez seleccionado el modelo ARIMA final, se puede estimar el impacto de la intervención**.**

<u>Incorporación de una serie de control</u>

Incluir una serie de control en el análisis de ITS mejora la inferencia causal, ya que las ITS no pueden excluir la posibilidad de que cualquier cambio observado se deba a la intervención de interés u otra cointervención o evento. Una serie de control es aquella que no se ve afectada por la intervención; la selección de un control apropiado se describe en otro lugar. Al igual que en ITS con regresión segmentada, incluir una serie de control implica ejecutar un modelo ARIMA para la serie de interés y, por separado, para la serie de control. Si se observa un cambio en la serie de intervención, pero no en la serie de control, esto proporciona evidencia de que el impacto fue específico de la intervención.

Tamaño de muestra

No existe una regla definitiva sobre cuántos puntos de tiempo se necesitan para aplicar un modelo ARIMA. Aunque comúnmente se menciona que se requieren al menos 50 puntos de tiempo, esta cifra proviene de una afirmación de **Box y Jenkins** que no está respaldada empíricamente ni ha sido probada de manera formal. Cuanto más variables y ruidosos sean los datos, mayor será el número de observaciones necesarias para distinguir los patrones reales del ruido. En situaciones no complicadas, ARIMA puede ser efectivo incluso con series de tiempo cortas, siempre y cuando haya suficientes puntos de tiempo para estimar los parámetros. Si hay estacionalidad presente, es importante contar con suficientes puntos de tiempo para identificar correctamente los efectos estacionales y tener en cuenta las diferencias estacionales.

VI. RESULTADOS

```
# Cargar librerias
library(astsa)
library(forecast)
library(dplyr)
library(zoo)

# Leer data
quet <- read.csv("quet.csv")
head(quet)  #visualizar 5 primeras observaciones
```

```
> head(quet)
     month dispensings
1 1-Jan-11       16831
2 1-Feb-11       17234
```

```
3 1-Mar-11       20546
4 1-Apr-11       19226
5 1-May-11       21136
6 1-Jun-11       20939
```

```
# Convertir los datos de dispensaciones en un objeto de serie temporal
quet.ts <- ts(quet[,2], frequency=12, start=c(2011,1)) # Visualizar la serie temporal completa
quet.ts

# Graficar la serie temporal de dispensaciones
plot(quet.ts, ylim=c(0,40000), type='l', col="blue", xlab="Mes", ylab="Dispensaciones")

# Agregar una línea vertical para indicar la fecha de intervención (1 de enero de 2014)
abline(v=2014, col="gray", lty="dashed", lwd=2)
```

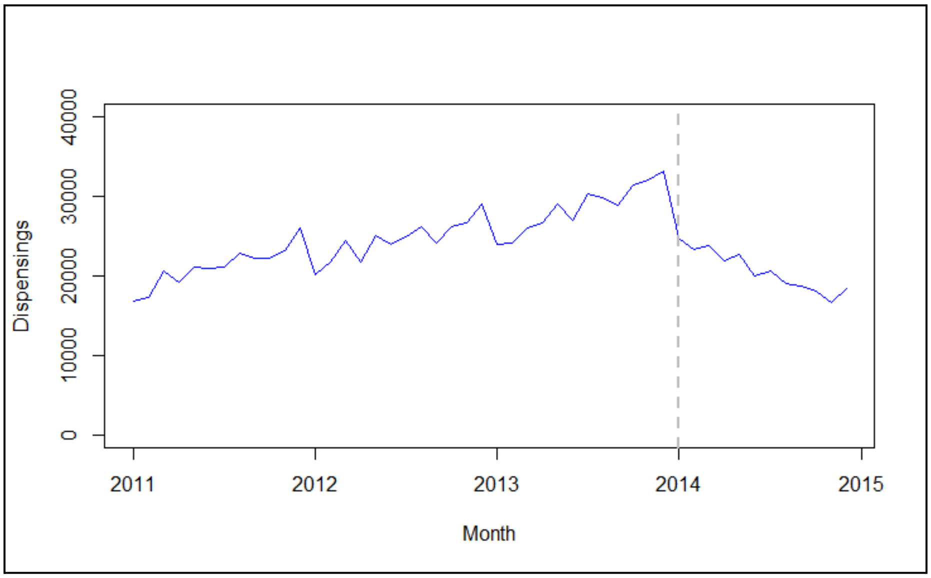

Figura 6. Gráfico de la serie temporal para visualizar la tendencia y posibles patrones estacionales. Se agregó una línea vertical en el año 2014 para marcar el punto de intervención.

```
# Visualizar las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) de los datos sin diferenciar
acf2(quet.ts, max.lag=24)

> acf2(quet.ts, max.lag=24)
     [,1] [,2] [,3]  [,4] [,5]  [,6] [,7]  [,8]  [,9] [,10] [,11] [,12] [,13]
[,14]
```

```
ACF  0.79 0.67 0.58  0.46 0.39  0.22 0.16  0.04 -0.03 -0.13 -0.18 -0.11 -0.22
-0.21
PACF 0.79 0.12 0.06 -0.10 0.04 -0.28 0.13 -0.23  0.08 -0.25  0.20  0.10 -0.23
0.07
     [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22] [,23] [,24]
ACF  -0.22 -0.24 -0.22 -0.27 -0.23 -0.25 -0.24 -0.27 -0.27  -0.2
PACF -0.08 -0.03 -0.08 -0.03 -0.06  0.01 -0.10 -0.03  0.01   0.0
```

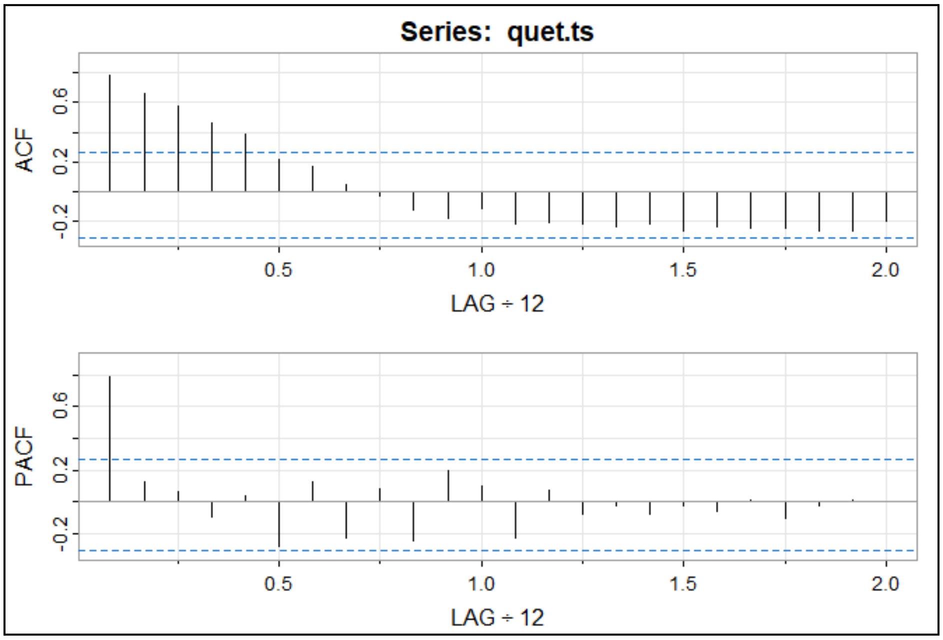

Figura 7. Grafico de las funciones ACF y PACF para identificar patrones de autocorrelación en los datos originales, lo que ayuda a determinar los parámetros del modelo ARIMA.

```
# Visualizar las funciones ACF y PACF de los datos diferenciados estacionalmente
acf2(diff(diff(quet.ts,12)), max.lag=24)

> acf2(diff(diff(quet.ts,12)), max.lag=24)
      [,1]  [,2] [,3]  [,4] [,5]  [,6]  [,7]  [,8]  [,9] [,10] [,11] [,12]
ACF  -0.04 -0.03 0.53 -0.09 0.11  0.25 -0.09  0.00  0.02 -0.11  0.16 -0.22
PACF -0.04 -0.03 0.53 -0.07 0.18 -0.05  0.02 -0.19 -0.11 -0.13  0.28 -0.34
     [,13] [,14] [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22] [,23] [,24]
ACF  -0.19  0.25 -0.32 -0.06  0.19 -0.29 -0.10  0.09 -0.12 -0.06 -0.03 -0.16
PACF  0.02  0.06 -0.03  0.02  0.03  0.00 -0.17 -0.06  0.06  0.00 -0.08 -0.14
```

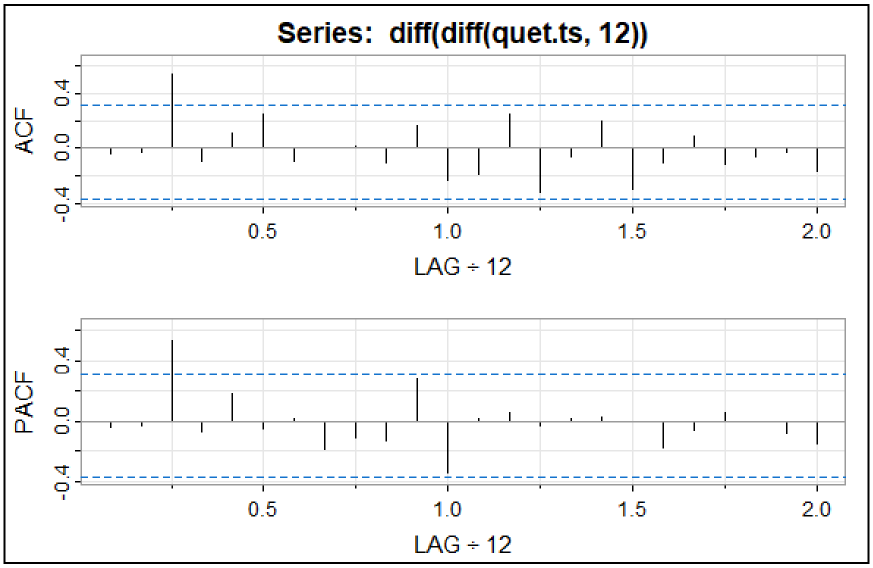

Figura 7. Gráfico de las funciones ACF y PACF luego de aplicar diferencias estacionales y no estacionales para estabilizar la media de la serie y eliminar patrones de estacionalidad y tendencia.

```
# Crear una variable que representa un cambio abrupto (escalón) a partir de enero de 2014
step <- as.numeric(time(quet.ts) >= 2014) # Toma el valor 1 desde 2014 en adelante

# Mostrar la variable 'step'
step

> step
 [1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 1 1 1 1 1 1 1 1 1 1 1
```

Se crea la variable step que modela un cambio instantáneo en el nivel de la serie a partir de 2014, asignando 0 antes de 2014 y 1 desde 2014 en adelante.

```
# Crear una variable que representa un cambio gradual (rampa) a partir de enero de 2014
ramp <- append(rep(0,36), seq(1,12,1)) # Valores de 1 a 12 para los meses de 2014

# Mostrar la variable 'ramp'
ramp

> ramp
 [1]  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  0  0  0  0  1  2  3  4  5  6  7  8  9 10 11 12
```

La variable ramp modela un cambio en la pendiente de la serie a partir de 2014, aumentando linealmente de 1 a 12 durante los meses de ese año.

```
# Utilizar un algoritmo automatizado para identificar los mejores parámetros del modelo ARIMA
# Especificando diferencias de primer orden y diferencias estacionales de primer orden
model1 <- auto.arima(quet.ts, seasonal=TRUE, xreg=cbind(step, ramp), max.d=1, max.D=1, stepwise=FALSE, trace=TRUE)

# Verificar los residuales del modelo para asegurarse de que sean ruido blanco
checkresiduals(model1)

> checkresiduals(model1)

	Ljung-Box test

data:  Residuals from Regression with ARIMA(2,1,0)(0,1,1)[12] errors
Q* = 9.5692, df = 7, p-value = 0.2143

Model df: 3.   Total lags used: 10
```

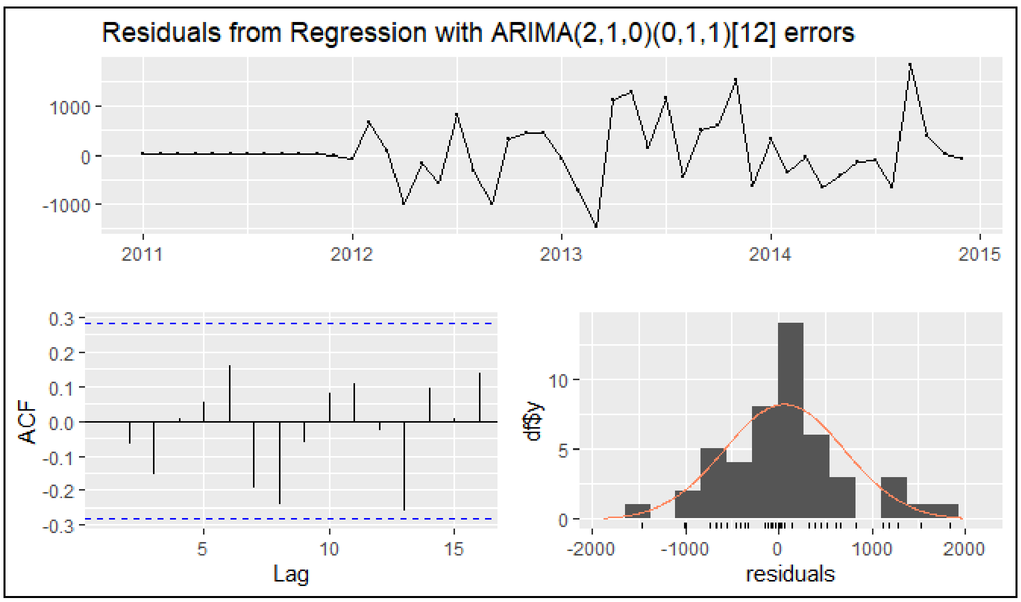

Figura 8. Gráfico de las funciones ACF y PACF luego de aplicar diferencias estacionales y no estacionales para estabilizar la media de la serie y eliminar patrones de estacionalidad y tendencia.

```
# Realizar la prueba de Ljung-Box con 24 rezagos
```

```
Box.test(model1$residuals, lag = 24, type = "Ljung-Box")

> Box.test(model1$residuals, lag = 24, type = "Ljung-Box")

      Box-Ljung test

data:  model1$residuals
X-squared = 23.388, df = 24, p-value = 0.497
```

La prueba de Ljung-Box indica que no hay autocorrelación en los residuos, lo que sugiere un buen ajuste del modelo.

```
# Resumen del modelo ajustado y estimación de intervalos de confianza para los parámetros
summary(model1)
confint(model1)

> summary(model1)
Series: quet.ts
Regression with ARIMA(2,1,0)(0,1,1)[12] errors

Coefficients:
         ar1      ar2     sma1        step        ramp
      -0.873  -0.6731  -0.6069  -3284.7792  -1396.6523
s.e.   0.124   0.1259   0.3872    602.3351    106.6327

sigma^2 = 648828:  log likelihood = -284.45
AIC=580.89   AICc=583.89   BIC=590.23

Training set error measures:
                   ME     RMSE      MAE       MPE     MAPE       MASE
ACF1
Training set 58.95388 636.8028 431.7604 0.1832229 1.744306 0.08366324
-0.0001290895
> confint(model1)
             2.5 %        97.5 %
ar1      -1.1159656    -0.6300580
ar2      -0.9198333    -0.4264438
sma1     -1.3658758     0.1519989
step -4465.3343533 -2104.2240434
ramp -1605.6485769 -1187.6559452
```

- El coeficiente asociado a step es -3284.7792, lo que indica una disminución abrupta promedio de aproximadamente 3,285 dispensaciones a partir de 2014.
- El coeficiente asociado a ramp es -1396.6523, lo que sugiere una disminución adicional promedio de aproximadamente 1,397 dispensaciones por mes durante 2014.
- Los intervalos de confianza para estos coeficientes no incluyen cero, confirmando su significancia estadística.

```
# Ajustar el modelo ARIMA a los datos hasta diciembre de 2013 para pronosticar el contrafactual
model2 <- Arima(window(quet.ts, end=c(2013,12)), order=c(2,1,0),
seasonal=list(order=c(0,1,1), period=12))

# Realizar pronósticos para los 12 meses posteriores a la intervención
fc <- forecast(model2, h=12)

# Convertir los pronósticos en una serie temporal
fc.ts <- ts(as.numeric(fc$mean), start=c(2014,1), frequency=12)
```

Se ajusta un modelo ARIMA utilizando únicamente los datos previos a la intervención (hasta 2013) y se generan pronósticos para 2014, que representan el escenario contrafactual sin intervención.

```
# Combinar los datos observados con los pronósticos contrafactuales
quet.ts.2 <- ts.union(quet.ts, fc.ts)
```

Se crea una serie temporal que contiene tanto los datos observados como los pronósticos, facilitando la comparación entre lo que ocurrió y lo que se habría esperado sin intervención.

```
# Graficar los datos observados y los pronósticos contrafactuales
plot(quet.ts.2, type="l", plot.type="s", col=c('blue','red'), xlab="Month", ylab="Dispensings",
linetype=c("solid","dashed"), ylim=c(0,40000))

# Corregir las advertencias eliminando el parámetro incorrecto
# Agregar una línea vertical para indicar la intervención
abline(v=2014, lty="dashed", col="gray")

# Agregar puntos a las series para mejorar la visualización
points(time(quet.ts.2), quet.ts.2[,1], pch=19, col='blue')
points(time(quet.ts.2), quet.ts.2[,2], pch=19, col='red')
```

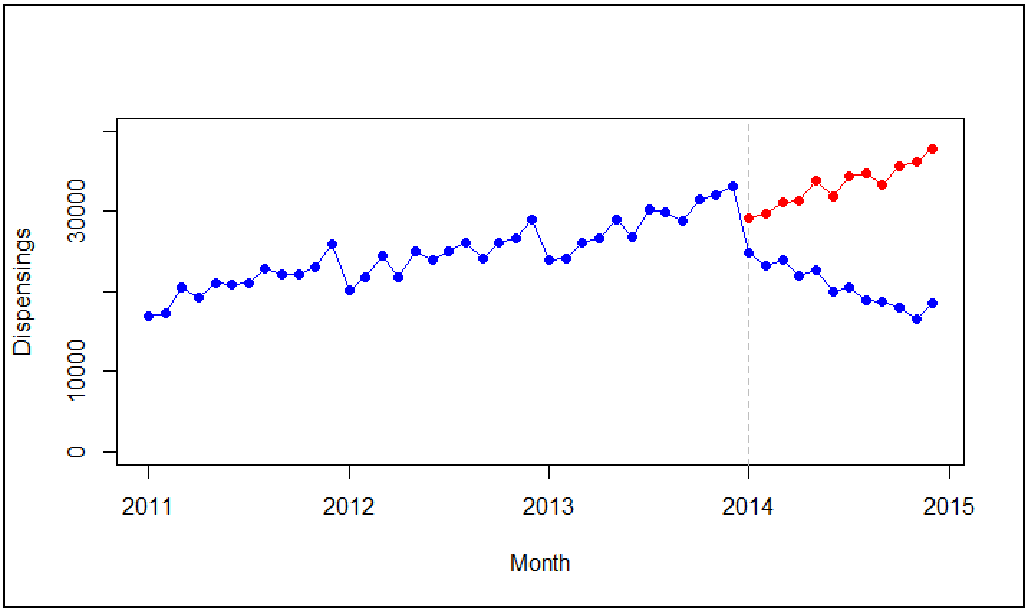

Figura 9. Gráfico de la serie observada en azul y los pronósticos contrafactuales en rojo. Se agrega una línea vertical en 2014 para marcar la intervención y se añaden puntos para resaltar los datos mensuales.

VII. CONCLUSIONES

1. El análisis muestra que la intervención de 2014 tuvo un impacto negativo significativo en las dispensaciones.
2. La disminución abrupta y la tendencia decreciente durante 2014 sugieren la influencia de factores externos, como las políticas de subsidio de medicamentos.
3. El modelo ARIMA con variables exógenas fue adecuado para cuantificar y evaluar el efecto de la intervención.
4. Se controlaron factores como la autocorrelación y la estacionalidad, lo que permitió un análisis más preciso.
5. El modelo capturó tanto un cambio inmediato como una disminución sostenida en las dispensaciones.
6. El ARIMA es una herramienta útil para la evaluación de políticas de salud cuando los métodos convencionales, como la regresión segmentada, no son suficientes debido a la complejidad de los datos.

VII. REFERENCIAS

Schaffer, A. L., Dobbins, T. A., & Pearson, S.-A. (2021). *Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: A guide for evaluating large-scale health interventions*. *BMC Medical Research Methodology, 21*(58). https://doi.org/10.1186/s12874-021-01235-8

Chatfield, C. (2004). *The analysis of time series: An introduction* (5th ed.). Chapman & Hall/CRC.

Bono Cabré, R. (1994). *Diseños de series temporales interrumpidas: Técnicas alternativas de análisis* [Tesis doctoral, Universidad de Barcelona].