



PONTIFICIA UNIVERSIDAD  
CATÓLICA DE CHILE

PROYECTO DE MAGÍSTER

Comparación de la metodología  
de Box - Jenkins y el método  
de Holt - Winters para el  
pronóstico de series de tiempo  
univariado

*Asesora: Prof. Constanza Prado*

*Autor: Lic. Fiorella Alexandra López Real*

Diciembre, 2021  
*Santiago, Chile*



# Índice

<b>1. Agradecimientos</b>	<b>4</b>
<b>2. Introducción</b>	<b>6</b>
2.1. Introducción . . . . .	6
2.2. Planteamiento del problema . . . . .	7
<b>3. Objetivos</b>	<b>7</b>
3.1. Objetivo General . . . . .	7
3.2. Objetivos Específicos . . . . .	8
<b>4. Antecedentes</b>	<b>8</b>
4.1. Antecedentes . . . . .	8
<b>5. Metodología</b>	<b>12</b>
5.1. Definición de variables . . . . .	12
5.2. Diseño de investigación . . . . .	12
5.3. Tipo de investigación . . . . .	12
5.4. Método de análisis de datos . . . . .	12
5.5. Metodología de Box-Jenkins (ARIMA) . . . . .	13
5.6. Método de Holt-Winters . . . . .	17
5.7. Ajuste y Pronóstico . . . . .	19
<b>6. Resultados</b>	<b>21</b>
6.1. Metodología de Box-Jenkins (ARIMA) . . . . .	21
6.2. Método de Holt - Winters . . . . .	26
6.3. Comparación de métodos . . . . .	27
<b>7. Conclusión</b>	<b>29</b>

# 1. Agradecimientos

Agradezco a Dios por haberme permitido llegar hasta este punto y haberme dado salud para cumplir mis objetivos.

Agradezco a mi madre Rosario por haberme apoyado en todo momento, por su motivación constante, por su confianza y por su inmenso amor.

Agradezco a mi padre Gabriel por demostrarme cariño en todo momento, por sus palabras de aliento, sus enseñanzas y los valores que me inculcó junto con mi madre.

Agradezco a mi hermano Gabriel por ser mi ejemplo de superación y por demostrarme cariño. Gracias a mi hermana Cinthia por escucharme siempre y a mi hermana Andrea por su apoyo.

Agradezco a mi sobrina Valentina por su cariño y momentos divertidos.

Gracias a Danny por su comprensión, paciencia, apoyo y su amor incondicional.

Agradezco a la Pontificia Universidad Católica de Chile y en especial a la Escuela de Estadística por haberme acogido 2 años y formarme como Magíster.

Agradezco también a todos los que me brindaron su apoyo para la realización de esta investigación.

## Resumen

En el presente estudio se realiza la comparación entre la Metodología de Box-Jenkins y el método de Holt-Winters, para determinar la capacidad predictiva, llevándose a cabo este estudio a partir de una serie de tiempo estacional y con tendencia, procedente de la serie Tráfico de pasajeros a nivel nacional vía aérea en Perú en el periodo enero 2005-diciembre 2019. Se da a conocer mediante este trabajo la teoría de ambas técnicas utilizadas y los fundamentos básicos de las series de tiempo como método de predicción. Finalmente, con ayuda de los errores de pronóstico MSE, RMSE, MAD y MAPE se logró encontrar que la mejor metodología que predice a la serie de tiempo Tráfico de pasajeros a nivel nacional vía aérea es la de Box-Jenkins.

**Palabras clave:** Metodología, método, serie de tiempo, tráfico de pasajeros, Box-Jenkins, Holt-Winters.

## **2. Introducción**

### **2.1. Introducción**

En esta investigación, se abordará dos tipos de enfoques de pronóstico que son la metodología de Box-Jenkins y el método de Holt-Winters, con el objetivo de comparar ambas técnicas, quedándose con la que tenga la mejor capacidad de predicción y de esta manera sea más eficiente para tomar decisiones con respecto a los datos correspondientes al Tráfico de pasajeros vía aérea a nivel nacional enero 2005 a diciembre 2019.

Se presentará la teoría correspondiente y los pasos a seguir tanto para la metodología de Box-Jenkins como para el método de Holt-Winters. Finalmente, se aplicará estas dos técnicas, a la serie de tiempo indicada, comparando los errores de pronóstico obtenidos de la metodología de Box-Jenkins como del método de Holt-Winters, hallando así el mejor modelo y técnica para pronosticar este tipo de series de tiempo univariada.

## 2.2. Planteamiento del problema

Actualmente, existen muchas técnicas para pronosticar una serie de tiempo univariado, como la metodología de Box Jenkins, más conocida como Arima y el método de Holt - Winters, siendo unas de las teorías más relevantes no se les da la importancia que merecen estas dos técnicas al momento de analizar el comportamiento de una serie de tiempo univariada porque no toman en cuenta todos los pasos a seguir para lograr un buen ajuste de modelo bajo estas técnicas.

La estadística clásica utiliza la metodología de Box-Jenkins, con frecuencia para poder predecir series de tiempo. Los pasos que se realizan dependen del tipo de dato a estudiar ya que estos necesariamente deben cumplir con ser estacionarios. Por otro lado, se tiene el método de Holt - Winters, que pertenece a la familia de series temporales de alisado exponencial, que a diferencia de otras técnicas el método de Holt-Winters, puede adaptarse fácilmente a cambios como es la tendencia y los patrones estacionales gracias a sus constantes de suavización  $\alpha$ ,  $\beta$  y  $\gamma$ . En comparación con los modelos ARIMA se considera que es más rápido el cálculo del pronóstico sin desmerecer el aporte que genera cada una de ellas al momento de predecir es por ello que se debe tener conocimiento de cada requerimiento, supuesto y paso a seguir ya que se podría estar entregando un pronóstico erróneo que se utilizará tal vez para la toma de decisiones.

Lo que se quiere lograr con este estudio es mostrar las bondades que tiene cada una de estas técnicas para pronosticar una serie de tiempo univariada mostrando también los requisitos aplicándolo a una serie de tiempo real Tráfico de pasajeros vía aérea a nivel nacional enero 2005 - diciembre 2019 y a la vez comparar ambas técnicas quedándose con la mejor que pronostique la serie de tiempo en estudio.

## 3. Objetivos

### 3.1. Objetivo General

Determinar la mejor técnica de predicción entre la Metodología de Box-Jenkins y el método de Holt-Winters, para la serie de tiempo Tráfico de pasajeros vía aérea a nivel nacional enero 2005 - diciembre 2019.

## 3.2. Objetivos Específicos

- Realizar el pronóstico de la serie de tiempo Tráfico de pasajeros vía aérea a nivel nacional con la metodología de Box-Jenkins.
- Realizar el pronóstico de la serie de tiempo Tráfico de pasajeros vía aérea a nivel nacional con el método de Holt-Winters.
- Comparar la metodología de Box-Jenkins con el método de Holt-Winters para el pronóstico de la serie de tiempo Pasajeros vía aérea a nivel nacional.

## 4. Antecedentes

### 4.1. Antecedentes

El transporte aéreo, tiene una función práctica, la cual verifica el cumplimiento de la regulación de las líneas aéreas comerciales y el funcionamiento adecuado de aeropuertos en un doble esfuerzo tanto para promover la industria turística como garantizar la seguridad de los pasajeros de ello tiene como función el Ministerios de Transporte y Comunicaciones.

El Tráfico de pasajeros a nivel nacional ha crecido considerablemente los últimos años, en el año 2015, de enero a diciembre, el tráfico aéreo en vuelos nacionales fue de 10,012,906 pasajeros, lo que representó un incremento de 11.9 % respecto a similar período del 2014, cuando se movilizaron 8,950,165 pasajeros.

Según informó la Dirección General de Aeronáutica Civil (DGAC), sólo en el mes de diciembre se movilizaron 856,628 pasajeros, lo que representó un incremento de 10.3 %, respecto al mismo mes del 2014, cuando el número de personas transportadas fue de 776,681. « En los meses de diciembre de los años 2006 al 2015, el comportamiento del tráfico de pasajeros a nivel nacional ha seguido una tendencia creciente en todo el período. El valor del año 2015 es el mayor alcanzado en el período, aunque la mayor tasa de crecimiento se produjo en el año 2010 (31.1 %) » , indicó la entidad dependiente del Ministerio de Transportes y Comunicaciones (MTC). **Fuente:** Portal del Turismo Enero 2016.

Es de esperar que la demanda por el servicio de transporte aéreo de vuelos nacionales se siga incrementando, el estado peruano busca promover el turismo a nivel nacional, si hablamos económicamente los pasajes para este año se a reducido hasta un 15 % esperando que más peruanos realicen turismo por el Perú, en ese sentido es necesario e importante que las líneas aéreas mejoren su capacidad de brindar servicios de calidad y a su vez confiable para sus pasajeros. A continuación, se muestra las estadísticas de los vuelos nacionales, internacionales y participación de las líneas aéreas de los años 2009 - 2014.



Figura 1: Tráfico a Nivel Nacional 2009-2014. Este gráfico nos muestra como está creciendo la demanda de pasajeros a nivel nacional, vemos que en el año 2014 se obtuvo la mayor concentración de Tráfico de pasajeros al interior del país. **Fuente:** Diario El Comercio enero 2015.



Figura 2: Tráfico a Nivel Internacional 2009-2013. El siguiente gráfico muestra la demanda a nivel internacional lo cual se aprecia que la mayor concentración de demanda se obtuvo el año 2013, con el 7.49 millones de pasajeros. **Fuente:** Diario El Comercio Enero 2015.

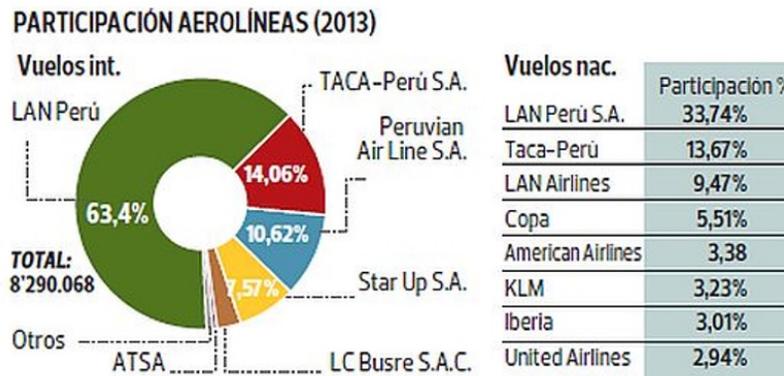


Figura 3: Participación de Aerolíneas 2013. El siguiente gráfico muestra la participación de aerolíneas para el año 2013, lo cual se aprecia que la mayor demanda de pasajeros la obtiene la línea aérea Lan Perú S.A. con el 33.74 %. **Fuente:** Diario El Comercio Enero 2015.

El tráfico aéreo a nivel nacional ascendió a 5,280,936 pasajeros de enero a mayo del 2019 en Perú, lo que representó un incremento de 5 % respecto a similar período del 2018, cuando se movilizaron 5,027,662 pasajeros, informó la Dirección General de Aeronáutica Civil (DGAC). Sólo en el mes de mayo se movilizaron 1,175,297 pasajeros, lo que representó un incremento de 10.5 % respecto al mismo mes del 2018, cuando el número de personas transportadas fue de 1,063,354.

En los meses de mayo de los años 2010 al 2019, el comportamiento del tráfico nacional de pasajeros ha seguido una tendencia creciente en todo el período. El valor del año 2019 es el mayor alcanzado, aunque la mayor tasa de crecimiento se produjo en el año 2013 (16.3 %). Durante mayo del 2019, las ciudades con mayor tráfico fueron Cusco con 331,397 pasajeros (variación positiva de 3.6 % respecto a mayo del 2018), Arequipa con 178,037 pasajeros (7.6 %), Piura con 94,358 pasajeros (17.5 %) e Iquitos con 88 398 pasajeros (8.9 %). **Fuente:** Diario Gestión Julio 2019.

De todas las empresas que operaron en los meses de enero a abril del 2019, Latam Airlines Perú tuvo la más alta participación dentro del tráfico aéreo nacional al transportar 3,270,042 pasajeros, lo que representó el 61.9 % del mercado nacional. Seguido por Viva Airlines Perú con 690,765 pasajeros (13.1 %), Peruvian Airlines con 619,069 pasajeros (11.7 %), Avianca Perú con 334,016 pasajeros (10.2 %) y Star Perú con 101,665 (1.9 %). Sky Airlines Perú en solo dos meses ha transportado 147,232 pasajeros. **Fuente:** Diario Gestión Julio 2019.

## 5. Metodología

### 5.1. Definición de variables

La variable de interés es la serie de tiempo tráfico de pasajeros vía aérea a nivel nacional en Perú que se obtuvo del conteo mensual de entradas y salidas de pasajeros a nivel nacional vía aérea.

### 5.2. Diseño de investigación

El diseño de investigación para el presente estudio es no experimental y de corte transversal ya que se consiguió la base de datos a través de la página oficial del Ministerio de Transporte y Comunicaciones, encontrando un total de 180 registros de tráfico de pasajeros vía aérea del periodo enero 2005-diciembre 2019, y se logró analizar esta serie con presencia de estacionalidad y tendencia.

### 5.3. Tipo de investigación

La presente investigación es de tipo descriptiva y con ella se analiza el comportamiento de la serie de tiempo constituida por el número de pasajeros que viajaron por vía aérea a nivel nacional entre el periodo de enero 2005 a diciembre 2019, este análisis se realiza mediante las técnicas de Box-Jenkins y Holt-Winters, para elegir la que explique mejor el comportamiento y pronóstico de la serie en estudio.

### 5.4. Método de análisis de datos

Para el análisis del pronóstico de series de tiempo univariado se plantea dos técnicas: la metodología de Box-Jenkins, la cual usará los diferentes gráficos y el test de Duckey - Fuller, la hipótesis será rechazada si  $t_\mu$  es menor o igual al estadístico t - student.

$$t_\mu = \frac{\hat{\rho} - 1}{\hat{s}_p} \quad (1)$$

Donde  $\hat{s}_p$  es la desviación típica del estimador de mínimos cuadrados y  $\hat{\rho}$  coeficiente de correlación estimado.

Si la varianza es no constante se aplicará la transformación de Box - Cox (W), que esta dada por:

$$W = \begin{cases} \frac{Y^{\lambda-1}}{\lambda} & , \text{ si } \lambda \text{ diferente a } 0 \\ \ln Y & , \text{ si } \lambda \text{ es igual a } 0 \end{cases} \quad (2)$$

Donde Y es la serie de tiempo original.

Se continua con los correlogramas para encontrar el orden del modelo, y para la estimación de los parámetros, se aplica máximo verósimil de esta manera se halla el pronóstico.

Para el método de Holt - Winters, se propone los coeficientes  $\alpha$ ,  $\beta$  y  $\gamma$ , para el caso de que la serie de tiempo esté afectada por la tendencia y la estacionalidad, finalmente se pretende comparar estas dos técnicas con los errores de pronóstico, quedándose con la que mejor pronostique la serie de tiempo en estudio.

## 5.5. Metodología de Box-Jenkins (ARIMA)

### Series de tiempo

«Una serie de tiempo es el resultado de observar los valores de una variable a lo largo del tiempo en intervalos regulares (cada día, cada mes, cada año, etc.)». (Peña, 2005, p.18).

**Componentes de una serie de tiempo** Para poder identificar estos patrones es conveniente saber que una serie de tiempo consiste en varios componentes.

- **Componente Tendencia T (t):** es el componente de largo plazo que representa el crecimiento o disminución en la serie sobre un periodo amplio. También puede definirse como cambios en la media.
- **Componente Ciclo C (t):** se refiere a movimientos hacia arriba y abajo en forma de onda alrededor de la tendencia. Los patrones cíclicos tienden a repetirse en los datos aproximadamente cada dos, tres o más años.
- **Componente Estacional S (t):** se refiere al patrón periódico que ocurre y se repite cada determinado tiempo. En el caso de series mensuales, el componente estacional mide la variabilidad de las series cada enero, febrero, etc.

- **Componente irregular A (t):** se refiere a la parte de la serie de tiempo que no sigue un patrón regular, ni reconocible. Tal componente representa «lo que queda» en una serie de tiempo después de que la tendencia, el ciclo y la variación estacional han sido explicadas.

**Uso de operadores y polinómios de retraso** El primer operador que se mencionara se llama operador de retraso, que se denotara por la letra B, operador que se define:

$$BZ_t = Z_t - 1 \quad (3)$$

Donde: B es el operador de retraso,  $Z_t - 1$  es la serie de tiempo en estudio en este caso el tráfico de pasajeros vía aérea.

### Procesos autoregresivos (AR)

- **Modelo AR (p)**

Como caso general de un proceso autorregresivo se procede a considerar el proceso AR(p) y se define:

$$Z_t = \phi_1 Z(t-1) + \phi_2 Z(t-2) + \dots + \phi_p Z(t-p) + a_t \quad (4)$$

Un proceso estacionario AR(p) será estacionario si y solo si las raíces de la ecuación característica

$$1 - \phi_1 x - \phi_2 x^2 - \dots - \phi_p x^p = 0 \quad (5)$$

Se encuentra fuera del círculo unitario

### Función de autocorelación

- **Función de autocorrelación Simple (ACF):**

Se tiene los coeficientes de autocorrelación AR(p) verifican la siguiente ecuación en diferencias:

$$\rho_k = \phi_1 \rho(k-1) + \dots + \phi_p \rho(k-p) \quad (6)$$

En las ecuaciones anteriores se ha visto casos particulares de esta ecuación para  $p = 1$  y  $p = 2$ . Se puede concluir que los coeficientes de autocorrelación satisfacen la misma ecuación en el proceso.

$$\phi_p(B)\rho_k = 0 \quad (7)$$

■ **Función de autocorrelación parcial (PACF):**

Llamaremos función de autocorrelación parcial (PACF) a la representación de coeficientes de autocorrelación parcial en función del retardo. Por lo tanto, si ajustamos la familia de regresiones.

$$\begin{aligned} Z_t &= \alpha_{11}Z(t-1) + \eta(kt) \\ &\dots \\ Z_t &= \alpha_{11}Z(t-1) + \dots + \alpha_k Z(t-k) + \eta(kt) \end{aligned}$$

**Modelos Media Móvil MA(q)**

En general, un proceso estocástico se dirá que sigue un esquema de promedio móvil de orden  $q \geq 1$  si se representa de manera:

$$Z_t = a_t - \theta_1 a(t-1) - \theta_2 a(t-1) - \dots - \theta_q a(t-q) \quad (8)$$

Con  $Z_t = Z_t - \mu$  donde  $\mu$  es el nivel del proceso,  $\theta_1, \theta_2, \dots, \theta_q$ , son los parámetros de promedios móviles y  $a_t$  es un proceso de ruido blanco con media cero y varianza constante. Como ya se sabe todo proceso de medias móviles son estacionarios y en particular se observa que ni la media, ni la varianza, ni la covarianza del proceso MA(q) dependen del tiempo.

$$E(Z_t) = E(Z_t - \mu) = 0 \quad (9)$$

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)\sigma_a^2 \quad (10)$$

Para que esta expresión de  $\gamma_k$ , sea valido en general, se define  $\theta_0 = \theta(q+1) = \theta(q+2) = \dots = 0$ . De aquí se obtiene la función de autocorrelación.

## Modelos ARMA

Una generalización de los modelos AR y MA previamente descritos, consiste en combinar ambas clases de modelos para obtener lo que se conoce como modelos autorregresivos y de promedios móviles (ARMA). El proceso ARMA(p,q) se representa mediante:

$$\phi(B)Z_t = \theta(B)a_t \quad (11)$$

En donde  $\phi(B)$  y  $\theta(B)$  son polinómios de retraso p y q respectivamente,  $a_t$ , es un proceso de ruido blanco y  $Z_t$  es la serie de desviaciones de variable  $Z_t$  respecto a su nivel  $\mu$ .

Esta generalización surge del hecho que las series de tiempo que se observan en la práctica muchas veces presentan características, tanto de AR como MA.

### ▪ Modelo ARMA(p,q)

El caso general de un proceso ARMA(p,q) se representa:

$$\phi(B)Z_t = \theta(B)a_t \quad (12)$$

En donde los polinomios  $\phi(B)$  y  $\theta(B)$  son de orden p y q respectivamente

$$Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (13)$$

Para que este proceso sea estacionario se requiere que las raíces de  $\phi(x) = 0$ , estén fuera del círculo unitario, y para que sea invertible la condición es que las raíces de la ecuación  $\theta(x) = 0$ , se encuentren también fuera del círculo unitario; si ambos casos ocurren, entonces el proceso admite las representaciones AR y MA.

## Modelos ARIMA

Los modelos autorregresivos e integrados de promedios móviles (ARIMA) pueden ser vistos como una generalización ARMA previamente estudiados.

El modelo ARIMA permite gran flexibilidad de la representación del modelo ARMA puesto que en realidad lo que se hace al aplicar diferencia es eliminar la posible tendencia polinomial de orden  $d$  presente en la serie de tiempo en estudio. El modelo ARIMA se define:

$$\phi(B)\Delta^d Z_t = \theta(B)a_t \quad (14)$$

$$d \geq 1$$

Para  $Z_t$  en donde  $a_t$  es un proceso de ruido blanco y  $\Delta^d$  es el termino integrado.

## Construcción de modelos ARIMA

En esta parte se presenta la estrategia de construcción de modelos para series de tiempo univariado desarrollado por Box-Jenkins (1970), dicha estrategia consta de 3 etapas fundamentales:

1. Identificación
2. Estimación de los parámetros y selección del modelo
3. Diagnósis del modelo y predicción

### 5.6. Método de Holt-Winters

El método de atenuación exponencial lineal y estacional de tres parámetros de Winters es una extensión del modelo de Holt, que pudiera reducir el error de pronóstico. El método aditivo de Holt-Winters, se usa para series temporales con variación estacional constante, y el método multiplicativo de Holt-Winters, se aplica en series temporales con variación estacional creciente. El método de Holt-Winters, se basa en tres ecuaciones suavizadas: una para el nivel, una para la tendencia y una para la estacionalidad, se usa tres parámetros de suavización para cada ecuación:  $\alpha$ ,  $\beta$  y  $\gamma$ . Para la construcción del modelo tanto aditivo como multiplicativo existen tres etapas importantes:

1. **Inicialización:** La razón de conseguir valores iniciales para los métodos de suavización exponencial es que los modelos son ecuaciones recursivas, por lo tanto, se necesita empezar por algún lado. Para inicializar el método de pronóstico de Holt-Winters, se debe inicializar los índices de niveles, tendencia, y las estacionalidades.

2. **Optimización:** Se habrá logrado obtener los parámetros óptimos cuando estos reduzcan al MSE al mínimo. Es posible utilizar alguna otra medida de error de predicción como el MAE o MAPE, pero el MSE tiende a ser más fácil para trabajar este método.
3. **Ajuste y Pronóstico:** La fase de ajuste es un proceso de cálculos antes de llegar a la etapa de pronóstico, en ambas etapas se utilizan las ecuaciones presentadas anteriormente correspondientes a cada modelo, y finalmente se halla los pronósticos que por lo general es el objetivo principal de estudiar series de tiempo.

## 5.7. Ajuste y Pronóstico

La fase de ajuste es un proceso de cálculos antes de llegar a la etapa de pronóstico, en ambas etapas se utilizan las ecuaciones presentadas anteriormente correspondientes a cada modelo, y finalmente se halla los pronósticos que por lo general es el objetivo principal de estudiar series de tiempo.

**Error de pronóstico:** Se ha ideado diversos métodos para resumir los errores generados por una técnica particular de pronóstico en este caso para Box-Jenkins y Holt-Winters. Notación básica de pronóstico:

- $Y_t$  : Valor de una serie de tiempo en el periodo t.
- $\hat{Y}_t$ : Valor del pronóstico para  $Y_t$ .
- $e_t = Y_t - \hat{Y}_t$ : Residual o error de pronóstico.

Entre los métodos para evaluar la técnica de pronóstico se tiene:

### 1. Desviación absoluta media (MAD):

Mide la precisión del pronóstico mediante el promedio de la magnitud de los errores del pronóstico. La MAD, resulta de gran utilidad cuando se desea medir el error del pronóstico en las unidades de la serie original.

$$MAD = \frac{\sum_{t=1}^n |Y_t - \hat{Y}_t|}{n} \quad (15)$$

### 2. Error cuadrático medio (MSE):

Es otro método para evaluar la técnica del pronóstico, el MSE, penaliza errores mayores, lo que reduce posibilidad de grandes extremos.

$$MSE = \frac{\sum_{t=1}^n (Y_t - \hat{Y}_t)^2}{n} \quad (16)$$

### 3. Porcentaje de error medio absoluto (MAPE):

Indica qué tan grande son los errores del pronóstico comparado con los valores reales de la serie de tiempo. También se utiliza para medir la precisión de la misma técnica u otra.

$$MAPE = \frac{\sum_{t=1}^n |Y_t - \hat{Y}_t|}{n \bar{Y}_t} \quad (17)$$

#### 4. Porcentaje medio del error (MPE):

Permite determinar si el método de pronóstico está sesgado. El resultado es un porcentaje cercano a cero. Si el resultado es un porcentaje negativo grande, el método del pronóstico está sobrestimado de manera consistente. Si el resultado es un porcentaje positivo alto comparado con otros métodos, el método de pronóstico está subestimado de manera consistente.

$$MPE = \frac{\sum_{t=1}^n (Y_t - \hat{Y}_t)}{n \bar{Y}_t} \quad (18)$$

Estas cuatro mediciones de precisión de un pronóstico se utilizan de la siguiente manera:

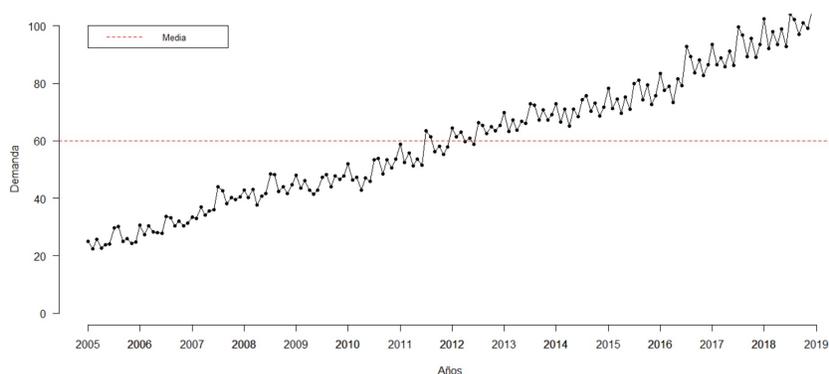
- La comparación de precisión de dos técnicas diferentes.
- La medición de la utilidad o confiabilidad de la técnica.
- La búsqueda de una técnica óptima.

## 6. Resultados

Se tiene la serie de tiempo real tráfico de pasajeros vía aérea en Perú, para entrenar el modelo de Box - Jenkins y Holt - Winters, se usará el periodo enero 2005 - Diciembre 2018 y para probar el ajuste de los modelos se usará el periodo 2019, en meses.

### 6.1. Metodología de Box-Jenkins (ARIMA)

- **Análisis descriptivo**



**Fuente:** Elaboración propia.

Figura 4: Tráfico de pasajeros a nivel nacional vía aérea, Perú, enero 2005 - diciembre 2019. Muestra el comportamiento del tráfico de pasajeros vía aérea a nivel nacional de manera mensual en el periodo enero 2005 a diciembre 2019, en el aeropuerto internacional Jorge Chávez, Perú, se observa la presencia de la componente estacional cada 12 meses, también se aprecia la competente tendencia creciente, con media y varianza no constante. El promedio de pasajeros a lo largo de este periodo es de 627,321, pasajeros mensuales, la demanda mínima fue en febrero 2005 con 223,546, pasajeros y la demanda máxima fue en enero 2019, con 1,119,424, pasajeros.

- **Ajuste de modelo**

Se probaron 4 modelos de series de tiempo con la metodología ARIMA, para la demanda de pasajeros vía aérea, encontrando el modelo SARIMA de

orden  $(0,1,1) \times (2,1,2)$  con estacionalidad cada 12 meses, el mejor indicador de criterio de información de Akaike (AIC) fue igual 3,475, siendo el menor entre los modelos propuestos, finalmente se encontró la ecuación para este modelo:

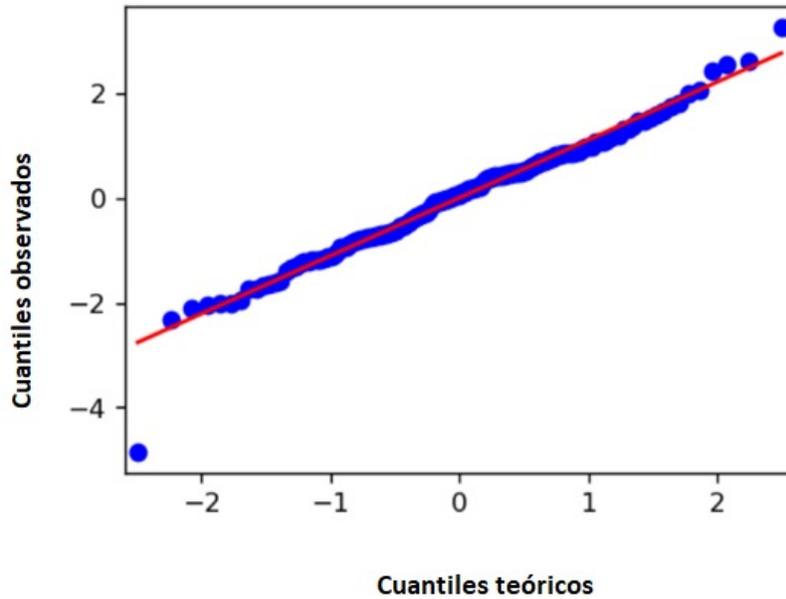
$$w_t = (1 - 0.2B^1)^1 \times (1 + 1.17^{12} - 1.52B^{12} - 0.55B^{24} + 0.75^{24})^1 \times 3.60.$$

(19)

- **Comprobación de supuestos**

Para probar que el modelo propuesto del tráfico aéreo a nivel nacional sigue un proceso estacionario se realizó el análisis de residuos.

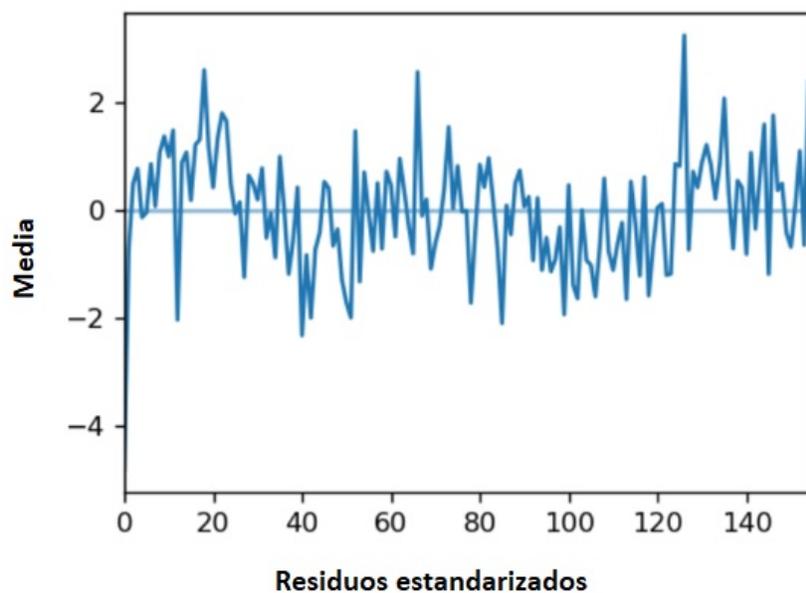
- **Normalidad**



**Fuente:** Elaboración propia.

Figura 5: Q-Q plot de los residuos. Para probar que los residuos siguen una distribución normal, se utilizó el estadístico Jarque Bera, con un p-valor mayor a 0.05 (2.18), Se encontró que los errores son significativos y se distribuyen normal, también se probó el supuesto a través del gráfico de Q-Q plot donde se aprecia que la mayoría de los errores pasan por la recta de la probabilidad acumulada lo que comprueba este supuesto.

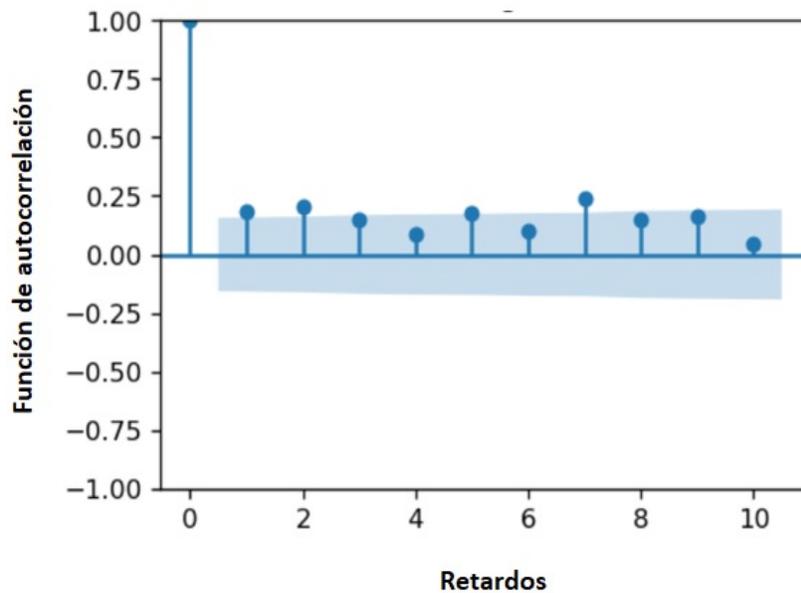
- **Homocedasticidad**



**Fuente:** Elaboración propia.

Figura 6: Homocedasticidad de los residuos estandarizados. Para probar que la varianza de los residuos sea constante se aplicó el estadístico Breuch - Pagan con resultado igual a 1.09, por lo tanto, no se rechaza  $H_0$ , los errores son homocedasticos es decir los residuos tienen varianza constante, gráficamente también se comprueba este supuesto.

- **Autocorrelación**



**Fuente:** Elaboración propia.

Figura 7: Correlograma. Para probar el supuesto de no autocorrelación de los residuos se hizo uso del estadístico Ljung-Box dando como resultado 2.23, por lo tanto, no se rechaza  $H_0$ , no existe autocorrelación en los residuos. También se realizó el gráfico de correlograma de los residuos apreciando que los retardos no exceden las bandas de confianza lo cual nos indica la no autocorrelación

#### ■ Pronóstico

Se encontró el mejor modelo que se ajusta a la serie de tiempo tráfico de pasajeros y se comprobó que los residuos siguen un proceso de ruido blanco, se logró pronosticar la demanda aérea nacional para el año 2019, en meses, apreciando que la demanda aumentará en un 6 % con respecto al año 2018, así mismo se espera el mismo comportamiento que años anteriores con respecto a tendencia y estacionalidad, para el mes de enero se pronostica el mes con mayor demanda aérea con 1,106,884, pasajeros a nivel nacional y en el mes de junio la mínima demanda con 991,059, pasajeros.

## 6.2. Método de Holt - Winters

### ▪ Inicialización

En esta primera etapa para desarrollar el método de Holt-Winters multiplicativo se empezará por la inicialización que consiste en encontrar los valores iniciales para poder hallar las ecuaciones de nivel, tendencia y estacionalidad.

Valores iniciales	
Nivel ( $L_t$ )	3.60
Tendencia ( $b_t$ )	13.70
Estacionalidad ( $S_t$ )	-1.34

Tabla 1: Valores iniciales.

### ▪ Optimización

La segunda etapa consta de encontrar los parámetros Parámetros óptimos que ayuden a pronosticar la serie de tiempo Tráfico de Pasajeros a nivel nacional. Luego de reiteradas iteraciones se logró encontrar que los parámetros óptimos que hacen más pequeños los errores de pronóstico y son:

Parámetros óptimos	
Alpha ( $\alpha$ )	0.60
Beta ( $\beta$ )	0.07
Gamma ( $\gamma$ )	0.30

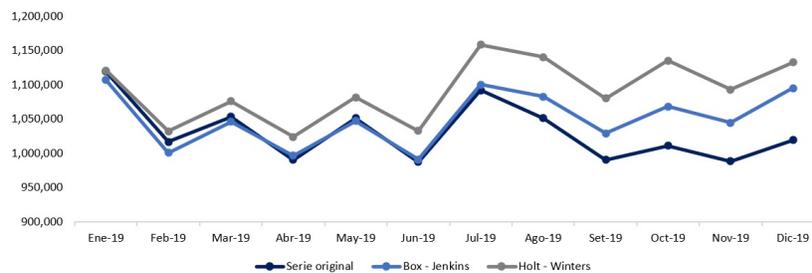
Tabla 2: Parámetros óptimos.

### ▪ Ajuste y pronóstico

Se encontró los valores iniciales y los parámetros óptimos se logró pronosticar la demanda de pasajeros vía aérea a nivel nacional para el año 2019, en meses, Se observa que según el método de Holt-Winters, la demanda crecerá en un 11 % con respecto al año 2018, Se espera el mismo comportamiento de años

anteriores con respecto a tendencia creciente mes a mes y estacionalidad cada 12 meses. Para el mes de julio se pronóstica la mayor demanda con 1,158,725 pasajeros a nivel nacional y en el mes de mayo la menor demanda con 1,024,027 pasajeros.

### 6.3. Comparación de métodos



**Fuente:** Elaboración propia.

Figura 8: Comparación de metodologías. Se Muestra la comparación de la serie original con el pronóstico bajo la metodología de Box-Jenkins y el método de Holt-Winters, para el año 2019, en meses, observando que el método de Holt-Winters, pronostica un 6 % más de demanda que el tráfico aéreo real y la metodología de Box - Jenkins pronostica un 2 % más que los datos originales. Se aprecia gráficamente que la metodología de Box-Jenkins, es la que mejor técnica que se ajusta al pronóstico de series de tiempo de tráfico de pasajeros vía aérea a nivel nacional.

#### ■ Medidas de error de pronóstico

Para la elección de la mejor metodología que predice la serie de tiempo tráfico aéreo se realizó las medidas de error de pronóstico.

La raíz del error cuadrático medio (RMSE) para el caso de Box-Jenkins, la desviación estándar de los errores de pronóstico es 35,567, siendo el más pequeño comparado con el método de Holt-Winters, 73,299. La desviación media absoluta (MAD) indicó que el tamaño del error de pronóstico para la metodología de Box-Jenkins es 26,432, y para el método de Holt-Winters, el error mide 61,345.

La media absoluta del porcentaje del error (MAPE) mide el tamaño del error de pronóstico en porcentaje para el caso de la metodología de Box-Jenkins, el error mide 2 %, para el caso del método de Holt-Winter, mide 6 % el error siendo más pequeño la metodología de Box-Jenkins.

Finalmente, se encontró que la metodología de Box-Jenkins, devuelve los errores de pronósticos más pequeños comparándolo con el método de Holt - Winters, por lo tanto la mejor técnica que pronostica la serie de tiempo Tráfico de pasajeros a nivel nacional vía aérea es la metodología de Box - Jenkins.

Medidas de error de pronóstico		
Medidas de error	Box-Jenkins	Holt-Winters
RMSE	35,567	73,299
MAD	26,432	61,345
MAPE	2 %	6 %

Tabla 3: Medidas de error de pronóstico

## 7. Conclusión

- Se pronosticó la serie de tiempo Tráfico de pasajeros a nivel nacional vía aérea con la metodología de Box - Jenkins teniendo como el mejor modelo: SARIMA (0,0,1) x (2,1,0), que cumplió con todas las exigencias de esta metodología. También se probó el modelo para los siguientes 12 meses, dando como resultado que la demanda de pasajeros aumentará en un 6 % con respecto al año 2018, Se espera un comportamiento similar a los años anteriores con respecto a tendencia y estacionalidad. Para el mes de enero 2019, se pronostica el mes con mayor demanda de pasajeros vía aérea con 1,106,884, pasajeros a nivel nacional y en el mes de junio la mínima demanda con 991,059, turistas nacionales.
- Habiendo obtenido los valores iniciales y los parámetros óptimos, se logró pronosticar la demanda de pasajeros vía aérea a nivel nacional para el año 2019, en meses, donde se observó que según el método de Holt-Winters, la demanda crecerá en un 11 % con respecto al año 2018, esperando el mismo comportamiento de años anteriores con respecto a tendencia creciente mes a mes y estacionalidad cada 12 meses. Para el mes de julio, se pronosticó la mayor demanda con 1,158,725, pasajeros a nivel nacional y en el mes de mayo la menor demanda con 1,024,027, pasajeros.
- Se encontró que la mejor metodología que pronosticó la serie de tiempo Tráfico de pasajeros a nivel nacional vía aérea para periodo enero 2005 - diciembre 2019, es la metodología de Box-Jenkins, que en comparación con el método de Holt-Winters, esta obtuvo los errores de pronóstico más pequeños.

## Referencias

- [1] BOX, G. P., JENKINS, G. M, REINSEL, G. C AND LJUNG G. M, *Time series analysis forecasting and control*, Fifth edition, Wiley Publishing Switzerland, Canada, Canada, 2016.
- [2] BROCKWELL, P. J. AND DAVIS, R. A, *Introduction to time series and forecasting*, Third edition, Springer International Publishing Switzerland, USA, New York, 2016.
- [3] PEÑA, D., *Análisis de series temporales*, Primera edición, Alianza Editorial S.A, España, Madrid, 2005.
- [4] GUERRERO, V., *Análisis estadístico de series de tiempo*, Segunda edición, Internacional Thomson Editores S.A, México, DF, 2003.
- [5] GUJARATI, D. Y PORTER, D, *Econometría*, Quinta edición, McGraw-Hill Interamericana Editores, S.A., México, DF, 2010.
- [6] MCKINNEY, W., *Python for data analysis*, First edition, O'Reilly Media Inc., USA, California, 2012.