

UNIVERSIDAD NACIONAL DE SAN ANTONIO ABAD DEL CUSCO
FACULTAD DE CIENCIAS QUÍMICAS, FÍSICAS Y MATEMÁTICAS
ESCUELA PROFESIONAL DE MATEMÁTICA



TESIS

**COMPARACIÓN DE MODELOS HETEROCEDÁSTICOS
CONDICIONALES Y METODOLOGÍA BOX-JENKINS
APLICADO AL PRECIO DEL PETRÓLEO EN LA REGIÓN DE
CUSCO 2011-2023**

PRESENTADO POR:

Br. SHYRLEY ROMANI ALCARRAZ

Br. ROBINSON SANCHEZ TORRES

**PARA OPTAR AL TÍTULO PROFESIONAL DE
LICENCIADO(A) EN MATEMÁTICA MENCIÓN
ESTADÍSTICA**

ASESOR:

Dr. ALFREDO VALENCIA TOLEDO

CUSCO - PERÚ

2025

INFORME DE ORIGINALIDAD
(Aprobado por Resolución Nro.CU-303-2020-UNSAAC)

El que suscribe, el **Asesor** del trabajo de investigación/tesis titulada: **COMPARACIÓN DE MODELOS HETEROCEDÁSTICOS CONDICIONALES Y METODOLOGÍA BOX JENKINS APLICADO AL PRECIO DEL PETRÓLEO EN LA REGIÓN DE CUSCO 2011-2023**. Presentado por: SHYRLEY ROMANI ALCARRAZ con DNI Nro.: 71621405, presentado por ROBINSON SANCHEZ TORRES con DNI Nro.: 63093692, para optar el título profesional académico de LICENCIADO(A) EN MATEMÁTICA MENCIÓN ESTADÍSTICA.

Informo que el trabajo de investigación ha sido sometido a revisión por 4 veces, mediante el Software Antiplagio, conforme al Art. 6° del Reglamento para Uso de Sistema Antiplagio de la UNSAAC y de la evaluación de originalidad se tiene un porcentaje de 6%.

Evaluación y acciones de reporte de coincidencia para trabajos de investigación conducentes a grado académico o título profesional, tesis.

Porcentaje	Evaluación y acciones	Marque con una (X)
Del 1 al 10%	No se considera plagio.	X
Del 11% al 30%	Devolver al usuario para las correcciones.	
Mayor a 31%	El responsable de la revisión del documento emite un informe al inmediato jerárquico, quien a su vez eleva el informa a la autoridad académica para que tome acciones correspondientes. Sin perjuicio de sanciones administrativas que correspondan de acuerdo a Ley.	

Por tanto, en mi condición de asesor, firmo el presente informe en señal de conformidad y adjunto la primera pagina del reporte de Sistema Antiplagio

Cusco, 09 de Julio de 2025


.....
Firma

Post firma: Alfredo Valencia Toledo

Nro. DNI. 43162177

ORCID del Asesor 0000-0001-6505-9634

Se adjunta:

1. Reporte generado por el Sistema Antiplagio
2. Enlace del reporte Generado por el Sistema Antiplagio: <trn:oid::27259:472785977>

SHYRLEY ROMANI ALCARRAZ ROBINSON SANCHE...

COMPARACIÓN DE MODELOS HETEROCEDÁSTICOS CONDICIONALES Y METODOLOGÍA BOX-JENKINS APLICAD...

 Universidad Nacional San Antonio Abad del Cusco

Detalles del documento

Identificador de la entrega

trn:oid::27259:472785977

Fecha de entrega

9 jul 2025, 10:29 a.m. GMT-5

Fecha de descarga

9 jul 2025, 10:47 a.m. GMT-5

Nombre de archivo

COMPARACIÓN DE MODELOS HETEROCEDÁSTICOS CONDICIONALES Y METODOLOGÍA BOX-JENKI....pdf

Tamaño de archivo

5.4 MB

142 Páginas

25.847 Palabras

131.346 Caracteres

6% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- ▶ Bibliografía
- ▶ Texto citado
- ▶ Texto mencionado
- ▶ Coincidencias menores (menos de 10 palabras)

Exclusiones

- ▶ N.º de coincidencias excluidas

Fuentes principales

- 4%  Fuentes de Internet
- 1%  Publicaciones
- 5%  Trabajos entregados (trabajos del estudiante)

Marcas de integridad

N.º de alertas de integridad para revisión

No se han detectado manipulaciones de texto sospechosas.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.

PRESENTACIÓN

**Señor Decano de la Facultad de Ciencias Químicas, Físicas y Matemáticas,
Señor Director de la Escuela Profesional de Matemática,
Señores Docentes Miembros del Jurado:**

De conformidad con las normas y el reglamento de grados y títulos de la Escuela Profesional de Matemática, tenemos el honor de someter a su consideración el presente trabajo de tesis, titulado:

**"COMPARACIÓN DE MODELOS HETEROCEDÁSTICOS
CONDICIONALES Y METODOLOGÍA BOX-JENKINS APLICADO AL
PRECIO DEL PETRÓLEO EN LA REGIÓN CUSCO 2011-2023"**

El objetivo de este trabajo es Optar al Título Profesional de Licenciado en Matemática Mención Estadística. Esperamos que el presente trabajo de investigación exponga de manera clara los conceptos abordados y sirva como referencia para futuras investigaciones en este campo.

Agradecemos de antemano su atención y valoración de este esfuerzo académico.

DEDICATORIA

Dedico esta tesis a mis padres, quienes con su amor y apoyo incondicional han sido mi mayor fuente de fortaleza a lo largo de este viaje académico. Su confianza en mí y sus constantes palabras de aliento han sido fundamentales para alcanzar esta meta.

A mi asesor, cuya orientación experta y paciencia durante el proceso de investigación han sido cruciales para el desarrollo de este trabajo. Su dedicación y compromiso han enriquecido mi experiencia y aprendizaje.

Con eterna gratitud

Shyrley Romani Alcarraz

Dedico esta tesis a mis padres y seres queridos, quienes con su comprensión y guía incondicional, han sido una fuente constante de apoyo y motivación en cada paso de este proceso. Su amor y sabiduría han sido fundamentales para la realización de este proyecto.

Esta investigación es un testimonio del amor y la profunda gratitud que siento hacia mi familia, quienes han sido mi apoyo incondicional y mi razón de ser. Cada palabra escrita en estas páginas refleja la dedicación, el esfuerzo y la fe que hemos compartido juntos en este camino. Sin su aliento constante y su apoyo, este logro no habría sido posible.

Con eterna gratitud

Robinson Sanchez Torres

AGRADECIMIENTO

Quiero expresar mi más sincero agradecimiento a mi familia, cuyo apoyo incondicional ha sido el pilar fundamental en cada paso de este proceso. Su comprensión, paciencia y aliento constante me han impulsado a seguir adelante, incluso en los momentos más desafiantes.

A mi compañero **Robinson Sanchez Torres**, por su colaboración y por estar siempre dispuesto a compartir sus conocimientos y experiencias. Su dedicación y esfuerzo han sido cruciales para el desarrollo y éxito de este proyecto.

Y a mi asesor **Alfredo Valencia Toledo**, por su guía experta y por brindarme el apoyo necesario para avanzar en esta investigación. Sus valiosos consejos y observaciones han enriquecido mi trabajo y han sido esenciales para alcanzar los objetivos propuestos.

Shyrley Romani Alcarraz

Quiero expresar mi más sincero agradecimiento a la **Universidad Nacional de San Antonio Abad del Cusco** por brindarme la oportunidad de formar parte de la Escuela Profesional de Matemática con Mención Estadística. Este viaje académico ha sido profundamente enriquecedor, contribuyendo significativamente a mi desarrollo profesional y personal.

Mi gratitud más profunda va hacia mi asesor, **Alfredo Valencia Toledo**, por su orientación experta, paciencia y dedicación inquebrantable.

Mis más sinceros agradecimientos a mi amiga, colega, compañera **Shyrley Romani Alcarraz** que con sus aportes y constante motivación han sido fundamentales para la culminación exitosa de esta investigación.

Robinson Sanchez Torres

Índice general

PRESENTACIÓN	IV
DEDICATORIA	V
AGRADECIMIENTO	VII
ÍNDICE GENERAL	XII
ÍNDICE DE TABLAS	XV
ÍNDICE DE FIGURAS	XVII
RESUMEN	XVIII
RESUMO	XIX
INTRODUCCIÓN	XX
1. PLANTEAMIENTO DEL PROBLEMA	1
1.1. Descripción de la situación problemática	1
1.2. Formulación del problema	3
1.2.1. Problema general	3
1.2.2. Problemas específicos	3

	x
1.3. Justificación de la investigación	4
1.4. Objetivo de la investigación	5
1.4.1. Objetivo general	5
1.4.2. Objetivos específicos	6
2. MARCO TEÓRICO CONCEPTUAL	7
2.1. Antecedentes	7
2.1.1. Antecedentes internacional	7
2.1.2. Antecedentes nacional	8
2.2. Bases teóricas	10
2.2.1. Historia de las series de tiempo	10
2.2.2. Variable aleatoria	11
2.2.3. Procesos estocásticos	11
2.2.4. Series de tiempo	13
2.2.5. El petróleo	49
2.3. Marco conceptual	50
3. HIPÓTESIS Y VARIABLES	52
3.1. Hipótesis	52
3.1.1. Hipótesis general	52
3.1.2. Hipótesis específicas	52
3.2. Operacionalización de variables	53
4. METODOLOGÍA	54
4.1. Tipo de investigación	54
4.2. Enfoque de la investigación	54

5.5. Resultado de la hipótesis específico 3	93
5.5.1. Pronóstico del modelo ARIMA(1, 1, 2) para la serie de prueba . . .	94
5.5.2. Pronóstico del modelo ARIMA(2, 1, 1) para la serie de prueba . . .	96
5.5.3. Pronóstico del modelo GARCH(2, 1)	99
5.6. Resultados del objetivo general	102
DISCUSIÓN	103
CONCLUSIONES	106
RECOMENDACIONES	108
REFERENCIAS BIBLIOGRÁFICAS	110
ANEXOS	115
.1. Matriz de Consistencia	115
.2. Solicitud de datos a Osinerming y INEI	117
.3. Base de datos de la Institución Nacional de Estadística e Informática . .	118
.4. Validación de datos de Institución Nacional de Estadística e Informática .	119

Lista de Tablas

2.1. Comparación de los aspectos entre un modelos matemático y un modelo estadístico	17
3.1. Operacionalización de variables para el precio de petróleo en soles por galón 2011-2023	53
5.1. Análisis descriptivo de la serie precio del petróleo (soles por galón) 2011-2022	65
5.2. Resultados del test de tendencia Mann-Kendall para el precio del petróleo en soles por galón	69
5.3. Resultados del test de Kruskal-Wallis para el precio del petróleo en soles por galón	70
5.4. Resultados del test de Breusch-Pagan para el precio del petróleo en soles por galón	71
5.5. Resultados de la prueba de normalidad de Lilliefors (Kolmogorov-Smirnov) para el precio del petróleo	72
5.6. Resultados de la prueba de Dickey-Fuller para el precio del petróleo en soles por galón	73

5.7. Resultados de la prueba de Augmented Dickey-Fuller para la serie diferenciada del precio del petróleo	74
5.8. Valores del criterio de información Akaike (AIC), Bayesiano (BIC) y MA-PE para diferentes modelos ARIMA	77
5.9. Parámetros estimado del modelo ARIMA(1, 1, 2) para la serie precio del petróleo	78
5.10. Parámetros estimados del modelo ARIMA(2, 1, 1) para la serie entrenamiento precio del petróleo	80
5.11. Resultados de la prueba de normalidad de (Kolmogorov-Smirnov) para los residuos del modelo ARIMA(1, 1, 2)	81
5.12. Resultados del test de Box-Ljung para los residuos del modelo ARIMA(1, 1, 2)	82
5.13. Resultados de la prueba de Augmented Dickey-Fuller para los residuos del modelo ARIMA(1, 1, 2)	83
5.14. Resultados de la prueba t Student para los residuos del modelo ARIMA(1, 1, 2)	84
5.15. Resultados de la prueba de normalidad de (Kolmogorov-Smirnov) para los residuos del modelo ARIMA(2, 1, 1)	85
5.16. Resultados del test de Box-Ljung para los residuos del modelo ARIMA(2, 1, 1)	86
5.17. Resultados de la prueba de Augmented Dickey-Fuller para los residuos del modelo ARIMA(2, 1, 1)	87
5.18. Resultados de la prueba t Student para los residuos del modelo ARIMA(2, 1, 1)	88

5.19. Resultados de la prueba para los efectos ARCH para los residuos del modelo $ARIMA(2, 1, 1)$	90
5.20. Coeficientes y el criterio de información de Akaike (AIC) de los Modelos GARCH)	92
5.21. Criterios de selección de los modelos GARCH para la serie precio del petróleo	92
5.22. Datos de Prueba y Pronóstico para el Año 2023 según el modelo $ARIMA(1, 1, 2)$	95
5.23. Datos de Prueba y Pronóstico para el Año 2023 según el modelo $ARIMA(2, 1, 1)$	97
5.24. Comparación de Métricas de Modelos $ARIMA(1,1,2)$ y $ARIMA(2,1,1)$ en la diferencia entre la data de prueba y data pronosticada	98
5.25. Datos de Prueba y Pronóstico para el Año 2023 según el modelo GARCH(2, 1)	101

Índice de figuras

2.1. Esquema a seguir para el desarrollo adecuado de la metodología Box-Jenkins	42
5.1. Serie temporal del precio del petróleo en soles por galón en la región de Cusco 2011-2023	61
5.2. Descomposición clásica de la serie precio del petróleo en soles por galón en la región de Cusco 2011-2023	64
5.3. Representación de la serie temporal del precio del petróleo en soles por galón 2011-2022	66
5.4. Gráfico de boxplot de la data de entrenamiento del precio de petróleo en (soles por galón)	67
5.5. Gráfico de box plot para la serie de entrenamiento del precio de petróleo en (soles por galón)	68
5.6. Gráfico de autocorrelación (ACP) para la serie de entrenamiento (soles por galón)	75
5.7. Gráfico de autocorrelación parcial (PACP) para la serie de entrenamiento (soles por galón)	76

5.8. Gráfico de autocorrelación Simple (ACF) para los residuos del modelo ARIMA(1, 1, 2)	82
5.9. Gráfico de autocorrelación Simple (ACF) para los residuos del modelo ARIMA(2, 1, 1)	86
5.10. Visualización de los errores al cuadrado de la serie precio del petróleo en soles por galón	91
5.11. Gráfico de comparación para la serie original, el modelo ajustado y el pronóstico del modelo ARIMA(1, 1, 2)	94
5.12. Gráfico de comparación para la serie original, el modelo ajustado y el pronóstico del modelo ARIMA(2, 1, 1)	96
5.13. Gráfico de ajuste del modelo GARCH(2,1) y el intervalo de pronóstico y variación del coeficiente gamma	99
5.14. Gráfico de ajuste del modelo GARCH(2,1) a la serie precio del petróleo .	100

RESUMEN

La investigación titulada "Comparación de Modelos Heterocedásticos Condicionales y Metodología Box-Jenkins Aplicado al Precio del Petróleo en la Región Cusco 2011-2023" tiene como objetivo principal determinar el mejor modelo para pronosticar el precio del petróleo en la Región Cusco durante el periodo 2011-2023, considerando la aplicación de modelos heterocedásticos condicionales y la metodología Box-Jenkins. La investigación sigue un enfoque cuantitativo, de tipo aplicado, con un alcance descriptivo y diseño no experimental longitudinal. La población de estudio corresponde a los precios del petróleo entre enero de 2011 y diciembre de 2023. Los datos son obtenidos mediante solicitudes al Instituto Nacional de Estadística e Informática (INEI) y OSINERGMIN. Los resultados revelan que el modelo GARCH (2,1) proporciona un pronóstico más preciso, especialmente en series con alta volatilidad y heterocedasticidad, superando al modelo ARIMA (2, 1, 1), que, aunque tiene un buen ajuste inicial, no logra capturar la volatilidad inherente del mercado con la misma efectividad. Por lo tanto, se concluye que el modelo GARCH es más eficaz para predecir fluctuaciones en mercados con características volátiles como el del petróleo.

Palabras Clave: Series de tiempo, Modelos heterocedásticos, Metodología de Box Jenkins, Pronóstico, Precio del petróleo.

RESUMO

A pesquisa intitulada "Comparação de Modelos Heterocedásticos Condicionais e Metodologia Box-Jenkins Aplicados ao Preço do Petróleo na Região Cusco 2011-2023" tem como objetivo principal determinar o melhor modelo para prever o preço do petróleo na Região Cusco durante o período de 2011 a 2023, considerando a aplicação de modelos heterocedásticos condicionais e a metodologia Box-Jenkins. A pesquisa segue uma abordagem quantitativa, de tipo aplicado, com um alcance descritivo e design não experimental longitudinal. A população do estudo corresponde aos preços do petróleo entre janeiro de 2011 e dezembro de 2023. Os dados são obtidos por meio de solicitações ao Instituto Nacional de Estatística e Informática (INEI) e à OSINERGMIN. Os resultados revelam que o modelo GARCH (2,1) fornece uma previsão mais precisa, especialmente em séries com alta volatilidade e heterocedasticidade, superando o modelo ARIMA (2, 1, 1), que, embora tenha um bom ajuste inicial, não consegue capturar a volatilidade inerente ao mercado com a mesma eficácia. Portanto, conclui-se que o modelo GARCH é mais eficaz para prever flutuações em mercados com características voláteis, como o do petróleo.

Palavras-chave: Séries temporais, Modelos heterocedásticos, Metodología de Box-Jenkins, Previsão, Forecasting, Preço do petróleo.

INTRODUCCIÓN

El precio del petróleo es un factor crucial en la economía global, ya que afecta directamente los costos de producción y la estabilidad de los mercados. Según datos de la Agencia Internacional de Energía (AIE), las fluctuaciones en los precios del petróleo han mostrado variaciones de hasta un 60 % durante períodos de alta volatilidad provocados por conflictos a nivel mundial el cual resultó en efecto negativo para los diversos sectores económicos (Industria, construcción, agricultura, etc.). En el contexto nacional, Perú no es ajeno a estas fluctuaciones, ya que el petróleo es una fuente energética esencial para diversas actividades económicas. De acuerdo con el Instituto Nacional de Estadística e Informática (INEI), el sector de hidrocarburos en Perú experimentó una variabilidad en los precios de hasta un 50 % en los últimos 20 años. Esta volatilidad ha afectado especialmente a las industrias en la región Cusco, donde el abastecimiento de combustibles es vital para el desarrollo de sectores clave como la construcción y el turismo, entre otros, generando un impacto significativo en la planificación económica regional.

El precio del petróleo viene a ser información de vital importancia para la toma de decisiones de múltiples empresas, la cual, a lo largo del tiempo presenta fluctuaciones o comportamientos que a simple vista son indetectables, imposibilitando tomar decisiones adecuadas de esta información, por lo que se emplean técnicas y metodologías de las series de tiempo para poder analizar adecuadamente los datos y brindar conclusiones

adecuadas para tomar decisiones preventivas al cambio en los precios del petróleo.

El objetivo de la presente investigación es analizar y comparar dos métodos estadístico, los Modelos Heterocedásticos Condicionales (GARCH) y la Metodología Box-Jenkins (ARIMA), para pronosticar el precio del petróleo en la región Cusco durante el período 2011-2023. Mediante la aplicación de estos modelos, se busca identificar el modelo más preciso y así proporcionar herramientas que permitan a los sectores económicos locales anticipar las variaciones del mercado, mejorando así la toma de decisiones en un entorno cada vez más incierto. Esta investigación busca brindar nociones para el uso de modelos efectivos en el pronóstico del precio del petróleo en beneficio de múltiples industrias dependientes de este recurso, e incluso en la sociedad ya que en general los ciudadanos hacen un gran uso de los medios de transporte, es por ello que conocer el comportamiento del precio del petróleo a través del tiempo es importante, ya que ayuda a tomar decisiones preventivas tanto para gobiernos y empresas.

El desarrollo del trabajo presentado se organiza en diferentes capítulos que buscan ofrecer una estructura clara, secuenciada y comprensible para la investigación. **Capítulo I** Planteamiento del problema, contempla la descripción problemática, formulación de problemas, justificación de la investigación y los objetivos. **Capítulo II** Marco teórico conceptual, donde se plasman los antecedentes, las bases teóricas y el marco conceptual. **Capítulo III** Hipótesis y variables definidas en la operacionalización de variables. **Capítulo IV** Metodología donde se describe que tipo, enfoque, alcance y diseño adoptó la investigación, además de la población de estudio, las técnicas para recolectar los datos y técnicas utilizadas en el contraste de hipótesis. **Capítulo V** Resultados donde se muestra lo alcanzado respecto a los objetivos. Finalmente; se realiza la **Discusión**, las **Conclusiones** respecto a los objetivos y las **Recomendaciones**.

Capítulo 1

PLANTEAMIENTO DEL PROBLEMA

1.1. Descripción de la situación problemática

Según [Real Academia Española \(2023a\)](#) el petróleo es un "líquido natural oleaginoso e inflamable, constituido por una mezcla de hidrocarburos, que se extrae de lechos geológicos continentales o marítimos y del que se obtienen productos utilizables con fines energéticos o industriales, como la gasolina, el queroseno o el gasóleo".

El petróleo a nivel mundial ha experimentado fluctuaciones del precio debido a diversos factores. Entre ellas, se encuentra la recesión en la economía mundial, las alteraciones en la economía China debido a los brotes de la enfermedad por Coronavirus (COVID-19) y el desarrollo de la guerra entre Rusia y Ucrania ([Agencia Internacional de Energía, 2022](#)). A principios de Julio de 2023 se cotizó en los 78.43\$ por barril. Actualmente en América, Europa y Asia, el precio diario del petróleo ha ido aumentando en 0.87\$, lo que representa un incremento del 1.12% en comparación con meses anteriores. Sin embar-

go, en los últimos 12 meses el precio del petróleo ha experimentado una disminución del 29.12 % en el plano mundial ([Organización de Países Exportadores de Petróleo, 2023](#)).

En el Perú el petróleo es sumamente importante porque el 68 % de la energía comercial depende de este compuesto y sus derivados, esto significa que los sectores dependen de este recurso. Además, la situación de crisis política y social en el país ha dado lugar a una notable volatilidad en el precio del petróleo. En los últimos años, se ha evidenciado una inestabilidad considerable tanto en los aumentos como en las disminuciones del precio. Es importante resaltar que en los últimos años, se ha evidenciado un descenso constante en la producción de petróleo en el Perú; esta reducción ha llevado a un aumento en el precio promedio pasando de $S/9.76$ a $S/14.10$ por galon para los años 2016 – 2021 respectivamente; generando incertidumbre e inestabilidad en la economía peruana, en las empresas y los sectores correspondientes ([Ministerio de Energía y Minas, 2016](#)).

En la región de Cusco, el precio del petróleo enfrenta varios desafíos comenzando por la regularización del suministro, el cual requiere de un tiempo considerable. Además, la falta de una planta de refinación de petróleo en la región agrega dificultades en el suministro ([Tamayo et al., 2015](#)). La planta se ubica en el distrito de Mollendo en la provincia de Islay en la región Arequipa lo que implica tiempo y costos para el transporte del petróleo. Otro de los factores que influye en la variación del precio del petróleo en la región Cusco son los problemas sociales recurrentes lo que conlleva al cierre de carreteras por consiguiente al desabastecimiento y escasez del petróleo, asimismo, estos problemas sociales pueden llevar a los micro proveedores a aumentar desproporcionadamente su precio. En el año 2022 el precio del petróleo promedio osciló entre $S/8.5$ y $S/9.0$ por litro, sin embargo durante períodos de escasez o interrupción en el suministro del petróleo

se registraron incrementos del precio de hasta $S/15.0$ por litro (Ministerio de Energía y Minas, 2016).

En la tesis de Suárez Dosantos et al. (2022) se rescata la importancia de la metodología de Box-Jenkins el cual es altamente efectiva debido a su capacidad predictiva comparado con los modelos de suavización exponencial, modelos de aprendizaje automatico (Machine Learning) quienes presentaron limitaciones en su eficiencia y efectividad al pronósticar.

En la presente tesis se centra en la comparación de la disponibilidad y asignación de recursos y predicción de los modelos heterocedasticos condicionados frente a los modelos con la metodología de Box-Jenkins, aplicado al precio de petróleo, enero del 2011 a diciembre del 2023, en la región Cusco; para comparar los modelos se utilizó la selección de modelos del criterio de información Akaike (AIC), criterio de información Bayesiano (BIC), error del porcentaje medio absoluto (MAPE) entre otros.

1.2. Formulación del problema

1.2.1. Problema general

¿Qué modelo ofrece mayor predicción y robustez para la serie precio del petróleo: la aplicación de modelos heterocedásticos condicionales o la metodología Box-Jenkins, durante el período 2011-2023, Región Cusco?

1.2.2. Problemas específicos

- ¿Cuál es el modelo con la metodología de Box-Jenkins que presenta un mejor ajuste de la serie precio del pretróleo, región Cusco, para el periodo enero 2011 a

diciembre 2022?

- ¿Cuál es el modelo heterocedástico condicionado que presenta mejor ajuste de la serie precio del petróleo, región Cusco, para el periodo enero 2011 a diciembre 2022?
- ¿Qué enfoque ofrece un mejor pronóstico para el precio del petróleo: los modelos heterocedásticos condicionales o los modelos basados en la metodología Box-Jenkins, en la región Cusco, durante el período de enero a diciembre de 2023?

1.3. Justificación de la investigación

1. Justificación social

El pronóstico del precio del petróleo en Cusco es crucial para diversos sectores, como el energético, económico y social. Contar con una perspectiva precisa y confiable sobre la evolución futura de los precios del petróleo permite a las empresas planificar de manera estratégica, tomar decisiones acertadas en cuanto a inversiones, gestionar riesgos y mejorar sus operaciones. Además, tanto los gobiernos como los inversores y consumidores se ven impactados por las fluctuaciones en el precio del petróleo, por lo que disponer de predicciones confiables es fundamental para todos ellos.

2. Justificación económica

En el Perú, el petróleo tiene una gran relevancia ya que el 68% de la energía comercial depende de este recurso y sus derivados. Esto implica que muchos sectores están directamente ligados a su disponibilidad. Cuando el precio del petróleo sube, impacta en los costos de sus derivados, lo que genera un incremento en los precios

del transporte público y privado, dado que los combustibles derivados del petróleo representan una parte significativa de los gastos de operación de vehículos y flotas. Además, el aumento del petróleo también repercute en los costos de producción de las empresas y en el costo de vida, ya que muchos procesos industriales dependen de la energía generada por productos petrolíferos. Este estudio nos permitirá analizar si el precio del petróleo sube o baja con el tiempo, ayudándonos a estar preparados ante cualquier eventualidad.

3. Justificación teórica

Se emplean modelos de series de tiempo, los cuales son modelos estadísticos especializados en el análisis y pronóstico de datos secuenciales. Estos modelos son ampliamente reconocidos por su capacidad para identificar patrones, tendencias y estacionalidades en los datos históricos, lo que los convierte en una herramienta clave para prever el precio del petróleo.

Se ha seleccionado el período de enero de 2011 a diciembre de 2023 debido a su extensa cobertura temporal, lo que permite evaluar tanto las tendencias a largo plazo como los patrones más recientes en el precio del petróleo en la región de Cusco. Esta selección proporciona una visión completa y actualizada de la evolución del precio, aspecto fundamental para generar pronósticos confiables.

1.4. Objetivo de la investigación

1.4.1. Objetivo general

Determinar el mejor modelo para pronosticar al precio del petróleo en la Región Cusco durante el período 2011-2023, considerando la aplicación de modelos heterocedásticos

condicionales y la metodología Box-Jenkins.

1.4.2. Objetivos específicos

- Identificar el modelo de la metodología de Box-Jenkins que presenta un mejor ajuste de la serie precio del pretróleo para el periodo enero 2011 a diciembre 2022, región Cusco.
- Identificar el modelo heterocedástico condicionado que presenta un mejor ajuste de la serie precio del pretróleo para el periodo enero 2011 a diciembre 2022, región Cusco.
- Comparar qué enfoque ofrece un mejor pronóstico para el precio del petróleo en la region Cusco de enero a diciembre de 2023: los modelos heterocedásticos condicionados o modelos basados en la metodología Box-Jenkins.

Capítulo 2

MARCO TEÓRICO CONCEPTUAL

2.1. Antecedentes

2.1.1. Antecedentes internacional

Lazaridis (2021) en su tesis doctoral intitulada "Análisis del Precio del crudo del petróleo y aplicación de métodos para su predicción", presenta como mejor modelo autorregresivo integrado de medias móviles estacional, SARIMA(1, 1, 2)X(1, 0, 0)[12], para la serie precio mensual del oro, plata y T-Note, para el periodo de enero 2000-diciembre 2020, con una tasa de error absoluta promedio MAPE 0.2229, que generó pronóstico dentro del rango esperado para el precio del petróleo.

Ortiz (2017) en su trabajo de investigación intitulada, "Modelado y pronóstico: análisis de precios de Petróleo Crudo(WTI) Utilizando Series Temporales", presenta un análisis detallado de los precios del crudo tipo WTI utilizando modelos de series de tiempo. El estudio incluye el modelado y la predicción de estos precios, obteniendo un Modelo Integrado de Medias Móviles Autorregresivas (ARIMA) mediante la metodología de Box-Jenkins, complementado con modelos híbridos como el GARCH. Con la ayuda de los criterios AIC y BIC, se identificó que el modelo híbrido ARIMA-APARCH(1,1) era el más

adecuado para los precios diarios del crudo del petróleo WTI, tomando como referencia los datos de la Energy Information Administration (EIA) entre el 2 de enero de 1986 y el 31 de octubre de 2016. Este modelo es capaz de captar fenómenos como el efecto palanca, la memoria larga y la agrupación de la volatilidad presentes en el comportamiento del precio del crudo.

Villalba (2019) en su tesis doctoral presenta, una aproximación basada en modelos VAR en el análisis y predicción del precio del petróleo, cuyo objetivo es demostrar la relación estadística existente entre el precio del petróleo y distintos índices bursátiles y commodities, mediante la construcción de un modelo de series de tiempo. Donde se obtuvo el modelos VAR para el cual fue necesario realizar las estimación mediante MCO y por separado; dichos modelos generan en algunos casos predicciones mejores que los modelos de ecuaciones simultáneas más complejas. El pronóstico se realizó para el año 2019 a partir de un modelo econométrico. La investigación está basada en la cotización del Brent entre los años 2015 al 2018, la cual consta de la periodicidad mensual y se toma como valor de referencia a las últimas cotizaciones de cada mes, donde se afirma que los valores pronosticados no difieren significativamente de los valores reales, se evidencia que las variaciones de la serie no son captadas por el modelo en la totalidad, el coeficiente de determinación es 81.91 % el cual refleja que las variables expuestas explican las variables dependientes en una gran medida.

2.1.2. Antecedentes nacional

En su proyecto de investigación Castillo (2014) titulada "Modelación de la Volatilidad del índice general de la bolsa de valores Lima, período 2009-2011" cuyo objetivo se propuso analizar la volatilidad diaria del Índice General de la Bolsa de Valores de

Lima (IGBVL) para el período 2009-2011, utilizando modelos de varianza condicional como ARCH y GARCH, empleando el software econométrico EVIEWS. La investigación concluyó que el modelo GARCH (1,1) resultó ser el más adecuado, destacándose por presentar los menores valores en los criterios de información de Akaike (AIC) y Schwarz en comparación con otros modelos evaluados. Además, se utilizó un modelo AR (1) para ajustar la media del RIGBVL antes de la aplicación del GARCH (1,1).

[Aquino \(2018\)](#), en su proyecto de investigación titulada "Efecto del precio internacional del petróleo y del tipo de cambio en la importancia del Diésel de Petroperú", en los resultados se determina el efecto del precio internacional del petróleo junto con el tipo de cambio sobre el volumen de importación de Ultra Low Sulfur Diesel de la empresa Petroperú entre el periodo 2014 al 2016; llegando a la conclusión que la empresa Petroperú reacciona cuando se incrementa el precio internacional en una unidad aumentando así el precio del barril. Utilizaron métodos de series de tiempo.

[Lurento \(2021\)](#), en su proyecto de investigación titulada "Modelamiento Económico de la inflación en el Perú, período 2000-2019" se buscó determinar un modelo econométrico confiable para ajustar la inflación en el Perú, utilizando datos mensuales del BCRP desde 2000 hasta 2019. Se evaluaron varias especificaciones econométricas, incluyendo el modelo ARIMA, el modelo de inflación con crédito y tasas de interés, el modelo con brechas y el modelo clásico de inflación. Los resultados indicaron que el modelo ARIMA(5,1,0) era el más adecuado para la economía peruana. Este modelo identifica como determinantes significativos la variación de la inflación, la brecha del producto, las brechas de precios de importaciones y bienes de consumo, la brecha del precio del petróleo, la oferta monetaria, la variación del tipo de cambio y el nivel de salarios. Estos resultados proporcionan a los responsables de la política monetaria una

herramienta valiosa para entender y gestionar mejor la inflación, ajustando políticas para alcanzar la estabilidad de precios.

En su tesis de magister [Chavez \(2021\)](#) titulada, "Volatilidad sectorial en la bolsa de valores de Lima: estructura GARCH, 2015-2020", cuyo objetivo es determinar la evolución de la volatilidad diaria de cada sector económico que cotiza en la Bolsa de Valores de Lima, en el período 2015-2020. En este estudio, se analizaron los índices diarios de cinco sectores económicos (Consumo, Financiero, Industrial, Minero y Servicios Públicos). Se verificaron las estructuras GARCH específicas para cada sector: GARCH (1,1) en el sector consumo, GARCH (2,1) en el sector financiero, GARCH (1,1) en el sector industrial, GARCH (2,1) en el sector minero y GARCH (1,1) en el sector de servicios públicos.

2.2. Bases teóricas

2.2.1. Historia de las series de tiempo

El análisis de las series de tiempo ha jugado un rol significativo en la ciencia desde tiempos antiguos. Los astrónomos babilónicos fueron de los primeros en utilizar series de tiempo para predecir eventos astronómicos. Posteriormente, en 1900, Evgenij y Yule desarrollaron los procesos de medias móviles y autorregresivos para representar series de tiempo. Años más tarde, en 1919, Warren M. Persons introdujo la idea de descomponer una serie de tiempo en componentes que dependen de distintas causas, avanzando aún más el estudio de este campo. [Box et al. \(1970\)](#) publican un libro "Análisis de series de tiempo: pronóstico y control", el que al día de hoy es conocido como metodología de Box-Jenkins y en 1975 Granger Newbold mostraron que los pronósticos sencillos de métodos a menudo eran mejores que los pronósticos basados en modelos econométricos grandes

el cual impulsó aún más la popularidad de la metodología de Box-Jenkins. Por otro lado, el estudio de los modelos de volatilidad experimentó una revolución con el trabajo de [Engle \(1982\)](#), quien introdujo una nueva clase de procesos estocásticos conocidos como modelos ARCH. Estos modelos se caracterizan por tener una varianza condicional no constante y dependiente del cuadrado de los errores pasados. Posteriormente, los modelos ARCH fueron extendidos por [Bollerslev \(1986\)](#) quien propuso los modelos GARCH.

2.2.2. Variable aleatoria

Para un espacio de eventos elementales S , una variable aleatoria es una función cuyo dominio en el espacio de eventos y rango es algún subconjunto de los números reales ([Abarca, 2012](#)); la notación está dada de la siguiente forma:

$$X(s) = x, \quad s \in S$$

Esto significa que x es el valor asociado por X con el evento elemental s

2.2.3. Procesos estocásticos

"Un proceso estocástico es una sucesión de variables aleatorias; ordenadas y equidistantes cronológicamente, referidas a una o a varias características de una unidad observable en diferentes momentos."([Mauricio, 2007](#))

Un proceso estocástico está definido de la siguiente forma:

$$\dots, Y_{-1}, Y_0, Y_1, \dots; (Y_t : t = 0, \pm 1, \pm 2, \dots)$$

donde: (Y_t) , es una variable aleatoria que da referencia al valor observable en el momento t

Paseo aleatorio

Es un proceso estocástico (Y_t) tal que:

$$Y_t = \mu + Y_{t-1} + a_t$$

donde:

a_t : es un ruido blanco

μ : media de las observaciones

Procesos estocásticos estacionarios

Un proceso estocástico (Y_z) es estacionario si las distribuciones conjuntas de probabilidad asociadas con un vector (Y_1, Y_2, \dots, Y_k) son idénticas a las asociadas con el vector $(Y_{1+h}, Y_{2+h}, \dots, Y_{k+h})$ que se obtiene de una traslación temporal.

donde:

h : tiempo adicional no negativo, es decir, $h > 0$

Ruido blanco

Un ruido blanco es un proceso estocástico en el cual los valores son independientes para diferentes instantes de tiempo t y están distribuidos de manera idéntica a lo largo del tiempo. Este proceso tiene una media de cero y una varianza constante, y se representa por a_t .

$$a_t \sim N(0, \sigma^2), \quad \text{cov}(a_{(t_i)}, a_{(s_j)}) = 0 \quad \forall t_i \neq s_j$$

Una variable a_t (ruido blanco), consta de las siguientes propiedades:

- $E[a_t] = 0, \quad \forall t$

- $E[a_t]^2 = \sigma^2, \quad \forall t$
- $\text{cov}[a_{(t_i)}, a_{(s_j)}] = E[a_{(t_i)}, a_{(s_j)}] = 0, \quad \forall t_i \neq s_j$

2.2.4. Series de tiempo

Una de serie de tiempo: *“Es un conjunto de observaciones ordenadas según una característica cuantitativa de un fenómeno individual en diferentes momentos del tiempo, en el cual las observaciones son realizadas”* (Anderson, 1976).

“Una serie temporal es una secuencia de datos que consiste en N datos ordenados y tomados a intervalos regulares a lo largo del tiempo, registrando la evolución de una característica específica en una entidad observada en distintos momentos cronológicos” (Mauricio, 2007).

Sea una sucesión infinita de variables aleatorias ordenadas $\dots, Y_{-1}, Y_0, Y_1, Y_2, \dots; (Y_t : t = 0, \pm 1, \pm 2, \dots); (Y_t)$, donde Y_t es una variable aleatoria escalar referida a la unidad observable considerada en el momento t . Si se dispone de un conjunto finito de estas variables, Y_1, Y_2, \dots, Y_t . Se dice que esta sucesión de observaciones (realizaciones) forma una serie temporal.

Componentes de una serie temporal

El análisis preliminar de una serie constituye el primer paso a seguir a la hora de estudiar una serie temporal. Esta fase nos permite detectar las características más importantes de una serie, tales como su tendencia (creciente o decreciente), la existencia de ciclos, presencia de valores atípicos, etc. (Catalán, 2004).

a) **Tendencia** (T)

Representa la dirección general de cambio de una variable, puede ser creciente, decreciente o mantenerse estable.

b) **Componente cíclico o variación cíclica** (C)

Se refiere a los movimientos oscilantes y recurrentes que se presentan en torno a la tendencia principal.

c) **Componente estacional o variación estacional** (E)

Son las fluctuaciones regulares que ocurren en periodos menores a un año y se repiten de manera periódica a lo largo del tiempo.

d) **Componente aleatorio o variación residual** (a)

Abarca los cambios en la serie que no pueden explicarse mediante leyes conocidas y suelen ser el resultado de variaciones aleatorias o irregulares.

Descomposición tradicional de la serie de tiempo

La definición clásica del modelo de series de tiempo, indica que la serie Y_1, \dots, Y_t , puede ser expresada en sus componentes:

- Modelo aditivo: $Y_t = T_t + E_t + C_t + a_t$
- Modelo multiplicativo: $Y_t = T_t * E_t * C_t * a_t$
- Modelo mixto: $Y_t = T_t * E_t * C_t + a_t$

donde:

Y_t : Observaciones, T_t : Componente de tendencia, E_t : Componente estacional, C_t : Componente cíclica, a_t : Componente aleatoria

Análisis de series de tiempo

“El análisis de series de tiempo está dedicado al estudio de series; por lo general, los datos de dichas series son independientes, pero están correlacionados; se puede decir que existe una relación entre observaciones contiguas” (Anderson, 1976).

El análisis de series de tiempo consiste en una descripción (generalmente matemática) de los movimientos y componentes presentes:

- **Descripción:** Al tener una serie de tiempo, el primer paso en el análisis es graficar los datos y obtener medidas descriptivas simples de las propiedades principales de la serie.
- **Explicación:** Cuando las observaciones son tomadas sobre dos o más variables, es posible usar la variación en una serie para explicar la variación en las otras series.
- **Predicción:** Dada una serie de tiempo se intenta predecir los valores futuros de la serie. Este es el objetivo más frecuente en el análisis de series de tiempo.
- **Control:** Si una serie de tiempo se genera por mediciones de calidad de un proceso, el objetivo del análisis puede ser el control del proceso.

La necesidad de pronosticar

El pronóstico es tanto un arte como una ciencia que busca predecir eventos futuros. Aunque la intuición del investigador sigue siendo importante, quienes realizan esta tarea cuentan con métodos y herramientas avanzadas y sofisticadas. Desde tiempos anteriores a Cristo, el pronóstico ha estado asociado con fenómenos mágicos, supersticiones y fortuna. Sin embargo, con el avance tecnológico, la previsión se ha convertido en una práctica común en diversas áreas del conocimiento, incluyendo empresas y organizaciones modernas. Esto se debe a que todas las entidades deben enfrentar condiciones futuras sobre las que tienen información imperfecta. Para el pronóstico se realiza la siguiente pregunta: ¿por qué es necesario pronosticar? la respuesta es que en todas las organizaciones, existe un contexto de incertidumbre en el cual se deben tomar decisiones que impactan el futuro de la entidad, a pesar de esta situación ([Hanke, 1995](#)).

▪ Técnica de pronóstico cuantitativas

Se utiliza cuando hay suficientes datos históricos disponibles y cuando los datos al ser analizados son representativos de un futuro desconocido, se apoyan en modelos cuantitativos o modelos matemáticos con datos del pasado, bajo ciertos supuestos que son de suma importancia para el futuro ([Hanke, 1995](#)).

Modelo

Un modelo es una expresión formalizada de una teoría, o la representación matemática de los datos observados. En el análisis estadístico un modelo es expresado en símbolos, en forma matemática ([Kendall and Buckland, 1980](#)).

Modelos estadísticos

Un modelo estadístico es una representación formal de un sistema real, diseñado para mejorar la comprensión, realizar predicciones y facilitar su gestión. Estos modelos pueden manifestarse como representaciones físicas, análogas o simbólicas. Los modelos matemáticos, en particular, están compuestos por un conjunto de símbolos y relaciones lógico-matemáticas ([Felicísimo, 2009](#)).

Modelo vs Modelo Estadístico

Tabla 2.1:

Comparación de los aspectos entre un modelos matemático y un modelo estadístico

Aspecto	Modelo	Modelo estadístico
Definición	Representación abstracta o formal de un fenómeno o sistema real.	Representación formal con componentes aleatorios para describir y analizar datos.
Naturaleza	Puede ser conceptual, físico, simbólico o matemático.	Basado en estructuras matemáticas y probabilísticas.
Presencia del azar	No necesariamente incluye el azar o la incertidumbre.	El azar es un componente esencial.
Campo de aplicación	Se utiliza en diversas disciplinas (física, economía, ingeniería, etc.).	Se aplica en estadística y ciencia de datos.
Ejemplo	Modelo del sistema solar, modelo económico general.	Modelo de regresión lineal, modelo ARIMA, modelos log-lineales, etc.

Modelos de series de tiempo

Los modelos de series de tiempo se centran principalmente en la predicción y se basan exclusivamente en el comportamiento pasado de la variable de interés para elaborar pronósticos. Encontramos inmersos dos tipos de modelos:

- Modelo determinístico:** Se trata de métodos de extrapolación sencillos en los que no se hace referencia a las fuentes o naturaleza de la aleatoriedad subyacente en la serie ([Hurtado, 2008](#)).
- Modelo estocástico:** Se basan en la descripción simplificada del proceso aleatorio subyacente en la serie. En términos sencillos, se parte del supuesto de que la serie de datos observada, representada por $Y_1, Y_2, Y_3, \dots, Y_t$, proviene de un conjunto de variables aleatorias con una distribución conjunta complicada de determinar. Por lo tanto, se construyen modelos aproximados que resulten útiles para realizar predicciones futuras ([Hurtado, 2008](#)).

Modelos de series de tiempo univariante

Un modelo univariante para un proceso estocástico univariante o escalar (Y_t) consiste en un conjunto de suposiciones acerca de las propiedades teóricas de las distribuciones de probabilidad conjuntas, marginales o condicionales de los componentes del proceso (Y_t) del que se supone procede una serie temporal observada $(Y_t)_{t=1}^N$. (Mauricio, 2007).

Las técnicas más rigurosas para la predicción univariante son las denominadas técnicas o modelos Box-Jenkins, o modelos ARIMA, pues las técnicas Box-Jenkins constituyen un conjunto más amplio, dentro del cual los modelos ARIMA univariantes son solo una parte.

- a) Modelo univariante no integrado
- b) Modelo univariante integrado

Estacionariedad

"Se dice estacionario si su media y varianza son constantes en el tiempo y si el valor de la covarianza entre 2 períodos dependen sólo de la distancia entre los tiempos" (Rodriguez, 2022).

donde:

$E(Y_t)$: es constante para todo valor de t

$\text{Var}(Y_t)$ es constante para todo valor de t

$\text{Cov}(Y_t, Y_{t+h})$ sólo depende de h para cualquier t

Modelos lineales estacionarios

- a) **Modelos autorregresivos.**- "Los modelos autorregresivos (AR) expresan (Y_t) como una función lineal de cierto número de valores anteriores reales de (Y_t) " (Hanke, 2010).

- **Modelo AR(p).**- Un modelo autorregresivo de orden p , o abreviadamente un modelo AR(p), se define de la siguiente forma:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + a_t$$

donde:

$\phi_0, \phi_1, \dots, \phi_p$: son constantes y pertenecen a los reales (\mathbb{R})

$Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$: son los valores pasados de la serie

a_t : es ruido blanco

Esperanza del modelo AR(p)

Se aplica el valor esperado a la ecuación AR(p):

$$E(Y_t) = \phi_0 + \phi_1 E(Y_{t-1}) + \dots + \phi_p E(Y_{t-p}) + E(a_t)$$

de donde se obtiene:

$$\mu = \phi_0 + \phi_1 \mu + \dots + \phi_p \mu + 0$$

depejando tenemos: $E(Y_t) = \mu = \frac{\phi_0}{1 - \phi_1 - \phi_2 - \dots - \phi_p}$; $\phi_1 - \phi_2 - \dots - \phi_p \neq 1$

despejamos $\phi_0 = \mu(1 - \phi_1 - \phi_2 - \dots - \phi_p)$ y reemplazamos al modelo AR(p),

de donde se obtiene:

$$Y_t = \mu(1 - \phi_1 - \phi_2 - \dots - \phi_p) + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + a_t$$

Varianza del Modelo AR(p).- La varianza de un modelo autorregresivo de orden p se define como:

$$\text{Var}(Y_t) = \gamma_0 = \text{Var}(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + a_t)$$

Suponiendo que a_t es ruido blanco con varianza σ_t^2 , y que la varianza de una

constante es cero, se tiene:

$$\text{Var}(Y_t) = \gamma_0 = \phi_1^2 \text{Var}(Y_{t-1}) + \phi_2^2 \text{Var}(Y_{t-2}) + \cdots + \phi_p^2 \text{Var}(Y_{t-p}) + \sigma_t^2$$

Cuando $\text{Var}(Y_t) = \text{Var}(Y_{t-k})$,

se obtiene:

$$\gamma_0 = \frac{\sigma_a^2}{1 - \sum_{i=1}^p \phi_i^2}$$

Autocovarianza y Autocorrelación.- La covarianza entre Y_t y Y_{t-k} se define como:

$$\gamma_k = \text{Cov}(Y_t, Y_{t-k}) = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \cdots + \phi_p \gamma_{k-p}$$

La autocorrelación de orden k es:

$$\rho_k = \frac{\gamma_k}{\sigma_y^2}$$

- **Modelo AR(1).**- Un modelo autorregresivo de orden 1, se define de la siguiente forma:

$$AR(1) : Y_t = \phi_0 + \phi_1 Y_{t-1} + a_t$$

Esperanza del modelo AR(1)

$$E(Y_t) = \phi_0 + \phi_1 E(Y_{t-1}) + E(a_t)$$

de donde se tiene:

$$\mu = \phi_0 + \phi_1 \mu + 0$$

despejamos μ , y se obtiene:

$$E(Y_t) = \mu = \frac{\phi_0}{1 - \phi_1}$$

realizando el siguiente arreglo, tenemos: $\mu(1 - \phi_1) = \phi_0$

reemplazamos al modelo $AR(1)$,

$$Y_t = \mu(1 - \phi_1) + \phi_1 Y_{t-1} + a_t$$

$$Y_t - \mu = \phi_1(1 - \mu) + a_t$$

Varianza del modelo AR(1).- La varianza de un modelo autorregresivo de

orden 1 se define de la siguiente forma:

$$\text{Var}(Y_t) = \gamma_0 = \text{Var}(\phi_1 Y_{t-1} + a_t)$$

Dado que a_t es ruido blanco con varianza σ_a^2 , y la varianza de una constante es cero, se tiene:

$$\text{Var}(Y_t) = \gamma_0 = \phi_1^2 \text{Var}(Y_{t-1}) + \sigma_a^2$$

Cuando $\text{Var}(Y_t) = \text{Var}(Y_{t-1})$, se obtiene:

$$\gamma_0 = \frac{\sigma_a^2}{1 - \phi_1^2}$$

Autocovarianza y Autocorrelación.- La covarianza entre Y_t y Y_{t-k} se define como:

$$\gamma_k = \text{Cov}(Y_t, Y_{t-k}) = \phi_1^k \text{Var}(Y_t)$$

La autocorrelación de orden k es:

$$\rho_k = \phi_1^k$$

- **Modelo AR(2).**- Un modelo autorregresivo de orden 2, se define de la siguiente forma:

$$AR(2) : Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t$$

Esperanza del modelo AR(2)

Se aplica el valor esperado a la ecuación del modelo AR(2):

$$E(Y_t) = \phi_0 + \phi_1 E(Y_{t-1}) + \phi_2 E(Y_{t-2}) + E(a_t)$$

Dado que $E(a_t) = 0$, se tiene:

$$\mu = \phi_0 + \phi_1 \mu + \phi_2 \mu$$

despejando μ , obtenemos:

$$\mu = \frac{\phi_0}{1 - \phi_1 - \phi_2}$$

$$\phi_0 = \mu(1 - \phi_1 - \phi_2)$$

reemplazamos ϕ_0 por $\mu(1 - \phi_1 - \phi_2)$:

$$Y_t = \mu(1 - \phi_1 - \phi_2) + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t$$

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + a_t$$

Varianza del modelo AR(2)

La esperanza del modelo es:

$$\text{Var}(Y_t) = \gamma_0 = \text{Var}(\phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t)$$

Dado que a_t es ruido blanco con varianza σ_t^2 , y la varianza de una constante es cero, se tiene:

$$\text{Var}(Y_t) = \gamma_0 = \phi_1^2 \text{Var}(Y_{t-1}) + \phi_2^2 \text{Var}(Y_{t-2}) \sigma_t^2$$

Cuando $\text{Var}(Y_t) = \text{Var}(Y_{t-1})$, se obtiene:

$$\gamma_0 = \frac{\sigma_t^2}{1 - \phi_1^2 - \phi_2^2}$$

Autocovarianza del Modelo AR(2)

Autocovarianza $\text{Cov}(Y_t, Y_{t-1})$

Multiplicamos ambos lados de $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t$ por Y_{t-1} y tomamos el valor esperado:

$$E(Y_t Y_{t-1}) = \phi_1 E(Y_{t-1}^2) + \phi_2 E(Y_{t-2} Y_{t-1}) + E(a_t Y_{t-1})$$

dado que:

$$E(a_t Y_{t-1}) = 0, E(Y_{t-2} Y_{t-1}) = \text{Cov}(Y_{t-1}, Y_{t-2}) = \gamma_0 \text{ y}$$

$$E(Y_{t-1}^2) = \text{Var}(Y_{t-1}) = \gamma_0 :$$

$$\text{Cov}(Y_t, Y_{t-1}) = \phi_1 \text{Var}(Y_{t-1}) + \phi_2 \text{Cov}(Y_{t-1}, Y_{t-2})$$

Observa

$$\text{Var}(Y_{t-1}) = \text{Var}(Y_t) \text{ y } \text{Cov}(Y_{t-1}, Y_{t-2}) = \text{Cov}(Y_t, Y_{t-1}) :$$

$$\text{Cov}(Y_t, Y_{t-1}) = \phi_1 \text{Var}(Y_t) + \phi_2 \text{Cov}(Y_t, Y_{t-1})$$

$$\gamma_1 = \phi_1 \gamma_0 + \phi_2 \gamma_1$$

resolviendo para $\text{Cov}(Y_t, Y_{t-1})$:

$$\gamma_1 = \frac{\phi_1 \gamma_0}{1 - \phi_2}$$

Autocovarianza $\text{Cov}(Y_t, Y_{t-2})$

Multiplicamos ambos lados de $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + a_t$ por Y_{t-2} y tomamos el valor esperado:

$$E(Y_t Y_{t-2}) = \phi_1 E(Y_{t-1} Y_{t-2}) + \phi_2 E(Y_{t-2}^2) + E(a_t Y_{t-2})$$

dado que:

$$(E(a_t Y_{t-2}) = 0) \text{ y } (E(Y_{t-1} Y_{t-2}) = \text{Cov}(Y_{t-2}, Y_{t-1}))$$

$$\text{Cov}(Y_t, Y_{t-2}) = \phi_1 \text{Cov}(Y_{t-2}, Y_{t-1}) + \phi_2 \text{Var}(Y_{t-2})$$

Observa

$$\text{Var}(Y_{t-2}) = \text{Var}(Y_t) \text{ y } \text{Cov}(Y_{t-2}, Y_{t-1}) = \text{Cov}(Y_t, Y_{t-1})$$

$$\text{Cov}(Y_t, Y_{t-2}) = \phi_1 \text{Cov}(Y_t, Y_{t-1}) + \phi_2 \text{Var}(Y_t)$$

$$\gamma_2 = \phi_1 \gamma_1 + \phi_2 \gamma_0$$

Autocorrelación del Modelo AR(2): La autocorrelación de un modelo

AR(p) se define como:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

para $k = 1$:

$$\rho_1 = \frac{\gamma_1}{\gamma_0}$$

para $k = 2$:

$$\rho_2 = \frac{\gamma_2}{\gamma_0}$$

b) Modelos de medias móviles

"Un modelo de medias móviles explica el valor de una determinada variable en un período t en función de término independiente y una sucesión de errores correspondientes a períodos precedentes, ponderados convenientemente" [Uriel \(1985\)](#).

- **Modelo MA(q).**- Un modelo de medias móviles de orden q , $MA(q)$, se de-

fine de la siguiente forma:

$$Y_t = \mu + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}$$

Donde:

Y_t : Variable dependiente

$\theta_1, \theta_2, \dots, \theta_q$: Coeficiente de ponderación

μ : Valor medio

a_t : Residuo o error, también conocido como "ruido blanco"

a_{t-1}, \dots, a_{t-q} : Valores previos residuales

Esperanza del modelo $MA(q)$

$$E(Y_t) = \cancel{E(a_t)} + \theta_1 \cancel{E(a_{t-1})} + \dots + \theta_q \cancel{E(a_{t-q})} + \mu$$

$$E(Y_t) = \mu$$

Varianza del modelo $MA(q)$

$$\text{Var}(Y_t) = \gamma_0 = E((Y_t - \mu)^2)$$

$$\text{Var}(Y_t) = E[\mu + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} - \mu]^2$$

$$\text{Var}(Y_t) = E[a_t^2 + \theta_1^2 a_{t-1}^2 + \dots + \theta_q^2 a_{t-q}^2 + \cancel{2a_t \theta_1 a_{t-1}} + \dots + \cancel{2a_t \theta_q a_{t-q}} \dots]$$

$$\text{Var}(Y_t) = E(a_t^2) + \theta_1^2 E(a_{t-1}^2) + \dots + \theta_q^2 E(a_{t-q}^2)$$

$$\text{Var}(Y_t) = \sigma^2 + \theta_1^2 \sigma^2 + \dots + \theta_q^2 \sigma^2$$

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma^2$$

- **Modelo MA(1)**.- Un modelo de medias móviles de orden (1), se define de la siguiente forma:

$$Y_t = \mu + a_t + \theta_1 a_{t-1}$$

Esperanza del modelo MA(1)

$$E(Y_t) = \mu + \cancel{E(a_t)} + \theta_1 \cancel{E(a_{t-1})}$$

$$E(Y_t) = \mu$$

Varianza del modelo MA(1)

$$\text{Var}(Y_t) = \gamma_0 = E((Y_t - \mu)^2)$$

$$\text{Var}(Y_t) = E[\mu + a_t + \theta_1 a_{t-1} - \mu]^2$$

$$\text{Var}(Y_t) = E[+a_t + \theta_1 a_{t-1}]^2$$

$$\text{Var}(Y_t) = E[a_t^2 + \theta_1^2 a_{t-1}^2 + 2a_t \theta_1 a_{t-1}]$$

$$\text{Var}(Y_t) = E(a_t^2) + \theta_1^2 E(a_{t-1}^2) + \cancel{2\theta_1 E(a_t a_{t-1})}]$$

$$\text{Var}(Y_t) = \sigma^2 + \theta_1^2 \sigma^2$$

$$\gamma_0 = (1 + \theta_1^2) \sigma^2$$

- **Modelo MA(2).**- Un modelo de medias móviles de orden(2), esta definida de la siguiente forma:

$$Y_t = \mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}$$

Esperanza del modelo MA(2)

$$E(Y_t) = \mu + \cancel{E(a_t)} + \theta_1 \cancel{E(a_{t-1})} + \theta_2 \cancel{E(a_{t-2})}$$

$$E(Y_t) = \mu$$

Varianza del modelo MA(2)

$$\text{Var}(Y_t) = \gamma_0 = E((Y_t - \mu)^2)$$

$$\text{Var}(Y_t) = E(\mu + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} - \mu)^2$$

$$\text{Var}(Y_t) = E(a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2})^2 \text{Var}(Y_t) = E(a_t^2 + \theta_1^2 a_{t-1}^2 + \theta_2^2 a_{t-2}^2 + 2a_t \theta_1 a_{t-1} +$$

$$2a_t\theta_2a_{t-2} + 2\theta_1a_{t-1}a_t\theta_2a_{t-2})$$

$$\text{Var}(Y_t) = E(a_t^2) + \theta_1^2 E(a_{t-1}^2) + \theta_2^2 E(a_{t-2}^2) + \cancel{2\theta_1 E(a_t a_{t-1})} + \cancel{2\theta_2 E(a_t a_{t-2})} + \cancel{2\theta_1\theta_2 E(a_{t-1} a_{t-2})}$$

$$\text{Var}(Y_t) = \sigma^2 + \theta_1^2\sigma^2 + \theta_2^2\sigma^2$$

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma^2$$

c) **Modelos autorregresivo de medias móviles(ARMA)** Es una extensión natural de los modelos $AR(p)$ y $MA(q)$ es un modelo que incluye términos autorregresivos como de medias móviles.

- **Modelo ARMA(p, q).**- Está definido de la siguiente forma:

$$Y_t = \mu + \phi_0 + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}$$

Donde:

Y_t : Variable dependiente

$\theta_1, \theta_2, \dots, \theta_q$: Coeficiente de ponderación

μ : Valor medio

a_t : Residuo o error, también conocido como "ruido blanco"

$\phi_0, \phi_1, \dots, \phi_p$: Son constantes y pertenecen a los reales (\mathbb{R})

$Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$: Son los valores pasados de la serie

a_{t-1}, \dots, a_{t-q} : Valores previos residuales

Esperanza ARMA(p, q)

Tomando el valor esperado.

$$E(Y_t) = E(\mu + \phi_0 + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q})$$

al aplicar la linealidad de la esperanza y recordar que $E(a_t) = 0$:

$$E(Y_t) = \mu + \phi_0 + \phi_1 E(Y_{t-1}) + \cdots + \phi_p E(Y_{t-p})$$

dado que estamos en un estado estacionario, la esperanza de Y_t es la misma para todos los tiempos, así que:

$$E(Y_t) = E(Y_{t-1}) = E(Y_{t-2}) = \cdots = E(Y) = \mu$$

sustituyendo esto en la ecuación, obtenemos:

$$E(Y) = \mu + \phi_0 + (\phi_1 + \cdots + \phi_p)E(Y)$$

resolviendo para $E(Y)$, tenemos:

$$E(Y_t) = \frac{\mu + \phi_0}{1 - \phi_1 - \phi_2 - \cdots - \phi_p}$$

Varianza del modelo $ARMA(p, q)$

La varianza del modelo $ARMA(p, q)$ es una medida de la dispersión de los valores de Y_t en torno a su media. Dado que la esperanza de Y_t es μ en estado estacionario, podemos escribir:

$$\text{Var}(Y_t) = \gamma_0 = E[(Y_t - \mu)^2]$$

sustituyendo Y_t en la ecuación anterior:

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \cdots + \phi_p(Y_{t-p} - \mu) + a_t + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q}$$

al ser los errores a_t ruido blanco, sus varianzas se comportan de la siguiente manera:

$$\text{Var}(a_t) = \sigma_t^2 \quad \text{y} \quad E[a_t a_{t-k}] = 0 \quad \text{para} \quad k \neq 0$$

sustituyendo esta propiedad de los residuos en la ecuación:

$$\text{Var}(Y_t) = \text{Var}(\phi_1(Y_{t-1} - \mu) + \cdots + \phi_p(Y_{t-p} - \mu) + a_t + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q})$$

dado que la varianza de una suma de términos es la suma de las varianzas, siempre que no haya covarianzas, tenemos:

$$\text{Var}(Y_t) = \phi_1^2 \text{Var}(Y_{t-1}) + \cdots + \phi_p^2 \text{Var}(Y_{t-p}) + \sigma^2(1 + \theta_1^2 + \cdots + \theta_q^2)$$

en estado estacionario, la varianza de Y_t es constante, es decir,

$$\text{Var}(Y_t) = \text{Var}(Y_{t-1}) = \dots = \text{Var}(Y).$$

entonces, podemos escribir:

$$\text{Var}(Y) = \sigma^2 \frac{1+\theta_1^2+\dots+\theta_q^2}{1-\phi_1^2-\dots-\phi_p^2}$$

por lo tanto, la varianza del modelo $ARMA(p, q)$ es:

$$\gamma_0 = \sigma^2 \frac{1+\theta_1^2+\dots+\theta_q^2}{1-\phi_1^2-\dots-\phi_p^2}$$

- **Modelo $ARMA(1, 1)$.**- El modelo está dado de de la siguiente forma:

$$Y_t = \mu + \phi_0 + \phi_1 Y_{t-1} + a_t + \theta_1 a_{t-1}$$

Esperanza del modelo $ARMA(1, 1)$

Tomamos el valor esperado a ambos lados del modelo y recordamos que la esperanza de los errores a_t es cero, ya que es ruido blanco, es decir:

$$E(a_t) = 0$$

entonces, al aplicar la esperanza a la ecuación Y_t , obtenemos:

$$E(Y_t) = E(\mu + \phi_0 + \phi_1 Y_{t-1} + a_t + \theta_1 a_{t-1})$$

al aplicar la linealidad de la esperanza y recordar que

$$E(a_t) = 0 \text{ y } E(a_{t-1}) = 0:$$

$$E(Y_t) = \mu + \phi_0 + \phi_1 E(Y_{t-1})$$

dado que es un proceso estacionario, la esperanza de Y_t es la misma para todos los tiempos, así que:

$$E(Y_t) = E(Y_{t-1}) = E(Y)$$

sustituyendo esto en la ecuación, obtenemos:

$$E(Y) = \mu + \phi_0 + \phi_1 E(Y)$$

resolviendo para $E(Y)$, tenemos:

$$E(Y) - \phi_1 E(Y) = \mu + \phi_0$$

factorizando $E(Y)$ en el lado izquierdo:

$$(EY)(1 - \phi_1) = \mu + \phi_0$$

finalmente, despejamos $E(Y)$:

$$E(Y) = \frac{\mu + \phi_0}{1 - \phi_1}$$

Por lo tanto, la esperanza del modelo $ARMA(1, 1)$ es:

$$E(Y_t) = \frac{\mu + \phi_0}{1 - \phi_1}$$

Varianza del modelo $ARMA(1, 1)$

$$Var(Y_t) = \gamma_0 = \frac{(1 + \theta_1^2 + 2\phi_1\theta_1)\sigma^2}{1 - \phi_1^2}$$

Autocovarianza

$$\gamma_1 = E(\hat{Y}_t, \hat{Y}_{t-1}) = \phi_1\gamma_0 + \theta_1\sigma^2$$

$$\gamma_2 = E(\hat{Y}_t, \hat{Y}_{t-2}) = \phi_1^2\gamma_0 + \phi_1\theta_1\sigma^2 \quad \gamma_j = \phi_1^2\gamma_{j-1} \quad \forall j \geq 2$$

Autocorrelación

$$\rho_0 = \frac{\gamma_0}{\gamma_0} = 1$$

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\phi_1\gamma_0 + \theta_1\sigma^2}{\frac{\sigma^2(1 + \theta_1^2 + 2\phi_1\theta_1)}{1 - \phi_1^2}} = \phi_1 + \frac{\theta_1(1 - \phi_1^2)}{1 + \theta_1^2 + 2\phi_1\theta_1}$$

$$\rho_2 = \frac{\gamma_2}{\gamma_0} = \dots \quad \rho_j = \phi_1\rho_{j-1} \quad \forall j \geq 2$$

Modelos lineales no estacionarios

a) Modelo de medias móviles autorregresivo integrados (ARIMA)

EL modelo $ARIMA(p, d, q)$ denominado proceso autorregresivo integrado de medias móviles de orden p, d, q el cual permite describir una secuencias de observaciones después de que hayan pasado por la diferenciación d veces; también considerado una generalización de los modelos $AR(p)$ y $MA(q)$.

- **Modelo $ARIMA(p, d, q)$.**- Donde p : es número específico de periodos desfasados, q : número de valores previos y d : número de diferenciaciones; la

ecuación está definido de la siguiente forma:

$$\Delta^d Y_t = \phi_0 + \phi_1 \Delta^d Y_{t-1} + \dots + \phi_p \Delta^d Y_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}$$

Donde:

$\Delta^d Y_t$: Diferencia de orden d

$\theta_1, \theta_2, \dots, \theta_q$: Número de valores previos de residuo a_t

a_t : error aleatorio, también conocido como "ruido blanco"

$\phi_0, \phi_1, \dots, \phi_p$: Número de retardos para la variable Y

$Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$: Son los valores pasados de la serie

a_{t-1}, \dots, a_{t-q} : Valores previos residuales

Esperanza del modelo $ARIMA(p, d, q)$

Tomamos la esperanza de ambos lados de la ecuación. Sabemos que a_t es ruido blanco, por lo que $E[a_t] = 0$.

$$E(\Delta^d Y_t) = E(\phi_0 + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q})$$

aplicando la propiedad de linealidad de la esperanza:

$$E(\Delta^d Y_t) = \phi_0 + \phi_1 E(Y_{t-1}) + \dots + \phi_p E(Y_{t-p})$$

en estado estacionario, la esperanza de Y_t es la misma para todos los t , es decir,

$$E(Y_t) = E(Y_{t-1}) = \dots = E(Y) = \mu$$

Sustituyendo:

$$0 = \phi_0 + \phi_1 \mu + \dots + \phi_p \mu$$

factorizando μ :

$$0 = \phi_0 + \mu(\phi_1 + \dots + \phi_p)$$

despejando μ :

$$\mu = -\frac{\phi_0}{\phi_1 + \dots + \phi_p}:$$

por lo tanto, la esperanza del modelo $ARIMA(p, d, q)$ es:

$$E(Y) = -\frac{\phi_0}{\phi_1 + \dots + \phi_p}$$

Varianza del modelo $ARIMA(p, d, q)$

La varianza de la serie Y_t en un modelo $ARIMA$ depende de los términos de ruido blanco a_t , cuya varianza es σ^2 . Para calcular la varianza de $\Delta^d Y_t$, comenzamos con la varianza de la primera diferencia:

$$\text{Var}(\Delta Y_t) = \text{Var}(Y_t - Y_{t-1})$$

usamos la propiedad de la varianza de la diferencia:

$$\text{Var}(\Delta Y_t) = \text{Var}(Y_t) + \text{Var}(Y_{t-1}) - 2\text{Cov}(Y_t, Y_{t-1})$$

en estado estacionario, la varianza de Y_t y Y_{t-1} es la misma, es decir, $\text{Var}(Y_t) = \text{Var}(Y)$, por lo tanto:

$$\text{Var}(\Delta Y_t) = 2\text{Var}(Y) - 2\text{Cov}(Y_t, Y_{t-1})$$

Autocovarianza del modelo $ARIMA(p, d, q)$

La autocovarianza entre Y_t y Y_{t-k} se define como:

$$\gamma_k = \text{Cov}(Y_t, Y_{t-k}) = E[(Y_t - \mu)(Y_{t-k} - \mu)]$$

En un modelo $ARIMA$, la autocovarianza depende de los términos ϕ_i y θ_i , así como de la estructura del proceso autorregresivo y de medias móviles.

Autocorrelación en el modelo $ARIMA(p, d, q)$

La autocorrelación es la razón entre la autocovarianza y la varianza:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

donde

$$\gamma_0 = \text{Var}(Y_t) \text{ es la varianza de la serie.}$$

Para el caso $ARIMA(1, 1, 1)$, la autocorrelación de orden 1 sería: $\rho_1 =$

$$\frac{\gamma_1}{\gamma_0} = \frac{\phi_1(1+\theta_1)}{1-\phi_1^2}$$

Función de autocorrelación Simple

La Función de Autocorrelación Simple (FAS) es una herramienta de análisis estadístico que nos permite determinar el nivel de autocorrelación presente en los datos y en qué retardos, k , se manifiesta esta autocorrelación. Su utilidad radica en medir la inercia o tendencia de una serie de tiempo, es decir, cuantificar el grado de dependencia que existe entre las observaciones actuales y las observaciones de k periodos anteriores (Mauricio, 2007).

Para estimar los coeficientes de autocorrelación está dado por la siguiente ecuación:

$$\hat{\rho}_{(k)} = \frac{\sum_{t=k+1}^T [(x_t - \hat{\mu})(x_{t-k} - \hat{\mu})]}{\sum_{t=1}^T (x_t - \hat{\mu})^2}$$

$$0 < t \leq T \text{ y } -1 \leq \hat{\rho}_{(k)} \leq 1$$

De donde:

T: Número de periodos disponibles, en la parte superior se encuentra la covarianza de una observación con su pasado, en la parte inferior es la varianza de la observación presente.

x_t : Observación presente

$\hat{\mu}$: Media estimada

La función de autocorrelación de manera general está dada por:

$$\rho_{(t-1,t)} = \frac{\gamma_{(t-1,t)}}{\rho_{t-1}\rho_t}$$

donde $\rho_{(t-1,t)}$ representa el grado de dependencia lineal existente entre X_{t-1} y X_t , y su valor oscila entre -1 y 1.

Función de autocorrelación parcial

La función de autocorrelación parcial (PACF) es una medida que evalúa la correlación

entre observaciones de una serie de tiempo que están distanciadas por k unidades de tiempo (Y_t y Y_{t-k}), después de ajustarse para la presencia de los demás términos de desfase más corto ($Y_{t-1}, Y_{t-2}, \dots, Y_{t-k-1}$).

La diferencia con la función de autocorrelación simple es que, mide la autocorrelación o relación lineal entre valores que se encuentran en k intervalos tomando en cuenta los valores de los intervalos intermedios. La siguiente ecuación está dada por:

$$\alpha_{(s,t)} = \frac{\text{Cov}(Y_{(s)} - \hat{Y}_s^{(s,t)}, Y_{(t)} - \hat{Y}_t^{(s,t)})}{\sqrt{\text{Var}(Y_{(s)} - \hat{Y}_s^{(s,t)}) \cdot \text{Var}(Y_{(t)} - \hat{Y}_t^{(s,t)})}}$$

Donde:

$\hat{Y}_s^{(s,t)}$ es el predictor lineal construido a partir de las variables medidas en el intervalo de tiempo comprendido entre s y t .

$\alpha_{(s,t)}$ es una medida del grado de dependencia lineal existente entre $Y_{(s)}$ y $Y_{(t)}$.

Transformación de Box-Cox

"Es una técnica utilizada para estabilizar la varianza y hacer que las variables se ajusten mejor al supuesto de normalidad. Esta transformación debe aplicar una función matemática a la variable original, la cual está parametrizada por un valor lambda λ " (Box, 1964).

La fórmula general de la transformación de Box-Cox es la siguiente:

$$y^{(\lambda)} = \begin{cases} \frac{y^{\lambda}-1}{\lambda} & \text{si } \lambda \neq 0 \\ \log(y) & \text{si } \lambda = 0 \end{cases}$$

Donde:

- $y^{(\lambda)}$ representa el valor transformado de la variable original.
- y es el valor de la variable original.

- λ es el parámetro que controla la transformación. Puede tomar cualquier valor real, incluyendo el cero.

Criterio de selección de modelos

- a) **El Akaike's Information Criterion (AIC)**.- "El Criterio de Información de Akaike (AIC) es una técnica de selección de modelos utilizada en el análisis de datos, propuesta por Hirotugu Akaike en 1974. Su propósito es comparar diferentes modelos y determinar cuál de ellos es el más apropiado para describir y predecir los datos" (Schwarz, 1978); el AIC se calcula utilizando la siguiente fórmula:

$$AIC = -2 \ln(L) + 2k$$

donde $\ln(L)$ es el logaritmo de la función de verosimilitud del modelo y k es el número de parámetros estimados en el modelo.

En general, se prefiere un modelo con un valor de AIC más bajo, lo que indica un mejor equilibrio entre ajuste y complejidad.

- b) **El criterio de información bayesiano (BIC)** .- Es una medida utilizada para seleccionar el mejor modelo, fue propuesto en 1978 por Gideon E. Schwarz.

La fórmula para el BIC es la siguiente:

$$BIC = -2 \ln(L) + k \ln(n)$$

donde $\ln(L)$ es el logaritmo de la verosimilitud del modelo, k es el número de parámetros estimados en el modelo y n es el número de observaciones en los datos. En general, se prefiere un modelo con un valor de BIC más bajo, ya que

indica un mejor equilibrio entre ajuste y complejidad. La elección del modelo final está dado por un conjunto de criterios y el conocimiento del dominio específico (Schwarz, 1978).

- c) **Error Medio Absoluto Porcentual o MAPE.**- Es una métrica utilizada para medir la precisión de un modelo de predicción. Se expresa en términos porcentuales y se define como el promedio de las desviaciones absolutas porcentuales entre los valores reales y los valores predichos (Bergmeir and Benítez, 2012).

La fórmula para del MAPE es la siguiente:

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100$$

Donde:

- Y_t es el valor real en el período t .
- \hat{Y}_t es el valor predicho en el período t .
- n es el número total de observaciones.

Metodología de Box-Jenkins

La metodología se caracteriza porque no asume ningún patrón en los datos históricos de la serie a pronosticar ya que utiliza enfoque iterativo de identificación de un modelo, luego el modelo se compara con los datos históricos para ver si describe la serie con exactitud. Dicho modelo presenta buen ajuste cuando los residuos son generalmente pequeños y están distribuidos aleatoriamente; si en caso el modelo elegido no es satisfactorio, el proceso se repite nuevamente hasta encontrar un modelo satisfactorio (Hanke, 2005).

Etapas de la metodología Box-Jenkins

a) **Análisis de la serie temporal.**- El proceso inicia con la recopilación de la serie de tiempo, la cual debe estar compuesta por datos ordenados cronológicamente. Para comprender la naturaleza de la serie, se realiza un análisis descriptivo y gráfico visual, que permite detectar patrones, tendencias, estacionalidades, rupturas estructurales o valores atípicos.

- Histogramas y diagramas de dispersión, para observar la distribución de los datos.
- Gráficos de línea para visualizar el comportamiento temporal.
- Descomposición aditiva o multiplicativa, que permite visualizar los componentes de tendencia, estacionalidad y residual.
- Autocorrelación (ACF) y autocorrelación parcial (PACF), que ayudan a evaluar la dependencia temporal.

b) **Evaluación de la estacionariedad.**- Una serie es estacionaria cuando sus propiedades estadísticas (media, varianza y covarianza) se mantienen constantes en el tiempo. Esta condición es esencial para la aplicabilidad del enfoque Box-Jenkins. Se evalúa mediante:

- Pruebas de raíz unitaria, como la prueba ADF (Augmented Dickey-Fuller).
- Análisis visual de la serie y de las funciones ACF y PACF.
- Si la serie es estacionaria, se continúa con la identificación del modelo.
- Si no lo es, se procede a la transformación de los datos.

c) **Transformación de datos.**- Cuando la serie no presenta estacionariedad en la varianza, se recurre a transformaciones como la de Box-Cox, que consiste en aplicar

una función de potencia al conjunto de datos, con el objetivo de estabilizar la varianza.

d) **Verificación de la estacionariedad de los datos transformados.**- Después de la primera transformación, se comprueba si el proceso ha logrado la estacionariedad. Esto se hace mediante pruebas estadísticas y análisis gráfico.

- Si el proceso es estacionario, se continúa con la identificación del modelo.
- Si no lo es, se aplican nuevas transformaciones.

e) **Identificación del Modelo y Estimación de Parámetros.**- Una vez estacionarizada la serie, se identifican los posibles modelos ARIMA mediante el análisis de las funciones ACF y PACF. Se determina el orden de los componentes autoregresivos (AR), de medias móviles (MA) y de diferenciación (I).

Método de estimación: Máxima Verosimilitud Condicional

La estimación de los parámetros se realiza mediante la Máxima Verosimilitud Condicional (Conditional Maximum Likelihood). Según [Hamilton \(2020\)](#), este método consiste en maximizar la probabilidad condicional de los datos observados dados los valores pasados. A diferencia de la máxima verosimilitud exacta, la versión condicional es computacionalmente más simple y adecuada cuando se dispone de datos iniciales para condicionar la estimación.

Ventajas del método:

- Utiliza la estructura de dependencia temporal de la serie.
- Permite obtener estimadores consistentes y eficientes.

- Es compatible con la metodología Box-Jenkins, donde los modelos ARIMA se ajustan condicionalmente a las observaciones iniciales.

Estimación de los parámetros para el Modelo ARIMA(1,1,1)

1. Modelo ARIMA(1,1,1)

La forma del modelo es:

$$\Delta y_t = y_t - y_{t-1} = \phi(y_{t-1} - y_{t-2}) + \theta a_{t-1} + a_t$$

Es decir, la serie diferenciada sigue un modelo ARMA(1,1).

2. Expresión del error condicional

De la ecuación del modelo:

$$w_t = \Delta y_t = \phi w_{t-1} + \theta a_{t-1} + a_t$$

Despejamos a_t :

$$a_t = w_t - \phi w_{t-1} - \theta a_{t-1}$$

Como $w_t = y_t - y_{t-1}$, entonces:

$$a_t = (y_t - y_{t-1}) - \phi(y_{t-1} - y_{t-2}) - \theta a_{t-1}$$

Esta es la expresión condicional del error en función de los datos observados y errores pasados.

3. Supuesto de distribución

Se asume:

$$a_t \sim \mathcal{N}(0, \sigma^2)$$

Por lo tanto, la función de densidad condicional de cada error es:

$$f(a_t) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{a_t^2}{2\sigma^2}\right)$$

4. Función de verosimilitud condicional

La verosimilitud condicional para todos los datos desde $t = 2$ hasta $t = n$ es:

$$L(\phi, \theta, \sigma^2) = \prod_{t=2}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{a_t^2}{2\sigma^2}\right)$$

Tomando logaritmo para facilitar la estimación:

$$\log L = -\frac{n-1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=2}^n a_t^2$$

Sustituyendo la expresión de a_t :

$$a_t = y_t - y_{t-1} - \phi(y_{t-1} - y_{t-2}) - \theta a_{t-1}$$

Entonces:

$$\log L(\phi, \theta, \sigma^2) = -\frac{n-1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=2}^n [y_t - y_{t-1} - \phi(y_{t-1} - y_{t-2}) - \theta a_{t-1}]^2$$

5. Método de estimación

Para obtener los estimadores:

$$(\hat{\phi}, \hat{\theta}, \hat{\sigma}^2) = \arg \max_{\phi, \theta, \sigma^2} \log L$$

En la práctica:

- Se inicializa $a_1 = 0$ o por alguna aproximación.
- Se itera sobre los valores de los parámetros hasta encontrar el máximo de $\log L$.

f) **Evaluación y diagnóstico del modelo.**- Una vez estimado el modelo, se realiza su evaluación. Esta etapa incluye dos componentes fundamentales: el análisis de residuos y la comparación entre modelos candidatos.

Evaluación de los residuos

- Prueba de Normalidad.
- Prueba de Independencia.
- Prueba de Homogeneidad y Heterocedasticidad.
- Prueba de Estacionariedad.
- Prueba de T de Student.

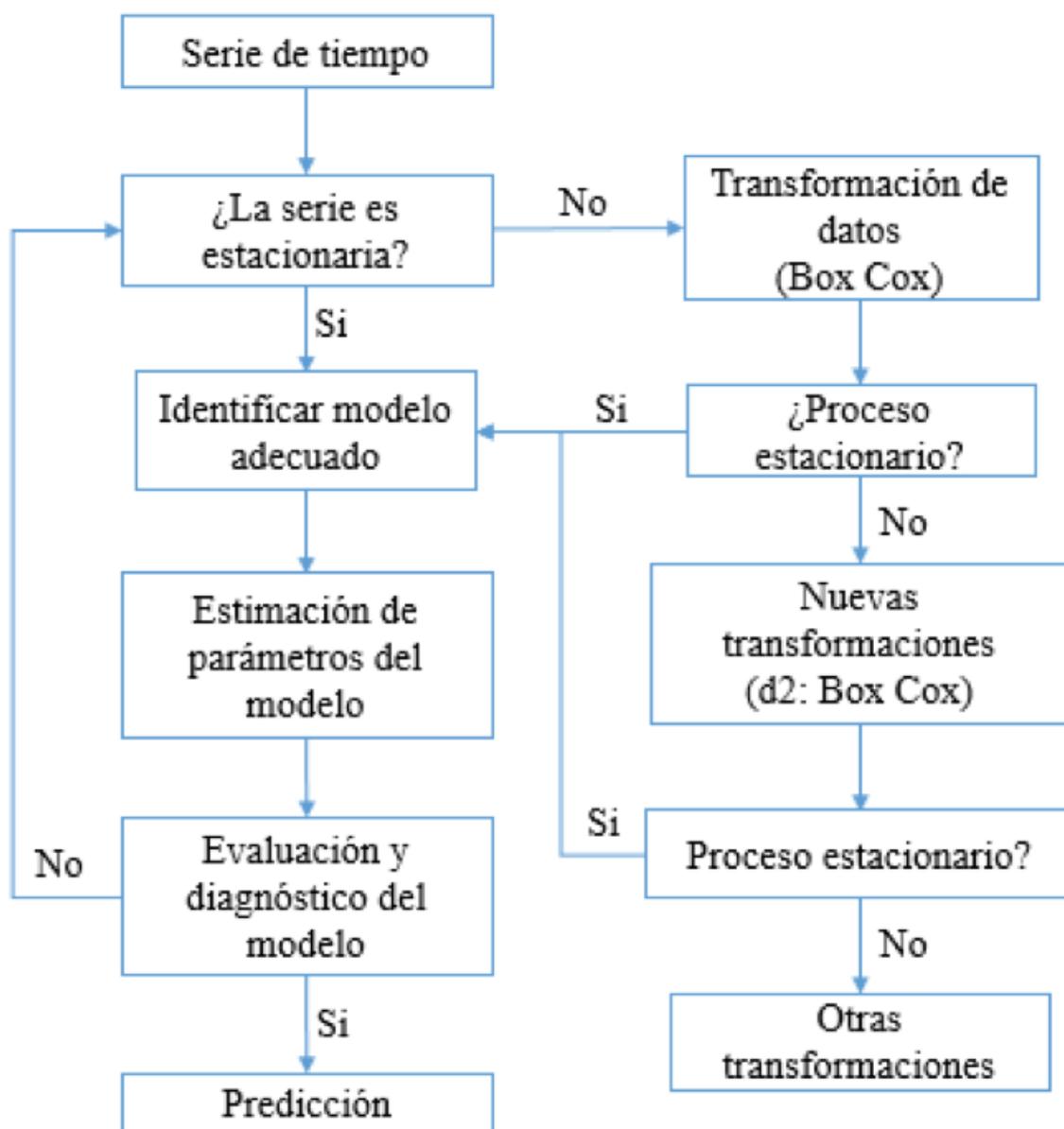
Criterios de evaluación y selección

- AIC (Akaike Information Criterion): penaliza la complejidad del modelo. Un menor AIC indica mejor ajuste..

- BIC (Bayesian Information Criterion): similar al AIC pero con una penalización más fuerte por el número de parámetros.
- MAE (Mean Absolute Error): promedio del valor absoluto del error.
- RMSE (Root Mean Squared Error): raíz del promedio del error cuadrático, sensible a errores grandes.
- MAPE (Mean Absolute Percentage Error): error promedio expresado en porcentaje.

Con un modelo validado, se procede a realizar predicciones sobre futuros valores de la serie. Estas predicciones se acompañan de intervalos de confianza, que cuantifican la incertidumbre. Las predicciones pueden ser puntuales o en horizonte de tiempo múltiple, dependiendo del objetivo del análisis.

Figura 2.1:
Esquema a seguir para el desarrollo adecuado de la metodología Box-Jenkins



Fuente. Elaborado por *Uribe et al. (2017)*, en su trabajo titulado "Incidencia de las políticas públicas de empleo sobre la desocupación en Colombia: un análisis de intervención para el periodo 2002 - 2014"

Modelos no estacionarios heterocedásticos Modelo ARCH (Autoregressive Conditional Heteroscedasticity)

Es un modelo utilizado para analizar y pronosticar la volatilidad de una serie temporal, el modelo asume que σ^2 cambia a travez del tiempo y depende de los valores del pasado. En otras palabras, la volatilidad de la serie temporal es autocorrelacionada y depende de su propia historia (Engle, 1982).

Modelo de la Serie Temporal

La serie temporal y_t se modela como:

$$y_t = \mu + \epsilon_t$$

$$\epsilon_t = a_t \sigma_t$$

donde:

- y_t : Es el valor observado de la serie temporal en el tiempo t .
- ϵ_t : Es el residuo o termino error en el tiempo t
- a_t : es el ruido blanco en el tiempo t , con media 0 y varianza finita.
- σ_t : Es la desviación estándar condicional en el tiempo t , que mide la volatilidad condicional.

Para garantizar la estimación de los parámetros del modelo, es necesario asumir que el término del error sigue una distribución normal estandar, con ello la expresión del modelo se reduce a:

$$y_t = a_t \sigma_t$$

donde:

- y_t : Es el valor observado de la serie temporal en el tiempo t .
- a_t : Es el término de error estandarizado que sigue una distribución normal estándar, $a_t \sim \mathcal{N}(0, 1)$.
- σ_t : Es la desviación estándar condicional en el tiempo t , que mide la volatilidad condicional.

Esta expresión implica que la serie temporal observada y_t se puede obtener multiplicando el error estandarizado a_t por la desviación estándar condicional σ_t .

Modelo para la Varianza Condicional

La varianza condicional σ_t^2 se modela como:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2$$

donde:

- σ_t^2 : Es la varianza condicional en el tiempo t .
- α_0 : Es el parámetro de intercepto, que representa el nivel base de la varianza.
- α_i : Son los parámetros del modelo que capturan la influencia de los errores pasados al cuadrado en la varianza condicional actual.

- ϵ_{t-i}^2 : Son los errores al cuadrado en los períodos pasados (rezagos), que depende de a_t .

Esta fórmula indica que la varianza condicional en el tiempo t se calcula como una función lineal de los errores pasados al cuadrado y un término constante α_0 .

Distribución de los Errores

Se asume que los errores a_t siguen una distribución normal con media 0 y varianza

σ_t^2 :

$$\epsilon_t \sim \mathcal{N}(0, \sigma_t^2)$$

- ϵ_t : Es el término de error estandarizado que sigue una distribución normal estándar, $a_t \sim \mathcal{N}(0, 1)$.
- σ_t : Es la desviación estándar condicional en el tiempo t , que mide la volatilidad condicional.

A partir de estos supuestos que surgen del proceso *ARCH*, es que los modelos *ARCH*(q) se forman como un modelo que se realiza a partir de la varianza condicional cuadrática.

Modelo GARCH (Generalized Autoregressive Conditional Heteroskedasticity)

Es una extensión del modelo ARCH que permite capturar de manera más precisa la volatilidad condicional en una serie temporal. Fue propuesto por [Bollerslev \(1986\)](#).

El modelo GARCH se define matemáticamente de la siguiente manera:

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2$$

Donde:

σ_t^2 es la varianza condicional en el tiempo t .

ω es el término de constante o intercepto.

α_i y β_j son coeficientes que representan la influencia de los errores pasados (ϵ_{t-i}^2) y las varianzas pasadas (σ_{t-j}^2) en la varianza condicional actual.

p y q son los órdenes del modelo, que indican cuántos términos de errores pasados y términos de varianzas pasadas se incluyen, respectivamente.

El modelo GARCH permite capturar la persistencia en la volatilidad, ya que los términos de varianzas pasadas (σ_{t-j}^2) influyen en la varianza condicional actual.

La estimación de los coeficientes α_i , β_j y ω en el modelo GARCH se realiza mediante métodos de estimación como la máxima verosimilitud. Una vez estimados, el modelo GARCH permite obtener pronósticos de la varianza condicional futura, lo que implica pronosticar la volatilidad futura de la serie temporal.

Métodos de estimación para el modelo ARCH y GARCH

El método de estimación de máxima verosimilitud para modelos ARCH y GARCH consiste en encontrar los valores de los parámetros que maximizan la función de verosimilitud condicional dada la información pasada de la serie. Este método parte del supuesto de que los errores o residuos de la serie siguen una distribución condicional normal, cuya varianza cambia a lo largo del tiempo según una estructura autorregresiva (en ARCH) o una combinación autorregresiva y de medias móviles

(en GARCH). La función de verosimilitud se construye sobre los residuos estandarizados y las varianzas condicionales estimadas, y se optimiza numéricamente, dado que no suele tener una solución analítica directa. El proceso de estimación se realiza generalmente mediante algoritmos iterativos como el de Newton-Raphson u otros algoritmos. Este método tiene como particularidad que incorpora la heterocedasticidad directamente en el modelo, permitiendo capturar la volatilidad variable a lo largo del tiempo. A través de este enfoque, se logra ajustar modelos ARCH y GARCH a datos financieros o económicos con mayor precisión, ya que la estimación por máxima verosimilitud proporciona estimadores consistentes, eficientes y asintóticamente normales bajo supuestos adecuados, garantizando así una representación realista de los procesos de varianza condicional (Wulff, 2017).

1. Modelo GARCH(1,1)

La estructura del modelo es:

$$\epsilon_t = a_t \sigma_t^2, \quad a_t \sim \mathcal{N}(0, 1)$$

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

2. Supuesto de distribución

Se asume que:

$$\epsilon_t \sim \mathcal{N}(0, \sigma_t^2)$$

Por lo tanto, la función de densidad condicional es:

$$f(\epsilon_t | \mathcal{F}_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{\epsilon_t^2}{2\sigma_t^2}\right)$$

3. Función de verosimilitud condicional

La log-verosimilitud condicional total es:

$$\log L(\omega, \alpha_1, \beta_1) = \sum_{t=2}^n \left[-\frac{1}{2} \log(2\pi\sigma_t^2) - \frac{\epsilon_t^2}{2\sigma_t^2} \right]$$

Donde:

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

4. Método de estimación

Se estiman los parámetros $(\omega, \alpha_1, \beta_1)$ maximizando la log-verosimilitud condicional. Para ello:

- Se inicializa σ_1^2 (por ejemplo, con la varianza muestral).
- Se calcula recursivamente σ_t^2 y ϵ_t .
- Se maximiza $\log L$ con algoritmos numéricos como BFGS o Newton-Raphson.
- **Paso 1: Inicialización.** Se escoge un valor inicial para σ_1^2 , que puede ser la varianza muestral de la serie de errores o un valor arbitrario positivo.
- **Paso 2: Cálculo recursivo.** Se calcula recursivamente la varianza condicional σ_t^2 para $t = 2, 3, \dots, n$, aplicando la ecuación del modelo:

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Donde ϵ_t son los residuos del modelo de medias (por ejemplo, residuos ARI-MA).

- **Paso 3: Sustitución en la log-verosimilitud.** Con los valores de σ_t^2 calculados, se evalúa la función de log-verosimilitud condicional:

$$\log L(\omega, \alpha_1, \beta_1) = \sum_{t=2}^n \left[-\frac{1}{2} \log(2\pi\sigma_t^2) - \frac{\epsilon_t^2}{2\sigma_t^2} \right]$$

- **Paso 4: Optimización.** Se maximiza numéricamente $\log L$ respecto a los parámetros ω , α_1 y β_1 , imponiendo las restricciones:

$$\omega > 0, \quad \alpha_1 \geq 0, \quad \beta_1 \geq 0, \quad \alpha_1 + \beta_1 < 1$$

Estas garantizan positividad y estacionariedad condicional del modelo.

Este proceso se implementa mediante lenguajes estadísticos como R, usando métodos de optimización iterativa.

2.2.5. El petróleo

¿Qué es el Petróleo?

El petróleo es una mezcla natural de hidrocarburo de origen orgánico, principalmente de restos de organismo marino y vegetales, que se han acumulado bajo tierra durante millones de años y han sido sometidos a procesos de presión y temperatura; es un recurso no renovable de gran importancia económica y energética. Es una mezcla natural de hidrocarburos de origen orgánico, principalmente de restos de organismos marinos y vegetales, que se han acumulado bajo tierra durante millones de años y han sido sometidos a procesos de gran importancia económica y energética.

2.3. Marco conceptual

- **Autocorrelación**, “ Correlación que existe entre una variable anterior uno o más periodos consigo misma” (Jenkins et al., 2011).
- **Coeficiente de correlación**, “Mide la magnitud de la relación lineal de dos variables entre sí” (Jenkins et al., 2011).
- **Componente cíclico**, “movimiento parecido a las ondas alrededor de la tendencia que se presenta” (Hanke, 2010).
- **Componente estacional**, “Patrón de cambios que se repite en periodos de tiempo con la misma dimensión” (Hanke, 2010).
- **Componente de tendencia**, “Componente de largo plazo que muestra el comportamiento creciente o decreciente en la serie de tiempo durante un periodo de tiempo amplio” (Hanke, 2010).
- **Autocorrelación simple**, “Representa la duración y la intensidad de la memoria de la serie de tiempo” (Mauricio, 2007).
- **Autoorrelación parcial**, “Representa la estimación mediante métodos como MCO (Mínimos Cuadrado Ordinarios) o MV(Máxima verosimilitud de una serie de tiempo) de los parámetros” (Mauricio, 2007).
- **Series de tiempo**, “Se llama Series de Tiempo a un conjunto de observaciones sobre valores que toma una variable (cuantitativa) en diferentes momentos del tiempo” (Ríos, 2008).
- **Procesos estocásticos**, “Un proceso estocástico es una secuencia de variables aleatorias, ordenadas y equidistantes cronológicamente, referidas a una (proceso univariante o escalar) o a varias (proceso multivariante o vectorial)

características de una unidad observable en diferentes momentos” ([Mauricio, 2007](#)).

- **Precio**, “Valor pecuniario en que se estima algo” ([Real Academia Española, 2023a](#)).
- **Petróleo**, “Líquido natural oleaginoso e inflamable, constituido por una mezcla de hidrocarburos, que se extrae de lechos geológicos continentales o marítimos y del que se obtienen productos utilizables con fines energéticos o industriales, como la gasolina, el queroseno o el gasóleo” ([Real Academia Española, 2023a](#)).

Capítulo 3

HIPÓTESIS Y VARIABLES

3.1. Hipótesis

3.1.1. Hipótesis general

El modelo $GARCH(2, 1)$ presenta mejor capacidad predictiva a la serie precio del petróleo frente a los modelos de la metodología de Box-Jenkins durante el período 2011-2023, región Cusco

3.1.2. Hipótesis específicas

- El modelo $ARIMA(2, 1, 1)$ presenta mejor ajuste a la serie precio del petróleo para el período enero 2011 a diciembre 2022, región Cusco
- El modelo $GARCH$ presenta mejor ajuste a la serie precio del petróleo para el período enero 2011 a diciembre 2022, región Cusco
- El enfoque de modelos heterocedásticos condicionados permite un pronóstico más eficiente a la serie precio del petróleo para el período de enero a diciembre

de 2023, región Cusco

3.2. Operacionalización de variables

Tabla 3.1:

Operacionalización de variables para el precio de petróleo en soles por galón 2011-2023

Variable	Definición conceptual	Definición operacional	Indicadores	Escala
Precio del petróleo	Es el valor monetario al cual se negocia el barril de petróleo en los mercados internacionales.(Santaella, 2021)	Es la cotización monetaria del crudo en los mercados internacionales	Precio promedio mensual	Razón
Tiempo	Es la magnitud que nos ayuda a ordenar sucesos, estableciendo un pasado, un presente y un futuro.(Real Academia Española, 2023b)	Es el intervalo entre las observaciones consecutivas	Periodo de tiempo en meses	Razón

Capítulo 4

METODOLOGÍA

4.1. Tipo de investigación

"El tipo de investigación es aplicada, aprovecha los conocimientos logrados por la investigación básica o teórica para el conocimiento y solución de problemas inmediatos" ([Sánchez Carlessi et al., 2018](#)).

El tipo de investigación desarrollada fue aplicado, ya que se utilizó el conocimiento obtenido a partir de investigaciones teóricas y básicas para abordar un problema específico y real, como lo es la predicción del precio del petróleo en la región de Cusco, 2011-2023.

4.2. Enfoque de la investigación

Según [Hernández et al. \(2014\)](#), el enfoque cuantitativo se basa en la recolección y análisis de datos numéricos para describir y explicar fenómenos. En esta investigación, se adoptó este enfoque porque permitió medir de manera precisa las variables

relacionadas con el precio del petróleo en la región de Cusco, 2011-2023. El uso de métodos estadísticos avanzados, como los modelos heterocedásticos condicionales y la metodología Box-Jenkins, facilitó el establecimiento de relaciones causales y la generalización de los resultados, asegurando una predicción confiable del comportamiento futuro.

4.3. Alcance de la investigación

La investigación que se llevó a cabo para el presente estudio, es descriptiva. Según [Hernández et al. \(2014\)](#), este busca especificar propiedades y características importantes de cualquier fenómeno que se analice. Describe tendencias de un grupo o población. Se aplicó este enfoque porque permitió medir de manera precisa las variaciones del precio del petróleo en la región de Cusco, utilizando modelos econométricos como los heterocedásticos condicionales y la metodología Box-Jenkins. Además, se prepararon relaciones estadísticas entre las variables estudiadas para generar predicciones confiables.

4.4. Diseño de investigación

El diseño de esta investigación es no experimental de corte longitudinal, ya que el investigador no interviene de manera directa en la manipulación de variables. Se analizan las variaciones del precio del petróleo en la región de Cusco, 2011-2023, utilizando series de tiempo para realizar predicciones ([Hernández et al., 2014](#)).

4.5. Población de estudio

La población de este estudio es el precio del petróleo en la región Cusco 2011-2023.

4.6. Técnicas de recolección de datos

La data se recopiló mediante una solicitud que fue enviada al Instituto Nacional de Estadística e Informática (INEI) y OSINERMIG sobre el Precio del Petróleo, para así poder hacer una comparación y validación de los datos, dichos datos fueron solicitados para la región Cusco 2011-2023.

4.7. Técnicas para demostrar la verdad o falsedad de las hipótesis planteadas

- **Prueba de normalidad de Kolmogorov-Smirnov** Es una prueba no paramétrica, la cual permite verificar si los residuos siguen o no una distribución normal.
 - Ho: Los residuos proceden de una distribución normal
 - H1: Los residuos no proceden de una distribución normal

El estadístico de dicha prueba está dado por, sin dejar de lado que la prueba está trabajando con un nivel de significancia de $\alpha = 0.05$:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \begin{cases} 1, & \text{si } y_n < x \\ 0, & \text{en caso contrario} \end{cases} \quad (4.1)$$

De donde: x : las observaciones de la muestra

- **Prueba de Mann Kendall** Se utiliza una prueba de tendencia del tipo Mann-Kendall para determinar si existe o no una tendencia en los datos de la serie de tiempo. Es una prueba no paramétrica, lo que significa que no se realiza ninguna suposición subyacente a la normalidad. Se calcula la prueba de Mann-Kendall para detectar una tendencia ascendente o descendente en los datos recogidos en el tiempo. Se realiza un ajuste para las observaciones empatadas en esta prueba no paramétrica. Debe tener por lo menos 10 observaciones para que la aproximación a la normal sea apropiada
 - Ho: No hay tendencia presente en los residuos
 - H1: Existe tendencia presente en los residuos

$$S_s = \sum_{i=2}^n \sum_{j=1}^{i-1} \text{signo}(x_i - x_j) \quad (4.2)$$

$$\text{Var}(S_s) = \frac{n(n-1)(2n+5)}{18} \quad (4.3)$$

$$Z_s = \frac{S_s + m_c}{\sqrt{\text{Var}(S_s)}} \quad (4.4)$$

- **Prueba de Kruskal**

El contraste de [Kruskal and Wallis \(1952\)](#) es un análisis basado en rangos, pero en este caso el rango es el conjunto de la serie de datos. Este contraste se realiza mediante una frecuencia específica, en este caso mensual y la Hipótesis nula que se quiere contrastar es la misma que en el contraste anterior.

- H_0 : Ausencia de estacionalidad
- H_1 : No hay ausencia de estacionalidad

El estadístico de contraste con un nivel de significancia $\alpha = 0.05$ está dado por:

$$K = \frac{12}{N(N+1)} \sum_{i=1}^g \left(\frac{R_i^2}{n_i} - 3(N+1) \right) \quad (4.5)$$

De donde:

N : Cantidad total de datos

R : Es la suma de los rangos

n : Es el tiempo (años, meses) de la serie

g : El número de rango

■ Prueba de independencia LJUNG-BOX

Es un test muy utilizado (en series de tiempo) para contrastar la hipótesis de independencia. Se contrasta la hipótesis nula de que las primeras “m” autocorrelaciones son cero:

- $H_0 : \rho_1 = \rho_2 = \dots = \rho_n = 0$
- $H_1 : \rho_1 \neq 0$

El estadístico de contraste está dado por:

$$Q = n(n+2) \sum_{k=1}^m \frac{r_k^2}{n-k} \sim X_{(m-1)}^2 \quad (4.6)$$

Se rechaza H_0 , si el valor observado es grande $Q \geq X_{(m-1, 1-\alpha)}^2$

■ Prueba de Raíces Unitarias o Dicker-Fuller

Esta prueba indica si la serie es estacionaria o no. (Arce and Mahía, 2003).

$$\Delta Y_t = \alpha_0 + \gamma Y_{t-1} \sum_{i=1}^p \beta_i \Delta y_{t-i+1} + a_t$$

Donde:

$$\gamma = \sum_{i=1}^p \alpha_i - 1$$

$$\beta_i = \sum_{j=2}^p \alpha_j$$

Las hipótesis estadísticas de la prueba son:

- $H_0 : \gamma = 0$ (la serie tiene una raíz unitaria no estacionaria)
- $H_1 : \gamma < 0$ (la serie no tiene una raíz unitaria estacionaria)

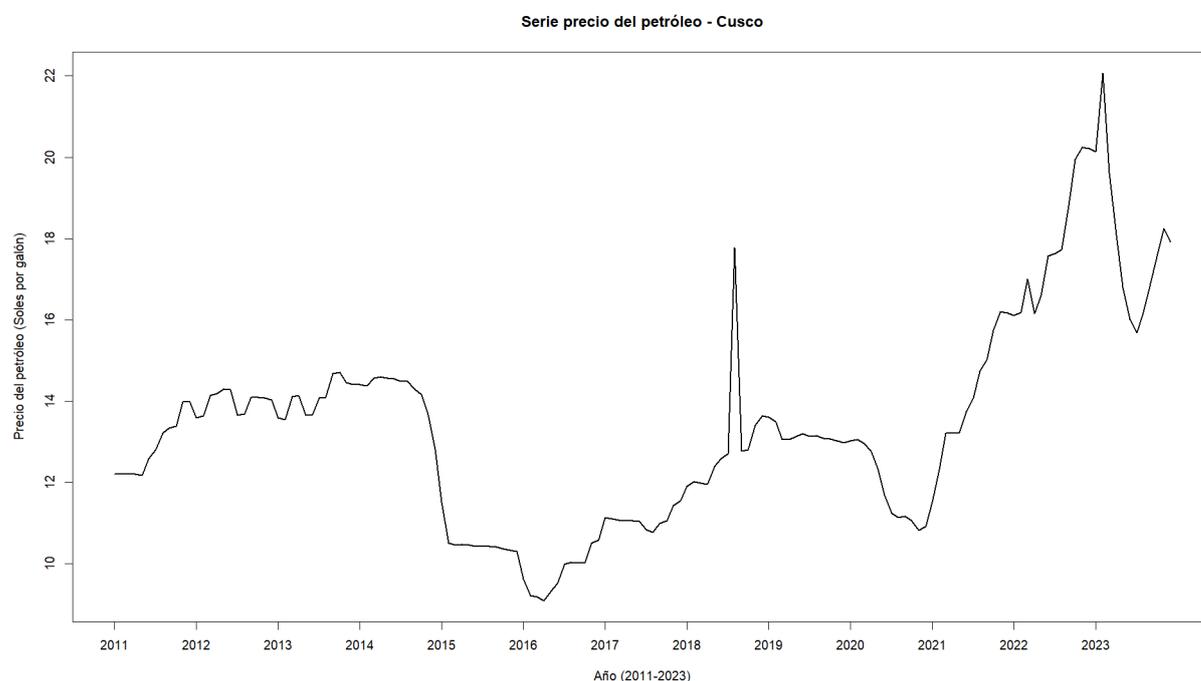
Capítulo 5

RESULTADOS

5.1. Análisis y descomposición clásica de la serie de tiempo

El análisis de series de tiempo permite examinar la evolución de un fenómeno a lo largo del tiempo, identificando patrones como tendencias, estacionalidad y variabilidad residual. Se presenta la serie temporal del precio del petróleo en soles por galón en la región de Cusco 2011-2023, seguida de su análisis de descomposición clásica.

Figura 5.1:
Serie temporal del precio del petróleo en soles por galón en la región de Cusco 2011-2023



Interpretación

La **Figura 5.1** presenta la evolución del precio del petróleo en soles por galón en la región de Cusco entre enero de 2011 y diciembre de 2023. A lo largo de este periodo, la serie muestra variaciones importantes asociadas a hechos económicos, geopolíticos y sociales ocurridos a nivel internacional, nacional y regional. Entre los años 2011 y mediados de 2014, el precio del petróleo en Cusco mantiene una trayectoria relativamente estable con ligeros incrementos. Este comportamiento coincide con un contexto internacional caracterizado por precios elevados del crudo debido a la demanda sostenida en economías industrializadas y emergentes, así como a la estabilidad relativa en la producción de la Organización de Países Exportadores de Petróleo (OPEP). A partir del segundo semestre de 2014, se observa una disminución pronunciada en el precio del petróleo. Este descenso corresponde

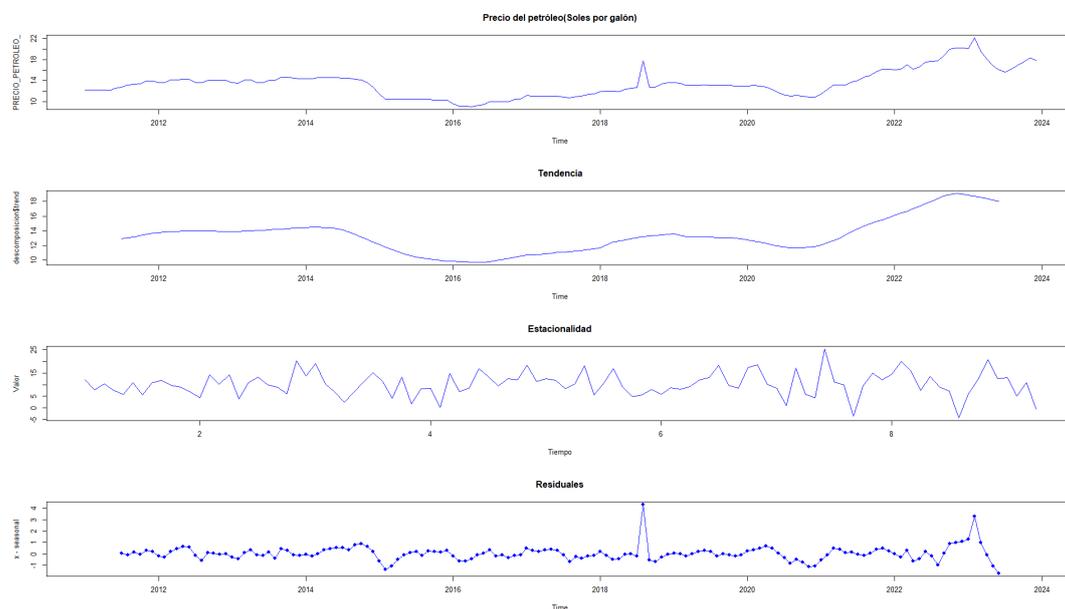
a la crisis petrolera mundial iniciada en 2014, cuando la OPEP decidió no reducir su producción ante el aumento de la extracción de petróleo no convencional (petróleo de esquisto) en Estados Unidos, lo que provocó una sobreoferta en el mercado global. Este fenómeno generó una reducción de los precios internacionales del crudo, cuyo impacto se refleja directamente en la serie de precios en Cusco.

Durante el periodo 2016-2018, el precio del petróleo en la región muestra una fase de recuperación moderada. En 2018 se identifica un incremento puntual con una caída posterior, lo cual corresponde con la incertidumbre generada por las tensiones comerciales entre Estados Unidos y China iniciadas en dicho año, lo que provocó fluctuaciones en los precios internacionales del petróleo y sus derivados. En el año 2020 se presenta la caída más pronunciada en la serie, correspondiente a la crisis sanitaria global por la COVID-19. La pandemia generó una contracción abrupta de la demanda mundial de petróleo debido a las restricciones de movilidad, suspensión de vuelos, paralización del transporte y de la actividad económica. En el Perú, se decretó el estado de emergencia sanitaria y la cuarentena nacional obligatoria desde marzo de 2020, lo que redujo drásticamente el consumo de combustibles. En la región Cusco, la paralización del sector turismo, transporte terrestre e infraestructura tuvo efectos significativos en el mercado local de hidrocarburos. Posteriormente, entre los años 2021 y 2022, se evidencia un incremento sostenido del precio del petróleo, en concordancia con la reactivación progresiva de la economía global. Este incremento se acentúa a partir del primer trimestre de 2022, cuando se produce el conflicto bélico entre Rusia y Ucrania, iniciado en febrero de dicho año. Rusia, como uno de los principales exportadores mundiales de petróleo y gas, enfrenta sanciones económicas impuestas por países de la

Unión Europea y Estados Unidos, lo que repercute en el abastecimiento global de hidrocarburos, generando aumentos significativos en los precios internacionales.

Durante 2022 y 2023, se presentan fluctuaciones relevantes en la serie, influenciadas por factores internos en el contexto peruano. En diciembre de 2022 se produce la vacancia presidencial de Pedro Castillo y la asunción de Dina Boluarte, lo que genera una crisis política acompañada de protestas sociales, bloqueos de carreteras y conflictos en diversas regiones del país, incluyendo Cusco. Estos eventos interrumpen el transporte de bienes, incluidos los combustibles, afectando los precios por escasez y costos logísticos elevados. En la segunda mitad del año 2023, el precio del petróleo en Cusco experimenta una disminución seguida de una recuperación parcial. Este comportamiento ocurre en un escenario de desaceleración económica global, marcado por el endurecimiento de las políticas monetarias en países como Estados Unidos, lo cual reduce la actividad económica y, con ello, la demanda de combustibles. A nivel regional, Cusco atraviesa aún las consecuencias económicas de los bloqueos y conflictos sociales del año anterior, afectando el comportamiento de la demanda local de petróleo.

Figura 5.2:
Descomposición clásica de la serie precio del petróleo en soles por galón en la región de Cusco 2011-2023



Interpretación

En la **Figura 5.2** se muestran los componentes de la serie. En primer lugar, se observa una tendencia ligeramente ascendente a partir del año 2022. Entre 2014 y 2021, se evidencia una alternancia entre tendencias decrecientes y crecientes en ciertos meses de esos años. No se observa estacionalidad en la serie, y se destaca un pico elevado a finales del año 2018 en los precios del petróleo dentro del componente de variabilidad residual.

5.2. Análisis descriptivo de la serie precio del petróleo

En la **Tabla 5.1** se presentan los estadísticos descriptivos correspondientes a la serie del precio del petróleo correspondiente a la data de entrenamiento de enero 2011 a

diciembre del 2022.

Tabla 5.1:
Análisis descriptivo de la serie precio del petróleo (soles por galón) 2011-2022

Estadísticos descriptivos	Valores
Total de observaciones	144
Precio mínimo (soles)	9.09
Precio máximo (soles)	20.24
Tasa mensual promedio	0.2468 %
Tasa anual promedio	3.0017 %
Primer cuartil (25 %)	11.1375
Segundo cuartil (50 %)	13.075
Tercer cuartil (75 %)	14.14
Rango intercuartil	3.0025

Nota. *Elaboración propia a partir de los datos obtenidos en la INEI*

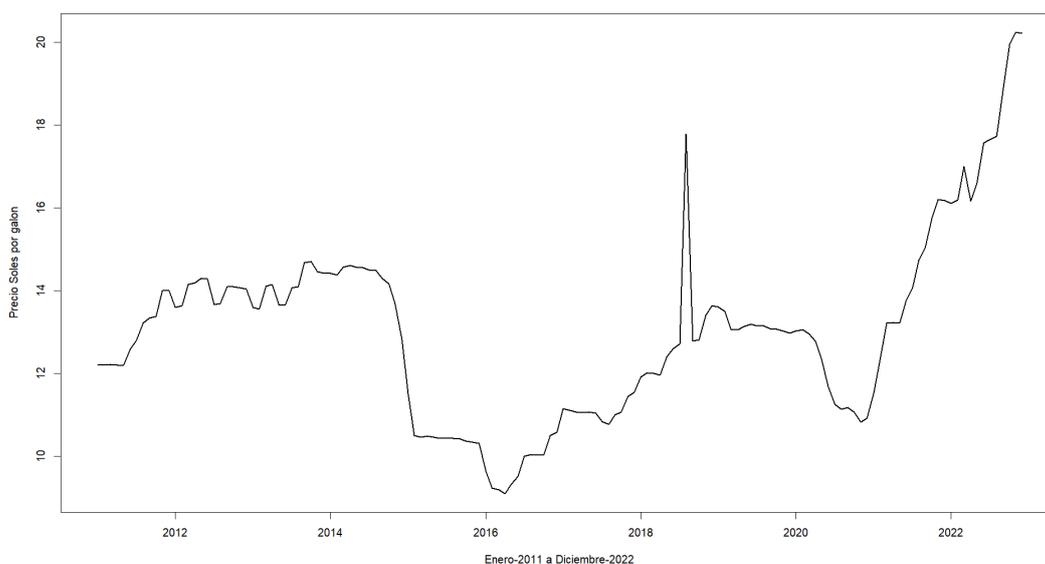
Interpretación

En la Tabla 5.1 se presentan 144 observaciones correspondientes a los precios del petróleo (expresados en soles por galón), recopilados por el INEI para la región del Cusco entre los años 2011 y 2022. El precio mínimo registrado fue de 9.09 soles, mientras que el máximo alcanzó los 20.24 soles, mostrando una considerable amplitud en la variación de precios durante el periodo. La tasa mensual promedio de variación fue de 0.2468 %, lo que equivale a una tasa anual promedio acumulada de 3.0017 %, indicando un crecimiento moderado y sostenido del precio del petróleo a lo largo de los años. Respecto a los valores cuartílicos, el primer cuartil (Q1) se ubicó en 11.1375 soles, lo cual significa que el 25 % de los precios se encontraron por debajo de este valor. El segundo cuartil (Q2) o mediana fue de 13.075 soles, representando el valor central de la distribución: el 50 % de las observaciones estuvieron por debajo de este precio y el otro 50 % por encima. Por su parte, el tercer cuartil (Q3) fue de 14.14 soles, indicando que el 75 % de los precios estuvieron por debajo o igual a ese valor, mientras que el 25 % restante se ubicó por encima. El rango intercuartílico (RIC), calculado como la diferencia entre el tercer y el

primer cuartil ($Q3 - Q1 = 3.0025$ soles), representa la dispersión del 50 % central de los datos. Este valor permite visualizar la concentración de precios en torno a la mediana, evidenciando que la mayoría de los precios se mantuvieron dentro de un intervalo moderado de variabilidad.

Figura 5.3:

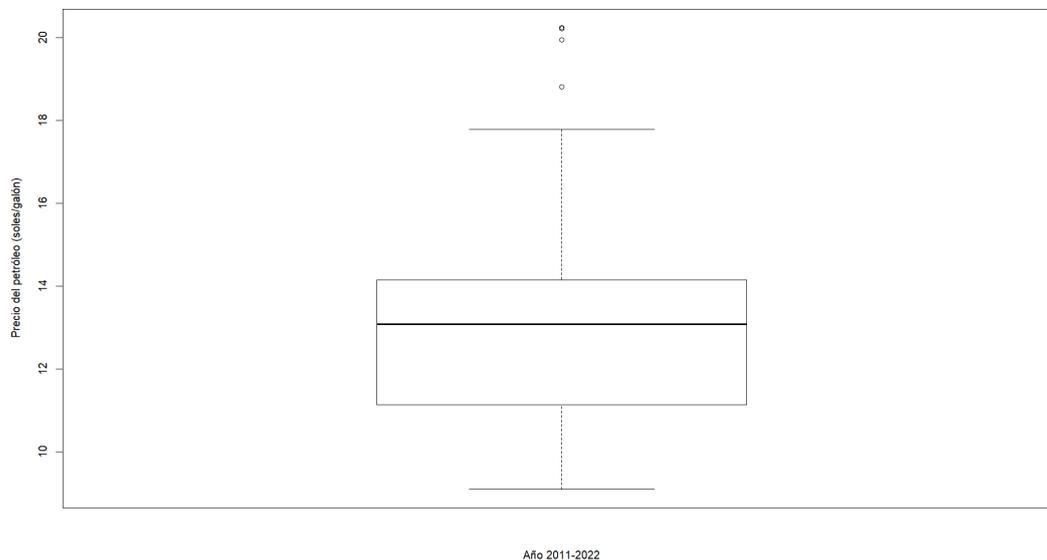
Representación de la serie temporal del precio del petróleo en soles por galón 2011-2022



Interpretación

En la **Figura 5.3** se observa los datos por cada mes desde el año 2011 hasta el año 2022; se detecta un tendencia creciente desde mediados del 2019, pero hacia años atrás vemos que los datos son muy variables y también aparentemente datos excesivamente altos y bajo, estos datos atípicos son pocos a través de toda la serie, mediante un análisis clásico de series de tiempo, se observa el componente de tendencia, un componente estacional en ciertas épocas y una variabilidad irregular.

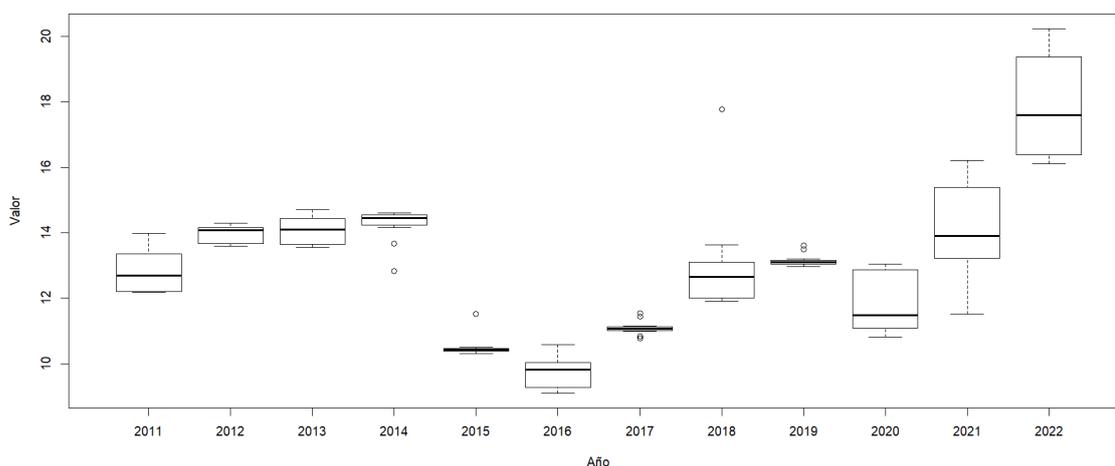
Figura 5.4:
Gráfico de *boxplot* de la data de entrenamiento del precio de petróleo en (soles por galón)



En la **Figura 5.4** se presenta el boxplot de los precios de petróleo en soles por galón, basado en la data de entrenamiento. El análisis de este gráfico proporciona una visión clara de la distribución de los precios. La mediana, representada por la línea horizontal dentro de la caja, se sitúa en aproximadamente 14 soles por galón, lo que indica el valor central de la distribución de los precios. El rango intercuartil abarca desde el primer cuartil (Q1) hasta el tercer cuartil (Q3), se extiende desde cerca a 12 hasta 16 soles por galón. Este rango encapsula el 50% central de los datos y demuestra la concentración más densa de los precios del petróleo. Los bigotes del boxplot se extienden más allá de la caja principal, llegando hasta los valores máximos y mínimos. Adicionalmente, se observan varios puntos por encima del bigote superior. Estos puntos representan valores atípicos, que son precios significativamente más altos en comparación con la mayoría de los datos. La presencia de estos valores atípicos sugiere que, aunque la mayoría de los precios

se encuentran dentro del rango normal identificado, existen ciertos casos donde los precios del petróleo pueden exceder notablemente lo usual.

Figura 5.5:
Gráfico de box plot para la serie de entrenamiento del precio de petróleo en (soles por galón)



Interpretación

En la **Figura 5.5** muestran los precios del petróleo en soles por galón para cada año desde 2011 hasta 2022 en la Región Cusco, se observa un incremento gradual y sostenido en los precios del petróleo, especialmente notorio a partir del 2020. Esto refleja una tendencia de crecimiento en los costos del petróleo en la región. La variabilidad y volatilidad de los precios también es evidente, con los años más recientes como el 2021 y 2022 mostrando cajas más grandes en el gráfico, indicando una mayor dispersión de los precios. Esto sugiere que los precios han sido más volátiles en estos años. Además, la presencia de valores atípicos en varios años destaca episodios de extrema volatilidad, donde los precios han sido excepcionalmente altos o bajos. Estos valores atípicos podrían estar relacionados con eventos específicos que impactaron el mercado del petróleo, como cambios en la política económica, conflictos geopolíticos o fluctuaciones en la demanda global.

5.3. Resultados para el objetivo específico 1

Para alcanzar el primer objetivo específico, que consiste en identificar el modelo de serie temporal con la metodología Box-Jenkins que mejor se ajusta a la serie de tiempo del precio del petróleo, se emplea la técnica de particionamiento de datos en conjuntos de entrenamiento y prueba. Esta estrategia permite evaluar los modelos en función de su capacidad de predicción. Los datos de entrenamiento corresponden al período 2011-2022, mientras que los datos de prueba abarcan desde enero de 2023 hasta diciembre de 2023. A continuación, se presenta el desarrollo detallado de la metodología Box-Jenkins.

5.3.1. Identificación de componentes de la serie de precio de petróleo

Paso a paso se menciona las pruebas indicadas a continuación se muestra la primera prueba:

Prueba de tendencia

H_0 : No hay tendencia presente en los datos.

H_1 : Existe tendencia presente en los datos.

Tabla 5.2:

Resultados del test de tendencia Mann-Kendall para el precio del petróleo en soles por galón

Estadísticas del test	Valores
z	2.1972
n	144
p-valor	0.02801

Nota. *Elaboración propia utilizando el software R-studio*

Interpretación

En la **Tabla 5.2** se muestra los resultados del test de Mann-Kendall arrojan un valor de $z = 2.1972$ y un p -valor de 0.02801. Dado que el p -valor es menor al nivel de significancia establecido de 0.05, se rechaza la hipótesis nula (H_0), que indica la ausencia de tendencia en los datos. Por lo tanto, se apoya la hipótesis alternativa (H_1), la cual sugiere la existencia de una tendencia en los datos del precio del petróleo en soles por galón para el período analizado.

Prueba de estacionalidad

H_0 : Ausencia de estacionalidad según la prueba de rangos de Kruskal-Wallis.

H_1 : Existe estacionalidad según la prueba de rangos de Kruskal-Wallis.

Tabla 5.3:

Resultados del test de Kruskal-Wallis para el precio del petróleo en soles por galón

Estadísticas del test	Valores
chi-cuadrado	143
n	143
p-valor	0.4843

Nota. *Elaboración propia utilizando el software R-studio*

Interpretación

En la **Tabla 5.3** se muestra los resultados del test de Kruskal-Wallis muestran un valor de chi-cuadrado de 143 con un p -valor de 0.4843. Dado que el p -valor es significativamente mayor que el nivel de significancia (0.05), no hay evidencia suficiente para rechazar la hipótesis nula (H_0). Por lo tanto, no se observa estacionalidad en el precio del petróleo en soles por galón, lo que sugiere que las diferencias en los precios no están asociadas con un patrón estacional significativo.

Prueba de heterocedasticidad

A continuación se contrastan las siguientes hipótesis:

H_0 : La varianza es constante (homocedasticidad) de acuerdo con la prueba de Breusch-Pagan.

H_1 : La varianza no es constante (heterocedasticidad) de acuerdo con la prueba de Breusch-Pagan.

Tabla 5.4:

Resultados del test de Breusch-Pagan para el precio del petróleo en soles por galón

Estadístico	Valor
Estadístico de Breusch-Pagan	14.375
Grados de libertad (df)	1
p-valor	0.0001498

Nota. *Elaboración propia utilizando el software R-studio*

Interpretación

En la **Tabla 5.4** se muestra los resultados del test de Breusch-Pagan indican un estadístico de 14.375 con un p-valor de 0.0001498. Dado que el p-valor es considerablemente menor que el nivel de significancia 0.05, se rechaza la hipótesis nula (H_0) que sugiere la ausencia de heterocedasticidad. Por lo tanto, hay evidencia suficiente para concluir que existe heterocedasticidad en el precio del petróleo en soles por galón, indicando que la varianza de los errores no es constante a lo largo de las observaciones. Además de haber identificado las componentes de la serie a continuación se evalúa la normalidad de serie.

Prueba de de normalidad

A continuación se contrastan las siguientes hipótesis:

H_0 : Los datos siguen una distribución normal de acuerdo con la prueba de Kolmogorov-Smirnov.

H_1 : Los datos no siguen una distribución normal de acuerdo con la prueba de Kolmogorov-Smirnov.

Tabla 5.5:
Resultados de la prueba de normalidad de Lilliefors (Kolmogorov-Smirnov) para el precio del petróleo

Estadístico	Valor
Estadístico de Kolmogorov	0.10437
p-valor	0.5867

Nota. *Elaboración propia utilizando el software R-studio*

Interpretación

En la **Tabla 5.5** se muestra la prueba de normalidad de Lilliefors (Kolmogorov-Smirnov) aplicada a los datos de precios del petróleo, el valor obtenido es 0.10437 y el p-valor es 0.5867. Dado que el p-valor es mayor que el umbral comúnmente utilizado de 0.05, no se rechaza la hipótesis nula de que los datos siguen una distribución normal. Por lo tanto, podemos inferir que los precios del petróleo en esta serie de datos no presentan una desviación significativa de la normalidad, sugiriendo que la distribución de precios es aproximadamente normal.

En resumen, los resultados sugieren que, aunque existe una tendencia significativa en los precios del petróleo, los datos no presentan estacionalidad, existe presencia de heterocedasticidad y presentan normalidad los datos.

5.3.2. Metodología de Box-Jenkins para la serie precio del petróleo

A continuación se procede a aplicar la metodología de Box-Jenkins para identificar el modelo que mejor se ajuste a la serie precio del petróleo.

5.3.2.1. Estacionariedad de la serie de entrenamiento precio del petróleo

En este apartado se realiza en contraste de hipótesis de la Dickey-Fuller Aumentada (ADF):

H_0 : La serie es no estacionaria.

H_1 : La serie es estacionaria.

Tabla 5.6: *Resultados de la prueba de Dickey-Fuller para el precio del petróleo en soles por galón*

Estadístico	Valor
Dickey-Fuller	-0.28437
Lag order	5
p-valor	0.99

Nota. *Elaboración propia utilizando el software R-studio*

Interpretación

En la **Tabla 5.6** se muestra la prueba de Dickey-Fuller aplicada a los datos del precio del petróleo, el estadístico Dickey-Fuller obtenido es de -0.28437, con un orden de rezago de 5 y un p-valor de 0.99. Dado que el p-valor es significativamente mayor que el umbral comúnmente utilizado de 0.05, no se rechaza la hipótesis nula de que la serie de precios del petróleo tiene una raíz unitaria. Esto sugiere que los precios no son estacionarios en su nivel original.

A partir de este hallazgo y por la presencia de la tendencia en la serie, la metodología de Box-Jenkins indica que se debe aplicar una diferenciación en su componente regular de la serie precio de petróleo.

5.3.2.2. Estacionariedad de la serie de entrenamiento diferenciada precio del petróleo

En este apartado se realiza en contraste de hipótesis de la Dickey-Fuller Aumentada (ADF) para la serie diferenciada del precio del petróleo:

H_0 : La serie es no estacionaria.

H_1 : La serie es estacionaria.

Tabla 5.7: Resultados de la prueba de Augmented Dickey-Fuller para la serie diferenciada del precio del petróleo

Estadístico	Valor
Dickey-Fuller	-3.8536
Lag order	5
p-valor	0.01842

Nota. Elaboración propia utilizando el software R-studio

Interpretación

En la **Tabla 5.7** En cuanto a la prueba de Augmented Dickey-Fuller aplicada a la serie diferenciada del precio del petróleo, el estadístico Dickey-Fuller obtenido es de -3.8536, con un orden de rezago de 5 y un p-valor de 0.01842. Dado que el p-valor es menor que el umbral comúnmente utilizado de 0.05, se rechaza la hipótesis nula de que la serie diferenciada tiene una raíz unitaria. Esto sugiere que la serie diferenciada de precios del petróleo es estacionaria.

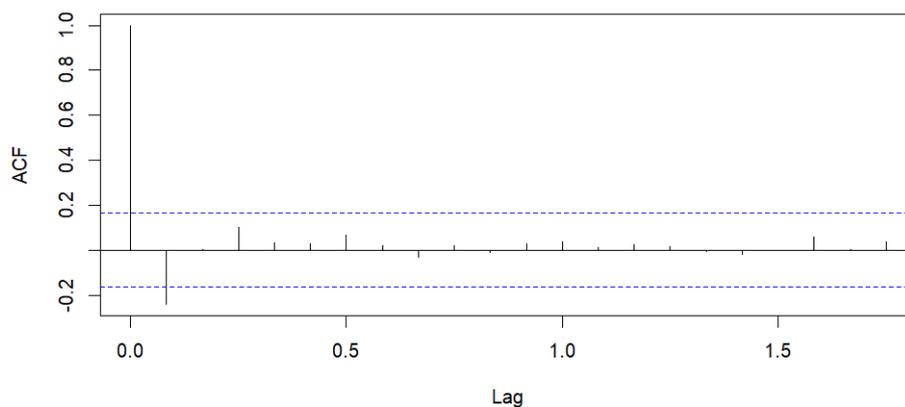
5.3.2.3. Identificación de los modelos

Para identificar los modelos se considera necesario los gráficos de autocorrelación simple (ACP) y autocorrelación parcial (PACF) estos gráficos nos ayudan a encontrar los lags o rezagos para identificar posibles términos AR (autoregresivos) y

MA (medias móviles) en el modelo ARIMA a continuación se muestra el gráfico de autocorrelación simple.

Figura 5.6:

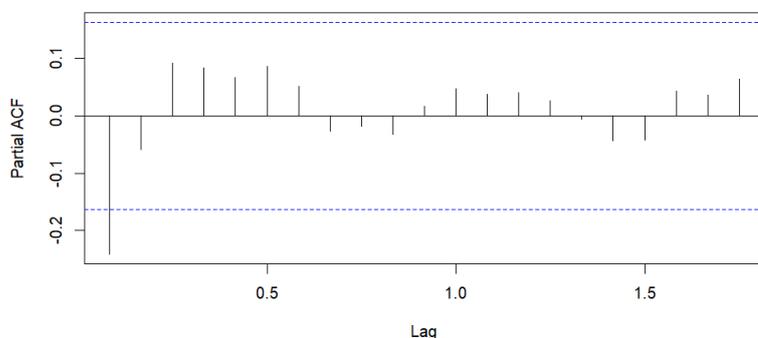
Gráfico de autocorrelación (ACP) para la serie de entrenamiento (soles por galón)



Interpretación

En la **Figura 5.6** se observa los lags o rezagos en el gráfico de autocorrelación, en el rezago 0 la autocorrelación es siempre 1 esto se debe a que está correlacionando la serie consigo misma en el mismo período, en cuanto al rezago 1, la autocorrelación es -0.242. Esto indica una correlación negativa entre la observación actual y la observación 1 período atrás. Sin embargo, es importante tener en cuenta que el valor es relativamente pequeño. Los valores de autocorrelación en otros rezagos son bastante pequeños y oscilan alrededor de cero, lo que sugiere una autocorrelación débil o insignificante en esos rezagos. En cuanto a la significancia, generalmente se considera que los rezagos con valores cuya magnitud es mayor a 0.2 o -0.2 son estadísticamente significativos. Basándonos en eso, podríamos decir que el rezago más significativo en tu serie diferenciada es el rezago 1 debido a su autocorrelación relativamente alta en términos de magnitud, aunque es negativa.

Figura 5.7:
Gráfico de autocorrelación parcial (PACP) para la serie de entrenamiento (soles por galón)



Interpretación

En la **Figura 5.7** se observa que el valor de autocorrelación parcial en el rezago 1 es -0.242, lo que sugiere una correlación negativa entre la observación actual y la observación un período atrás. Luego de observar los rezagos posibles a utilizar se consideró varios modelos ARIMA(p,d,q) de los cuales según el criterio se eligió al que tenía menor AIC: este modelo no solo se ajusta bien a los datos históricos, Además, es capaz de captar tanto las tendencias a largo plazo como las fluctuaciones a corto plazo. El componente autorregresivo de orden 1; indica que los precios del petróleo están fuertemente influenciados por los precios del período anterior, lo que refleja una inercia significativa en la serie temporal. Esto sugiere que los choques en el precio del petróleo tienden a tener efectos persistentes.

5.3.2.4. Selección de modelos de series de tiempo

En la tabla 5.8 se muestra resultados de diferentes modelos de series de tiempo para elegir el modelo que mejor se ajusta a los datos según los criterios AIC, BIC y MAPE.

Tabla 5.8:
Valores del criterio de información Akaike (AIC), Bayesiano (BIC) y MAPE para diferentes modelos ARIMA

Modelo	AIC	BIC	MAPE
ARIMA(1,1,2)	298.6676	305.3701	2.518
ARIMA(1,1,0)	299.3068	309.7012	2.542
ARIMA(0,1,1)	299.4444	312.3441	2.586
ARIMA(1,1,3)	300.0901	310.0304	2.528
ARIMA(2,1,2)	300.1367	310.5190	2.614
ARIMA(0,1,3)	300.4927	314.9044	2.615
ARIMA(0,1,2)	300.8126	314.4550	2.536
ARIMA(1,1,1)	301.1418	314.9509	2.609
ARIMA(2,1,3)	302.0775	319.8545	2.619
ARIMA(2,1,1)	302.6036	305.2325	2.515

Nota. Elaboración propia utilizando el software R-Studio

Interpretación

En la **Tabla 5.8** se muestra los valores del criterio de información Akaike (AIC), Bayesiano (BIC) y el Error Porcentual Absoluto Medio (MAPE) para diferentes modelos ARIMA. El modelo ARIMA(1,1,2) destaca por tener el valor más bajo de AIC (298.6676) y el segundo más bajo de BIC (305.3701), lo que indica un buen equilibrio entre ajuste y complejidad. Su MAPE de 2.518 lo posiciona como el segundo más bajo entre los modelos evaluados, sugiriendo una precisión razonable en las predicciones. Por otro lado, el modelo ARIMA(2,1,1) exhibe el segundo menor AIC (302.6036) y el BIC más bajo (305.2325), junto con el MAPE más bajo (2.515). Esto implica que, a pesar de su mayor complejidad, este modelo podría ofrecer una capacidad predictiva, lo cual es fundamental para la interpretación de los resultados.

Al tener en consideración el AIC, BIC y MAPE se evalúa ambos modelos debido a que presentan mejor ajuste para la serie de entrenamiento.

5.3.2.5. Estimación de parámetros del modelo ARIMA(1, 1, 2)

Una vez encontrado el modelo adecuado según el criterio del AIC y BIC entre los seleccionados pasamos a encontrar los parámetros del modelo ARIMA(1,1,2); a continuación se muestra la siguiente tabla:

Tabla 5.9:

Parámetros estimado del modelo ARIMA(1, 1, 2) para la serie precio del petróleo

Parámetro	Valor
ϕ_1 (ar1)	0.8738
θ_1 (ma1)	-1.156
θ_2 (ma2)	0.3246
σ^2	0.4463
Log likelihood	-145.3
AIC	298.67

Nota. *Elaboración propia utilizando el software R-studio*

Apartir de la **Tabla 5.9** se construye el modelo ARIMA(1,1,2) definido matemáticamente de la siguiente manera:

$$\Delta^1 Y_t = 0.8738 Y_{t-1} + a_t - 1.156 \cdot a_{t-1} + 0.3246 \cdot a_{t-2}$$

donde:

- $\Delta^1 Y_t$ = Es la diferenciación de primer orden de la serie temporal
- $\phi_1 = 0.8738$ es el parámetro autoregresivo de orden 1
- $\theta_1 = -1.156$ y $\theta_2 = 0.3246$ son los parámetros de media móvil de orden 1 y 2, respectivamente
- a_t es el término de error blanco con varianza σ^2

Interpretación

- El coeficiente ϕ_1 (ar1); es positivo (0.8738), lo que sugiere una correlación positiva entre el precio del petróleo en un período y su valor en el período

previo, lo que indica cierta tendencia a seguir una trayectoria similar. Esto significa que si el precio del petróleo ha estado aumentando en los últimos períodos, es más probable que continúe aumentando en el siguiente período

- Los coeficientes θ_1 (ma1) y θ_2 (ma2) de los términos de media móvil; resultan ser negativos (-1.1560 y 0.3246 respectivamente), Esto significa que los errores en las predicciones pasadas tienen un impacto negativo en las predicciones actuales. Si en el mes pasado hubo una sobreestimación en las predicciones del precio del petróleo, este coeficiente negativo indica que es probable que en el próximo mes las predicciones se ajusten hacia abajo para corregir este error, es decir, se espera una corrección negativa en las predicciones actuales basadas en los errores pasados. .
- $\sigma^2 = 0.4463$; esto significa que, en promedio, el precio del petróleo fluctúa alrededor de su tendencia en aproximadamente 0.4463 soles por galón.

5.3.2.6. Estimación de parámetros del modelo ARIMA(2, 1, 1)

Una vez encontrado el modelo adecuado según el criterio del menor MAPE entre los seleccionados pasamos a encontrar los parámetros del modelo ARIMA(2, 1, 1); a continuación se muestra la **Tabla 5.10**:

Tabla 5.10:

Parámetros estimados del modelo ARIMA(2, 1, 1) para la serie entrenamiento precio del petróleo

Parámetro	Valor
ϕ_1 (ar1)	-0.6948
ϕ_2 (ar2)	-0.1662
θ_1 (ma1)	0.4499
σ^2	0.4592
Log likelihood	-147.3
AIC	302.6

Nota. *Elaboración propia utilizando el software R-studio*

El modelo ARIMA(2,1,1) se define matemáticamente de la siguiente manera:

$$\Delta^1 Y_t = -0.6948 \cdot Y_{t-1} - 0.1662 \cdot Y_{t-2} + a_t - 0.4499 \cdot a_{t-1}$$

donde:

- $\Delta^1 Y_t$ es la diferenciación de primer orden de la serie temporal.
- $\phi_1 = -0.6948$ y $\phi_2 = -0.1662$ son los parámetros autoregresivos de orden 1 y 2, respectivamente.
- $\theta_1 = 0.4499$ es el parámetro de media móvil de orden 1.
- a_t es el término de error blanco con varianza σ^2 .

Interpretación

- Los coeficientes $\phi_1 = -0.6948$ y $\phi_2 = -0.1662$ indican una relación negativa entre los valores actuales del precio del petróleo y los valores de los dos períodos anteriores, donde un aumento en el pasado tiende a estar seguido por una disminución, y viceversa, dando a entender un comportamiento oscilatorio.

- El coeficiente θ_1 es positivo (0.4499), lo que implica que los errores en las predicciones pasadas tienen un impacto positivo en las predicciones actuales. Es decir, si hubo un error positivo en la predicción del período anterior, es probable que se observe un ajuste positivo en las predicciones actuales.
- $\sigma^2 = 0.4592$, indica que, en promedio, la serie de datos fluctúa alrededor de su tendencia en aproximadamente 0.4592 unidades.

5.3.2.7. Evaluación y diagnóstico del modelo ARIMA(1, 1, 2)

Validación del supuesto de Normalidad

H_0 : Los residuos siguen una distribución normal.

H_1 : Los residuos no siguen una distribución normal.

Tabla 5.11:

Resultados de la prueba de normalidad de (Kolmogorov-Smirnov) para los residuos del modelo ARIMA(1, 1, 2)

Estadístico	Valor
Estadístico de Kolmogorov	0.20399
p-valor	0.05

Interpretación

Según los resultados de la **Tabla 5.11** en el software R el valor p es de 0.05, con lo que se tiene suficiente evidencia para aceptar la hipótesis nula y rechazar la hipótesis alterna, tenemos suficiente evidencia para afirmar que los residuos siguen una distribución normal, ya que el valor p es igual que el nivel de significancia utilizado de 0.05.

Validación del supuesto de Independencia

H_0 : Los residuos son independiente, lo que significa que no existe autocorrelación significativa en los residuos hasta un determinado rezago h .

H_1 : La residuos no son independientes, es decir, existe al menos una autocorrelación significativa en los residuos hasta el rezago h .

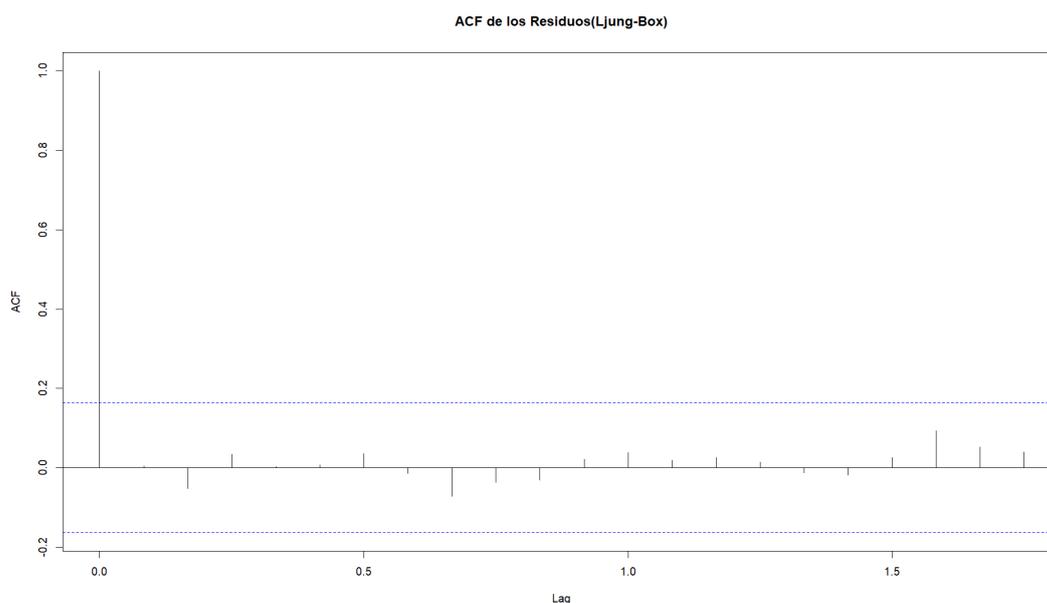
Tabla 5.12:

Resultados del test de Box-Ljung para los residuos del modelo ARIMA(1, 1, 2)

Estadístico	Valor
Estadístico de Box-Ljung	4.5704
Grados de libertad (df)	20
p-valor	0.9999

Figura 5.8:

Gráfico de autocorrelación Simple (ACF) para los residuos del modelo ARIMA(1, 1, 2)



Interpretación

Según los resultados de la **Tabla 5.12** del test de Ljung-Box para los residuos del modelo ARIMA(1,1,2) muestran un estadístico de 4.5704 con un p-valor de 0.9999. Dado que el p-valor es mucho mayor que el nivel de significancia 0.05, no hay evidencia suficiente para rechazar la hipótesis nula (H_0). Esto indica que

no existe autocorrelación significativa en los residuos hasta el rezago h , lo que sugiere que los residuos del modelo $ARIMA(1,1,2)$ son independientes y el modelo ha capturado adecuadamente las dependencias temporales en los datos. Según la **Figura 5.8** de autocorrelación simple (ACF) para los residuos también respalda esta conclusión al mostrar que las autocorrelaciones están dentro de los intervalos de confianza, lo que confirma la independencia de los residuos.

Validación del supuesto de Estacionariedad

H_0 : Los residuos del modelo $ARIMA(1, 1, 2)$ es no estacionaria.

H_1 : Los residuos del modelo $ARIMA(1, 1, 2)$ serie es estacionaria.

Tabla 5.13: *Resultados de la prueba de Augmented Dickey-Fuller para los residuos del modelo $ARIMA(1, 1, 2)$*

Estadístico	Valor
Dickey-Fuller	-1.8536
Lag order	1
p-valor	0.0234

Interpretación

Según los resultados de la **Tabla 5.13** se muestra la prueba de Augmented Dickey-Fuller para los residuos del modelo $ARIMA(1,1,2)$ muestran un estadístico de -1.8536 con un p-valor de 0.0234. Dado que el p-valor es menor que el nivel de significancia 0.05, se rechaza la hipótesis nula (H_0) que sugiere la presencia de una raíz unitaria en los residuos. Esto indica que los residuos del modelo $ARIMA(1,1,2)$ son estacionarios, ya que el p-valor sugiere que no hay una tendencia estacional en los residuos.

Prueba t de Student

H_0 : La media de los residuos es igual a 0.

H_1 : La media de los residuos no es igual a 0.

Tabla 5.14:

Resultados de la prueba t Student para los residuos del modelo ARIMA(1, 1, 2)

Estadístico	Valor
Estadístico t	0.93692
Grados de libertad (df)	143
p-valor	0.3504

Interpretación

Según los resultados de la **Tabla 5.14** Los resultados de la prueba t de Student para los residuos del modelo ARIMA(1,1,2) muestran un estadístico t de 0.93692 con un p-valor de 0.3504. Dado que el p-valor es significativamente mayor que el nivel de significancia 0.05, no hay evidencia suficiente para rechazar la hipótesis nula H_0 que establece que la media de los residuos es cero. Esto sugiere que no hay evidencia de sesgo en los residuos del modelo ARIMA(1,1,2), indicando que el modelo ha ajustado los datos de manera adecuada.

El análisis de los residuos del modelo ARIMA(1,1,2) muestra resultados positivos en cuanto a los supuestos fundamentales del modelo. La prueba de Kolmogorov-Smirnov indica que los residuos siguen una distribución normal, ya que el p-valor de 0.05 es igual al nivel de significancia utilizado, permitiendo aceptar la hipótesis nula de normalidad. Además, el test de Box-Ljung, con un p-valor de 0.9999, confirma que no hay autocorrelación significativa en los residuos, apoyando la independencia de los mismos. El gráfico de autocorrelación simple también respalda esta conclusión al mostrar residuos dentro de los intervalos de confianza. La prueba de Augmented Dickey-Fuller muestra un p-valor de 0.0234, lo que indica que los residuos son estacionarios, rechazando la presencia de una raíz unitaria. Finalmente, la prueba t de Student, con un p-valor de 0.3504, sugiere que no hay sesgo en

los residuos, confirmando que la media de los residuos es cero. En conjunto, estos resultados sugieren que el modelo ARIMA(1,1,2) se ajusta adecuadamente a los datos, cumpliendo con los supuestos de normalidad, independencia, y estacionariedad de los residuos, también se puede afirmar que los residuos del modelo son ruidos blanco $\mu = 0$ y Varianza constante.

5.3.2.8. Evaluación y diagnóstico del modelo ARIMA(2, 1, 1)

Validación del supuesto de Normalidad

H_0 : Los residuos siguen una distribución normal.

H_1 : Los residuos no siguen una distribución normal.

Tabla 5.15:

Resultados de la prueba de normalidad de (Kolmogorov-Smirnov) para los residuos del modelo ARIMA(2, 1, 1)

Estadístico	Valor
Estadístico de Kolmogorov	0.19771
p-valor	0.058

Interpretación

Según los resultados de la **Tabla 5.15**, Los resultados de la prueba de normalidad de Kolmogorov-Smirnov para los residuos del modelo ARIMA(2,1,1) indican un estadístico de 0.19771 y un p-valor de 0.058. Dado que el p-valor es mayor que el nivel de significancia de 0.05, no se rechaza la hipótesis nula H_0 , que establece que los residuos siguen una distribución normal. Aunque el p-valor está cerca del umbral de significancia, no hay suficiente evidencia para concluir que los residuos no siguen una distribución normal.

Validación del supuesto de Independencia

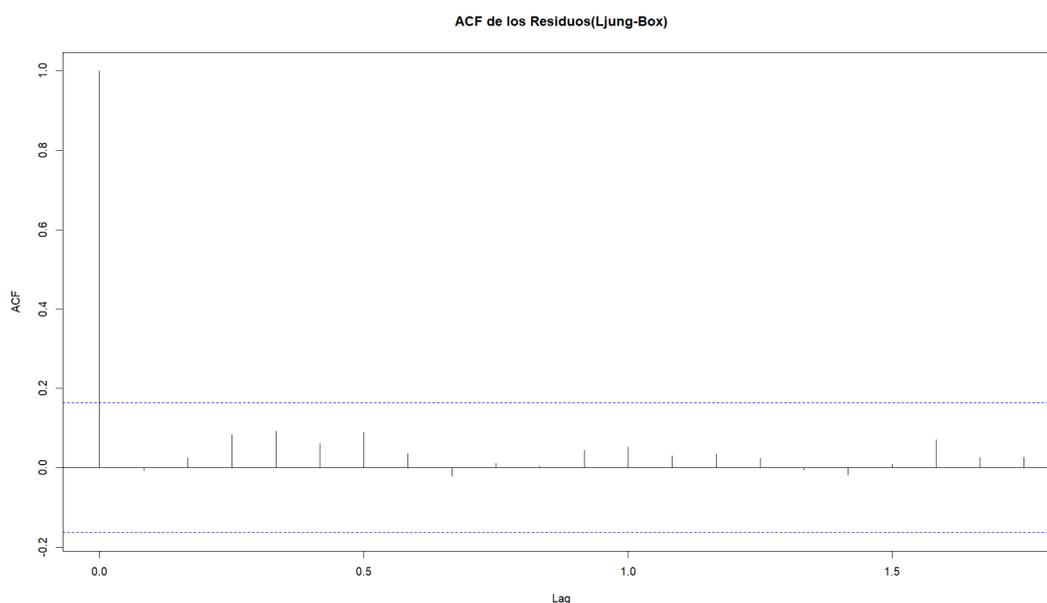
H_0 : La serie temporal es independiente, lo que significa que no existe autocorrelación significativa en los residuos hasta un determinado rezago h .

H_1 : La serie temporal no es independiente, es decir, existe al menos una autocorrelación significativa en los residuos hasta el rezago h .

Tabla 5.16:
Resultados del test de Box-Ljung para los residuos del modelo ARIMA(2, 1, 1)

Estadístico	Valor
Estadístico de Box-Ljung	0.10806
Grados de libertad (df)	2
p-valor	0.9474

Figura 5.9:
Gráfico de autocorrelación Simple (ACF) para los residuos del modelo ARIMA(2, 1, 1)



Interpretación

Según los resultados de la **Tabla 5.16**, del test de Box-Ljung para los residuos del modelo ARIMA(2,1,1) muestran un estadístico de 0.10806 y un p-valor de 0.9474.

Dado que el p-valor es mucho mayor que el nivel de significancia de 0.05, no se rechaza la hipótesis nula H_0 , que sostiene que los residuos no presentan autoco-

relación significativa hasta el rezago h . Esto indica que los residuos del modelo $ARIMA(2,1,1)$ son independientes, y que el modelo ha capturado adecuadamente las dependencias temporales en los datos. En la **Figura 5.9** de autocorrelación simple (ACF) para los residuos respalda esta conclusión, ya que muestra que las autocorrelaciones están dentro de los intervalos de confianza, confirmando la independencia de los residuos.

Validación del supuesto de Estacionariedad

H_0 : Los residuos del modelo $ARIMA(2, 1, 1)$ es no estacionaria.

H_1 : Los residuos del modelo $ARIMA(2, 1, 1)$ serie es estacionaria.

Tabla 5.17:

Resultados de la prueba de Augmented Dickey-Fuller para los residuos del modelo $ARIMA(2, 1, 1)$

Estadístico	Valor
Dickey-Fuller	-3.5955
Lag order	5
p-valor	0.03621

Interpretación

Según los resultados de la **Tabla 5.17**, de la prueba de Augmented Dickey-Fuller para los residuos del modelo $ARIMA(2,1,1)$ muestran un estadístico de -3.5955 y un p-valor de 0.03621. Dado que el p-valor es menor que el nivel de significancia de 0.05, se rechaza la hipótesis nula H_0 , que sugiere la presencia de una raíz unitaria en los residuos. Esto indica que los residuos del modelo $ARIMA(2,1,1)$ son estacionarios, confirmando que el modelo ha eliminado las tendencias estacionales y las unidades de raíz en los residuos, lo que es consistente con la adecuada especificación del modelo.

Prueba t de Student

H_0 : La media de los residuos es igual a 0.

H_1 : La media de los residuos no es igual a 0.

Tabla 5.18:

Resultados de la prueba t Student para los residuos del modelo ARIMA(2, 1, 1)

Estadístico	Valor
Estadístico t	1.274
Grados de libertad (df)	143
p-valor	0.2047

Interpretación

Según los resultados de la **Tabla 5.18**, de la prueba t de Student para los residuos del modelo ARIMA(2,1,1) muestran un estadístico t de 1.274 y un p-valor de 0.2047. Como el p-valor es significativamente mayor que el nivel de significancia de 0.05, no se puede rechazar la hipótesis nula H_0 , que establece que la media de los residuos es igual a cero. Esto sugiere que no hay evidencia de sesgo en los residuos del modelo ARIMA(2,1,1), indicando que el modelo ha ajustado adecuadamente los datos y no presenta sesgo sistemático en los residuos.

En resumen, el análisis del modelo ARIMA(2, 1, 1) revela resultados positivos en la evaluación de sus supuestos. La prueba de normalidad de Kolmogorov-Smirnov muestra un p-valor de 0.058, lo que indica que no hay evidencia suficiente para rechazar la hipótesis nula de que los residuos siguen una distribución normal, a pesar de estar cerca del umbral de significancia. La prueba de Box-Ljung, con un p-valor de 0.9474, sugiere que los residuos no presentan autocorrelación significativa, confirmando la independencia de los residuos y la adecuada captura de las dependencias temporales por parte del modelo. Finalmente, la prueba de Augmented Dickey-Fuller muestra un p-valor de 0.03621, indicando que los residuos son estacionarios y que el modelo ha eliminado adecuadamente las tendencias estacio-

nales y las raíces unitarias. Estos resultados sugieren que el modelo ARIMA(2, 1, 1) cumple con los supuestos de normalidad, independencia y estacionariedad, lo que respalda su adecuación para el análisis de la serie temporal. también se puede afirmar que los residuos del modelo son ruidos blanco debido a que presentan $\mu=0$ y Varianza constante.

5.4. Resultados del objetivo específico 2

Para alcanzar el segundo objetivo específico, identificaremos el modelo heterocedástico condicionado que presente mejor ajuste a la serie precio del petróleo para el período 2011 a 2022; a continuación se menciona el primero paso:

5.4.1. Prueba de inferencia para los efectos Autorregresivos Condicionados

A partir del análisis heterocedástico y evidenciando la variabilidad en los datos, reforzado también con el modelo ARIMA sin realizar la diferenciación se llega a tener evidencia clara de la variabilidad de los residuos, a pesar que se tiene suficiente evidencia para poder utilizar un modelo autorregresivo condicionado; se opta por la prueba de efectos ARCH a continuación se muestran las pruebas de hipótesis a contrastar:

H_0 : La serie de tiempo (precio del petróleo) no presente efectos autorregresivos condicionados heterocedástico.

H_1 : La serie de tiempo (precio del petróleo) presente efectos autorregresivos condicionados heterocedástico.

Tabla 5.19: *Resultados de la prueba para los efectos ARCH para los residuos del modelo ARIMA(2, 1, 1)*

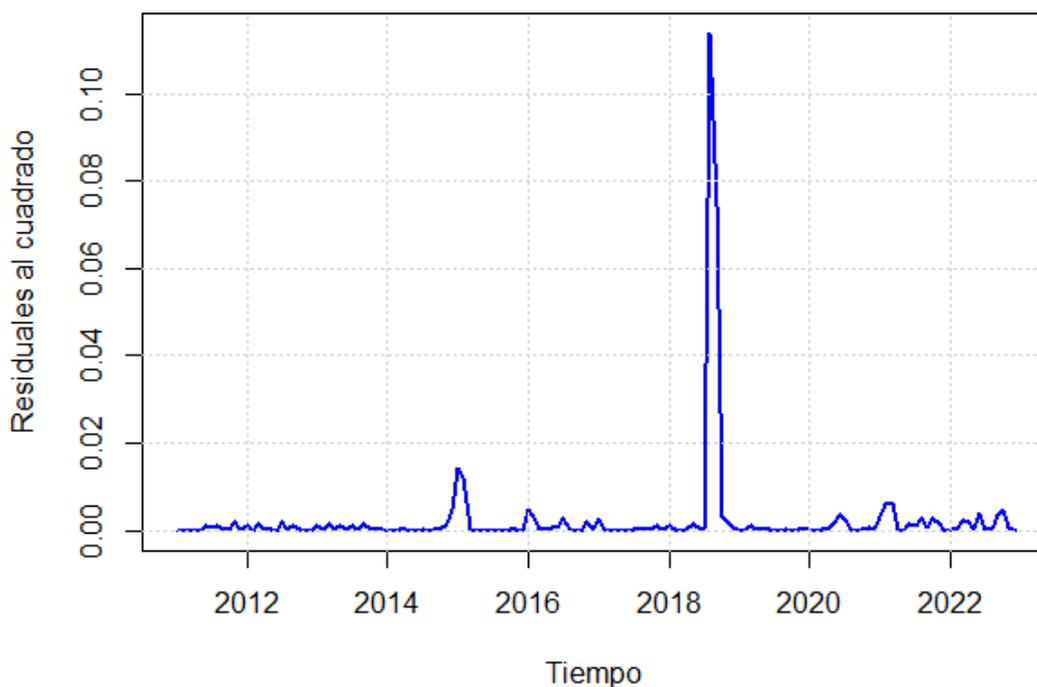
Estadístico	Valor
Chi-squared	29.608
Grados de libertad	1
p-valor	5.289e-08

Interpretación

Según los resultados de la **Tabla 5.19**, Los resultados de la prueba con la finalidad de detectar los efectos para la serie precio del petróleo mostraron evidencia sobre el estadístico Chi-cuadrado de 29.608 con 1 grado de libertad y un valor p prácticamente cero ($<5.289e-08$). Esto sugiere una fuerte evidencia en contra de la hipótesis nula de que no hay efectos de heterocedasticidad condicional autorregresiva (ARCH).

Dado que el valor p es extremadamente pequeño, esto indica que los residuos de la serie de tiempo exhiben una estructura de varianza que está relacionada con los cuadrados de los errores pasados.

Figura 5.10:
Visualización de los errores al cuadrado de la serie precio del petróleo en soles por galón



Interpretación

Según los resultados de la **Figura 5.10** se reafirma los hallazgos anteriores sobre los efectos de volatilidad y heterocedasticada presentes en la serie precio del petróleo en soles por galón.

5.4.2. Ajuste del modelo GARCH

Para la parte de ajuste del modelo GARCH se prueba distintos modelos teniendo en consideración los criterios de selección de modelos AIC, BIC y MAPE y MAE

Tabla 5.20:
Coeficientes y el criterio de información de Akaike (AIC) de los Modelos GARCH)

Modelo	ω_1	α_1	α_2	β_1	β_2
GARCH(1, 2)	0.001363	0.225027	0.000	0.000	0.000
GARCH(2, 1)	0.001363	0.225029	0.000000	0.0000	0.0000
GARCH(2, 2)	0.001369	0.232288	0.000000	0.000000	0.000000
GARCH(1, 1)	0.001363	0.225035	0.000000	0.000000	0.000000
GARCH(0, 1)	0.000003	0.999000	0.000000	0.000000	0.000000

Tabla 5.21:
Criterios de selección de los modelos GARCH para la serie precio del petróleo

Modelo	AIC	MAPE	MAE
GARCH(1, 2)	2.105644	56440.54	0.3437
GARCH(2, 1)	1.795596	32587.514	0.3250
GARCH(2, 2)	1.820001	3898.630	0.3410
GARCH(1, 1)	2.091756	56441.3822	0.3437
GARCH(0, 1)	14.56331	124854.2569	2.4697

Interpretación

Según los resultados de la Tabla 5.20 y la Tabla 5.21, el modelo GARCH(2, 1) se destaca como el mejor ajuste para la serie de precios del petróleo. En la Tabla 5.20, los coeficientes del modelo GARCH(2, 1) ($\omega_1 = 0.001363$, $\alpha_1 = 0.225029$, $\alpha_2 = 0.000000$, $\beta_1 = 0.0000$, $\beta_2 = 0.0000$) muestran valores que indican un ajuste adecuado sin valores excesivamente altos o bajos, lo que refleja estabilidad en el modelo.

Por otro lado, en la Tabla 5.21, el modelo GARCH(2, 1) presenta el mejor desempeño en términos de los criterios de selección, con el valor más bajo de AIC (1.795596), y el MAE más bajo (0.3250), señalando una menor diferencia entre los valores pronosticados y los reales. Además, el MAPE para el GARCH(2, 1) es significativamente menor (32,587.514) en comparación con los otros modelos

evaluados, lo que indica una mejor precisión en las predicciones.

En contraste, otros modelos como el GARCH(0, 1) tienen valores mucho más altos en AIC (14.56331), MAPE (124,854.2569) y MAE (2.4697), sugiriendo un ajuste deficiente y menor precisión en comparación con el GARCH(2, 1). Por lo tanto, el modelo GARCH(2, 1) no solo tiene coeficientes que sugieren estabilidad, sino que también demuestra valores adecuados en términos de AIC, MAPE y MAE, consolidándose como el modelo más adecuado para la serie de precios del petróleo según los criterios evaluados.

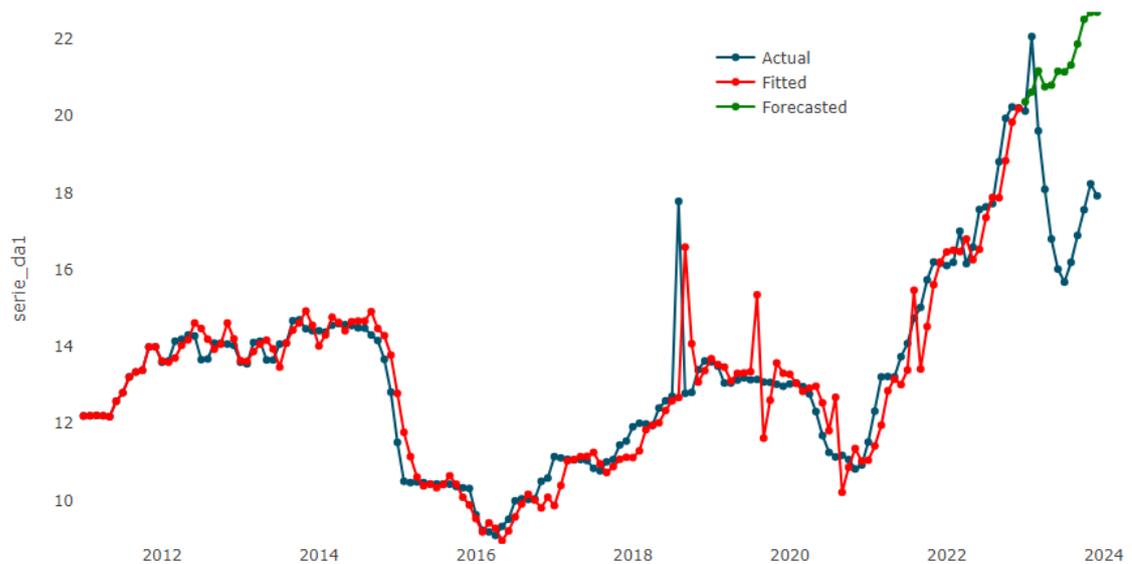
Con todos los pasos abarcados en este apartado se alcanza el objetivo específico 2 propuesto esto se debe a que se encontró el modelo heterocedásticos condicionado mejor ajustado a la serie precio del petróleo.

5.5. Resultado de la hipótesis específico 3

Para alcanzar el tercer objetivo específico, el cual consiste en comparar que enfoque ofrece un mejor pronóstico para el precio del petróleo en la región de Cusco 2011-2023; los modelos heterocedásticos condicionados o la metodología Box-Jenkins. Para ello se evalúan los pronósticos del modelo ARIMA(1, 1, 2), ARIMA(2, 1, 1) y GARCH(2,1) con la serie de prueba los doce meses del 2023; a continuación se muestran lo siguientes apartado:

5.5.1. Pronóstico del modelo ARIMA(1, 1, 2) para la serie de prueba

Figura 5.11:
Gráfico de comparación para la serie original, el modelo ajustado y el pronóstico del modelo ARIMA(1, 1, 2)



En la **Figura 5.11**, se muestra el ajuste del modelo ARIMA(1, 1, 2) a la serie de tiempo del precio del petróleo en soles por galón. El modelo ARIMA, que combina componentes autoregresivos (AR), de media móvil (MA), y de diferenciación (I) para hacer la serie estacionaria, se ha utilizado para capturar las características esenciales de la serie de tiempo, como las tendencias y las estructuras de correlación temporal.

Tabla 5.22:
Datos de Prueba y Pronóstico para el Año 2023 según el modelo ARIMA(1, 1, 2)

Año	Mes	Datos de Prueba	Datos del pronóstico	Diferencia
2023	Enero	20.13	20.51293	-0.38293
2023	Febrero	22.07	20.72947	1.34053
2023	Marzo	19.61	20.91868	-1.30868
2023	Abril	18.09	21.08402	-2.99402
2023	Mayo	16.80	21.22849	-4.42849
2023	Junio	16.02	21.35474	-5.33474
2023	Julio	15.68	21.46505	-5.78505
2023	Agosto	16.19	21.56144	-5.37144
2023	Setiembre	16.89	21.64566	-4.75566
2023	Octubre	17.56	21.71926	-4.15926
2023	Noviembre	18.24	21.78357	-3.54357
2023	Diciembre	17.92	21.83977	-3.91977

Interpretación

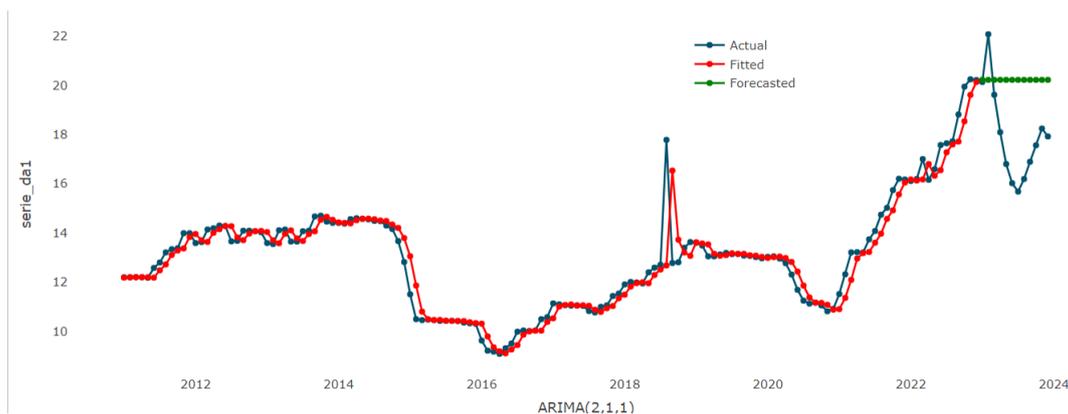
Según los resultados de la **Tabla 5.22**, se presentan los datos de prueba y los pronósticos generados por el modelo ARIMA(1, 1, 2) para el precio del petróleo por mes en el año 2023. La tabla muestra cómo el modelo intenta ajustarse a la serie temporal. Aunque el modelo captura la tendencia general, se observan discrepancias significativas en ciertos meses, donde el pronóstico presenta picos que no coinciden con los datos reales.

Los datos reales indican un descenso en los precios durante la primera mitad del año, seguido por un repunte hacia finales de 2023. Estos resultados sugieren que el modelo ARIMA(1, 1, 2) tiene limitaciones en la captura de cambios abruptos o inesperados en la serie temporal.

5.5.2. Pronóstico del modelo ARIMA(2, 1, 1) para la serie de prueba

Figura 5.12:

Gráfico de comparación para la serie original, el modelo ajustado y el pronóstico del modelo ARIMA(2, 1, 1)



Según los resultados de la **Figura 5.12**, se muestra el ajuste del modelo ARIMA(2, 1, 1) a la serie de tiempo del precio del petróleo en soles por galón. El modelo ARIMA, que combina componentes autoregresivos (AR), de media móvil (MA), y de diferenciación (I) para hacer la serie estacionaria, se ha utilizado para capturar las características esenciales de la serie de tiempo, como las tendencias y las estructuras de correlación temporal. Además, en la figura se puede observar el pronóstico realizado por el modelo ARIMA(2, 1, 1). Este pronóstico abarca los 12 meses correspondientes al año 2023, proporcionando una previsión de cómo podría evolucionar el precio del petróleo en el futuro cercano.

Tabla 5.23:
Datos de Prueba y Pronóstico para el Año 2023 según el modelo ARIMA(2, 1, 1)

Año	Mes	Datos de Prueba	Datos del pronóstico	Diferencia
2023	Enero	20.13	20.21747	-0,08747
2023	Febrero	22.07	20.21727	1,85273
2023	Marzo	19.61	20.21617	-0,60617
2023	Abril	18.09	20.21697	-2,12697
2023	Mayo	16.80	20.21659	-3,41672
2023	Junio	16.02	20.21672	-4,19672
2023	Julio	15.68	20.21669	-4,53669
2023	Agosto	16.19	20.21669	-4,02669
2023	Setiembre	16.89	20.21670	-3,3267
2023	Octubre	17.56	20.21669	-2,65669
2023	Noviembre	18.24	20.21669	-1,97669
2023	Diciembre	17.92	20.21669	-2,29669

Interpretación

Según los resultados de la **Tabla 5.23**, se presentan los datos de prueba y los pronósticos generados por el modelo ARIMA(2, 1, 1) para el precio del petróleo por mes en el año 2023. La tabla muestra cómo el modelo intenta ajustarse a la serie temporal. Se observan discrepancias significativas en ciertos meses, donde el pronóstico presenta picos que no coinciden con los datos reales. Los datos reales indican un descenso en los precios durante la primera mitad del año, seguido por un repunte hacia finales de 2023. Estos resultados sugieren que el modelo ARIMA(2, 1, 1) tiene limitaciones en la captura de cambios abruptos o inesperados en la serie temporal. El cuadro muestra la comparación entre los datos reales y los pronosticados por el modelo ARIMA(2, 1, 1) para el año 2023, destacando que, en algunos meses, como enero, la diferencia entre ambos es pequeña (-0,08747), lo que refleja un buen ajuste del modelo. Sin embargo, en otros meses, como junio (-4,19672), el error es mayor, lo que indica que el modelo subestimó los datos reales en esos casos. A lo largo del año, se observa una tendencia general donde

el modelo tiende a pronosticar valores ligeramente superiores a los reales.

Resultados para la selección del mejor modelo

Para elegir el mejor modelo se compara los residuos del modelo ARIMA(1,1,2) y ARIMA(2,1,1) utilizando la diferencia entre los datos de prueba y los datos de la serie pronosticada en ambos modelos. Para comparar estos residuos se toma en cuenta a los criterios del BIC, AIC, RMSE y MAE, los cuales se muestran a continuación

Tabla 5.24:

Comparación de Métricas de Modelos ARIMA(1,1,2) y ARIMA(2,1,1) en la diferencia entre la data de prueba y data pronosticada

Modelo	AIC	MAPE	RMSE
ARIMA(1, 1, 2)	43,03	3.752.445	1.115.848
ARIMA(2, 1, 1)	42,97	1.130.583	0.8279612

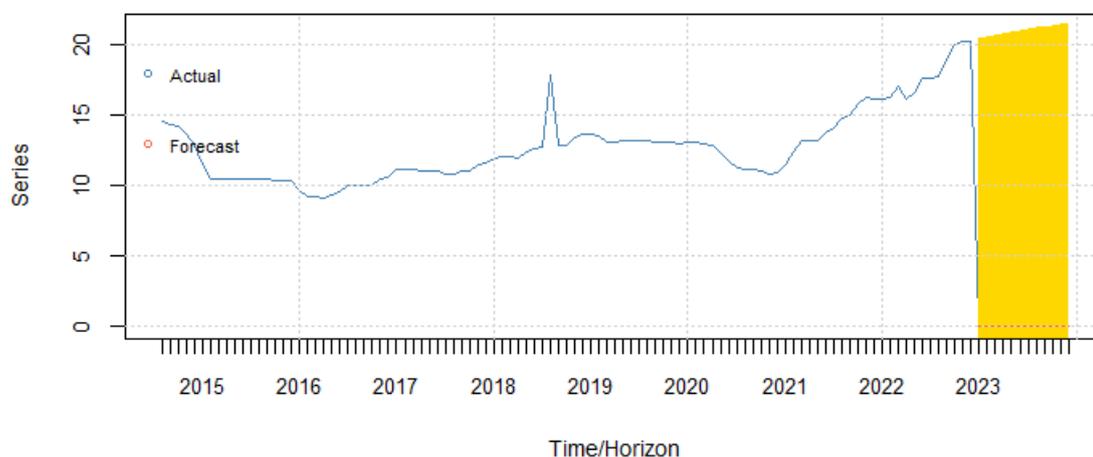
Según los resultados de la **Tabla 5.24**, la tabla comparativa de métricas de ajuste entre los modelos ARIMA(1, 1, 2) y ARIMA(2, 1, 1) proporciona información valiosa sobre su desempeño en la predicción de precios del petróleo. En términos de AIC, el modelo ARIMA(2, 1, 1) presenta un valor más bajo (42.97) en comparación con ARIMA(1, 1, 2) (43.03). Esto indica que el primer modelo tiene la capacidad de ajuste a los datos, sugiriendo una preferencia por su uso. Además, el Error Porcentual Absoluto Medio (MAPE) del modelo ARIMA(2, 1, 1) es significativamente menor (1.130.583) que el del modelo ARIMA(1, 1, 2) (3.752.445), lo que resalta la superioridad de ARIMA(2, 1, 1) en términos de precisión de pronósticos. Un MAPE más bajo significa que las predicciones de este modelo se acercan más a los valores reales. En cuanto al Error Cuadrático Medio (RMSE), ARIMA(2, 1, 1) presenta un RMSE de 0.8279612, que también es más bajo que el RMSE de 1.115.848 del modelo ARIMA(1, 1, 2). Esto indica que ARIMA(2, 1, 1) no solo es

más preciso, sino que también tiene un menor nivel de error en sus predicciones. Esto refuerza la conclusión de que ARIMA(2, 1, 1) es el modelo más eficaz en la predicción de los precios del petróleo.

5.5.3. Pronóstico del modelo GARCH(2, 1)

Figura 5.13:

Gráfico de ajuste del modelo GARCH(2,1) y el intervalo de pronóstico y variación del coeficiente gamma



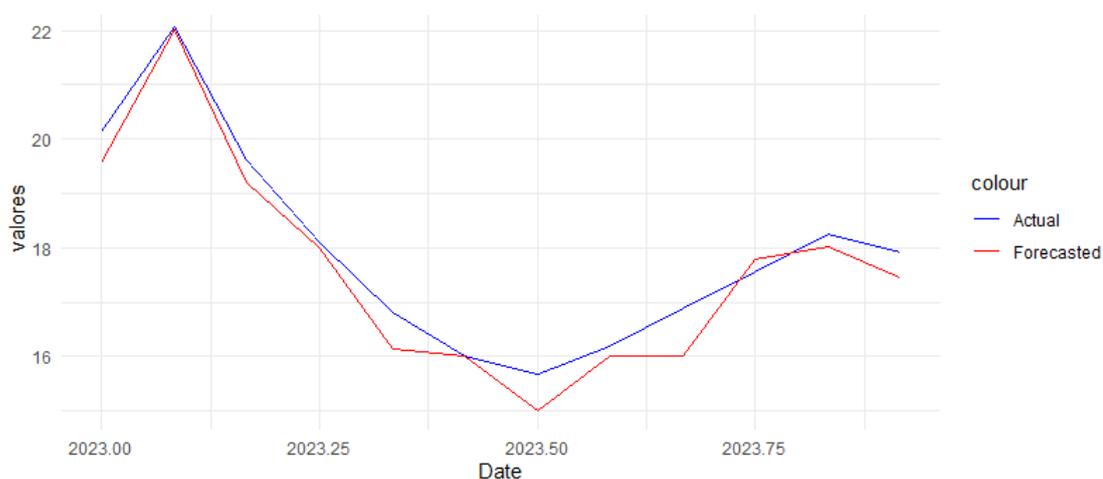
Interpretación

Según los resultados de la **Figura 5.13**, se observa el pronóstico en un intervalo hecho por el modelo GARCH(2,1) para la data de prueba, la cual corresponde a los 12 meses del año 2023. El intervalo de pronóstico muestra la evolución esperada del precio del petróleo a lo largo del tiempo, reflejando la capacidad del modelo GARCH de capturar la volatilidad y la heteroscedasticidad en los datos. El modelo GARCH (Generalized Autoregressive Conditional Heteroskedasticity) es especialmente adecuado para series temporales financieras debido a su habilidad para modelar cambios en la volatilidad.

El intervalo de confianza en el gráfico indica la posible variación en los precios del

petróleo, y cómo el coeficiente gamma dentro del modelo GARCH, que representa la reacción a cambios en la varianza, puede influir en las predicciones. La inclusión de un intervalo de confianza proporciona una medida de la incertidumbre del pronóstico, mostrando no solo una línea de pronóstico central, sino también los límites dentro de los cuales se espera que los precios futuros se encuentren con un cierto nivel de confianza.

Figura 5.14:
Gráfico de ajuste del modelo GARCH(2,1) a la serie precio del petróleo



En el siguiente gráfico se puede observar el pronóstico ajustado al modelo Garch(2,1) al precio del petróleo, también se aprecia que la curva de pronóstico con los datos para el año 2023 es similar, esto debido a que el modelo captura de mejor manera la volatilidad o el cambio en el precio del petróleo.

Tabla 5.25:
Datos de Prueba y Pronóstico para el Año 2023 según el modelo GARCH(2, 1)

Año	Mes	Datos de Prueba	Datos del Pronóstico	Diferencia
2023	Enero	20.13	19.56	0.57
2023	Febrero	22.07	22.01	0.06
2023	Marzo	19.61	19.21	0.40
2023	Abril	18.09	18.00	0.09
2023	Mayo	16.80	16.15	0.65
2023	Junio	16.02	16.00	0.02
2023	Julio	15.68	15.00	0.68
2023	Agosto	16.19	16.00	0.19
2023	Setiembre	16.89	16.00	0.89
2023	Octubre	17.56	17.78	-0.22
2023	Noviembre	18.24	18.01	0.23
2023	Diciembre	17.92	17.45	0.47

Interpretación

Según los resultados de la **Figura 5.14** y **Tabla 5.25**, se observan los resultados de los datos de prueba y los pronósticos para el año 2023, generados con un modelo GARCH(2, 1). También se presenta la diferencia entre los valores observados (datos de prueba) y los valores estimados por el modelo (datos del pronóstico). En términos generales, las diferencias son pequeñas, lo que indica que el modelo GARCH captura bien la dinámica de los datos. Los meses con las mayores diferencias son septiembre, con 0.89, y julio, con 0.68, lo que sugiere que el modelo subestimó los valores reales en esos meses. Por otro lado, en octubre, el modelo sobrestimó ligeramente, con una diferencia de -0.22. Los meses con mejor ajuste son junio, con una diferencia casi nula de 0.02, y febrero, con 0.06, lo que demuestra que el modelo hizo un pronóstico bastante preciso en estos meses. En general, el modelo GARCH(2, 1) parece seguir bien la tendencia general de los datos durante el año, con un ajuste confiable en la mayoría de los meses, aunque algunos periodos muestran una mayor volatilidad que el modelo no capturó completamente.

5.6. Resultados del objetivo general

Para responder el objetivo general, se presenta una síntesis de los principales hallazgos obtenidos en los apartados previos. En primer lugar, se identificó el modelo más adecuado para la serie del precio del petróleo utilizando la metodología Box-Jenkins, determinándose que el modelo ARIMA(2, 1, 1) proporciona el mejor ajuste.

Luego, se evaluaron modelos heterocedasticosd condicionales, encontrándose que el modelo GARCH(2, 1) ofrece un mejor desempeño en comparación con otros modelos de la misma clase. La selección de los modelos se fundamentó tanto en los criterios de información (AIC, BIC) como en el análisis de los residuos generados.

Después se realizó el pronóstico del precio del petróleo (en soles por galón) empleando tanto los modelos ARIMA como los modelos GARCH. Los resultados muestran que el modelo GARCH(2, 1) proporciona predicciones más precisas, al captar de manera más efectiva la variabilidad de la serie, en particular frente a fluctuaciones bruscas o leves en los precios.

A partir de los hallazgos desarrollados, se confirma que los modelos GARCH en particular el modelo GARCH(2,1) ofrece un mejor ajuste y pronóstico en comparación con los modelos ARIMA derivados de la metodología de Box-Jenkins para la serie del precio del petróleo.

DISCUSIÓN

El objetivo general de esta investigación consiste en determinar el modelo más adecuado para pronosticar el precio del petróleo en la Región Cusco durante el período 2011-2023, utilizando el lenguaje de programación R y su entorno de desarrollo integrado. Se aplicaron técnicas básicas de series temporales, dividiendo los datos en entrenamiento y datos de prueba para asegurar un buen ajuste y capacidad predictiva de los modelos Box y Arch. A través de un análisis riguroso, se determinó que los modelos heterocedásticos condicionales generalizados (GARCH) proporcionan un mejor ajuste y capacidad predictiva para la serie de precios del petróleo en comparación con la metodología Box-Jenkins. Estos hallazgos son consistentes con los resultados presentados en la tesis de **Ortiz (2017)**, donde se demostró que, ante series con datos atípicos o extremos, los modelos de alta volatilidad proporcionan mejores características para series económicas, como es el caso del precio del petróleo crudo (WTI). La elección del modelo GARCH se fundamenta en su capacidad para capturar la volatilidad cambiante a lo largo del tiempo, una característica frecuente en las series temporales de precios de materias primas. Los modelos GARCH permiten ajustar la varianza condicional a los datos históricos, lo que resulta en pronósticos más precisos en presencia de fluctuaciones extremas. Esta ventaja es crucial para establecer pronósticos que permitan la implementación de medidas preventivas y la salvaguarda de aspectos económicos de un sector,

empresa, región o país. Adicionalmente, la metodología de Box-Jenkins, aunque robusta y ampliamente utilizada, tiende a ser menos efectiva cuando se enfrenta a datos con alta volatilidad y presencia de outliers. Los modelos ARIMA, típicos de la metodología Box-Jenkins, suponen una varianza constante, lo cual puede resultar en pronósticos subóptimos en contextos de alta variabilidad.

Otro de los proyectos de investigación que refuerza nuestros hallazgos es el de **Lazaridis (2021)**, cuyo objetivo fue analizar y utilizar técnicas de pronóstico para el precio del crudo del petróleo. La conclusión de este estudio fue que los modelos provenientes de la metodología de Box-Jenkins, con componentes estacionales, son efectivos para series econométricas y ayudan a generar pronósticos adecuados. En relación con nuestros hallazgos, se destaca el adecuado ajuste del modelo ARIMA(2, 1, 1) para la serie de precios del petróleo en la Región Cusco, lo que indica su alta efectividad para generar predicciones y permitir la toma de decisiones informadas a nivel de sector, empresa, región o país. Aunque nuestro estudio identificó que los modelos heterocedásticos condicionales generalizados (GARCH) proporcionan un mejor pronóstico y ajuste en presencia de alta volatilidad y datos atípicos, es importante resaltar que los modelos ARIMA, particularmente con componentes estacionales, también pueden ofrecer una valiosa precisión en contextos donde la serie temporal presenta patrones estacionales claros. La complementariedad entre estos enfoques resalta la importancia de seleccionar el modelo más adecuado según las características específicas de la serie temporal analizada. Esto refuerza la idea de que, si bien los modelos GARCH son preferibles en situaciones de alta volatilidad, los modelos ARIMA estacionales no deben ser descartados, ya que pueden ofrecer un excelente ajuste y predicción en escenarios con patrones

estacionales bien definidos.

El proyecto de investigación de **Castillo (2015)**, que tuvo por objetivo el modelar volatilidad diaria del Índice General de la Bolsa de Valores de Lima (IGVBL), periodo 2009 – 2011; concluyó que los modelos de alta volatilidad son modelos adecuado para series con cambios constantes en los datos y también para series que presentan observaciones extremas tanto superior como inferior, estos resultados demuestran un afirmar a los hallazgos alcanzado en el objetivo específicos y el general debido a que cuando una serie tiene problemas de heterocedasticidad los modelos ARIMA tiene falencias o deben sufrir muchos cambios que perjudican a la inferencia final cosa que se vuelve más compleja el brindar un análisis mas directo para el sector al que esté brindandose la información de la predicción. Ambos estudios coinciden en que los modelos de alta volatilidad, como los GARCH, son más adecuados para series temporales con heterocedasticidad y observaciones extremas. Tanto en el pronóstico del precio del petróleo en la Región Cusco como en la volatilidad del IGVBL, los modelos GARCH demostraron un mejor desempeño comparado con los modelos tradicionales de Box-Jenkins. Ambas investigaciones enfrentaron el desafío de la heterocedasticidad y observaron que los modelos ARIMA requieren constantes ajustes para manejar la varianza no constante, lo que limita su eficacia. Además, en ambos contextos, las series contenían datos extremos que los modelos GARCH manejaron con mayor precisión. Esta similitud resalta la capacidad de los modelos GARCH para proporcionar predicciones más precisas y útiles para la toma de decisiones económicas informadas en entornos volátiles.

CONCLUSIONES

1. Con el uso de la metodología de Box-Jenkins, se concluye que el modelo ARIMA(2,1,1) es el más adecuado para la predicción de los precios del petróleo. Su selección se fundamenta en su menor valor en los criterios de información de Akaike (AIC) y bayesiano de Schwarz (BIC), así como en métricas de error como el MAPE y el MAE, que evidencian un mejor ajuste y menor variabilidad en los residuos. Además, su mayor precisión en los pronósticos lo posiciona como la opción más fiable frente a los demás modelos ARIMA evaluados.
2. Con el análisis de la serie de tiempo del precio del petróleo en la Región Cusco 2011-2022, se identificó que el modelo GARCH(2,1) es el que mejor se ajusta a la dinámica de los precios. Este modelo destaca por presentar los valores más bajos en los criterios de selección, como el AIC (1.795596), MAE (0.3250) y MAPE (32,587.514), evidenciando un ajuste más preciso y eficiente en la captura de la volatilidad de la serie. Además, los resultados de las pruebas de efectos ARCH confirman la presencia de heterocedasticidad en los datos, lo que justifica el uso de un modelo GARCH para su análisis. En conclusión, el objetivo específico 2 se ha cumplido satisfactoriamente, al identificar el modelo GARCH(2,1) como la opción más adecuada para describir la evolución y la volatilidad de los precios del petróleo en el período

de estudio.

3. El análisis de los modelos de pronóstico ARIMA(1,1,2), ARIMA(2,1,1) y GARCH(2,1) para la predicción del precio del petróleo en la región de Cusco durante el año 2023, se concluye que, aunque los modelos ARIMA logran un ajuste razonable, presentan limitaciones para capturar cambios abruptos en la tendencia de los precios. Al evaluar métricas como el AIC, BIC, RMSE y MAE, el modelo ARIMA(2,1,1) muestra un mejor desempeño en términos de precisión del pronóstico, con menores errores en ciertos períodos, lo que sugiere un ajuste más eficiente en comparación con ARIMA(1,1,2). No obstante, el modelo GARCH(2,1), diseñado para capturar la heterocedasticidad en series de tiempo, demuestra una mayor capacidad para modelar la volatilidad del precio del petróleo, lo que lo convierte en una opción más robusta para el análisis de fluctuaciones y variaciones en el tiempo.
4. Los pronósticos analizados y comparados mediante el modelo ARIMA(2,1,1) de la metodología Box-Jenkins y el modelo GARCH(2,1) de enfoque heterocedástico condicionado para el precio del petróleo en el período de enero a diciembre de 2023, se concluye que el modelo GARCH(2,1) ofrece un pronóstico más eficiente. Este resultado se fundamenta en su capacidad para capturar la volatilidad y la heterocedasticidad de la serie de tiempo, lo que permite generar predicciones más precisas y ajustadas a los datos observados. Así, el enfoque heterocedástico condicionado se consolida como una herramienta más adecuada para modelar la dinámica del precio del petróleo en contextos de alta variabilidad.

RECOMENDACIONES

1. Se recomienda utilizar el modelo ARIMA(2,1,1) para realizar pronósticos de corto y mediano plazo en series de tiempo con tendencia y sin alta volatilidad. Es recomendable complementar este modelo con análisis de diagnóstico para asegurar que los residuos sean ruido blanco y que no existan patrones no capturados.
2. Se recomienda que el modelo GARCH(2,1) demuestre ser el más adecuado para capturar la volatilidad de los precios del petróleo al ser usado en escenarios donde la variabilidad y los cambios abruptos en los precios sean frecuentes. Para una mejor interpretación de los resultados, es recomendable combinar el modelo GARCH con análisis cualitativos de los factores económicos y geopolíticos que afectan el precio del petróleo.
3. Se recomienda emplear los modelos ARIMA en combinación con modelos GARCH al analizar series de tiempo que presenten alta volatilidad, debido a que los modelos ARIMA, por sí solos, tienen limitaciones para capturar cambios abruptos en la tendencia de los precios. Asimismo, actualizar periódicamente los modelos y evaluar su desempeño utilizando datos recientes, con el fin de mantener la precisión y validez de los pronósticos.
4. Se recomienda utilizar el modelo GARCH(2,1) en contextos donde la volatilidad es un factor clave en la toma de decisiones, como en la planificación de

estrategias de inversión o en la gestión de riesgos financieros. Es recomendable validar la calidad de los pronósticos utilizando datos fuera de la muestra y compararlos con otros modelos predictivos para garantizar su confiabilidad en diferentes escenarios económicos.

Bibliografía

- Abarca, F. (2012). Presentaciones con teoría del curso ie0405-modelos probabilísticos de señales y sistemas.
- Agencia Internacional de Energía (2022). Informe del mercado del petróleo - noviembre de 2022. Technical report, Agencia Internacional de Energía, París. Licencia: CC BY 4.0.
- Anderson, O. (1976). *Time Series Analysis and Forecasting: The Box-Jenkins Approach*. Butterworth.
- Aquino, Eduardo Francesqui, R. (2018). Efecto del precio internacional del petróleo y del tipo de cambio en la importación de diesel de petroperú.
- Arce, R. and Mahía, R. (2003). Modelos arima. *Programa CITUS: Técnicas de Variables Financieras*, pages 5–6.
- Bergmeir, C. and Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191:192–213. Data Mining for Software Trustworthiness.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

- Box, George EP Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 26(2):211–243.
- Box, G., Jenkins, G., and Reinsel, G. (1970). Time series analysis: forecasting and control holden-day san francisco. *Box Time Series Analysis: Forecasting and Control Holden Day*1970.
- Castillo, J. E. (2014). Modelación de la volatilidad del índice general de la bolsa de valores de lima, periodo 2009-2011.
- Catalán, C. E. (2004). Series temporales. *Laboratorio de Estadística. Consejo Superior de Investigaciones Científicas (CSIC). España*, 54.
- Chavez, A. A. (2021). Volatilidad sectorial en la bolsa de valores de lima: Estructura garch, 2015-2020.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007.
- Felicísimo, A. M. (2009). Conceptos básicos, modelos y simulación. *Obtenido de http://www6.uniovi.es/~feli/CursoMDT/Tema_1.pdf*.
- Hamilton, J. D. (2020). *Time series analysis*. Princeton university press.
- Hanke, John E Reitsch, A. G. (1995). *Pronósticos en los negocios*. Prentice Hall Hispanoamericana.
- Hanke, John E Wichern, D. W. (2005). *Business forecasting*. Pearson Educación.
- Hanke, J. E. (2010). *Pronosticos en los negocios*. Pearson Educacion de Mexico, Naucalpan de Juarez (Mexico), 9^a ed. edition.

- Hernández, R., Fernández, C., and Baptista, P. (2014). *Metodología de la investigación*, volume 6. México Mc Graw-Hill.
- Hurtado, Carlos Ríos, G. (2008). *Series de tiempo*. Santiago: Universidad de Chile.
- Jenkins, G. M., Box, G. E., and Reinsel, G. C. (2011). *Time series analysis: forecasting and control*, volume 734. John Wiley & Sons.
- Kendall, M. G. and Buckland, W. R. (1980). *Diccionario de estadística*, volume 310.
- Kruskal, W. H. and Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260):583–621.
- Lazaridis, I. (2021). *Ánálisis del precio del petróleo ARGO y su aplicación de métodos modernos para su predcción*. PhD thesis, University of Piraeus (Greece).
- Lurento, L. (2021). *Modelamiento Econométrico de la Inflación en el Perú, período 2000-2019*. PhD thesis, Universidad Nacional del Altiplano.
- Mauricio, J. A. (2007). *Introducción al análisis de series temporales*. Universidad Complutense de Madrid.
- Ministerio de Energía y Minas (31-12-2016). *Libro anual de recursos de hidrocarburos*. Technical report.
- Organización de Países Exportadores de Petróleo (5-6-2023). *Opec bulletin*. Technical report.

- Ortiz, F. (2017). Pronóstico de precios de petróleo: una comparación entre modelos garch y redes neuronales diferenciales. *Investigación Económica*, 76(300):105–126.
- Real Academia Española (2023a). Diccionario de la lengua española(23a ed.). <https://dle.rae.es/petróleo>. Fecha de consulta: 13/06/2023.
- Real Academia Española (2023b). Diccionario de la lengua española(23a ed.). Definición 2.
- Ríos, Gonzalo Hurtado, C. (2008). Series de tiempo. *Universidad de Chile. Facultad de Ciencias Físicas y Matemáticas*, 52.
- Rodriguez, W. S. (12/02/2022). Análisis de series de tiempo en r. Disponible en: https://rstudio-pubs-static.s3.amazonaws.com/865045_6c24466e1bd949b58e63267cf8292f05.html. 08/07/2023.
- Sánchez Carlessi, H., Reyes Romero, C., and Mejía Sáenz, K. (2018). Manual de términos en investigación científica, tecnológica y humanística.
- Santaella, J. (2021). ¿Qué es el barril de Brent y cómo es la cotización de este petróleo? ECO3 MULTIMEDIA. URL: <https://economia3.com/que-es-barril-petroleo-brent/#:~:text=El%20precio%20del%20petr%C3%B3leo%20se,se%20comercializan%20a%20diferentes%20precios.>
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461 – 464.

- Suárez Dosantos, P. et al. (2022). *Comparación de métodos de predicción para series temporales*. PhD thesis, Universidad de Oviedo.
- Tamayo, J., Salvador, J., Vásquez, A., and De la Cruz, R. (2015). La industria de los hidrocarburos líquidos en el Perú: 20 años de aporte al desarrollo del país. *Osinergmin. Lima-Perú*.
- Uribe, J., Fajardo Ortiz, E., Romero, H., and Romero Valbuena, H. (2017). Incidencia de las políticas públicas de empleo sobre la desocupación en Colombia: un análisis de intervención para el periodo 2002 - 2014. *Revista Espacios*, 38:22.
- Uriel, E. (1985). *Análisis de series temporales, modelos Arima*. Madrid.
- Villalba, C. (2019). *Análisis y predicción del precio del petróleo: una aproximación basada en modelos VAR*. PhD thesis, Universidad del Salvador.
- Wulff, S. S. (2017). *Time series analysis: Forecasting and control*.

ANEXOS

.1. Matriz de Consistencia

Problemas	Objetivos	Hipotesis	Marco Teórico	Metodología
<p>Problema General ¿Qué modelo ofrece mayor predicción y robustez para la serie precio del petróleo: la aplicación de modelos heterocedásticos condicionales o la metodología Box-Jenkins, durante el período 2011-2023, Región Cusco?</p> <p>Problemas Específicos</p> <p>a) ¿Cuál es el modelo con la metodología de Box-Jenkins que presenta un mejor ajuste de la serie precio del petróleo, para el período enero 2011 a diciembre 2022, región Cusco?</p> <p>b) ¿Cuál es el modelo heterocedástico condicionado que presenta mejor ajuste de la serie precio del petróleo, para el período enero 2011 a diciembre 2022?</p> <p>c) ¿Qué enfoque ofrece un mejor pronóstico para el precio del petróleo: los modelos heterocedásticos condicionados o los modelos basados en la metodología Box-Jenkins, en la región Cusco, durante el período de enero a diciembre de 2023?</p>	<p>Objetivo General Determinar el mejor modelo para pronosticar al precio del petróleo en la Región Cusco durante el período 2011-2023, considerando la aplicación de modelos heterocedásticos condicionales y la metodología Box-Jenkins</p> <p>Objetivos Específicos</p> <p>a) Identificar el modelo de la metodología de Box-Jenkins que presenta mejor ajuste a la serie precio del petróleo para el período enero 2011 a diciembre 2022, región Cusco.</p> <p>b) Identificar el modelo heterocedásticos condicionados que presenta mejor ajuste a la serie precio del petróleo para el período enero 2011 a diciembre 2022, región Cusco.</p> <p>c) Comparar qué enfoque ofrece un mejor pronóstico para el precio del petróleo en la región Cusco de enero a diciembre de 2023: los modelos heterocedásticos condicionados o los modelos basados en la metodología Box-Jenkins.</p>	<p>Hipotesis General El modelo $GARCH(2, 1)$ presenta mejor capacidad predictiva en la serie precio del petróleo frente a los modelos de la metodología de Box-Jenkins durante el período 2011-2023, región Cusco</p> <p>Hipotesis Específicas</p> <p>a) El modelo ARIMA(2, 1, 1) presenta mejor ajuste a la serie precio del petróleo para el período enero 2011 a diciembre 2022, región Cusco</p> <p>b) El modelo GARCH-M presenta mejor ajuste a la serie precio del petróleo para el período enero 2011 a diciembre 2022, región Cusco</p> <p>c) El enfoque de los heterocedásticos condicionados permite un pronóstico más eficiente a la serie precio del petróleo para el período de enero a diciembre de 2023, región Cusco</p>	<p>Antecedentes Internacionales</p> <ul style="list-style-type: none"> ▪ Lazaridis, Ioannis. (2021) presenta, "Análisis del Precio del crudo del petróleo y aplicación de métodos para su predicción" ▪ Zgera, Cristina. (2016) presenta, "Modelado y pronóstico: análisis de precios de Petróleo Crudo(WTI) Utilizando Series Temporales" ▪ Villalba, Camila. (2019) presenta, "Análisis y predicción del precio del petróleo una aproximación basada en modelos VAR <p>Nacional</p> <ul style="list-style-type: none"> ▪ Aquino Villanueva & Francesqui Duran. (2018) presenta, "Efecto del precio internacional del petróleo y del tipo de cambio en la importancia del Diesel de Petroperú" 	<p>Enfoque de la Investigación Según Hernández et al. (2014), el enfoque es cuantitativo el cual es una aproximación de la investigación que se basa en la recolección y análisis de datos numéricos para describir y explicar fenómenos.</p> <p>Tipo de investigación "El tipo de investigación es aplicada, aprovecha los conocimientos logrados por la investigación básica o teórica para el conocimiento y solución de problemas inmediatos" Sánchez Carlessi et al. (2018) Afance de Investigación La investigación que se llevará a cabo para el presente estudio, será descriptiva. Según Hernández et al. (2014) este busca especificar propiedades y características importantes de cualquier fenómeno.</p> <p>Diseño de la Investigación Se llevará a cabo un diseño no experimental ya que el investigador no tendrá una participación en este estudio sobre la predicción del precio del petróleo mediante series de tiempo en la región de Cusco</p> <p>Población La población de este estudio es el precio del petróleo desde enero 2011 a febrero 2023, región Cusco.</p>

.2. Solicitud de datos a Osinerming y INEI

18/5/24, 10:39 a.m. Correo de Universidad Nacional de San Antonio Abad del Cusco - UNSAAC - Correo de Solicitud de Confirmación de Regi...



ROBINSON SANCHEZ TORRES <174958@unsaac.edu.pe>

Correo de Solicitud de Confirmación de Registro

Respuesta de INGRESO <transparencia@inei.gob.pe> 17 de agosto de 2023, 9:16
 Para: Respuesta de Solicitud del Portal de Transparencia <174958@unsaac.edu.pe>

Estimada/o USUARIO:ROBINSON SANCHEZ TORRES

Su Solicitud se grabo con Exito. Sus datos ingresados son:

FUNCIONARIO RESPONSABLE DE ENTREGAR LA INFORMACIÓN:	Milagros Canales Garrido
DATOS INGRESADOS:	
CÓDIGO DE REGISTRO:	TSR-001413-2023
FECHA INGRESO:	2023-08-17 09:16:39
NOMBRE:	ROBINSON SANCHEZ TORRES
TIPO DOCUMENTO:	D.N.I.
NUMERO DE DOCUMENTO:	63093692
PAIS:	Perú
DEPARTAMENTO:	Cusco
PROVINCIA:	CUSCO
DISTRITO:	SAN SEBASTIAN
URBANIZACIÓN:	Urb. Santa Beatriz
AV/CALLE/JR/PSJ.:	Apv. Jesus nuestro salvador
REFERENCIA DEL DOMICILIO:	Frente al aeropuerto, detrás de los paneles
N°/DPTO/INT:	3
TELÉFONO:	960357254
CORREO ELECTRÓNICO:	174958@unsaac.edu.pe
Autorizo que todo acto administrativo derivado del presente procedimiento, se me notifique en el correo electrónico (e-mail) consignado en el presente formulario. (Artículo 20º numeral 20.4 del Texto Único Ordenado de la Ley N° 27444-Ley del Procesamiento Administrativo General, aprobado por Decreto Supremo N° 004-2019-JUS)	SI
INFORMACIÓN SOLICITADA:	SOLICITO DATOS REFERENTES AL PRECIO DEL PETROLEO Y SUS DERIVADOS PARA EL DEPARTAMENTO DE CUSCO PARA EL PERIODO ENERO DEL AÑO 1999 HASTA AGOSTO DEL 2023
DEPENDENCIA DE LA CUAL SE REQUIERE LA INFORMACIÓN:	Dirección Técnica de Indicadores Económicos
FORMA DE ENTREGA DE LA INFORMACIÓN:	CORREO ELECTRÓNICO
ESPECIFICAR FORMA DE ENTREGA DE LA INFORMACIÓN:	
OBSERVACIONES:	

Las solicitudes de acceso a la información pública tienen un plazo de atención de diez (10) días útiles, contabilizados a partir del día siguiente de entregada.

.3. Base de datos de la Institución Nacional de Estadística e Informática

PRECIOS PROMEDIO AL CONSUMIDOR DEL PETRÓLEO Y SUS DERIVADOS EN CUSCO:
ENERO 2011 - FEBRERO 2023(Soles x galón)

AÑO	MES	PETRÓLEO DIÉSEL												
2011	Enero	12.20	2014	Enero	14.41	2017	Enero	11.14	2020	Enero	13.03	2023	Enero	20.13
2011	Febrero	12.20	2014	Febrero	14.38	2017	Febrero	11.10	2020	Febrero	13.05	2023	Febrero	22.07
2011	Marzo	12.21	2014	Marzo	14.56	2017	Marzo	11.07	2020	Marzo	12.96	2023	Marzo	19.61
2011	Abril	12.20	2014	Abril	14.60	2017	Abril	11.05	2020	Abril	12.77	2023	Abril	18.09
2011	Mayo	12.18	2014	Mayo	14.57	2017	Mayo	11.06	2020	Mayo	12.31	2023	Mayo	16.80
2011	Junio	12.58	2014	Junio	14.55	2017	Junio	11.04	2020	Junio	11.69	2023	Junio	16.02
2011	Julio	12.80	2014	Julio	14.49	2017	Julio	10.83	2020	Julio	11.25	2023	Julio	15.68
2011	Agosto	13.21	2014	Agosto	14.48	2017	Agosto	10.77	2020	Agosto	11.13	2023	Agosto	16.19
2011	Setiembre	13.34	2014	Setiembre	14.30	2017	Setiembre	11.00	2020	Setiembre	11.17	2023	Setiembre	16.89
2011	Octubre	13.38	2014	Octubre	14.16	2017	Octubre	11.06	2020	Octubre	11.06	2023	Octubre	17.56
2011	Noviembre	13.99	2014	Noviembre	13.67	2017	Noviembre	11.44	2020	Noviembre	10.82	2023	Noviembre	18.24
2011	Diciembre	13.99	2014	Diciembre	12.82	2017	Diciembre	11.54	2020	Diciembre	10.92	2023	Diciembre	17.92
2012	Enero	13.59	2015	Enero	11.51	2018	Enero	11.91	2021	Enero	11.52			
2012	Febrero	13.63	2015	Febrero	10.50	2018	Febrero	12.01	2021	Febrero	12.32			
2012	Marzo	14.14	2015	Marzo	10.46	2018	Marzo	11.99	2021	Marzo	13.21			
2012	Abril	14.19	2015	Abril	10.48	2018	Abril	11.95	2021	Abril	13.22			
2012	Mayo	14.30	2015	Mayo	10.46	2018	Mayo	12.40	2021	Mayo	13.21			
2012	Junio	14.28	2015	Junio	10.43	2018	Junio	12.59	2021	Junio	13.74			
2012	Julio	13.66	2015	Julio	10.43	2018	Julio	12.71	2021	Julio	14.08			
2012	Agosto	13.68	2015	Agosto	10.43	2018	Agosto	17.78	2021	Agosto	14.74			
2012	Setiembre	14.09	2015	Setiembre	10.42	2018	Setiembre	12.78	2021	Setiembre	15.02			
2012	Octubre	14.09	2015	Octubre	10.36	2018	Octubre	12.81	2021	Octubre	15.74			
2012	Noviembre	14.07	2015	Noviembre	10.33	2018	Noviembre	13.40	2021	Noviembre	16.20			
2012	Diciembre	14.03	2015	Diciembre	10.31	2018	Diciembre	13.63	2021	Diciembre	16.17			
2013	Enero	13.59	2016	Enero	9.63	2019	Enero	13.60	2022	Enero	16.11			
2013	Febrero	13.55	2016	Febrero	9.22	2019	Febrero	13.49	2022	Febrero	16.19			
2013	Marzo	14.11	2016	Marzo	9.18	2019	Marzo	13.05	2022	Marzo	17.00			
2013	Abril	14.14	2016	Abril	9.09	2019	Abril	13.05	2022	Abril	16.16			
2013	Mayo	13.65	2016	Mayo	9.32	2019	Mayo	13.13	2022	Mayo	16.59			
2013	Junio	13.65	2016	Junio	9.51	2019	Junio	13.19	2022	Junio	17.57			
2013	Julio	14.07	2016	Julio	9.99	2019	Julio	13.14	2022	Julio	17.64			
2013	Agosto	14.09	2016	Agosto	10.04	2019	Agosto	13.15	2022	Agosto	17.72			
2013	Setiembre	14.67	2016	Setiembre	10.02	2019	Setiembre	13.08	2022	Setiembre	18.81			
2013	Octubre	14.70	2016	Octubre	10.04	2019	Octubre	13.07	2022	Octubre	19.94			
2013	Noviembre	14.46	2016	Noviembre	10.50	2019	Noviembre	13.02	2022	Noviembre	20.24			
2013	Diciembre	14.41	2016	Diciembre	10.58	2019	Diciembre	12.97	2022	Diciembre	20.21			

^{1/} A partir del mes de agosto 2011, los precios corresponden al gasohol.

Nota: Cambia la denominación de la unidad monetaria "Nuevo Sol" a "Sol", desde el 15 de diciembre del 2015 - Ley N° 30381

FUENTE: INEI-DEIPRE

.4. Validación de datos de Institución Nacional de Estadística e Informática



PERÚ

Ministerio
de Energía y Minas

"Decenio de la Igualdad de Oportunidades para Mujeres y Hombres"
"Año de la unidad, la paz y el desarrollo"

INFORME TÉCNICO N° 047-2023-MINEM/DGH-DPTC

A : Raquel Irene Lazaro Castillo
Jefa de la Oficina de Administración Documentaria y Archivo Central (e)

De : Mishel Efraín Espinoza Ramos
Director de Procesamiento, Transporte y Comercialización de Hidrocarburos y Biocombustibles (d.t.)

Asunto : Solicitud de Acceso a la Información Pública

Referencia : Memorando N° 01967-2023/MINEM-SG-OADAC (Expediente N° 3462446)

Fecha : San Borja, 08 de marzo de 2024

Tengo el agrado de dirigirme a usted, en relación al documento de la referencia, mediante el cual vuestro Despacho al amparo de lo establecido en el Texto Único Ordenado de la Ley N° 27806, Ley de Transparencia y Acceso a la Información Pública, aprobado mediante Decreto Supremo N° 021-2019-JUS, trasladó a esta Dirección la solicitud de información del ciudadano Robinson Sanchez Torres, quien requiere la siguiente información: "SOLICITO DATOS REFERENTES AL PRECIO DEL PETROLEO Y SUS DERIVADOS PARA EL DEPARTAMENTO DE CUSCO PARA EL PERIODO ENERO DEL AÑO 1999 HASTA MARZO DEL 2023, CUYA FINALIDAD ES REALIZAR UN TRABAJO DE INVESTIGACIÓN CON RESPECTO AL PRECIO DEL PETROLEO Y SUS DERIVADOS PARA LA UNIVERSIDAD SAN ANTONIO ABAD DEL CUSCO", con la finalidad que esta Dirección realice la evaluación y atención correspondiente.

I. ANTECEDENTES

- 1.1. Mediante Memorando N° 01967-2023/MINEM-SG-OADAC de fecha 06 de marzo de 2023, la Oficina de Administración Documentaria y Archivo Central – OADAC remitió a la Dirección de Procesamiento, Transporte y Comercialización de Hidrocarburos y Biocombustibles la solicitud de información petitionada por el administrado.

II. BASE LEGAL

- 2.1. Constitución Política del Perú de 1993.
- 2.2. Decreto Supremo N° 021-2019-JUS, que aprueba el Texto Único Ordenado de la Ley N° 27806, Ley de Transparencia y Acceso a la Información Pública.
- 2.3. Decreto Supremo N° 072-2003-PCM, que aprueba el Reglamento de la Ley de Transparencia y Acceso a la Información.
- 2.4. Resolución Secretarial N° 054-2018-MEM/SEG, que aprueba la Directiva N° 011-2018-MEM/SEG, "Directiva Sobre Disposición para la Atención de Solicitud de Acceso a la Información Pública en el Ministerio de Energía y Minas".



PERÚ

Ministerio
de Energía y Minas

"Decenio de la Igualdad de Oportunidades para Mujeres y Hombres"
"Año de la unidad, la paz y el desarrollo"

III. ANÁLISIS

3.1. Al respecto, corresponde a esta Dirección efectuar el análisis de la Solicitud de Acceso a la Información Pública, conforme al marco legal aplicable para tales efectos, debiendo realizar de forma preclusiva el siguiente análisis:

- a) Calificación de Solicitud de Acceso a la Información Pública.
- b) Identificación de información solicitada y búsqueda en acervo documentario.
- c) Determinación de publicidad de información solicitada y no subsunción en causales de excepción al ejercicio de Derecho a la Información Pública.
- d) Conclusión según corresponda.

El referido análisis es preclusivo, de manera que se realizará de acuerdo a como corresponda para cada solicitud.

a) Calificación de Solicitud de Acceso al Información Pública

3.2. El numeral 2.5 del artículo 2 de la Constitución Política del Perú, respecto del Derecho a la información, establece:

*"Toda persona tiene derecho:
A solicitar sin expresión de causa la información que requiera y a recibirla de cualquier entidad pública, en el plazo legal, con el costo que suponga el pedido.
Se exceptúan las informaciones que afectan la intimidad personal y las que expresamente se excluyan por ley o por razones de seguridad nacional".*

- 3.3. Asimismo, el artículo 10 de Texto Único Ordenado de la Ley N° 27806, Ley de Transparencia y Acceso a la Información Pública aprobado mediante Decreto Supremo N° 021-2019-JUS, establece que las entidades de la Administración Pública tienen la obligación de proveer la información requerida si se refiere a la contenida en documentos escritos, fotografías, grabaciones, soporte magnético o digital, o en cualquier otro formato, siempre que haya sido creada u obtenida por ella o que se encuentre en su posesión o bajo su control.
- 3.4. El inciso d) del artículo 10 del Reglamento de la Ley de Transparencia y Acceso a la Información Pública, aprobado por el Decreto Supremo N° 072-2003-PCM, modificado por el Decreto Supremo N° 070-2013-PCM, establece que la solicitud debe contener la expresión concreta y precisa del pedido de información, así como cualquier otro dato que propicie la localización o facilite la búsqueda de la información solicitada.
- 3.5. Ahora bien, por su lado, el numeral 4.11 de la Directiva N° 011-2018-MEM/SEG "Directiva Sobre Disposición para la Atención de Solicitud de Acceso a la Información Pública en el Ministerio de Energía y Minas" establece los supuestos por los cuales se identifica no estar ante una Solicitud de Acceso a la Información Pública, en ese sentido, es pertinente que pueda verificarse si la solicitud ingresada por el administrado se encuentra subsumida o no en algunos de estos supuestos¹.

¹ No constituye Solicitud de Acceso a la Información Pública

4.11 No son considerados como Solicitud de Acceso a la Información Pública los siguientes supuestos:

4.11.1. El requerimiento de copias fedateadas o certificadas que habiliten el ejercicio de derechos preexistentes del administrado, siempre que no afecten derechos de terceros; pues se encuentra referido a un procedimiento de aprobación automática de conformidad con lo establecido en el



PERÚ

Ministerio
de Energía y Minas

"Decenio de la Igualdad de Oportunidades para Mujeres y Hombres"
"Año de la unidad, la paz y el desarrollo"

artículo 13 del Texto Único Ordenado de la Ley N° 27806, Ley de Transparencia y Acceso a la Información Pública, aprobado por Decreto Supremo N° 021-2019-JUS.

Es todo cuanto tengo que informar.

Atentamente,

Firmado digitalmente por ESPINOZA RAMOS
Mishel Efraín FAU 20131368829 soft
Entidad: Ministerio de Energía y Minas
Motivo: Firma del documento
Fecha: 2023/03/08 12:16:05-0500

Ing. Mishel Espinoza Ramos
Director de Procesamiento, Transporte Y
Comercialización de Hidrocarburos y
Biocombustibles (d.t.)



PERÚ

Ministerio
de Energía y Minas

"Decenio de la Igualdad de Oportunidades para Mujeres y Hombres"
"Año de la unidad, la paz y el desarrollo"

b) Identificación de información solicitada y búsqueda en acervo documentario

- 3.6. En esa línea, la información requerida por el administrado es: "SOLICITO DATOS REFERENTES AL PRECIO DEL PETROLEO Y SUS DERIVADOS PARA EL DEPARTAMENTO DE CUSCO PARA EL PERIODO ENERO DEL AÑO 1999 HASTA MARZO DEL 2023, CUYA FINALIDAD ES REALIZAR UN TRABAJO DE INVESTIGACIÓN CON RESPECTO AL PRECIO DEL PETROLEO Y SUS DERIVADOS PARA LA UNIVERSIDAD SAN ANTONIO ABAD DEL CUSCO".
- 3.7. Al respecto, se efectuó la búsqueda de la documentación solicitada en el acervo físico y digital de la Dirección General de Hidrocarburos, motivo por el cual, no es posible atender la presente solicitud; sin embargo, es preciso señalar que dicha información puede ser solicitada ante el INEI.
- 3.8. El artículo 13 del TUO de la Ley N° 27806 establece que la solicitud de acceso a la información pública no implica la obligación de las entidades de la Administración Pública de crear o producir información con la que no cuente o no tenga obligación de contar al momento de efectuarse el pedido.
- 3.9. Siendo ello así, cabe señalar que la administración pública desarrolla sus actividades, analiza y evalúa las solicitudes presentadas por los administrados y emite sus actos y decisiones en el marco del irrestricto respeto y observancia al Principio de Legalidad previsto en el Texto Único Ordenado de la Ley N° 27444, Ley del Procedimiento Administrativo General aprobado mediante Decreto Supremo N° 004-2019-JUS².
- 3.10. En ese sentido, estando a los fundamentos antes esgrimidos al no haberse ubicado la información en el acervo documentario de este Ministerio, no corresponde atender el pedido del administrado, conforme lo dispuesto en el artículo 13 del TUO de la Ley N° 27806.

CONCLUSIÓN

De lo expuesto se concluye:

Conforme a la búsqueda realizada en el acervo documentario de este Ministerio, corresponde denegar la solicitud de acceso a la información pública, en el marco de lo establecido en el

numeral 32.4 del artículo 32 del Texto Único Ordenado de la Ley N° 27444, Ley del Procedimiento Administrativo General (en adelante, TUO de la Ley N° 27444)

4.11.2. Las solicitudes efectuadas por entidades públicas, las mismas son tramitadas conforme lo dispuesto en los artículos 85 al 88 del TUO de la Ley N° 27444.

4.11.3. Las solicitudes que tengan por finalidad obtener un pronunciamiento expreso respecto a las materias que se encuentren bajo el ámbito de competencia de la entidad, de conformidad con lo dispuesto en el artículo 169 del TUO de la Ley N° 27444.

4.11.4. Las solicitudes de información presentadas por los administrados respecto de los procedimientos administrativos en trámite en los cuales son parte, en razón a que las mismas se tramitan de conformidad con lo dispuesto en el artículo 169 del TUO de la Ley N° 27444.

4.11.5. Lo requerimientos de información efectuados por el Poder Judicial, Ministerio Público, Congreso de la República, Policía Nacional y Defensoría del Pueblo en el ejercicio de sus atribuciones, dado que, los mismos son tramitados conforme a la normativa especial que les faculta pedir información.

4.11.6 La solicitud de información no implica la obligación de las entidades de la Administración Pública de crear o producir información con la que no cuente o no tenga obligación de contar al momento de efectuarse el pedido.

4.11.7 Esta ley no faculta que los solicitantes exijan evaluaciones o análisis de la información que posee el MEM.

² El numeral 1.1 del artículo IV del Título Preliminar del Texto Único Ordenado de la Ley N° 27444, Ley del Procedimiento Administrativo General aprobado mediante Decreto Supremo N° 021-2019 establece:

1.1 Principio de Legalidad. – Las autoridades administrativas deben actuar con respeto a la Constitución, la ley y al derecho, dentro de las facultades que le estén atribuidas y de acuerdo con los fines para los que les fueron conferidas.
(...)"



"Decenio de la Igualdad de Oportunidades para Mujeres y Hombres"
"Año de la unidad, la paz y el desarrollo"

INFORME N° 004-2023-INEI/DTIE-DEIPRE

Señora : **LILIA MONTOYA SÁNCHEZ**
Directora Técnica
Dirección Técnica de Indicadores Económicos

Asunto : Solicita precio del petróleo y sus derivados para el departamento de Cusco, periodo enero 1999 - marzo 2023

Referencia : Exp. N° 20230003544

Fecha : Lima, 16 de marzo de 2023

Tengo el agrado de dirigirme a usted, y en relación al asunto indicado, informarle lo siguiente:

ANTECEDENTES

Mediante el documento de la referencia, remitida por el Ministerio de Energía y Minas el Señor Robinson Sánchez Torres, quien en amparo de la Ley de Transparencia y Acceso de la Información Pública- LTAIP, solicita datos referentes al precio del petróleo y sus derivados para el departamento de Cusco para el periodo enero del año 1999 hasta marzo del 2023.

ANÁLISIS

En base a lo expresado en la solicitud, se ha procedido a recabar información de precios promedio al consumidor de la ciudad del Cusco del petróleo y sus derivados.

- Gasolina de 84 octanos desde enero 1999 - junio 2021
- Gasolina de 90 octanos, setiembre 2010 - febrero 2023
- Petróleo Diésel de enero 2011 - febrero 2023

Cabe indicar que, a partir del mes de agosto 2011, los precios de gasolinas de 84 y 90 octanos corresponden a gasohol de 84 y 90 octanos; así también la unidad monetaria cambia de "Nuevo sol" a "Sol" a partir 15 de diciembre 2015.

CONCLUSIONES

Los precios promedio que se tiene disponible del Índice de Precios al Consumidor (IPC) de la ciudad de Cusco, es como se indica en el cuadro adjunto:

- Gasolina 84 octanos para el periodo enero 1999 - julio 2011 y gasohol 84 octanos para el periodo agosto 2011- junio 2021.
- Gasolina 90 octanos para el periodo setiembre 2010 - julio 2011 y Gasohol de 90 octanos para el periodo agosto 2011- febrero 2023; y
- Petróleo diésel desde enero 2011 a febrero 2023.

Atentamente,

Firmado digitalmente por CORDOVA
CUELLAR Jesús Alberto FAU
INEI Dirección Ejecutiva
20131366981 soft
Módulo: Soy el autor del documento
Fecha: 21.03.2023 18:55:20 -05:00

INSTITUTO NACIONAL DE ESTADÍSTICA E INFORMÁTICA

www.gob.pe/inei
Av. Gral. Garzón 658, Jesús María / Lima - Perú
203 2640 / 652 0000
infoinei@inei.gob.pe



Gobierno del Perú