

Informe Práctica Universitaria

Informe Series de Tiempo sobre el rendimiento académico
6° año Politécnico común Ing. Navales (2007-2023)
y 1er año Mc.El. (2010-2023)
Apolinav



Alumno: Benjamín Acuña
Carrera: Ingeniería Civil Matemáticas
Universidad: Universidad Técnica Federico Santa
María
Fecha: Verano 2024



UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Carta de Presentación.

Aquí se presenta un Informe estadístico sobre las notas promedios semestrales de las asignaturas de la Escuela de Ingeniería (ESING) en la Academia Politécnica Naval (Apolinav), centrándose en los cursos “6° año Politécnico Común Ing. Navales” y “1er año Mc.El.” durante los períodos específicos de estudio.

Para el curso “6° año Politécnico Común Ing. Navales”, dirigido a oficiales de la Armada, el estudio abarca el período desde 2007 hasta 2023. Mientras que para el curso “1er año Mc.El.”, dirigido a gente de mar (GM), el período de análisis se extiende desde 2010 hasta 2023.

El objetivo de este informe es analizar las notas promedio semestrales de las asignaturas clave de dichos cursos. Para el curso “6° año Politécnico Común Ing. Navales”, se consideraron las asignaturas fundamentales como Análisis de Circuito I, Cálculo Avanzado, Ecuaciones Diferenciales Ordinarias y Física Mecánica y Ondas. Mientras tanto, para el curso “1er año Mc.El.”, se examinaron asignaturas cruciales como Circuitos Eléctricos I, Electricidad, Física Técnica I y Matemática Técnica I. Estas asignaturas corresponden al 1° semestre de cada año.

Este estudio proporciona un análisis detallado del rendimiento promedio semestral en estas asignaturas a lo largo de los períodos mencionados. Se incluyen gráficos de líneas que muestran la evolución del rendimiento académico en cada asignatura a lo largo del tiempo para ambos cursos. Además, se presentan correlogramas para examinar las relaciones de autocorrelación en los datos temporales, lo que proporciona una comprensión más profunda de los patrones de rendimiento a lo largo de los años. Y por último se propondrán modelos de series de tiempo y la estimación de sus parámetros.

Este informe, fue elaborado por Benjamín Acuña, alumno en práctica, tiene como objetivo proporcionar una herramienta valiosa para la toma de decisiones y la mejora continua en el ámbito académico, al destacar áreas de fortaleza y oportunidades de mejora en el desempeño estudiantil.

Índice

1. La Academia Politécnica Naval de Chile	4
2. Estudio del rendimiento académico de los alumnos del curso “6° año Politécnico común Ing. Navales” usando series de tiempo en las asignaturas:	5
2.1. Análisis de Circuitos Eléctricos I	5
2.1.1. Objetivos específicos:	6
2.1.2. Preguntas de investigación	6
2.1.3. Metodología	6
2.1.4. Estadística Descriptiva de las notas promedios semestrales de los estudiantes	7
2.1.5. Pruebas de estacionariedad	9
2.1.6. Selección de un modelo	11
2.1.7. Estimación de parámetros	11
2.1.8. Resumen y conclusión	11
2.2. Calculo Avanzado	13
2.2.1. Objetivos específicos:	13
2.2.2. Preguntas de investigación	13
2.2.3. Metodología	13
2.2.4. Estadística Descriptiva de las notas promedios semestrales de los estudiantes	14
2.2.5. Pruebas de estacionariedad	16
2.2.6. Selección de un modelo	18
2.2.7. Estimación de parámetros	18
2.2.8. Resumen y conclusión	19
2.3. Ecuaciones Diferenciales Ordinarias	20
2.3.1. Objetivos específicos:	20
2.3.2. Preguntas de investigación	20
2.3.3. Metodología	20
2.3.4. Estadística Descriptiva de las notas promedios semestrales de los estudiantes	21
2.3.5. Pruebas de estacionariedad	23
2.3.6. Selección de un modelo	25
2.3.7. Estimación de parámetros	25
2.3.8. Resumen y conclusión	26
2.4. Física Mecánica y Ondas	27
2.4.1. Objetivos específicos:	27
2.4.2. Preguntas de investigación	27
2.4.3. Metodología	27
2.4.4. Estadística Descriptiva de las notas promedios semestrales de los estudiantes	28
2.4.5. Pruebas de estacionariedad	30
2.4.6. Selección de un modelo	32
2.4.7. Estimación de parámetros	32
2.4.8. Resumen y conclusión	33
3. Estudio del rendimiento académico de los alumnos del curso “1er año Mc.El.” usando series de tiempo en las asignaturas:	34
3.1. Circuitos Eléctricos I	34
3.1.1. Objetivos específicos:	35
3.1.2. Preguntas de investigación	35

3.1.3.	Metodología	35
3.1.4.	Estadística Descriptiva de las notas promedios semestrales de los estudiantes	36
3.1.5.	Pruebas de estacionariedad	38
3.1.6.	Selección de un modelo	40
3.1.7.	Estimación de parámetros	40
3.1.8.	Resumen y conclusión	41
3.2.	Electricidad	42
3.2.1.	Objetivos específicos:	42
3.2.2.	Preguntas de investigación	42
3.2.3.	Metodología	42
3.2.4.	Estadística Descriptiva de las notas promedios semestrales de los estudiantes	43
3.2.5.	Pruebas de estacionariedad	45
3.2.6.	Selección de un modelo	47
3.2.7.	Estimación de parámetros	47
3.2.8.	Resumen y conclusión	48
3.3.	Física Técnica I	49
3.3.1.	Objetivos específicos:	49
3.3.2.	Preguntas de investigación	49
3.3.3.	Metodología	49
3.3.4.	Estadística Descriptiva de las notas promedios semestrales de los estudiantes	50
3.3.5.	Pruebas de estacionariedad	52
3.3.6.	Selección de un modelo	54
3.3.7.	Estimación de parámetros	54
3.3.8.	Resumen y conclusión	54
3.4.	Matemática Técnica I	56
3.4.1.	Objetivos específicos:	56
3.4.2.	Preguntas de investigación	56
3.4.3.	Metodología	56
3.4.4.	Estadística Descriptiva de las notas promedios semestrales de los estudiantes	57
3.4.5.	Pruebas de estacionariedad	59
3.4.6.	Selección de un modelo	61
3.4.7.	Estimación de parámetros	61
3.4.8.	Resumen y conclusión	62

1. La Academia Politécnica Naval de Chile

La Academia Politécnica Naval de Chile (Apolinav) es una institución de educación superior con una larga trayectoria en la formación de oficiales y personal de mar para la Armada de Chile. Fundada con el propósito de impartir una educación de excelencia en el ámbito naval, la Apolinav ha sido reconocida por su compromiso con la formación integral de sus estudiantes y su contribución al desarrollo y modernización de la Armada.

La estructura académica de la Apolinav se compone de diversas escuelas, cada una orientada a la formación de profesionales en áreas específicas de la ingeniería, la navegación y la tecnología naval. Entre estas escuelas se encuentran la Escuela de Ingeniería (ESING), la Escuela de Grumetes y la Escuela de Abastecimiento Naval, entre otras. Cada una de estas escuelas ofrece programas de estudio adaptados a las necesidades y requerimientos de los diferentes cuerpos de la Armada, proporcionando una formación especializada y de alta calidad.

Los cursos impartidos en la Apolinav varían en duración y contenido, dependiendo del perfil del estudiante y sus objetivos profesionales. Por ejemplo, los cursos dirigidos a oficiales de la Armada suelen tener una duración más prolongada y abarcar un amplio espectro de materias, mientras que los cursos destinados al personal de mar se centran en aspectos prácticos y técnicos relacionados con la navegación y el mantenimiento naval.

En este contexto, el presente informe se enfoca en el análisis de las notas promedio semestrales de dos cursos específicos impartidos en la Escuela de Ingeniería de la Apolinav: el “6° año Politécnico Común Ing. Navales” y el “1er año Mc.El.”. Estos cursos representan dos perfiles distintos de estudiantes, con diferentes necesidades y exigencias académicas, lo que permite un análisis comparativo del rendimiento estudiantil en distintos contextos.



2. Estudio del rendimiento académico de los alumnos del curso “6° año Politécnico común Ing. Navales” usando series de tiempo en las asignaturas:

El curso “6° año politécnico común Ing. Navales”, ofrecido por la Escuela de Ingeniería (ESING) de la Apolinav (Academia Politécnica Naval) para oficiales, desempeña un papel fundamental en la formación de futuros Ingenieros Navales. En este contexto, se realiza un exhaustivo análisis utilizando series de tiempo sobre una base de datos que incluye el promedio de 266 alumnos, considerando tanto aprobados como separados de curso, recopilados desde el año 2007 hasta el 2023. A partir de esta información, se calculan los promedios semestrales de las asignaturas más relevantes.

Este riguroso estudio se centra en las cuatro asignaturas principales del curso: Análisis de Circuitos Eléctricos I, Cálculo Avanzado, Ecuaciones Diferenciales Ordinarias y Física Mecánica y Ondas. Estas son las asignaturas que han mostrado mayor dificultad estos ultimo años.

La Figura 1 proporciona los datos necesarios para la realización de este estudio.

	ANALISIS DE CIRCUITOS ELECTRICOS I	CALCULO AVANZADO	ECUACIONES DIFERENCIALES ORDINARIAS	FISICA MECANICA Y ONDAS
AÑO				
2007	5.379412	4.882353	4.615294	5.251176
2008	4.677778	4.330000	4.285556	4.780000
2009	4.713846	3.590000	3.149615	5.107308
2010	4.892143	3.647857	3.482143	4.876429
2011	4.450769	3.245385	3.236154	4.490000
2012	4.656667	2.958667	2.931333	3.772667
2013	4.815000	3.529000	3.112000	3.868000
2014	5.365000	4.000000	4.190000	4.255000
2015	4.185238	3.919048	3.403333	3.739048
2016	4.100500	3.731500	3.605000	3.878500
2017	4.262174	3.829130	3.849130	4.056087
2018	3.831429	4.181429	3.227857	4.119286
2019	3.528095	3.447143	3.314762	3.862381
2020	5.067143	4.264286	3.581429	4.246429
2021	4.439444	3.532778	3.210000	4.488333
2022	4.190000	4.822000	3.373333	4.727333
2023	4.302857	4.735714	3.705714	4.850000

Figura 1: Base de datos con el promedio por asignatura correspondiente a los años de estudio.

2.1. Análisis de Circuitos Eléctricos I

Durante el análisis de la asignatura Análisis de Circuitos Eléctricos I, se examinaron los promedios obtenidos por semestre a lo largo de los 17 años de estudio. En promedio, los alumnos finalizaban la asignatura con una calificación de 4.52. Sin embargo, se observaron variaciones significativas en el rendimiento a lo largo de los años. En particular, en el año 2019 se registró el rendimiento más bajo como curso, con un promedio de 3.53. Por el contrario, en el año 2007 se

alcanzó el rendimiento más alto como curso, con un promedio de 5.38, destacando como un año excepcional en términos de desempeño académico en esta asignatura.

2.1.1. Objetivos específicos:

- Analizar la evolución del rendimiento estudiantil en la asignatura de Análisis de Circuitos Eléctricos I, dentro del marco del curso “6° año Politécnico común Ing. Navales”, impartido en la Escuela de Ingeniería.
- Elaborar un modelo adecuado que se ajuste con mayor precisión al desempeño de los estudiantes en la asignatura de Análisis de Circuitos Eléctricos I, perteneciente al curso “6° año Politécnico común Ing. Navales”, dictado en la Escuela de Ingeniería.

2.1.2. Preguntas de investigación

- ¿Se identifica alguna tendencia en el rendimiento de los estudiantes en la asignatura de Análisis de Circuitos Eléctricos I, dentro del curso “6° año Politécnico común Ing. Navales”, impartido en la Escuela de Ingeniería?
- ¿Cuál es el modelo más adecuado para describir el desempeño de los estudiantes en la asignatura de Análisis de Circuitos Eléctricos I, dentro del curso “6° año Politécnico común Ing. Navales”, dictado en la Escuela de Ingeniería?

2.1.3. Metodología

La metodología empleada en este estudio abarcó una serie de pasos fundamentales para el análisis exhaustivo de la serie temporal en cuestión. En primer lugar, se realizó un análisis descriptivo detallado de los datos, con el fin de comprender su comportamiento general, identificar posibles tendencias y visualizar patrones significativos. Se incluyó un gráfico de línea para representar visualmente la evolución de los datos, añadiendo una línea de tendencia para resaltar cualquier dirección en el conjunto de datos.

Posteriormente, se procedió a generar gráficos de las funciones de autocorrelación ACF y autocorrelación parcial PACF. Estas visualizaciones jugaron un papel crucial en la evaluación preliminar de la estacionariedad de los datos y en la identificación de posibles estructuras de autocorrelación. Estos correlogramas proporcionaron una comprensión visual inicial de la relación entre las observaciones en diferentes puntos temporales.

Con base en las observaciones obtenidas de los correlogramas, se llevó a cabo una evaluación más rigurosa de la estacionariedad de la serie temporal. Para ello, se realizaron pruebas de estacionariedad utilizando tanto el test ADF (Augmented Dickey-Fuller) como el test KPSS (Kwiatkowski, Phillips, Schmidt, Shin). Estas pruebas estadísticas sirvieron para reafirmar o refutar las hipótesis extraídas de los correlogramas, proporcionando una evaluación más sólida de la estacionariedad de la serie temporal.

Finalmente, una vez confirmada la estacionariedad de los datos, se procedió a la selección del modelo ARIMA más adecuado para el conjunto de datos analizados. Esto implicó la identificación de los parámetros óptimos del modelo ARIMA, aprovechando tanto la información derivada de los correlogramas como de las pruebas estadísticas de estacionariedad.

En resumen, la metodología empleada abarcó desde la exploración inicial de los datos y la representación gráfica, pasando por el análisis de la estacionariedad con correlogramas y pruebas estadísticas, hasta la selección y estimación de un modelo ARIMA apropiado. Este enfoque integral

garantizó un análisis completo y riguroso de la serie temporal, proporcionando una base sólida para su modelado y predicción.

2.1.4. Estadística Descriptiva de las notas promedios semestrales de los estudiantes

Para realizar un estudio efectivo de la asignatura en estudio, se inició reduciendo la base de datos para incluir únicamente las columnas relevantes (Figura 2). Posteriormente, se llevó a cabo un análisis exploratorio de datos centrado en el rendimiento promedio de los estudiantes en la asignatura de Análisis de Circuitos Eléctricos I a lo largo de 17 años consecutivos. Este análisis se realizó principalmente utilizando Python y técnicas de análisis de series de tiempo.

Durante este proceso, se realizaron cálculos para obtener estadísticas descriptivas iniciales relacionadas con el rendimiento promedio de los estudiantes en Análisis de Circuitos Eléctricos I. Luego, se procedió a la creación de gráficos de series de tiempo, los cuales permitieron visualizar la evolución del rendimiento promedio a lo largo de los años. Además, se realizó un análisis de tendencias para identificar patrones o cambios significativos en el rendimiento a lo largo del tiempo. Este enfoque proporcionó una comprensión más profunda del comportamiento de los estudiantes en la asignatura a lo largo del período de estudio.

ANALISIS DE CIRCUITOS ELECTRICOS I	
AÑO	
2007	5.379412
2008	4.677778
2009	4.713846
2010	4.892143
2011	4.450769
2012	4.656667
2013	4.815000
2014	5.365000
2015	4.185238
2016	4.100500
2017	4.262174
2018	3.831429
2019	3.528095
2020	5.067143
2021	4.439444
2022	4.190000
2023	4.302857

Figura 2: Base de datos Análisis de Circuitos Eléctricos I.

Se observó que la nota mínima promedio semestral de los estudiantes en la asignatura de Análisis de Circuitos Eléctricos I fue de 3.53 en el año 2019. Por otro lado, se registró una nota promedio anual máxima de 5.38 en el año 2007. En cuanto al desempeño promedio, este se situó en 4.52, con una desviación estándar de 0.50.

Datos	17
Promedio	4.521029
Desviación estándar	0.502399
Mínimo	3.528095
25 %	4.190000
50 %	4.450769
75 %	4.815000
Máximo	5.379412

Cuadro 1: Desempeño de los estudiantes en Análisis de Circuitos Eléctricos I.

La Figura 3 muestra las notas promedios del curso por año a lo largo del periodo. La ecuación de la tendencia, $y = -0,05x + 4,94$, representa una relación lineal entre el año (x) y el promedio de las notas de la asignatura de Análisis de Circuitos Eléctricos I (y). En este caso, el coeficiente de $-0,05$ indica la pendiente de la recta de tendencia, lo que sugiere una disminución lineal en el promedio de las notas a medida que pasa el tiempo. Por otro lado, el término independiente ($4,94$) representa el valor esperado del promedio de las notas en el año cero del estudio (en este caso, 2007 ajustado a $x = 0$).

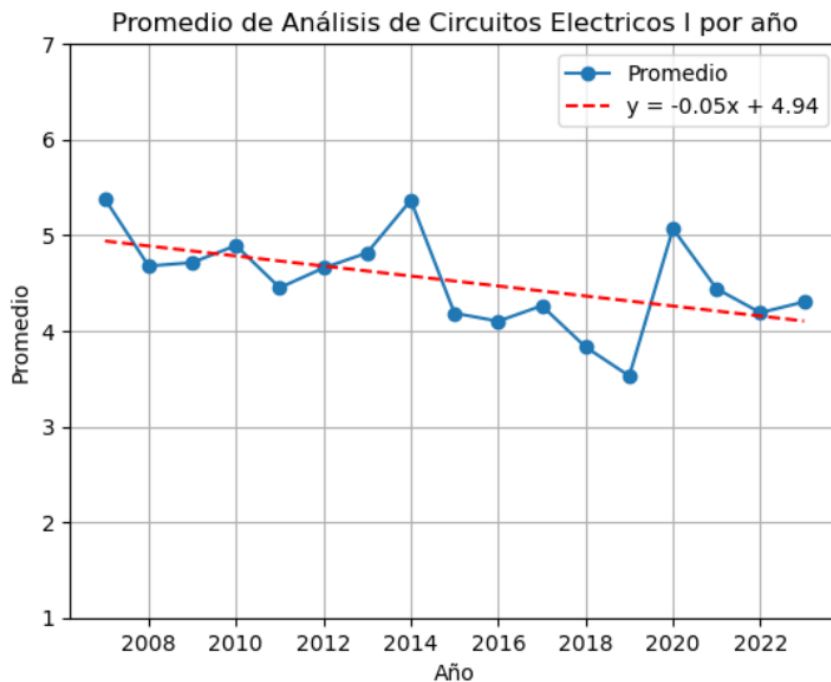


Figura 3: Serie de tiempo de las notas promedios semestrales de los estudiantes en Análisis de Circuitos Eléctricos I.

Al interpretar la ecuación, podemos concluir que la tendencia general a lo largo de los años muestra una disminución de aproximadamente 0,05 unidades en el promedio de las notas de la asignatura de Análisis de Circuitos Eléctricos I por año. Es decir, en promedio, se espera que el rendimiento de los estudiantes en esta asignatura disminuya en 0,05 puntos cada semestre. Sin embargo, es importante tener en cuenta que esta es una simplificación y que pueden existir otros factores no considerados en la ecuación que influyan en el rendimiento académico.

Es necesario analizar la naturaleza de los datos para seleccionar el modelo correcto y mostrar si se trata de datos cíclicos sin tendencia o con actividad estacional.

2.1.5. Pruebas de estacionariedad

Se examinaron los gráficos de la función de autocorrelación (ACF) y de la función de autocorrelación parcial (PACF) de las notas promedio semestrales de los estudiantes en Análisis de Circuitos Eléctricos I. Estos gráficos, mostrados en la Figura 4, proporcionaron información inicial sobre la estacionariedad de la serie temporal.

En los gráficos ACF y PACF, no se observaron picos significativos en ningún retraso, y además tenemos correlaciones positivas y negativas, lo que sugiere que la serie temporal podría ser estacionaria o que la autocorrelación ha disminuido. Esto indica una posible estacionariedad en la serie de datos, aunque se requieren pruebas estadísticas adicionales para confirmar esta observación.

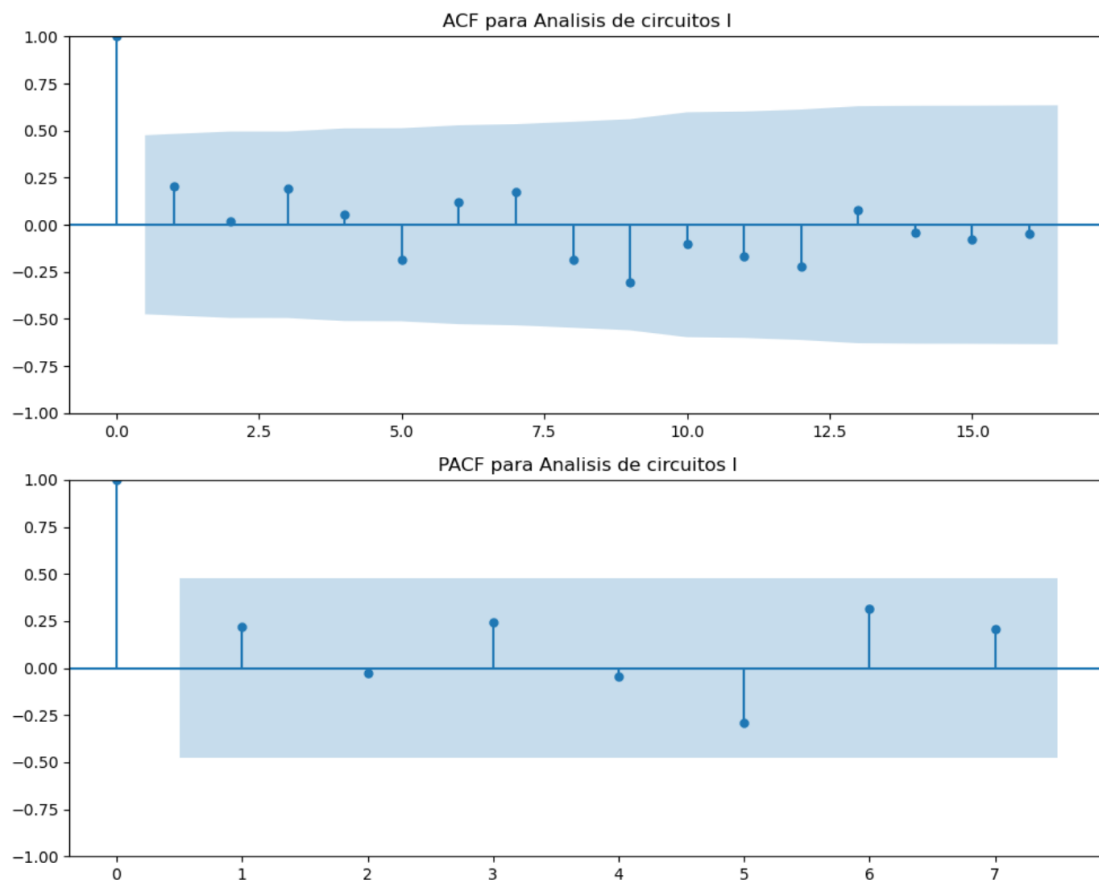


Figura 4: Gráfico ACF y PACF de las notas promedio semestrales de los estudiantes en Análisis de Circuitos Eléctricos I.

Con estas observaciones preliminares, se procedió a realizar las pruebas de estacionariedad ADF y KPSS para verificar la presencia de estacionariedad en la serie temporal.

Un ciclo estacionario tiene una media y una varianza constante en el tiempo y no hay patrones en el ciclo. Es necesario que la serie sea estacionaria antes de continuar con la estimación de un modelo ARIMA, ya que en este análisis se utilizó la prueba Augmented Dickey-Fuller (ADF) y la prueba de Kwiatkowski, Phillips, Schmidt y Shin (KPSS) para verificar la estacionariedad en el rendimiento académico de los estudiantes en métodos cuantitativos.

-
- Probamos la hipótesis para la prueba ADF de que;

- H0: La secuencia no es estacionaria.
- H1: La secuencia estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es estacionaria.

Parámetro	Valor
Estadística ADF	-3.35920
Valor-p	0.01243
Lags utilizados	0
Valor critico (1 %)	-3.92402
Valor critico (5 %)	-3.06850
Valor critico (10 %)	-2.67389

Cuadro 2: Resultado de la prueba augmented Dickey-Fuller.

- Probamos la hipótesis para la prueba KPSS de que;

- H0: La secuencia es estacionaria.
- H1: La secuencia no es estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es no estacionaria.

Parámetro	Valor
Estadística KPSS	0.45093
Valor-p	0.05520
Lags utilizados	1
Valor critico (1 %)	0.739
Valor critico (5 %)	0.463
Valor critico (10 %)	0.347

Cuadro 3: Resultado de la prueba KPSS.

Basándonos en los resultados de las pruebas de estacionariedad ADF y KPSS mostrados en el Cuadro 2 y Cuadro 3, podemos tomar decisiones sobre la estacionariedad de la secuencia.

Para la prueba ADF, la hipótesis nula (H0) establece que la secuencia no es estacionaria, mientras que la hipótesis alternativa (H1) sugiere lo contrario. Con un nivel de significancia del 5 %, el valor-p asociado a la prueba ADF es 0.01243, que es inferior al umbral crítico de 0.05. Por lo tanto, rechazamos la hipótesis nula (H0) y concluimos que la secuencia es estacionaria.

En contraste, para la prueba KPSS, la H0 postula que la secuencia es estacionaria, mientras que la H1 sugiere que no lo es. Aquí, con un nivel de significancia del 5 %, el valor-p asociado a la prueba KPSS es 0.05520, que excede el umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H0) y concluimos que la secuencia es estacionaria.

Estos resultados sugieren una consistencia en la estacionariedad de la secuencia, como se evidencia por la reafirmación de la estacionariedad a través de la prueba ADF y la falta de evidencia

suficiente para rechazar la estacionariedad a través de la prueba KPSS, confirmando así nuestra suposición realizada en el análisis de los gráficos ACF y PACF.

2.1.6. Selección de un modelo

Ahora procederemos a seleccionar el modelo ARIMA más apropiado para nuestra serie temporal. Utilizaremos las capacidades de Python para explorar y evaluar diversos modelos, con el fin de identificar aquel que mejor se ajuste a nuestras necesidades. Este análisis nos proporcionará la base necesaria para tomar decisiones informadas respecto a la elección del modelo más adecuado.

MODELO	AIC	BIC	HQ
ARIMA (1,0,1)	30.805	34.138	31.136
ARIMA (0,0,0)	27.809	29.475	27.975
ARIMA (1,0,0)	28.948	31.448	29.197

Cuadro 4: Identificación del modelo.

El modelo con los mínimos Criterios de Información de Akaike (AIC), Criterio de Información Bayesiano (BIC) y Hannan-Quinn (HQ) es el más adecuado para la secuencia. Por lo tanto, al probar todos los modelos competidores en el Cuadro 4, el modelo ARIMA (0, 0, 0) tiene los valores más bajos y, por lo tanto, el mejor modelo predictivo.

2.1.7. Estimación de parámetros

El Cuadro 5 a continuación muestra estimaciones de los parámetros del sistema ARIMA (0, 0, 0).

Tipo	Coefficiente	Error estándar	Valor Z	Valor-p
Constante	4.5210	0.118	38.246	0.000
sigma2 (σ^2)	0.2376	0.093	2.552	0.011

Cuadro 5: Muestra de la estimación de parámetros.

El modelo ARIMA(0,0,0) indica que estamos utilizando un modelo autorregresivo integrado de media móvil con cero componentes autorregresivas (AR), cero componentes de integración (I) y cero componentes de media móvil (MA). Esto significa que no hay dependencia de los valores pasados, ni tendencia ni error en los datos.

En cuanto a la estimación de parámetros, la constante tiene un valor de 4,5210 con un error estándar de 0,118. El valor Z es alto (38,246), lo que indica que la constante es estadísticamente significativa. Además, el valor-p asociado es muy bajo (0,000), lo que sugiere una alta significancia estadística.

En resumen, con un modelo ARIMA(0,0,0), se está utilizando una estimación constante para predecir la serie temporal de notas promedio semestrales en Análisis de Circuitos Eléctricos I. La constante estimada parece ser estadísticamente significativa, lo que sugiere que este modelo puede ser útil para predecir las notas con un nivel aceptable de precisión.

2.1.8. Resumen y conclusión

Los resultados del análisis revelaron una variabilidad significativa en el rendimiento de los estudiantes en la asignatura de Análisis de Circuitos Eléctricos I a lo largo de los 17 años de

estudio. En promedio, los alumnos finalizaron la asignatura con una calificación de 4.52, pero se observaron fluctuaciones destacadas en diferentes períodos. Por ejemplo, el año 2019 registró el rendimiento más bajo con un promedio de 3.53. Por el contrario, el año 2007 destacó como el de mejor desempeño, con un promedio de 5.38.

Los objetivos específicos de este estudio incluyeron analizar la evolución del rendimiento estudiantil en la asignatura y elaborar un modelo adecuado que se ajustara con precisión a dicho desempeño. Para lograr estos objetivos, se implementó una metodología integral que comenzó con un análisis descriptivo detallado de los datos, seguido de gráficos de las funciones de autocorrelación ACF y PACF para identificar estacionariedad. Además, se utilizó el test ADF y el test KPSS para respaldar las hipótesis obtenidas del análisis de los gráficos de autocorrelación.

Basándonos en los resultados de las pruebas estadísticas, se seleccionó el modelo ARIMA más apropiado utilizando los criterios de información AIC, BIC y HQ. Se encontró que el modelo ARIMA(0,0,0) era el más adecuado para la serie temporal estudiada. Este modelo implica una constante de 4.5210, que resultó ser estadísticamente significativa. Si bien este modelo simple no captura la dinámica compleja de la serie temporal, proporciona una base sólida para la predicción del rendimiento académico en la asignatura de Análisis de Circuitos Eléctricos I.

En resumen, este estudio proporcionó una visión completa del rendimiento estudiantil en la asignatura de Análisis de Circuitos Eléctricos I, desde un análisis descriptivo inicial hasta la selección y estimación de un modelo ARIMA adecuado. Los resultados obtenidos contribuyen a una mejor comprensión de los factores que influyen en el desempeño académico de los estudiantes en esta área específica y pueden servir como base para futuras investigaciones y acciones de mejora educativa.

2.2. Cálculo Avanzado

Durante el análisis de la asignatura de Cálculo Avanzado, se examinaron los promedios obtenidos por semestre a lo largo de los 17 años de estudio. En promedio, los alumnos finalizaban la asignatura con una calificación de 3.92. Sin embargo, se observaron variaciones significativas en el rendimiento a lo largo de los años. En particular, en el año 2012 se registró el rendimiento más bajo como curso, con un promedio de 2.96. Por el contrario, en el año 2007 se alcanzó el rendimiento más alto como curso, con un promedio de 4.88.

2.2.1. Objetivos específicos:

- Analizar la evolución del rendimiento estudiantil en la asignatura de Cálculo Avanzado, dentro del marco del curso “6° año Politécnico común Ing. Navales”, impartido en la Escuela de Ingeniería.
- Elaborar un modelo adecuado que se ajuste con mayor precisión al desempeño de los estudiantes en la asignatura de Cálculo Avanzado, perteneciente al curso “6° año Politécnico común Ing. Navales”, dictado en la Escuela de Ingeniería.

2.2.2. Preguntas de investigación

- ¿Se identifica alguna tendencia en el rendimiento de los estudiantes en la asignatura de Cálculo Avanzado, dentro del curso “6° año Politécnico común Ing. Navales”, impartido en la Escuela de Ingeniería?
- ¿Cuál es el modelo más adecuado para describir el desempeño de los estudiantes en la asignatura de Cálculo Avanzado, dentro del curso “6° año Politécnico común Ing. Navales”, dictado en la Escuela de Ingeniería?

2.2.3. Metodología

La metodología empleada en este estudio abarcó una serie de pasos fundamentales para el análisis exhaustivo de la serie temporal en cuestión. En primer lugar, se realizó un análisis descriptivo detallado de los datos, con el fin de comprender su comportamiento general, identificar posibles tendencias y visualizar patrones significativos. Se incluyó un gráfico de línea para representar visualmente la evolución de los datos, añadiendo una línea de tendencia para resaltar cualquier dirección en el conjunto de datos.

Posteriormente, se procedió a generar gráficos de las funciones de autocorrelación ACF y autocorrelación parcial PACF. Estas visualizaciones jugaron un papel crucial en la evaluación preliminar de la estacionariedad de los datos y en la identificación de posibles estructuras de autocorrelación. Estos correlogramas proporcionaron una comprensión visual inicial de la relación entre las observaciones en diferentes puntos temporales.

Con base en las observaciones obtenidas de los correlogramas, se llevó a cabo una evaluación más rigurosa de la estacionariedad de la serie temporal. Para ello, se realizaron pruebas de estacionariedad utilizando tanto el test ADF (Augmented Dickey-Fuller) como el test KPSS (Kwiatkowski, Phillips, Schmidt, Shin). Estas pruebas estadísticas sirvieron para reafirmar o refutar las hipótesis extraídas de los correlogramas, proporcionando una evaluación más sólida de la estacionariedad de la serie temporal.

Finalmente, una vez confirmada la estacionariedad de los datos, se procedió a la selección del modelo ARIMA más adecuado para el conjunto de datos analizados. Esto implicó la identificación

de los parámetros óptimos del modelo ARIMA, aprovechando tanto la información derivada de los correlogramas como de las pruebas estadísticas de estacionariedad.

En resumen, la metodología empleada abarcó desde la exploración inicial de los datos y la representación gráfica, pasando por el análisis de la estacionariedad con correlogramas y pruebas estadísticas, hasta la selección y estimación de un modelo ARIMA apropiado. Este enfoque integral garantizó un análisis completo y riguroso de la serie temporal, proporcionando una base sólida para su modelado y predicción.

2.2.4. Estadística Descriptiva de las notas promedios semestrales de los estudiantes

Para realizar un estudio efectivo de la asignatura en estudio, se inició reduciendo la base de datos para incluir únicamente las columnas relevantes (Figura 5). Posteriormente, se llevó a cabo un análisis exploratorio de datos centrado en el rendimiento promedio de los estudiantes en la asignatura de Calculo Avanzado a lo largo de 17 años consecutivos. Este análisis se realizó principalmente utilizando Python y técnicas de análisis de series de tiempo.

Durante este proceso, se realizaron cálculos para obtener estadísticas descriptivas iniciales relacionadas con el rendimiento promedio de los estudiantes en Calculo Avanzado. Luego, se procedió a la creación de gráficos de series de tiempo, los cuales permitieron visualizar la evolución del rendimiento promedio a lo largo de los años. Además, se realizó un análisis de tendencias para identificar patrones o cambios significativos en el rendimiento a lo largo del tiempo. Este enfoque proporcionó una comprensión más profunda del comportamiento de los estudiantes en la asignatura a lo largo del período de estudio.

CALCULO AVANZADO	
AÑO	
2007	4.882353
2008	4.330000
2009	3.590000
2010	3.647857
2011	3.245385
2012	2.958667
2013	3.529000
2014	4.000000
2015	3.919048
2016	3.731500
2017	3.829130
2018	4.181429
2019	3.447143
2020	4.264286
2021	3.532778
2022	4.822000
2023	4.735714

Figura 5: Base de datos Calculo Avanzado.

Se observó que la nota mínima promedio semestral de los estudiantes en la asignatura de Calculo Avanzado fue de 2.96 en el año 2012. Por otro lado, se registró una nota promedio anual máxima de 4.88 en el año 2007. En cuanto al desempeño promedio, este se situó en 3.92, con una desviación estándar de 0.55.

Datos	17
Promedio	3.920370
Desviación estándar	0.553338
Mínimo	2.958667
25 %	3.532778
50 %	3.829130
75 %	4.264286
Máximo	4.882353

Cuadro 6: Desempeño de los estudiantes en Calculo Avanzado.

La Figura 6 muestra las notas promedios del curso por año. La ecuación de la tendencia, $y = 0,02x + 3,73$, representa una relación lineal entre el año (x) y el promedio de las notas de la asignatura de Calculo Avanzado (y). En este caso, el coeficiente de 0,02 indica la pendiente de la recta de tendencia, lo que sugiere un aumento lineal en el promedio de las notas a medida que pasa el tiempo. Por otro lado, el término independiente (3,73) representa el valor esperado del promedio de las notas en el año cero del estudio (en este caso, 2007 ajustado a $x = 0$).

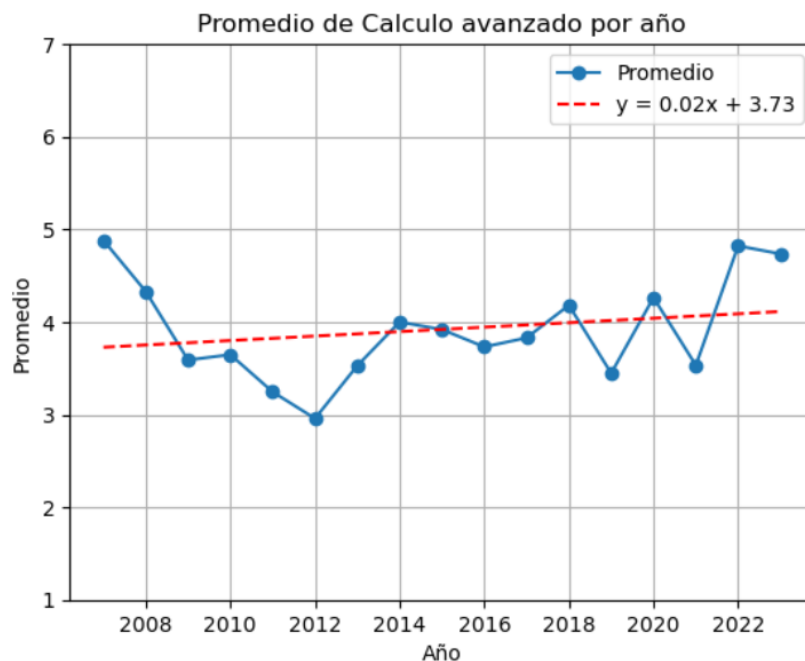


Figura 6: Serie de tiempo de las notas promedios semestrales de los estudiantes en Calculo Avanzado.

Al interpretar la ecuación, podemos concluir que la tendencia general a lo largo de los años muestra un aumento de aproximadamente 0.02 unidades en el promedio de las notas de la asignatura de Calculo Avanzado por año. Es decir, en promedio, se espera que el rendimiento de los estudiantes en esta asignatura aumente en 0.02 puntos cada semestre. Sin embargo, es importante

tener en cuenta que esta es una simplificación y que pueden existir otros factores no considerados en la ecuación que influyan en el rendimiento académico.

Es necesario analizar la naturaleza de los datos para seleccionar el modelo correcto y mostrar si se trata de datos cíclicos sin tendencia o con actividad estacional.

2.2.5. Pruebas de estacionariedad

Se examinaron los gráficos de la función de autocorrelación (ACF) y de la función de autocorrelación parcial (PACF) de las notas promedio semestrales de los estudiantes en Calculo Avanzado. Estos gráficos, mostrados en la Figura 7, proporcionaron información inicial sobre la estacionariedad de la serie temporal.

En los gráficos ACF y PACF, no se observaron picos significativos en ningún retraso, y además tenemos correlaciones positivas y negativas, lo que sugiere que la serie temporal podría ser estacionaria o que la autocorrelación ha disminuido. Esto indica una posible estacionariedad en la serie de datos, aunque se requieren pruebas estadísticas adicionales para confirmar esta observación.

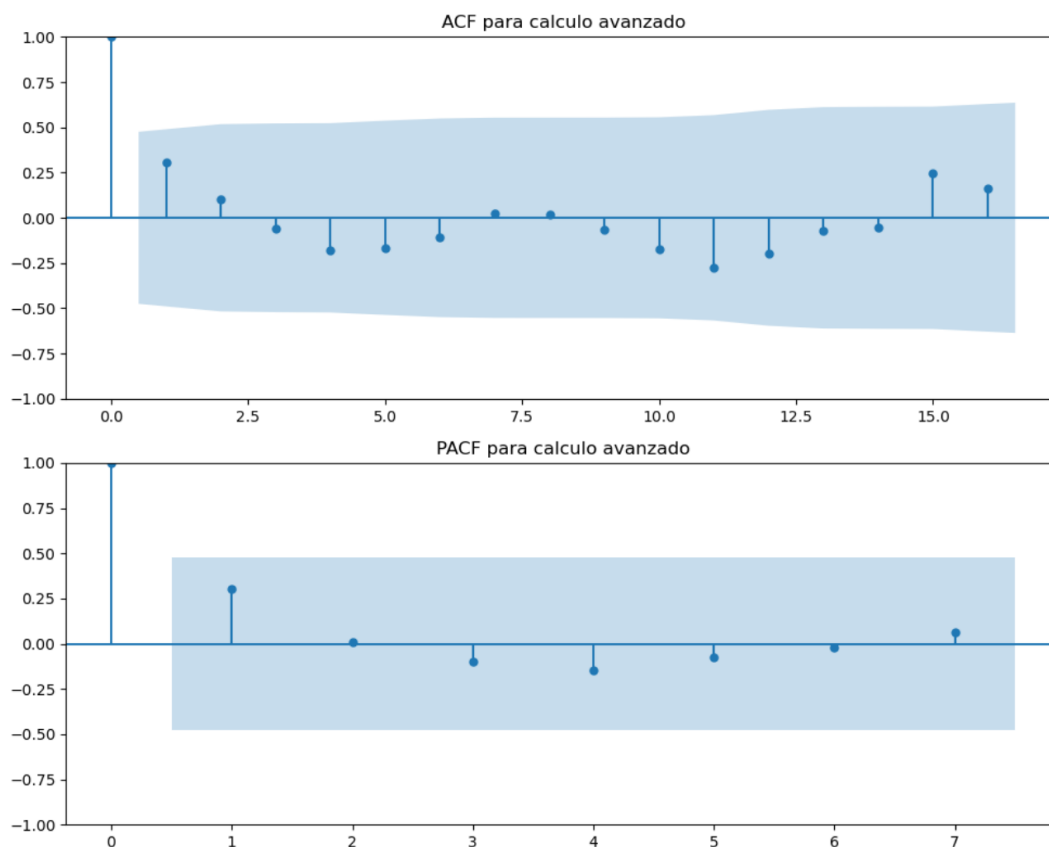


Figura 7: Gráfico ACF y PACF de las notas promedio semestrales de los estudiantes en Calculo Avanzado.

Con estas observaciones preliminares, se procedió a realizar las pruebas de estacionariedad ADF y KPSS para verificar la presencia de estacionariedad en la serie temporal.

Un ciclo estacionario tiene una media y una varianza constante en el tiempo y no hay patrones en el ciclo. Es necesario que la serie sea estacionaria antes de continuar con la estimación de un modelo ARIMA, ya que en este análisis se utilizó la prueba Augmented Dickey-Fuller (ADF) y la

prueba de Kwiatkowski, Phillips, Schmidt y Shin (KPSS) para verificar la estacionariedad en el rendimiento académico de los estudiantes en métodos cuantitativos.

- Probamos la hipótesis para la prueba ADF de que;
 - H0: La secuencia no es estacionaria.
 - H1: La secuencia estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es estacionaria.

Parámetro	Valor
Estadística ADF	0.60985
Valor-p	0.98786
Lags utilizados	6.00000
Valor critico (1 %)	-4.33157
Valor critico (5 %)	-3.23295
Valor critico (10 %)	-2.7487

Cuadro 7: Resultado de la prueba augmented Dickey-Fuller.

- Probamos la hipótesis para la prueba KPSS de que;
 - H0: La secuencia es estacionaria.
 - H1: Una secuencia no es estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es no estacionaria.

Parámetro	Valor
Estadística KPSS	0.21415
Valor-p	0.1
Lags utilizados	1.000
Valor critico (1 %)	0.739
Valor critico (5 %)	0.463
Valor critico (10 %)	0.347

Cuadro 8: Resultado de la prueba KPSS.

Basándonos en los resultados de las pruebas de estacionariedad ADF y KPSS mostrados en el Cuadro 7 y Cuadro 8, podemos tomar decisiones sobre la estacionariedad de la secuencia.

Para la prueba ADF, la hipótesis nula (H0) establece que la secuencia no es estacionaria, mientras que la hipótesis alternativa (H1) sugiere lo contrario. Con un nivel de significancia del 5 %, el valor-p asociado a la prueba ADF es 0.98786, que excede al umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H0) y concluimos que la secuencia no es estacionaria.

En contraste, para la prueba KPSS, la H0 postula que la secuencia es estacionaria, mientras que la H1 sugiere que no lo es. Aquí, con un nivel de significancia del 5 %, el valor-p asociado a la

prueba KPSS es 0.1, que excede el umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H_0) y concluimos que la secuencia es estacionaria.

Los resultados obtenidos a través de las pruebas de estacionariedad presentan una discrepancia significativa en cuanto a la naturaleza de la serie temporal analizada. Mientras que la prueba ADF sugiere consistentemente la estacionariedad de la secuencia, indicada por la no rechazo de la hipótesis nula, la prueba KPSS no proporciona evidencia suficiente para refutar la estacionariedad. Esta discrepancia plantea interrogantes sobre la verdadera naturaleza de la serie temporal. Esto se puede deber a que no poseemos suficientes datos para obtener una conclusión más clara.

A la luz de los resultados contradictorios de las pruebas de estacionariedad ADF y KPSS, se ha optado por asumir la estacionariedad de la serie temporal analizada. Aunque la prueba ADF no proporcionó suficiente evidencia para rechazar la hipótesis nula de no estacionariedad, los gráficos de autocorrelación ACF y PACF mostraron indicios de posible estacionariedad, lo que respalda nuestra decisión. Reconocemos que esta asunción puede ser discutible y que hay incertidumbre en la verdadera naturaleza de la serie temporal. Sin embargo, dadas las limitaciones de los datos y la falta de consenso entre las pruebas estadísticas, consideramos que es prudente proceder bajo la suposición de estacionariedad para los fines de este análisis.

2.2.6. Selección de un modelo

Ahora procederemos a seleccionar el modelo ARIMA más apropiado para nuestra serie temporal. Utilizaremos las capacidades de Python para explorar y evaluar diversos modelos, con el fin de identificar aquel que mejor se ajuste a nuestras necesidades. Este análisis nos proporcionará la base necesaria para tomar decisiones informadas respecto a la elección del modelo más adecuado.

MODELO	AIC	BIC	HQ
ARIMA (1,0,1)	32.588	35.921	32.919
ARIMA (0,0,0)	31.093	32.759	31.258
ARIMA (1,0,0)	30.683	33.182	30.931

Cuadro 9: Identificación del modelo.

El modelo con los mínimos Criterios de Información de Akaike (AIC), Criterio de Información Bayesiano (BIC) y Hannan-Quinn (HQ) es el más adecuado para la secuencia. Por lo tanto, al probar todos los modelos competidores en el Cuadro 9, el modelo ARIMA (1, 0, 0) tiene los valores más bajos y, por lo tanto, el mejor modelo predictivo.

2.2.7. Estimación de parámetros

El Cuadro 10 a continuación muestra estimaciones de los parámetros del sistema ARIMA (1, 0, 0).

Tipo	Coefficiente	Error estándar	Valor Z	Valor-p
Constante	3.9929	0.208	19.231	0.000
ar.L1	0.4305	0.276	1.560	0.119
sigma2 (σ^2)	0.2471	0.114	2.173	0.030

Cuadro 10: Muestra de la estimación de parámetros.

El modelo ARIMA(1, 0, 0) indica que estamos utilizando un modelo autorregresivo integrado de media móvil con una componente autorregresiva (AR) de orden 1, cero componentes de integración

(I) y cero componentes de media móvil (MA). Esto implica que hay dependencia lineal de un valor pasado en la serie temporal, pero no hay tendencia ni error en los datos.

En cuanto a la estimación de parámetros, la constante tiene un valor de 3,9929 con un error estándar de 0,208. El valor Z es alto (19,231), lo que indica que la constante es estadísticamente significativa. Además, el valor-p asociado es muy bajo (0,000), lo que sugiere una alta significancia estadística.

En resumen, con un modelo ARIMA(1, 0, 0), se está utilizando una estimación constante junto con la dependencia lineal de un valor pasado para predecir la serie temporal de notas promedio semestrales en Calculo Avanzado. La constante estimada parece ser estadísticamente significativa, lo que sugiere que este modelo puede ser útil para predecir las notas con un nivel aceptable de precisión. Además, este modelo ha sido seleccionado como el más adecuado según los criterios de información AIC, BIC y HQ, lo que respalda su idoneidad como modelo predictivo para la serie temporal analizada.

2.2.8. Resumen y conclusión

Los resultados del análisis revelaron una variabilidad significativa en el rendimiento de los estudiantes en la asignatura de Calculo Avanzado a lo largo de los 17 años de estudio. En promedio, los alumnos finalizaron la asignatura con una calificación de 3.92, pero se observaron fluctuaciones destacadas en diferentes períodos. Por ejemplo, el año 2012 registró el rendimiento más bajo con un promedio de 2.96. Por el contrario, el año 2007 destacó como el de mejor desempeño, con un promedio de 4.88.

Los objetivos específicos de este estudio incluyeron analizar la evolución del rendimiento estudiantil en la asignatura y elaborar un modelo adecuado que se ajustara con precisión a dicho desempeño. Para lograr estos objetivos, se implementó una metodología integral que comenzó con un análisis descriptivo detallado de los datos, seguido de gráficos de las funciones de autocorrelación ACF y PACF para identificar estacionariedad. Además, se utilizó el test ADF y el test KPSS para respaldar las hipótesis obtenidas del análisis de los gráficos de autocorrelación.

Basándonos en los resultados de las pruebas estadísticas, se seleccionó el modelo ARIMA más apropiado utilizando los criterios de información AIC, BIC y HQ. Se encontró que el modelo ARIMA(1,0,0) era el más adecuado para la serie temporal estudiada. Este modelo implica una constante de 3.9929, que resultó ser estadísticamente significativa. Si bien este modelo simple no captura la dinámica compleja de la serie temporal, proporciona una base sólida para la predicción del rendimiento académico en la asignatura de Calculo Avanzado.

En resumen, este estudio proporcionó una visión completa del rendimiento estudiantil en la asignatura de Calculo Avanzado, desde un análisis descriptivo inicial hasta la selección y estimación de un modelo ARIMA adecuado. Los resultados obtenidos contribuyen a una mejor comprensión de los factores que influyen en el desempeño académico de los estudiantes en esta área específica y pueden servir como base para futuras investigaciones y acciones de mejora educativa.

2.3. Ecuaciones Diferenciales Ordinarias

Durante el análisis de la asignatura de Ecuaciones Diferenciales Ordinarias, se examinaron los promedios obtenidos por semestre a lo largo de los 17 años de estudio. En promedio, los alumnos finalizaban la asignatura con una calificación de 3.55. Sin embargo, se observaron variaciones significativas en el rendimiento a lo largo de los años. En particular, en el año 2012 se registró el rendimiento más bajo como curso, con un promedio de 2.93. Por el contrario, en el año 2007 se alcanzó el rendimiento más alto como curso, con un promedio de 4.62.

2.3.1. Objetivos específicos:

- Analizar la evolución del rendimiento estudiantil en la asignatura de Ecuaciones Diferenciales Ordinarias, dentro del marco del curso “6° año Politécnico común Ing. Navales”, impartido en la Escuela de Ingeniería.
- Elaborar un modelo adecuado que se ajuste con mayor precisión al desempeño de los estudiantes en la asignatura de Ecuaciones Diferenciales Ordinarias, perteneciente al curso “6° año Politécnico común Ing. Navales”, dictado en la Escuela de Ingeniería.

2.3.2. Preguntas de investigación

- ¿Se identifica alguna tendencia en el rendimiento de los estudiantes en la asignatura de Ecuaciones Diferenciales Ordinarias, dentro del curso “6° año Politécnico común Ing. Navales”, impartido en la Escuela de Ingeniería?
- ¿Cuál es el modelo más adecuado para describir el desempeño de los estudiantes en la asignatura de Ecuaciones Diferenciales Ordinarias, dentro del curso “6° año Politécnico común Ing. Navales”, dictado en la Escuela de Ingeniería?

2.3.3. Metodología

La metodología empleada en este estudio abarcó una serie de pasos fundamentales para el análisis exhaustivo de la serie temporal en cuestión. En primer lugar, se realizó un análisis descriptivo detallado de los datos, con el fin de comprender su comportamiento general, identificar posibles tendencias y visualizar patrones significativos. Se incluyó un gráfico de línea para representar visualmente la evolución de los datos, añadiendo una línea de tendencia para resaltar cualquier dirección en el conjunto de datos.

Posteriormente, se procedió a generar gráficos de las funciones de autocorrelación ACF y autocorrelación parcial PACF. Estas visualizaciones jugaron un papel crucial en la evaluación preliminar de la estacionariedad de los datos y en la identificación de posibles estructuras de autocorrelación. Estos correlogramas proporcionaron una comprensión visual inicial de la relación entre las observaciones en diferentes puntos temporales.

Con base en las observaciones obtenidas de los correlogramas, se llevó a cabo una evaluación más rigurosa de la estacionariedad de la serie temporal. Para ello, se realizaron pruebas de estacionariedad utilizando tanto el test ADF (Augmented Dickey-Fuller) como el test KPSS (Kwiatkowski, Phillips, Schmidt, Shin). Estas pruebas estadísticas sirvieron para reafirmar o refutar las hipótesis extraídas de los correlogramas, proporcionando una evaluación más sólida de la estacionariedad de la serie temporal.

Finalmente, una vez confirmada la estacionariedad de los datos, se procedió a la selección del modelo ARIMA más adecuado para el conjunto de datos analizados. Esto implicó la identificación

de los parámetros óptimos del modelo ARIMA, aprovechando tanto la información derivada de los correlogramas como de las pruebas estadísticas de estacionariedad.

En resumen, la metodología empleada abarcó desde la exploración inicial de los datos y la representación gráfica, pasando por el análisis de la estacionariedad con correlogramas y pruebas estadísticas, hasta la selección y estimación de un modelo ARIMA apropiado. Este enfoque integral garantizó un análisis completo y riguroso de la serie temporal, proporcionando una base sólida para su modelado y predicción.

2.3.4. Estadística Descriptiva de las notas promedios semestrales de los estudiantes

Para realizar un estudio efectivo de la asignatura en estudio, se inició reduciendo la base de datos para incluir únicamente las columnas relevantes (Figura 8). Posteriormente, se llevó a cabo un análisis exploratorio de datos centrado en el rendimiento promedio de los estudiantes en la asignatura de Ecuaciones Diferenciales Ordinarias a lo largo de 17 años consecutivos. Este análisis se realizó principalmente utilizando Python y técnicas de análisis de series de tiempo.

Durante este proceso, se realizaron cálculos para obtener estadísticas descriptivas iniciales relacionadas con el rendimiento promedio de los estudiantes en Ecuaciones Diferenciales Ordinarias. Luego, se procedió a la creación de gráficos de series de tiempo, los cuales permitieron visualizar la evolución del rendimiento promedio a lo largo de los años. Además, se realizó un análisis de tendencias para identificar patrones o cambios significativos en el rendimiento a lo largo del tiempo. Este enfoque proporcionó una comprensión más profunda del comportamiento de los estudiantes en la asignatura a lo largo del período de estudio.

ECUACIONES DIFERENCIALES ORDINARIAS	
AÑO	
2007	4.615294
2008	4.285556
2009	3.149615
2010	3.482143
2011	3.236154
2012	2.931333
2013	3.112000
2014	4.190000
2015	3.403333
2016	3.605000
2017	3.849130
2018	3.227857
2019	3.314762
2020	3.581429
2021	3.210000
2022	3.373333
2023	3.705714

Figura 8: Base de datos Ecuaciones Diferenciales Ordinarias.

Se observó que la nota mínima promedio semestral de los estudiantes en la asignatura de Ecuaciones Diferenciales Ordinarias fue de 2.93 en el año 2012. Por otro lado, se registró una nota promedio anual máxima de 4.62 en el año 2007. En cuanto al desempeño promedio, este se situó en 3.55, con una desviación estándar de 0.46.

Datos	17
Promedio	3.545450
Desviación estándar	0.458825
Mínimo	2.931333
25 %	3.227857
50 %	3.403333
75 %	3.705714
Máximo	4.615294

Cuadro 11: Desempeño de los estudiantes en Ecuaciones Diferenciales Ordinarias.

La Figura 9 muestra las notas promedios del curso por año. La ecuación de la tendencia, $y = -0,03x + 3,76$, representa una relación lineal entre el año (x) y el promedio de las notas de la asignatura de Ecuaciones Diferenciales Ordinarias (y). En este caso, el coeficiente de $-0,03$ indica la pendiente de la recta de tendencia, lo que sugiere una disminución lineal en el promedio de las notas a medida que pasa el tiempo. Por otro lado, el término independiente ($3,76$) representa el valor esperado del promedio de las notas en el año cero del estudio (en este caso, 2007 ajustado a $x = 0$).

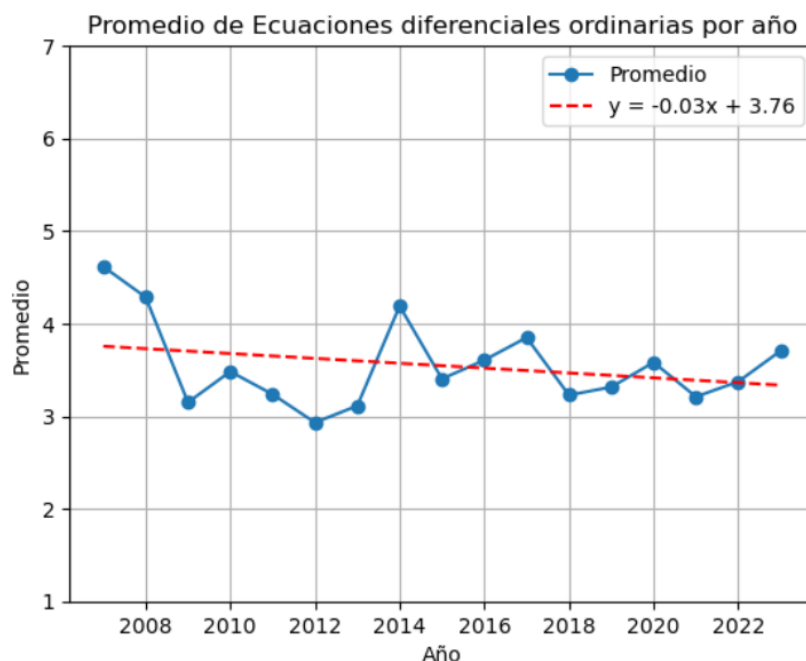


Figura 9: Serie de tiempo de las notas promedios semestrales de los estudiantes en Ecuaciones Diferenciales Ordinarias.

Al interpretar la ecuación, podemos concluir que la tendencia general a lo largo de los años muestra una disminución de aproximadamente 0.03 unidades en el promedio de las notas de la asignatura de Ecuaciones Diferenciales Ordinarias por año. Es decir, en promedio, se espera que

el rendimiento de los estudiantes en esta asignatura disminuya en 0.03 puntos cada semestre. Sin embargo, es importante tener en cuenta que esta es una simplificación y que pueden existir otros factores no considerados en la ecuación que influyan en el rendimiento académico.

Es necesario analizar la naturaleza de los datos para seleccionar el modelo correcto y mostrar si se trata de datos cíclicos sin tendencia o con actividad estacional.

2.3.5. Pruebas de estacionariedad

Se examinaron los gráficos de la función de autocorrelación (ACF) y de la función de autocorrelación parcial (PACF) de las notas promedio semestrales de los estudiantes en Ecuaciones Diferenciales Ordinarias. Estos gráficos, mostrados en la Figura 10, proporcionaron información inicial sobre la estacionariedad de la serie temporal.

En los gráficos ACF y PACF, no se observaron picos significativos en ningún retraso, y además tenemos correlaciones positivas y negativas, lo que sugiere que la serie temporal podría ser estacionaria o que la autocorrelación ha disminuido. Esto indica una posible estacionariedad en la serie de datos, aunque se requieren pruebas estadísticas adicionales para confirmar esta observación.

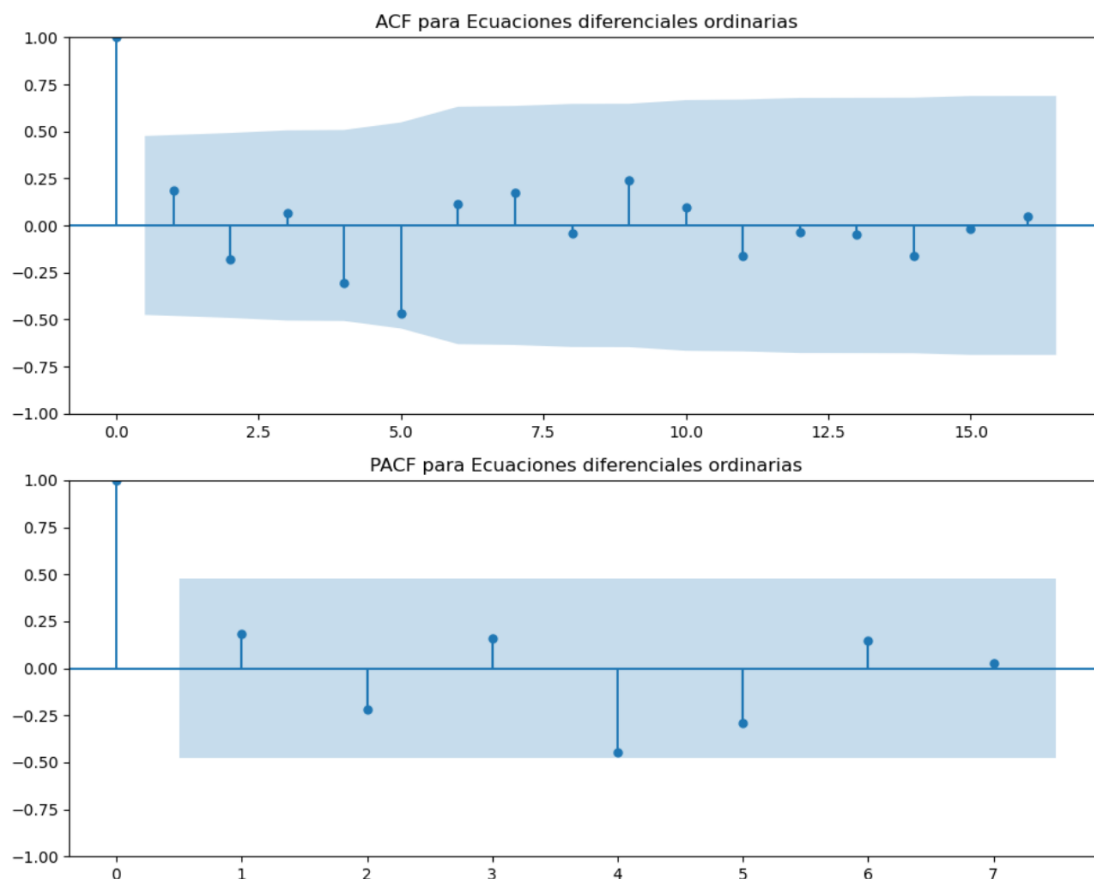


Figura 10: Gráfico ACF y PACF de las notas promedio semestrales de los estudiantes en Ecuaciones Diferenciales Ordinarias.

Con estas observaciones preliminares, se procedió a realizar las pruebas de estacionariedad ADF y KPSS para verificar la presencia de estacionariedad en la serie temporal.

Un ciclo estacionario tiene una media y una varianza constante en el tiempo y no hay patrones en el ciclo. Es necesario que la serie sea estacionaria antes de continuar con la estimación de un

modelo ARIMA, ya que en este análisis se utilizó la prueba Augmented Dickey-Fuller (ADF) y la prueba de Kwiatkowski, Phillips, Schmidt y Shin (KPSS) para verificar la estacionariedad en el rendimiento académico de los estudiantes en métodos cuantitativos.

- Probamos la hipótesis para la prueba ADF de que;
 - H0: La secuencia no es estacionaria.
 - H1: La secuencia estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es estacionaria.

Parámetro	Valor
Estadística ADF	-5.43283
Valor-p	2.88896e-06
Lags utilizados	4
Valor critico (1 %)	-4.13783
Valor critico (5 %)	-3.15497
Valor critico (10 %)	-2.71448

Cuadro 12: Resultado de la prueba augmented Dickey-Fuller.

- Probamos la hipótesis para la prueba KPSS de que;
 - H0: La secuencia es estacionaria.
 - H1: Una secuencia no es estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es no estacionaria.

Parámetro	Valor
Estadística KPSS	0.17311
Valor-p	0.1
Lags utilizados	1
Valor critico (1 %)	0.739
Valor critico (5 %)	0.463
Valor critico (10 %)	0.347

Cuadro 13: Resultado de la prueba KPSS.

Basándonos en los resultados de las pruebas de estacionariedad ADF y KPSS mostrados en el Cuadro 12 y Cuadro 13, podemos tomar decisiones sobre la estacionariedad de la secuencia.

Para la prueba ADF, la hipótesis nula (H0) establece que la secuencia no es estacionaria, mientras que la hipótesis alternativa (H1) sugiere lo contrario. Con un nivel de significancia del 5 %, el valor-p asociado a la prueba ADF es 2.88896e-06, que es inferior al umbral crítico de 0.05. Por lo tanto, rechazamos la hipótesis nula (H0) y concluimos que la secuencia es estacionaria.

En contraste, para la prueba KPSS, la H0 postula que la secuencia es estacionaria, mientras que la H1 sugiere que no lo es. Aquí, con un nivel de significancia del 5 %, el valor-p asociado a la

prueba KPSS es 0.1, que excede el umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H_0) y concluimos que la secuencia es estacionaria.

Estos resultados sugieren una consistencia en la estacionariedad de la secuencia, como se evidencia por la reafirmación de la estacionariedad a través de la prueba ADF y la falta de evidencia suficiente para rechazar la estacionariedad a través de la prueba KPSS, confirmando así nuestra suposición realizada en el análisis de los gráficos ACF y PACF.

2.3.6. Selección de un modelo

Ahora procederemos a seleccionar el modelo ARIMA más apropiado para nuestra serie temporal. Utilizaremos las capacidades de Python para explorar y evaluar diversos modelos, con el fin de identificar aquel que mejor se ajuste a nuestras necesidades. Este análisis nos proporcionará la base necesaria para tomar decisiones informadas respecto a la elección del modelo más adecuado.

MODELO	AIC	BIC	HQ
ARIMA (1,0,1)	26.630	29.963	26.961
ARIMA (0,0,0)	24.724	26.391	24.890
ARIMA (1,0,0)	25.848	28.348	26.097

Cuadro 14: Identificación del modelo.

El modelo con los mínimos Criterios de Información de Akaike (AIC), Criterio de Información Bayesiano (BIC) y Hannan-Quinn (HQ) es el más adecuado para la secuencia. Por lo tanto, al probar todos los modelos competidores en el Cuadro 14, el modelo ARIMA (0, 0, 0) tiene los valores más bajos y, por lo tanto, el mejor modelo predictivo.

2.3.7. Estimación de parámetros

El Cuadro 15 a continuación muestra estimaciones de los parámetros del sistema ARIMA (0, 0, 0).

Tipo	Coefficiente	Error estándar	Valor Z	Valor-p
Constante	3.5454	0.145	24.384	0.000
sigma2 (σ^2)	0.1981	0.091	2.187	0.029

Cuadro 15: Muestra de la estimación de parámetros.

El modelo ARIMA(0,0,0) indica que estamos utilizando un modelo autorregresivo integrado de media móvil con cero componentes autorregresivas (AR), cero componentes de integración (I) y cero componentes de media móvil (MA). Esto significa que no hay dependencia de los valores pasados, ni tendencia ni error en los datos.

En cuanto a la estimación de parámetros, la constante tiene un valor de 3,5454 con un error estándar de 0,145. El valor Z es alto (24,384), lo que indica que la constante es estadísticamente significativa. Además, el valor-p asociado es muy bajo (0,000), lo que sugiere una alta significancia estadística.

En resumen, con un modelo ARIMA(0,0,0), se está utilizando una estimación constante para predecir la serie temporal de notas promedio semestrales en Ecuaciones Diferenciales Ordinarias. La constante estimada parece ser estadísticamente significativa, lo que sugiere que este modelo puede ser útil para predecir las notas con un nivel aceptable de precisión.

2.3.8. Resumen y conclusión

Los resultados del análisis revelaron una variabilidad significativa en el rendimiento de los estudiantes en la asignatura de Ecuaciones Diferenciales Ordinarias a lo largo de los 17 años de estudio. En promedio, los alumnos finalizaron la asignatura con una calificación de 3.55, pero se observaron fluctuaciones destacadas en diferentes períodos. Por ejemplo, el año 2012 registró el rendimiento más bajo con un promedio de 2.93. Por el contrario, el año 2007 destacó como el de mejor desempeño, con un promedio de 4.62.

Los objetivos específicos de este estudio incluyeron analizar la evolución del rendimiento estudiantil en la asignatura y elaborar un modelo adecuado que se ajustara con precisión a dicho desempeño. Para lograr estos objetivos, se implementó una metodología integral que comenzó con un análisis descriptivo detallado de los datos, seguido de gráficos de las funciones de autocorrelación ACF y PACF para identificar estacionariedad. Además, se utilizó el test ADF y el test KPSS para respaldar las hipótesis obtenidas del análisis de los gráficos de autocorrelación.

Basándonos en los resultados de las pruebas estadísticas, se seleccionó el modelo ARIMA más apropiado utilizando los criterios de información AIC, BIC y HQ. Se encontró que el modelo ARIMA(0,0,0) era el más adecuado para la serie temporal estudiada. Este modelo implica una constante de 3.5454, que resultó ser estadísticamente significativa. Si bien este modelo simple no captura la dinámica compleja de la serie temporal, proporciona una base sólida para la predicción del rendimiento académico en la asignatura de Ecuaciones Diferenciales Ordinarias.

En resumen, este estudio proporcionó una visión completa del rendimiento estudiantil en la asignatura de Ecuaciones Diferenciales Ordinarias, desde un análisis descriptivo inicial hasta la selección y estimación de un modelo ARIMA adecuado. Los resultados obtenidos contribuyen a una mejor comprensión de los factores que influyen en el desempeño académico de los estudiantes en esta área específica y pueden servir como base para futuras investigaciones y acciones de mejora educativa.

2.4. Física Mecánica y Ondas

Durante el análisis de la asignatura de Física Mecánica y Ondas, se examinaron los promedios obtenidos por semestre a lo largo de los 17 años de estudio. En promedio, los alumnos finalizaban la asignatura con una calificación de 4.37. Sin embargo, se observaron variaciones significativas en el rendimiento a lo largo de los años. En particular, en el año 2015 se registró el rendimiento más bajo como curso, con un promedio de 3.74. Por el contrario, en el año 2007 se alcanzó el rendimiento más alto como curso, con un promedio de 5.25.

2.4.1. Objetivos específicos:

- Analizar la evolución del rendimiento estudiantil en la asignatura de Física Mecánica y Ondas, dentro del marco del curso “6° año Politécnico común Ing. Navales”, impartido en la Escuela de Ingeniería.
- Elaborar un modelo adecuado que se ajuste con mayor precisión al desempeño de los estudiantes en la asignatura de Física Mecánica y Ondas, perteneciente al curso “6° año Politécnico común Ing. Navales”, dictado en la Escuela de Ingeniería.

2.4.2. Preguntas de investigación

- ¿Se identifica alguna tendencia en el rendimiento de los estudiantes en la asignatura de Física Mecánica y Ondas, dentro del curso “6° año Politécnico común Ing. Navales”, impartido en la Escuela de Ingeniería?
- ¿Cuál es el modelo más adecuado para describir el desempeño de los estudiantes en la asignatura de Física Mecánica y Ondas, dentro del curso “6° año Politécnico común Ing. Navales”, dictado en la Escuela de Ingeniería?

2.4.3. Metodología

La metodología empleada en este estudio abarcó una serie de pasos fundamentales para el análisis exhaustivo de la serie temporal en cuestión. En primer lugar, se realizó un análisis descriptivo detallado de los datos, con el fin de comprender su comportamiento general, identificar posibles tendencias y visualizar patrones significativos. Se incluyó un gráfico de línea para representar visualmente la evolución de los datos, añadiendo una línea de tendencia para resaltar cualquier dirección en el conjunto de datos.

Posteriormente, se procedió a generar gráficos de las funciones de autocorrelación ACF y autocorrelación parcial PACF. Estas visualizaciones jugaron un papel crucial en la evaluación preliminar de la estacionariedad de los datos y en la identificación de posibles estructuras de autocorrelación. Estos correlogramas proporcionaron una comprensión visual inicial de la relación entre las observaciones en diferentes puntos temporales.

Con base en las observaciones obtenidas de los correlogramas, se llevó a cabo una evaluación más rigurosa de la estacionariedad de la serie temporal. Para ello, se realizaron pruebas de estacionariedad utilizando tanto el test ADF (Augmented Dickey-Fuller) como el test KPSS (Kwiatkowski, Phillips, Schmidt, Shin). Estas pruebas estadísticas sirvieron para reafirmar o refutar las hipótesis extraídas de los correlogramas, proporcionando una evaluación más sólida de la estacionariedad de la serie temporal.

Finalmente, una vez confirmada la estacionariedad de los datos, se procedió a la selección del modelo ARIMA más adecuado para el conjunto de datos analizados. Esto implicó la identificación

de los parámetros óptimos del modelo ARIMA, aprovechando tanto la información derivada de los correlogramas como de las pruebas estadísticas de estacionariedad.

En resumen, la metodología empleada abarcó desde la exploración inicial de los datos y la representación gráfica, pasando por el análisis de la estacionariedad con correlogramas y pruebas estadísticas, hasta la selección y estimación de un modelo ARIMA apropiado. Este enfoque integral garantizó un análisis completo y riguroso de la serie temporal, proporcionando una base sólida para su modelado y predicción.

2.4.4. Estadística Descriptiva de las notas promedios semestrales de los estudiantes

Para realizar un estudio efectivo de la asignatura en estudio, se inició reduciendo la base de datos para incluir únicamente las columnas relevantes (Figura 11). Posteriormente, se llevó a cabo un análisis exploratorio de datos centrado en el rendimiento promedio de los estudiantes en la asignatura de Física Mecánica y Ondas a lo largo de 17 años consecutivos. Este análisis se realizó principalmente utilizando Python y técnicas de análisis de series de tiempo.

Durante este proceso, se realizaron cálculos para obtener estadísticas descriptivas iniciales relacionadas con el rendimiento promedio de los estudiantes en Física Mecánica y Ondas. Luego, se procedió a la creación de gráficos de series de tiempo, los cuales permitieron visualizar la evolución del rendimiento promedio a lo largo de los años. Además, se realizó un análisis de tendencias para identificar patrones o cambios significativos en el rendimiento a lo largo del tiempo. Este enfoque proporcionó una comprensión más profunda del comportamiento de los estudiantes en la asignatura a lo largo del período de estudio.

FISICA MECANICA Y ONDAS	
AÑO	
2007	5.251176
2008	4.780000
2009	5.107308
2010	4.876429
2011	4.490000
2012	3.772667
2013	3.868000
2014	4.255000
2015	3.739048
2016	3.878500
2017	4.056087
2018	4.119286
2019	3.862381
2020	4.246429
2021	4.488333
2022	4.727333
2023	4.850000

Figura 11: Base de datos Física Mecánica y Ondas.

Se observó que la nota mínima promedio anual de los estudiantes en la asignatura de Física mecánica y ondas fue de 3.74 en el año 2015. Por otro lado, se registró una nota promedio semestral máxima de 5.25 en el año 2007. En cuanto al desempeño promedio, este se situó en 4.37, con una desviación estándar de 0.49.

Datos	17
Promedio	4.374587
Desviación estándar	0.489059
Mínimo	3.739048
25 %	3.3.87850
50 %	4.25500
75 %	4.78000
Máximo	5.25118

Cuadro 16: Desempeño de los estudiantes en Física mecánica y ondas.

La Figura 12 muestra las notas promedios del curso por año. La ecuación de la tendencia, $y = -0,03x + 4,61$, representa una relación lineal entre el año (x) y el promedio de las notas de la asignatura de Física Mecánica y Ondas (y). En este caso, el coeficiente de $-0,03$ indica la pendiente de la recta de tendencia, lo que sugiere una disminución lineal en el promedio de las notas a medida que pasa el tiempo. Por otro lado, el término independiente ($4,61$) representa el valor esperado del promedio de las notas en el año cero del estudio (en este caso, 2007 ajustado a $x = 0$).

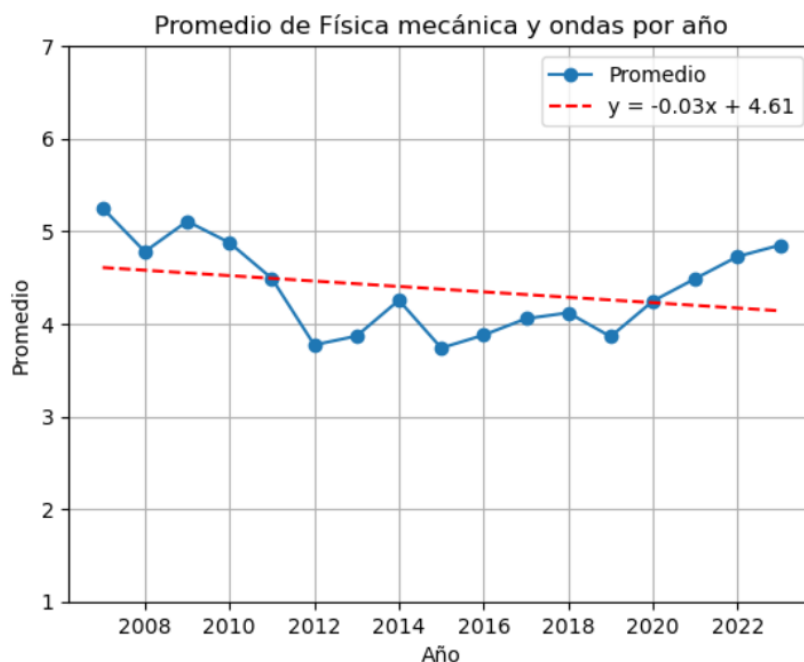


Figura 12: Serie de tiempo de las notas promedios semestrales de los estudiantes en Física Mecánica y Ondas.

Al interpretar la ecuación, podemos concluir que la tendencia general a lo largo de los años muestra una disminución de aproximadamente 0.03 unidades en el promedio de las notas de la

asignatura de Física Mecánica y Ondas por año. Es decir, en promedio, se espera que el rendimiento de los estudiantes en esta asignatura disminuya en 0.03 puntos cada semestre. Sin embargo, es importante tener en cuenta que esta es una simplificación y que pueden existir otros factores no considerados en la ecuación que influyan en el rendimiento académico.

Es necesario analizar la naturaleza de los datos para seleccionar el modelo correcto y mostrar si se trata de datos cíclicos sin tendencia o con actividad estacional.

2.4.5. Pruebas de estacionariedad

Se examinaron los gráficos de la función de autocorrelación (ACF) y de la función de autocorrelación parcial (PACF) de las notas promedio semestrales de los estudiantes en Física Mecánica y Ondas. Estos gráficos, mostrados en la Figura 13, proporcionaron información inicial sobre la estacionariedad de la serie temporal.

En los gráficos ACF y PACF, no se observaron picos significativos en ningún retraso, y además tenemos correlaciones positivas y negativas, lo que sugiere que la serie temporal podría ser estacionaria o que la autocorrelación ha disminuido. Esto indica una posible estacionariedad en la serie de datos, aunque se requieren pruebas estadísticas adicionales para confirmar esta observación.

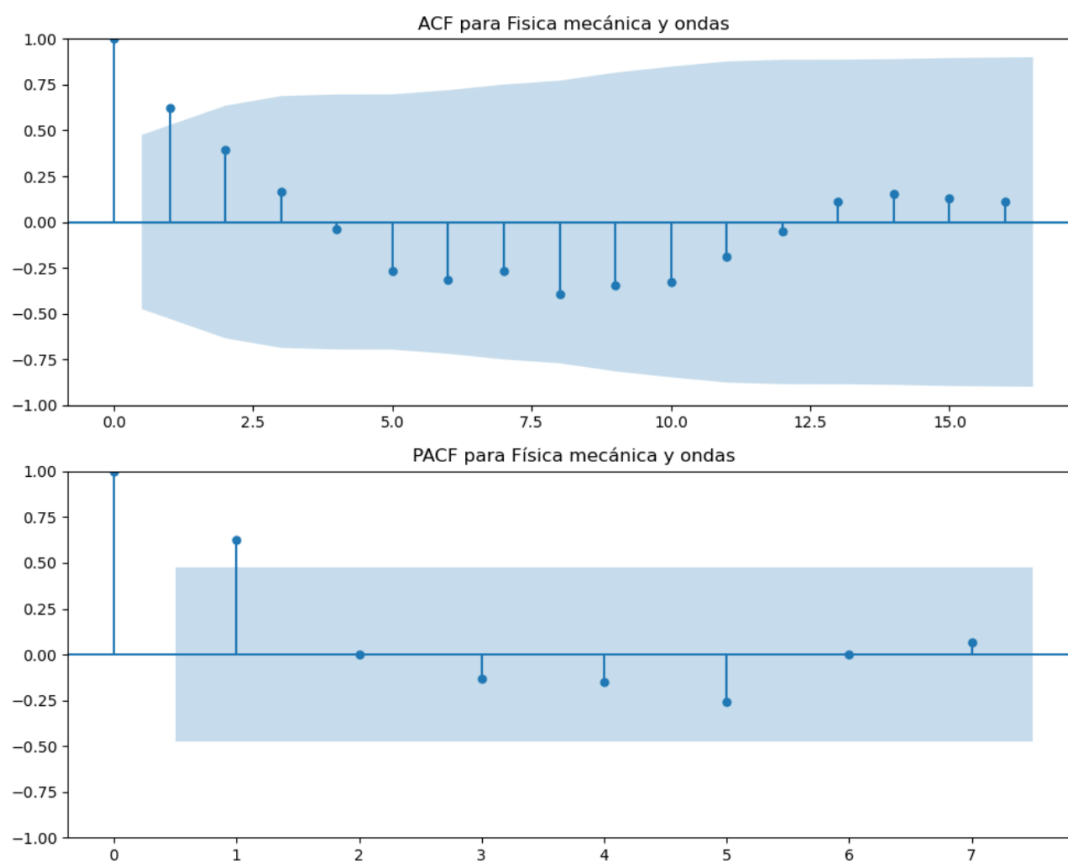


Figura 13: Gráfico ACF y PACF de las notas promedio semestrales de los estudiantes en Física Mecánica y Ondas.

Con estas observaciones preliminares, se procedió a realizar las pruebas de estacionariedad ADF y KPSS para verificar la presencia de estacionariedad en la serie temporal.

Un ciclo estacionario tiene una media y una varianza constante en el tiempo y no hay patrones

en el ciclo. Es necesario que la serie sea estacionaria antes de continuar con la estimación de un modelo ARIMA, ya que en este análisis se utilizó la prueba Augmented Dickey-Fuller (ADF) y la prueba de Kwiatkowski, Phillips, Schmidt y Shin (KPSS) para verificar la estacionariedad en el rendimiento académico de los estudiantes en métodos cuantitativos.

- Probamos la hipótesis para la prueba ADF de que;

- H0: La secuencia no es estacionaria.
- H1: La secuencia estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es estacionaria.

Parámetro	Valor
Estadística ADF	-0.29219
Valor-p	0.92660
Lags utilizados	5
Valor critico (1 %)	-4.22324
Valor critico (5 %)	-3.18937
Valor critico (10 %)	2.72984

Cuadro 17: Resultado de la prueba augmented Dickey-Fuller.

- Probamos la hipótesis para la prueba KPSS de que;

- H0: La secuencia es estacionaria.
- H1: Una secuencia no es estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es no estacionaria.

Parámetro	Valor
Estadística KPSS	0.29655
Valor-p	0.1
Lags utilizados	1
Valor critico (1 %)	0.739
Valor critico (5 %)	0.463
Valor critico (10 %)	0.347

Cuadro 18: Resultado de la prueba KPSS.

Basándonos en los resultados de las pruebas de estacionariedad ADF y KPSS mostrados en el Cuadro 17 y Cuadro 18, podemos tomar decisiones sobre la estacionariedad de la secuencia.

Para la prueba ADF, la hipótesis nula (H0) establece que la secuencia no es estacionaria, mientras que la hipótesis alternativa (H1) sugiere lo contrario. Con un nivel de significancia del 5 %, el valor-p asociado a la prueba ADF es 0.92660, que excede al umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H0) y concluimos que la secuencia es no estacionaria.

En contraste, para la prueba KPSS, la H_0 postula que la secuencia es estacionaria, mientras que la H_1 sugiere que no lo es. Aquí, con un nivel de significancia del 5 %, el valor-p asociado a la prueba KPSS es 0.1, que excede el umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H_0) y concluimos que la secuencia es estacionaria.

Los resultados obtenidos a través de las pruebas de estacionariedad presentan una discrepancia significativa en cuanto a la naturaleza de la serie temporal analizada. Mientras que la prueba ADF sugiere consistentemente la estacionariedad de la secuencia, indicada por la no rechazo de la hipótesis nula, la prueba KPSS no proporciona evidencia suficiente para refutar la estacionariedad. Esta discrepancia plantea interrogantes sobre la verdadera naturaleza de la serie temporal. Esto se puede deber a que no poseemos suficientes datos para obtener una conclusión más clara.

A la luz de los resultados contradictorios de las pruebas de estacionariedad ADF y KPSS, se ha optado por asumir la estacionariedad de la serie temporal analizada. Aunque la prueba ADF no proporcionó suficiente evidencia para rechazar la hipótesis nula de no estacionariedad, los gráficos de autocorrelación ACF y PACF mostraron indicios de posible estacionariedad, lo que respalda nuestra decisión. Reconocemos que esta asunción puede ser discutible y que hay incertidumbre en la verdadera naturaleza de la serie temporal. Sin embargo, dadas las limitaciones de los datos y la falta de consenso entre las pruebas estadísticas, consideramos que es prudente proceder bajo la suposición de estacionariedad para los fines de este análisis..

2.4.6. Selección de un modelo

Ahora procederemos a seleccionar el modelo ARIMA más apropiado para nuestra serie temporal. Utilizaremos las capacidades de Python para explorar y evaluar diversos modelos, con el fin de identificar aquel que mejor se ajuste a nuestras necesidades. Este análisis nos proporcionará la base necesaria para tomar decisiones informadas respecto a la elección del modelo más adecuado.

MODELO	AIC	BIC	HQ
ARIMA (1,0,1)	18.673	22.006	19.004
ARIMA (0,0,0)	26.894	28.560	27.060
ARIMA (1,0,0)	16.711	19.211	16.960

Cuadro 19: Identificación del modelo.

El modelo con los mínimos Criterios de Información de Akaike (AIC), Criterio de Información Bayesiano (BIC) y Hannan-Quinn (HQ) es el más adecuado para la secuencia. Por lo tanto, al probar todos los modelos competidores en el Cuadro 19, el modelo ARIMA (1, 0, 0) tiene los valores más bajos y, por lo tanto, el mejor modelo predictivo.

2.4.7. Estimación de parámetros

El Cuadro 20 a continuación muestra estimaciones de los parámetros del sistema ARIMA (1, 0, 0).

Tipo	Coeficiente	Error estándar	Valor Z	Valor-p
Constante	4.5731	0.302	15.154	0.000
ar.L1	0.7794	0.248	3.145	0.002
sigma2 (σ^2)	0.1041	0.046	2.261	0.024

Cuadro 20: Muestra de la estimación de parámetros.

El modelo ARIMA(1, 0, 0) indica que estamos utilizando un modelo autorregresivo integrado de media móvil con una componente autorregresiva (AR) de orden 1, cero componentes de integración (I) y cero componentes de media móvil (MA). Esto implica que hay dependencia lineal de un valor pasado en la serie temporal, pero no hay tendencia ni error en los datos.

En cuanto a la estimación de parámetros, la constante tiene un valor de 4,5731 con un error estándar de 0,302. El valor Z es alto (15,154), lo que indica que la constante es estadísticamente significativa. Además, el valor-p asociado es muy bajo (0,000), lo que sugiere una alta significancia estadística.

En resumen, con un modelo ARIMA(1, 0, 0), se está utilizando una estimación constante junto con la dependencia lineal de un valor pasado para predecir la serie temporal de notas promedio semestrales en Física Mecánica y Ondas. La constante estimada parece ser estadísticamente significativa, lo que sugiere que este modelo puede ser útil para predecir las notas con un nivel aceptable de precisión. Además, este modelo ha sido seleccionado como el más adecuado según los criterios de información AIC, BIC y HQ, lo que respalda su idoneidad como modelo predictivo para la serie temporal analizada.

2.4.8. Resumen y conclusión

Los resultados del análisis revelaron una variabilidad significativa en el rendimiento de los estudiantes en la asignatura de Física Mecánica y Ondas a lo largo de los 17 años de estudio. En promedio, los alumnos finalizaron la asignatura con una calificación de 4.37, pero se observaron fluctuaciones destacadas en diferentes períodos. Por ejemplo, el año 2015 registró el rendimiento más bajo con un promedio de 3.73. Por el contrario, el año 2007 destacó como el de mejor desempeño, con un promedio de 5.25.

Los objetivos específicos de este estudio incluyeron analizar la evolución del rendimiento estudiantil en la asignatura y elaborar un modelo adecuado que se ajustara con precisión a dicho desempeño. Para lograr estos objetivos, se implementó una metodología integral que comenzó con un análisis descriptivo detallado de los datos, seguido de gráficos de las funciones de autocorrelación ACF y PACF para identificar estacionariedad. Además, se utilizó el test ADF y el test KPSS para respaldar las hipótesis obtenidas del análisis de los gráficos de autocorrelación.

Basándonos en los resultados de las pruebas estadísticas, se seleccionó el modelo ARIMA más apropiado utilizando los criterios de información AIC, BIC y HQ. Se encontró que el modelo ARIMA(1,0,0) era el más adecuado para la serie temporal estudiada. Este modelo implica una constante de 4.5731, que resultó ser estadísticamente significativa. Si bien este modelo simple no captura la dinámica compleja de la serie temporal, proporciona una base sólida para la predicción del rendimiento académico en la asignatura de Física Mecánica y Ondas.

En resumen, este estudio proporcionó una visión completa del rendimiento estudiantil en la asignatura de Física Mecánica y Ondas, desde un análisis descriptivo inicial hasta la selección y estimación de un modelo ARIMA adecuado. Los resultados obtenidos contribuyen a una mejor comprensión de los factores que influyen en el desempeño académico de los estudiantes en esta área específica y pueden servir como base para futuras investigaciones y acciones de mejora educativa.

3. Estudio del rendimiento académico de los alumnos del curso “1er año Mc.El.” usando series de tiempo en las asignaturas:

El curso “1er año Mc.El.”, ofrecido por la Escuela de Ingeniería (ESING) de la Apolinav (Academia Politécnica Naval) para gente de mar, desempeña un papel fundamental en la formación de futuros mecánicos eléctricos. En este contexto, se realiza un exhaustivo análisis utilizando series de tiempo sobre una base de datos que incluye el promedio de 584 alumnos, considerando tanto aprobados como separados de curso, recopilados desde el año 2010 hasta el 2023. A partir de esta información, se calculan los promedios semestrales de las asignaturas más relevantes.

Este riguroso estudio se centra en las cuatro asignaturas principales del curso: Circuitos Eléctricos I, Electricidad, Física Técnica I y Matemática Técnica I. Estas son las asignaturas que han mostrado mayor dificultad estos ultimo años.

La Figura 14 proporciona los datos necesarios para la realización de este estudio.

	CIRCUITOS ELECTRICOS I	ELECTRICIDAD	FISICA TECNICA	MATEMATICA TECNICA I
AÑO				
2010	4.454706	5.025882	4.853824	6.083824
2011	4.876410	5.337179	5.039487	5.742051
2012	4.748462	5.149487	5.005641	5.833333
2013	5.373684	4.965789	4.711053	5.696842
2014	5.003077	4.975128	5.453077	5.676667
2015	4.739318	5.174091	5.022045	5.553636
2016	4.401860	4.730233	4.826744	5.690698
2017	4.444000	5.463556	5.158000	6.006000
2018	4.464231	5.191154	5.440577	5.592692
2019	4.792500	5.297143	5.691786	5.487321
2020	4.316471	5.308039	5.254314	5.019412
2021	4.818772	5.158421	5.761404	5.818070
2022	5.002083	5.356250	5.849167	6.102083
2023	4.923478	4.670000	5.540870	5.455652

Figura 14: Base de datos con el promedio por asignatura correspondiente a los años de estudio.

3.1. Circuitos Eléctricos I

Durante el análisis de la asignatura Circuitos Eléctricos I, se examinaron los promedios obtenidos por semestre a lo largo de los 14 años de estudio. En promedio, los alumnos finalizaban la asignatura con una calificación de 4.74. Sin embargo, se observaron variaciones significativas en el rendimiento a lo largo de los años. En particular, en el año 2020 se registró el rendimiento

más bajo como curso, con un promedio de 4.32. Por el contrario, en el año 2013 se alcanzó el rendimiento más alto como curso, con un promedio de 5.37, destacando como un año excepcional en términos de desempeño académico en esta asignatura.

3.1.1. Objetivos específicos:

- Analizar la evolución del rendimiento estudiantil en la asignatura de Circuitos Eléctricos I, dentro del marco del curso “1er año Mc.El.”, impartido en la Escuela de Ingeniería.
- Elaborar un modelo adecuado que se ajuste con mayor precisión al desempeño de los estudiantes en la asignatura de Circuitos Eléctricos I, perteneciente al curso “1er año Mc.El.”, dictado en la Escuela de Ingeniería.

3.1.2. Preguntas de investigación

- ¿Se identifica alguna tendencia en el rendimiento de los estudiantes en la asignatura de Circuitos Eléctricos I, dentro del curso “1er año Mc.El.”, impartido en la Escuela de Ingeniería?
- ¿Cuál es el modelo más adecuado para describir el desempeño de los estudiantes en la asignatura de Circuitos Eléctricos I, dentro del curso “1er año Mc.El.”, dictado en la Escuela de Ingeniería?

3.1.3. Metodología

La metodología empleada en este estudio abarcó una serie de pasos fundamentales para el análisis exhaustivo de la serie temporal en cuestión. En primer lugar, se realizó un análisis descriptivo detallado de los datos, con el fin de comprender su comportamiento general, identificar posibles tendencias y visualizar patrones significativos. Se incluyó un gráfico de línea para representar visualmente la evolución de los datos, añadiendo una línea de tendencia para resaltar cualquier dirección en el conjunto de datos.

Posteriormente, se procedió a generar gráficos de las funciones de autocorrelación ACF y autocorrelación parcial PACF. Estas visualizaciones jugaron un papel crucial en la evaluación preliminar de la estacionariedad de los datos y en la identificación de posibles estructuras de autocorrelación. Estos correlogramas proporcionaron una comprensión visual inicial de la relación entre las observaciones en diferentes puntos temporales.

Con base en las observaciones obtenidas de los correlogramas, se llevó a cabo una evaluación más rigurosa de la estacionariedad de la serie temporal. Para ello, se realizaron pruebas de estacionariedad utilizando tanto el test ADF (Augmented Dickey-Fuller) como el test KPSS (Kwiatkowski, Phillips, Schmidt, Shin). Estas pruebas estadísticas sirvieron para reafirmar o refutar las hipótesis extraídas de los correlogramas, proporcionando una evaluación más sólida de la estacionariedad de la serie temporal.

Finalmente, una vez confirmada la estacionariedad de los datos, se procedió a la selección del modelo ARIMA más adecuado para el conjunto de datos analizados. Esto implicó la identificación de los parámetros óptimos del modelo ARIMA, aprovechando tanto la información derivada de los correlogramas como de las pruebas estadísticas de estacionariedad.

En resumen, la metodología empleada abarcó desde la exploración inicial de los datos y la representación gráfica, pasando por el análisis de la estacionariedad con correlogramas y pruebas estadísticas, hasta la selección y estimación de un modelo ARIMA apropiado. Este enfoque integral

garantizó un análisis completo y riguroso de la serie temporal, proporcionando una base sólida para su modelado y predicción.

3.1.4. Estadística Descriptiva de las notas promedios semestrales de los estudiantes

Para realizar un estudio efectivo de la asignatura en estudio, se inició reduciendo la base de datos para incluir únicamente las columnas relevantes (Figura 15). Posteriormente, se llevó a cabo un análisis exploratorio de datos centrado en el rendimiento promedio de los estudiantes en la asignatura de Circuitos Eléctricos I a lo largo de 14 años consecutivos. Este análisis se realizó principalmente utilizando Python y técnicas de análisis de series de tiempo.

Durante este proceso, se realizaron cálculos para obtener estadísticas descriptivas iniciales relacionadas con el rendimiento promedio de los estudiantes de Circuitos Eléctricos I. Luego, se procedió a la creación de gráficos de series de tiempo, los cuales permitieron visualizar la evolución del rendimiento promedio a lo largo de los años. Además, se realizó un análisis de tendencias para identificar patrones o cambios significativos en el rendimiento a lo largo del tiempo. Este enfoque proporcionó una comprensión más profunda del comportamiento de los estudiantes en la asignatura a lo largo del período de estudio.

CIRCUITOS ELECTRICOS I	
AÑO	
2010	4.454706
2011	4.876410
2012	4.748462
2013	5.373684
2014	5.003077
2015	4.739318
2016	4.401860
2017	4.444000
2018	4.464231
2019	4.792500
2020	4.316471
2021	4.818772
2022	5.002083
2023	4.923478

Figura 15: Base de datos de Circuitos Eléctricos I.

Se observó que la nota mínima promedio anual de los estudiantes en la asignatura de Circuitos Eléctricos I fue de 4.32 en el año 2020. Por otro lado, se registró una nota promedio semestral máxima de 5.37 en el año 2013. En cuanto al desempeño promedio, este se situó en 4.74, con una desviación estándar de 0.30.

Datos	14
Promedio	4.739932
Desviación estándar	0.296068
Mínimo	4.316471
25 %	4.457087
50 %	4.770481
75 %	4.911711
Máximo	5.373684

Cuadro 21: Desempeño de los estudiantes en Circuitos Eléctricos I.

La Figura 16 muestra las notas promedio del curso por semestre a lo largo del periodo. La ecuación de la tendencia, $y = -0,00x + 4,76$, describe una relación lineal entre el año (x) y el promedio de las calificaciones en la asignatura de Circuitos Eléctricos I (y). En esta instancia, el coeficiente $-0,00$ representa la pendiente de la línea de tendencia, indicando una tendencia casi horizontal en el promedio de calificaciones a lo largo del tiempo. Por otro lado, el término independiente ($4,76$) señala el valor esperado del promedio de calificaciones en el año inicial del estudio (2010 en este caso, ajustado a $x = 0$).

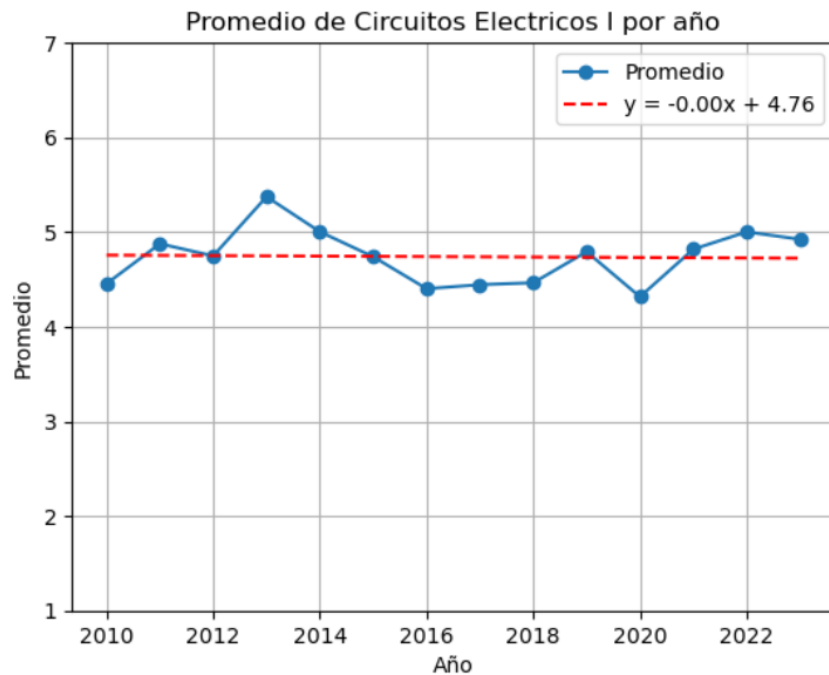


Figura 16: Serie de tiempo de las notas promedios semestrales de los estudiantes de Circuitos Eléctricos I.

Al interpretar la ecuación, podemos concluir que la tendencia general a lo largo de los años muestra una estabilidad en el promedio de las notas de la asignatura de Circuitos Eléctricos I. Con una pendiente de 0 en la ecuación, no hay un cambio promedio esperado en el rendimiento de los estudiantes en esta asignatura a lo largo del tiempo. Sin embargo, es fundamental tener en cuenta que esta es una simplificación y que pueden existir otros factores no considerados en la ecuación que influyan en el rendimiento académico.

Es necesario analizar la naturaleza de los datos para seleccionar el modelo correcto y mostrar si se trata de datos cíclicos sin tendencia o con actividad estacional.

3.1.5. Pruebas de estacionariedad

Se examinaron los gráficos de la función de autocorrelación (ACF) y de la función de autocorrelación parcial (PACF) de las notas promedio semestrales de los estudiantes en Circuitos Eléctricos I. Estos gráficos, mostrados en la Figura 17, proporcionaron información inicial sobre la estacionariedad de la serie temporal.

En los gráficos ACF y PACF, no se observaron picos significativos en ningún retraso, y además tenemos correlaciones positivas y negativas, lo que sugiere que la serie temporal podría ser estacionaria o que la autocorrelación ha disminuido. Esto indica una posible estacionariedad en la serie de datos, aunque se requieren pruebas estadísticas adicionales para confirmar esta observación.

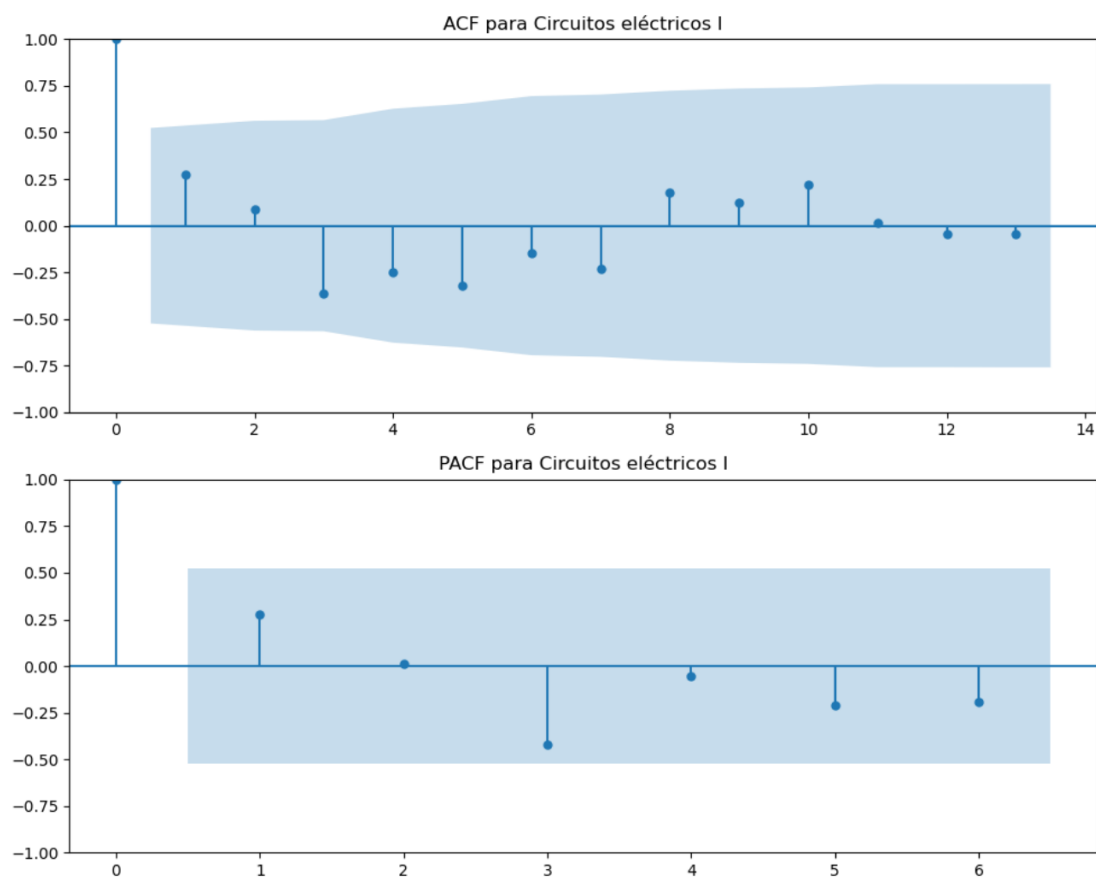


Figura 17: Gráfico ACF y PACF de las notas promedio semestrales de los estudiantes en Circuitos Eléctricos I.

Con estas observaciones preliminares, se procedió a realizar las pruebas de estacionariedad ADF y KPSS para verificar la presencia de estacionariedad en la serie temporal.

Un ciclo estacionario tiene una media y una varianza constante en el tiempo y no hay patrones en el ciclo. Es necesario que la serie sea estacionaria antes de continuar con la estimación de un modelo ARIMA, ya que en este análisis se utilizó la prueba Augmented Dickey-Fuller (ADF) y la prueba de Kwiatkowski, Phillips, Schmidt y Shin (KPSS) para verificar la estacionariedad en el rendimiento académico de los estudiantes en métodos cuantitativos.

-
- Probamos la hipótesis para la prueba ADF de que;

- H0: La secuencia no es estacionaria.
- H1: La secuencia estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es estacionaria.

Parámetro	Valor
Estadística ADF	-2.52842
Valor-p	0.10871
Lags utilizados	0
Valor critico (1 %)	-4.06885
Valor critico (5 %)	-3.12715
Valor critico (10 %)	-2.70173

Cuadro 22: Resultado de la prueba augmented Dickey-Fuller.

- Probamos la hipótesis para la prueba KPSS de que;

- H0: La secuencia es estacionaria.
- H1: La secuencia no es estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es no estacionaria.

Parámetro	Valor
Estadística KPSS	0.11011
Valor-p	0.1
Lags utilizados	1
Valor critico (1 %)	0.739
Valor critico (5 %)	0.463
Valor critico (10 %)	0.347

Cuadro 23: Resultado de la prueba KPSS.

Basándonos en los resultados de las pruebas de estacionariedad ADF y KPSS mostrados en el Cuadro 22 y Cuadro 23, podemos tomar decisiones sobre la estacionariedad de la secuencia.

Para la prueba ADF, la hipótesis nula (H0) establece que la secuencia no es estacionaria, mientras que la hipótesis alternativa (H1) sugiere lo contrario. Con un nivel de significancia del 5 %, el valor-p asociado a la prueba ADF es 0.10871, que es superior al umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H0) y concluimos que la secuencia es no estacionaria.

En contraste, para la prueba KPSS, la H0 postula que la secuencia es estacionaria, mientras que la H1 sugiere que no lo es. Aquí, con un nivel de significancia del 5 %, el valor-p asociado a la prueba KPSS es 0.1, que excede el umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H0) y concluimos que la secuencia es estacionaria.

Los resultados obtenidos a través de las pruebas de estacionariedad presentan una discrepancia significativa en cuanto a la naturaleza de la serie temporal analizada. Mientras que la prueba

ADF sugiere consistentemente la estacionariedad de la secuencia, indicada por la no rechazo de la hipótesis nula, la prueba KPSS no proporciona evidencia suficiente para refutar la estacionariedad. Esta discrepancia plantea interrogantes sobre la verdadera naturaleza de la serie temporal. Esto se puede deber a que no poseemos suficientes datos para obtener una conclusión más clara.

A la luz de los resultados contradictorios de las pruebas de estacionariedad ADF y KPSS, se ha optado por asumir la estacionariedad de la serie temporal analizada. Aunque la prueba ADF no proporcionó suficiente evidencia para rechazar la hipótesis nula de no estacionariedad, los gráficos de autocorrelación ACF y PACF mostraron indicios de posible estacionariedad, lo que respalda nuestra decisión. Reconocemos que esta asunción puede ser discutible y que hay incertidumbre en la verdadera naturaleza de la serie temporal. Sin embargo, dadas las limitaciones de los datos y la falta de consenso entre las pruebas estadísticas, consideramos que es prudente proceder bajo la suposición de estacionariedad para los fines de este análisis.

3.1.6. Selección de un modelo

Ahora procederemos a seleccionar el modelo ARIMA más apropiado para nuestra serie temporal. Utilizaremos las capacidades de Python para explorar y evaluar diversos modelos, con el fin de identificar aquel que mejor se ajuste a nuestras necesidades. Este análisis nos proporcionará la base necesaria para tomar decisiones informadas respecto a la elección del modelo más adecuado.

MODELO	AIC	BIC	HQ
ARIMA (1,0,1)	11.461	14.018	11.225
ARIMA (0,0,0)	8.612	9.890	8.494
ARIMA (1,0,0)	9.461	11.379	9.284

Cuadro 24: Identificación del modelo.

El modelo con los mínimos Criterios de Información de Akaike (AIC), Criterio de Información Bayesiano (BIC) y Hannan-Quinn (HQ) es el más adecuado para la secuencia. Por lo tanto, al probar todos los modelos competidores en el Cuadro 24, el modelo ARIMA (0, 0, 0) tiene los valores más bajos y, por lo tanto, el mejor modelo predicativo.

3.1.7. Estimación de parámetros

El Cuadro 25 a continuación muestra estimaciones de los parámetros del sistema ARIMA (0, 0, 0).

Tipo	Coefficiente	Error estándar	Valor Z	Valor-p
Constante	4.7399	0.080	59.602	0.000
sigma2 (σ^2)	0.0814	0.036	2.242	0.025

Cuadro 25: Muestra de la estimación de parámetros.

El modelo ARIMA(0,0,0) indica que estamos utilizando un modelo autorregresivo integrado de media móvil con cero componentes autorregresivas (AR), cero componentes de integración (I) y cero componentes de media móvil (MA). Esto significa que no hay dependencia de los valores pasados, ni tendencia ni error en los datos.

En cuanto a la estimación de parámetros, la constante tiene un valor de 4,7399 con un error estándar de 0,080. El valor Z es alto (59,602), lo que indica que la constante es estadísticamente

significativa. Además, el valor-p asociado es muy bajo (0,000), lo que sugiere una alta significancia estadística.

En resumen, con un modelo ARIMA(0,0,0), se está utilizando una estimación constante para predecir la serie temporal de notas promedio semestrales en Circuitos Eléctricos I. La constante estimada parece ser estadísticamente significativa, lo que sugiere que este modelo puede ser útil para predecir las notas con un nivel aceptable de precisión.

3.1.8. Resumen y conclusión

Los resultados del análisis revelaron una variabilidad significativa en el rendimiento de los estudiantes en la asignatura de Circuitos Eléctricos I a lo largo de los 14 años de estudio. En promedio, los alumnos finalizaron la asignatura con una calificación de 4.74, pero se observaron fluctuaciones destacadas en diferentes períodos. Por ejemplo, el año 2020 registró el rendimiento más bajo con un promedio de 4.32. Por el contrario, el año 2013 destacó como el de mejor desempeño, con un promedio de 5.37.

Los objetivos específicos de este estudio incluyeron analizar la evolución del rendimiento estudiantil en la asignatura y elaborar un modelo adecuado que se ajustara con precisión a dicho desempeño. Para lograr estos objetivos, se implementó una metodología integral que comenzó con un análisis descriptivo detallado de los datos, seguido de gráficos de las funciones de autocorrelación ACF y PACF para identificar estacionariedad. Además, se utilizó el test ADF y el test KPSS para respaldar las hipótesis obtenidas del análisis de los gráficos de autocorrelación.

Basándonos en los resultados de las pruebas estadísticas, se seleccionó el modelo ARIMA más apropiado utilizando los criterios de información AIC, BIC y HQ. Se encontró que el modelo ARIMA(0,0,0) era el más adecuado para la serie temporal estudiada. Este modelo implica una constante de 4.7399, que resultó ser estadísticamente significativa. Si bien este modelo simple no captura la dinámica compleja de la serie temporal, proporciona una base sólida para la predicción del rendimiento académico en la asignatura de Circuitos Eléctricos I.

En resumen, este estudio proporcionó una visión completa del rendimiento estudiantil en la asignatura Circuitos Eléctricos I, desde un análisis descriptivo inicial hasta la selección y estimación de un modelo ARIMA adecuado. Los resultados obtenidos contribuyen a una mejor comprensión de los factores que influyen en el desempeño académico de los estudiantes en esta área específica y pueden servir como base para futuras investigaciones y acciones de mejora educativa.

3.2. Electricidad

Durante el análisis de la asignatura Electricidad, se examinaron los promedios obtenidos por semestre a lo largo de los 14 años de estudio. En promedio, los alumnos finalizaban la asignatura con una calificación de 5.13. Sin embargo, se observaron variaciones significativas en el rendimiento a lo largo de los años. En particular, en el año 2023 se registró el rendimiento más bajo como curso, con un promedio de 4.67. Por el contrario, en el año 2017 se alcanzó el rendimiento más alto como curso, con un promedio de 5.46, destacando como un año excepcional en términos de desempeño académico en esta asignatura.

3.2.1. Objetivos específicos:

- Analizar la evolución del rendimiento estudiantil en la asignatura de Electricidad, dentro del marco del curso “1er año Mc.El.”, impartido en la Escuela de Ingeniería.
- Elaborar un modelo adecuado que se ajuste con mayor precisión al desempeño de los estudiantes en la asignatura de Electricidad, perteneciente al curso “1er año Mc.El.”, dictado en la Escuela de Ingeniería.

3.2.2. Preguntas de investigación

- ¿Se identifica alguna tendencia en el rendimiento de los estudiantes en la asignatura de Electricidad, dentro del curso “1er año Mc.El.”, impartido en la Escuela de Ingeniería?
- ¿Cuál es el modelo más adecuado para describir el desempeño de los estudiantes en la asignatura de Electricidad, dentro del curso “1er año Mc.El.”, dictado en la Escuela de Ingeniería?

3.2.3. Metodología

La metodología empleada en este estudio abarcó una serie de pasos fundamentales para el análisis exhaustivo de la serie temporal en cuestión. En primer lugar, se realizó un análisis descriptivo detallado de los datos, con el fin de comprender su comportamiento general, identificar posibles tendencias y visualizar patrones significativos. Se incluyó un gráfico de línea para representar visualmente la evolución de los datos, añadiendo una línea de tendencia para resaltar cualquier dirección en el conjunto de datos.

Posteriormente, se procedió a generar gráficos de las funciones de autocorrelación ACF y autocorrelación parcial PACF. Estas visualizaciones jugaron un papel crucial en la evaluación preliminar de la estacionariedad de los datos y en la identificación de posibles estructuras de autocorrelación. Estos correlogramas proporcionaron una comprensión visual inicial de la relación entre las observaciones en diferentes puntos temporales.

Con base en las observaciones obtenidas de los correlogramas, se llevó a cabo una evaluación más rigurosa de la estacionariedad de la serie temporal. Para ello, se realizaron pruebas de estacionariedad utilizando tanto el test ADF (Augmented Dickey-Fuller) como el test KPSS (Kwiatkowski, Phillips, Schmidt, Shin). Estas pruebas estadísticas sirvieron para reafirmar o refutar las hipótesis extraídas de los correlogramas, proporcionando una evaluación más sólida de la estacionariedad de la serie temporal.

Finalmente, una vez confirmada la estacionariedad de los datos, se procedió a la selección del modelo ARIMA más adecuado para el conjunto de datos analizados. Esto implicó la identificación

de los parámetros óptimos del modelo ARIMA, aprovechando tanto la información derivada de los correlogramas como de las pruebas estadísticas de estacionariedad.

En resumen, la metodología empleada abarcó desde la exploración inicial de los datos y la representación gráfica, pasando por el análisis de la estacionariedad con correlogramas y pruebas estadísticas, hasta la selección y estimación de un modelo ARIMA apropiado. Este enfoque integral garantizó un análisis completo y riguroso de la serie temporal, proporcionando una base sólida para su modelado y predicción.

3.2.4. Estadística Descriptiva de las notas promedios semestrales de los estudiantes

Para realizar un estudio efectivo de la asignatura en estudio, se inició reduciendo la base de datos para incluir únicamente las columnas relevantes (Figura 18). Posteriormente, se llevó a cabo un análisis exploratorio de datos centrado en el rendimiento promedio de los estudiantes en la asignatura de Electricidad a lo largo de 14 años consecutivos. Este análisis se realizó principalmente utilizando Python y técnicas de análisis de series de tiempo.

Durante este proceso, se realizaron cálculos para obtener estadísticas descriptivas iniciales relacionadas con el rendimiento promedio de los estudiantes de Electricidad. Luego, se procedió a la creación de gráficos de series de tiempo, los cuales permitieron visualizar la evolución del rendimiento promedio a lo largo de los años. Además, se realizó un análisis de tendencias para identificar patrones o cambios significativos en el rendimiento a lo largo del tiempo. Este enfoque proporcionó una comprensión más profunda del comportamiento de los estudiantes en la asignatura a lo largo del período de estudio.

ELECTRICIDAD	
AÑO	
2010	5.025882
2011	5.337179
2012	5.149487
2013	4.965789
2014	4.975128
2015	5.174091
2016	4.730233
2017	5.463556
2018	5.191154
2019	5.297143
2020	5.308039
2021	5.158421
2022	5.356250
2023	4.670000

Figura 18: Base de datos de Electricidad.

Se observó que la nota mínima promedio anual de los estudiantes en la asignatura de Electricidad fue de 4.67 en el año 2023. Por otro lado, se registró una nota promedio semestral máxima de

Datos	14
Promedio	5.128739
Desviación estándar	0.232885
Mínimo	4.670000
25 %	4.987817
50 %	5.166256
75 %	5.305315
Máximo	5.463556

Cuadro 26: Desempeño de los estudiantes de Electricidad.

5.46 en el año 2017. En cuanto al desempeño promedio, este se situó en 5.13, con una desviación estándar de 0.23.

La Figura 19 muestra las notas promedio del curso por semestre a lo largo del periodo. La ecuación de la tendencia, $y = 0,00x + 5,12$, describe una relación lineal entre el año (x) y el promedio de las calificaciones en la asignatura de Electricidad (y). En esta instancia, el coeficiente 0,00 representa la pendiente de la línea de tendencia, indicando una tendencia casi horizontal en el promedio de calificaciones a lo largo del tiempo. Por otro lado, el término independiente (5,12) señala el valor esperado del promedio de calificaciones en el año inicial del estudio (2010 en este caso, ajustado a $x = 0$).

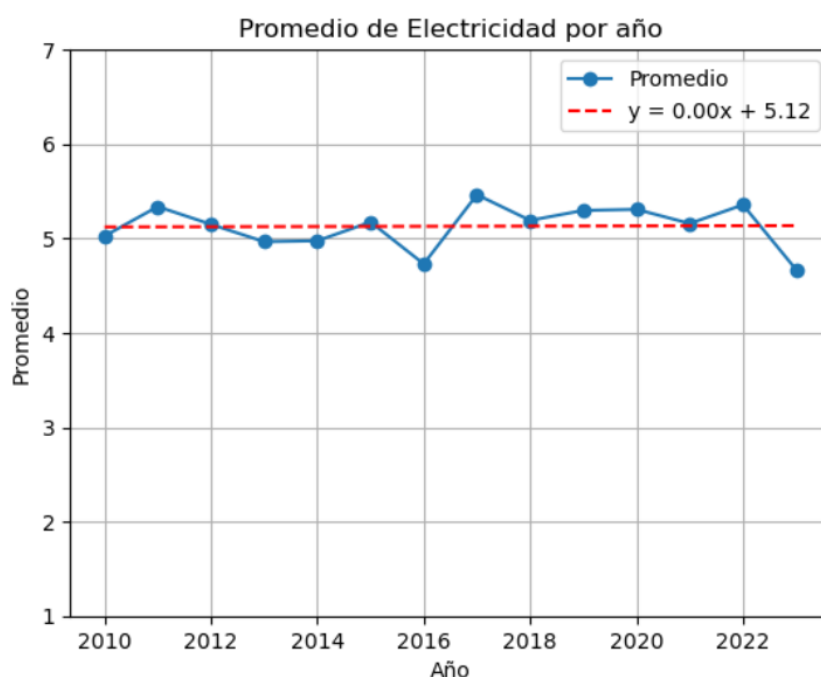


Figura 19: Serie de tiempo de las notas promedios semestrales de los estudiantes en Electricidad.

Al interpretar la ecuación, podemos concluir que la tendencia general a lo largo de los años muestra una estabilidad en el promedio de las notas de la asignatura de Electricidad. Con una pendiente de 0 en la ecuación, no hay un cambio promedio esperado en el rendimiento de los estudiantes en esta asignatura a lo largo del tiempo. Sin embargo, es fundamental tener en cuenta que esta es una simplificación y que pueden existir otros factores no considerados en la ecuación.

que influyan en el rendimiento académico.

Es necesario analizar la naturaleza de los datos para seleccionar el modelo correcto y mostrar si se trata de datos cíclicos sin tendencia o con actividad estacional.

3.2.5. Pruebas de estacionariedad

Se examinaron los gráficos de la función de autocorrelación (ACF) y de la función de autocorrelación parcial (PACF) de las notas promedio semestrales de los estudiantes en Electricidad. Estos gráficos, mostrados en la Figura 20, proporcionaron información inicial sobre la estacionariedad de la serie temporal.

En los gráficos ACF y PACF, no se observaron picos significativos en ningún retraso, y además tenemos correlaciones positivas y negativas, lo que sugiere que la serie temporal podría ser estacionaria o que la autocorrelación ha disminuido. Esto indica una posible estacionariedad en la serie de datos, aunque se requieren pruebas estadísticas adicionales para confirmar esta observación.

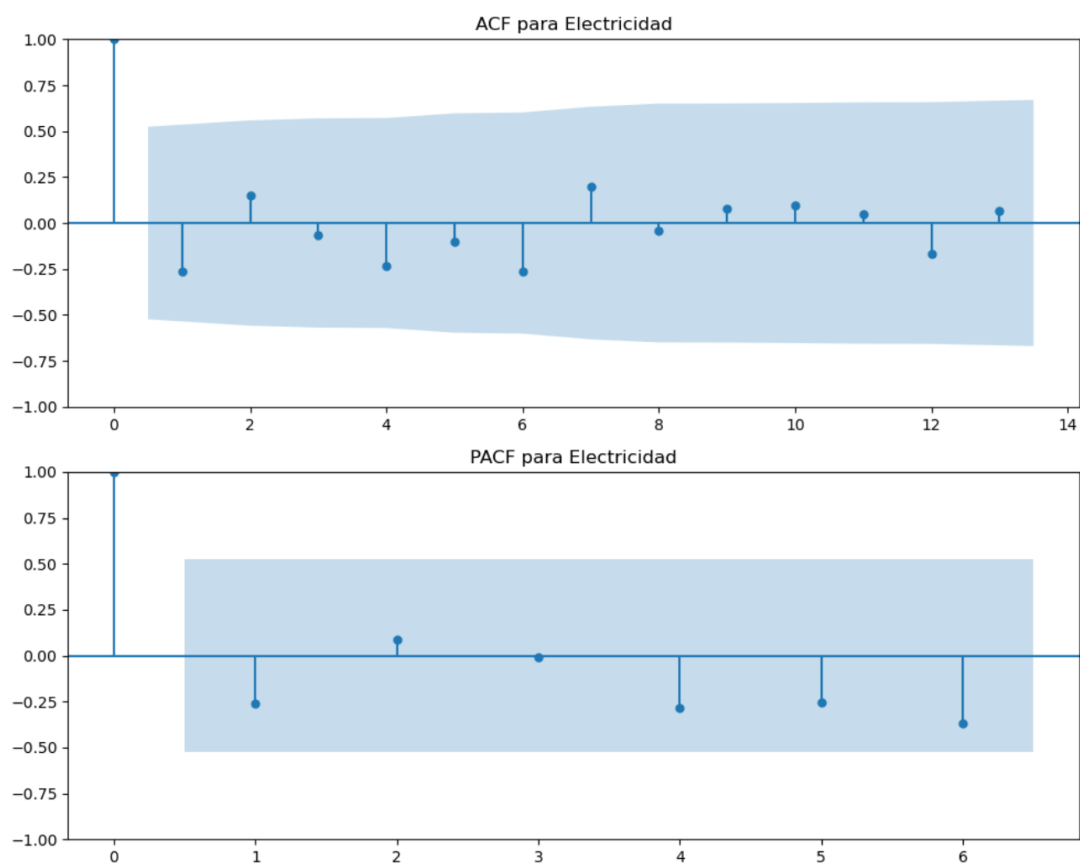


Figura 20: Gráfico ACF y PACF de las notas promedio semestrales de los estudiantes en Electricidad.

Con estas observaciones preliminares, se procedió a realizar las pruebas de estacionariedad ADF y KPSS para verificar la presencia de estacionariedad en la serie temporal.

Un ciclo estacionario tiene una media y una varianza constante en el tiempo y no hay patrones en el ciclo. Es necesario que la serie sea estacionaria antes de continuar con la estimación de un modelo ARIMA, ya que en este análisis se utilizó la prueba Augmented Dickey-Fuller (ADF) y la

prueba de Kwiatkowski, Phillips, Schmidt y Shin (KPSS) para verificar la estacionariedad en el rendimiento académico de los estudiantes en métodos cuantitativos.

- Probamos la hipótesis para la prueba ADF de que;
 - H0: La secuencia no es estacionaria.
 - H1: La secuencia estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es estacionaria.

Parámetro	Valor
Estadística ADF	-1.91123
Valor-p	0.32680
Lags utilizados	5
Valor critico (1 %)	-4.66518
Valor critico (5 %)	-3.36719
Valor critico (10 %)	2.80296

Cuadro 27: Resultado de la prueba augmented Dickey-Fuller.

- Probamos la hipótesis para la prueba KPSS de que;
 - H0: La secuencia es estacionaria.
 - H1: La secuencia no es estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es no estacionaria.

Parámetro	Valor
Estadística KPSS	0.10287
Valor-p	0.1
Lags utilizados	2
Valor critico (1 %)	0.739
Valor critico (5 %)	0.463
Valor critico (10 %)	0.347

Cuadro 28: Resultado de la prueba KPSS.

Basándonos en los resultados de las pruebas de estacionariedad ADF y KPSS mostrados en el Cuadro 27 y Cuadro 28, podemos tomar decisiones sobre la estacionariedad de la secuencia.

Para la prueba ADF, la hipótesis nula (H0) establece que la secuencia no es estacionaria, mientras que la hipótesis alternativa (H1) sugiere lo contrario. Con un nivel de significancia del 5 %, el valor-p asociado a la prueba ADF es 0.32680, que es superior al umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H0) y concluimos que la secuencia es no estacionaria.

En contraste, para la prueba KPSS, la H0 postula que la secuencia es estacionaria, mientras que la H1 sugiere que no lo es. Aquí, con un nivel de significancia del 5 %, el valor-p asociado a la

prueba KPSS es 0.1, que excede el umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H_0) y concluimos que la secuencia es estacionaria.

Los resultados obtenidos a través de las pruebas de estacionariedad presentan una discrepancia significativa en cuanto a la naturaleza de la serie temporal analizada. Mientras que la prueba ADF sugiere consistentemente la estacionariedad de la secuencia, indicada por la no rechazo de la hipótesis nula, la prueba KPSS no proporciona evidencia suficiente para refutar la estacionariedad. Esta discrepancia plantea interrogantes sobre la verdadera naturaleza de la serie temporal. Esto se puede deber a que no poseemos suficientes datos para obtener una conclusión más clara.

A la luz de los resultados contradictorios de las pruebas de estacionariedad ADF y KPSS, se ha optado por asumir la estacionariedad de la serie temporal analizada. Aunque la prueba ADF no proporcionó suficiente evidencia para rechazar la hipótesis nula de no estacionariedad, los gráficos de autocorrelación ACF y PACF mostraron indicios de posible estacionariedad, lo que respalda nuestra decisión. Reconocemos que esta asunción puede ser discutible y que hay incertidumbre en la verdadera naturaleza de la serie temporal. Sin embargo, dadas las limitaciones de los datos y la falta de consenso entre las pruebas estadísticas, consideramos que es prudente proceder bajo la suposición de estacionariedad para los fines de este análisis.

3.2.6. Selección de un modelo

Ahora procederemos a seleccionar el modelo ARIMA más apropiado para nuestra serie temporal. Utilizaremos las capacidades de Python para explorar y evaluar diversos modelos, con el fin de identificar aquel que mejor se ajuste a nuestras necesidades. Este análisis nos proporcionará la base necesaria para tomar decisiones informadas respecto a la elección del modelo más adecuado.

MODELO	AIC	BIC	HQ
ARIMA (1,0,1)	4.337	6.893	4.100
ARIMA (0,0,0)	1.891	3.169	1.773
ARIMA (1,0,0)	2.494	4.412	2.317

Cuadro 29: Identificación del modelo.

El modelo con los mínimos Criterios de Información de Akaike (AIC), Criterio de Información Bayesiano (BIC) y Hannan-Quinn (HQ) es el más adecuado para la secuencia. Por lo tanto, al probar todos los modelos competidores en el Cuadro 29, el modelo ARIMA (0, 0, 0) tiene los valores más bajos y, por lo tanto, el mejor modelo predictivo.

3.2.7. Estimación de parámetros

El Cuadro 30 a continuación muestra estimaciones de los parámetros del sistema ARIMA (0, 0, 0).

Tipo	Coefficiente	Error estándar	Valor Z	Valor-p
Constante	5.1287	0.070	73.625	0.000
sigma2 (σ^2)	0.0504	0.025	1.993	0.046

Cuadro 30: Muestra de la estimación de parámetros.

El modelo ARIMA(0,0,0) indica que estamos utilizando un modelo autorregresivo integrado de media móvil con cero componentes autorregresivos (AR), cero componentes de integración (I)

y cero componentes de media móvil (MA). Esto significa que no hay dependencia de los valores pasados, ni tendencia ni error en los datos.

En cuanto a la estimación de parámetros, la constante tiene un valor de 5,1287 con un error estándar de 0,070. El valor Z es alto (73,625), lo que indica que la constante es estadísticamente significativa. Además, el valor-p asociado es muy bajo (0,000), lo que sugiere una alta significancia estadística.

En resumen, con un modelo ARIMA(0,0,0), se está utilizando una estimación constante para predecir la serie temporal de notas promedio semestrales en Electricidad. La constante estimada parece ser estadísticamente significativa, lo que sugiere que este modelo puede ser útil para predecir las notas con un nivel aceptable de precisión.

3.2.8. Resumen y conclusión

Los resultados del análisis revelaron una variabilidad significativa en el rendimiento de los estudiantes en la asignatura de Electricidad a lo largo de los 14 años de estudio. En promedio, los alumnos finalizaron la asignatura con una calificación de 5.13, pero se observaron fluctuaciones destacadas en diferentes períodos. Por ejemplo, el año 2023 registró el rendimiento más bajo con un promedio de 4.67. Por el contrario, el año 2017 destacó como el de mejor desempeño, con un promedio de 5.46.

Los objetivos específicos de este estudio incluyeron analizar la evolución del rendimiento estudiantil en la asignatura y elaborar un modelo adecuado que se ajustara con precisión a dicho desempeño. Para lograr estos objetivos, se implementó una metodología integral que comenzó con un análisis descriptivo detallado de los datos, seguido de gráficos de las funciones de autocorrelación ACF y PACF para identificar estacionariedad. Además, se utilizó el test ADF y el test KPSS para respaldar las hipótesis obtenidas del análisis de los gráficos de autocorrelación.

Basándonos en los resultados de las pruebas estadísticas, se seleccionó el modelo ARIMA más apropiado utilizando los criterios de información AIC, BIC y HQ. Se encontró que el modelo ARIMA(0,0,0) era el más adecuado para la serie temporal estudiada. Este modelo implica una constante de 5.1287, que resultó ser estadísticamente significativa. Si bien este modelo simple no captura la dinámica compleja de la serie temporal, proporciona una base sólida para la predicción del rendimiento académico en la asignatura de Electricidad.

En resumen, este estudio proporcionó una visión completa del rendimiento estudiantil en la asignatura Electricidad, desde un análisis descriptivo inicial hasta la selección y estimación de un modelo ARIMA adecuado. Los resultados obtenidos contribuyen a una mejor comprensión de los factores que influyen en el desempeño académico de los estudiantes en esta área específica y pueden servir como base para futuras investigaciones y acciones de mejora educativa.

3.3. Física Técnica I

Durante el análisis de la asignatura Física Técnica I, se examinaron los promedios obtenidos por semestre a lo largo de los 14 años de estudio. En promedio, los alumnos finalizaban la asignatura con una calificación de 5.26. Sin embargo, se observaron variaciones significativas en el rendimiento a lo largo de los años. En particular, en el año 2013 se registró el rendimiento más bajo como curso, con un promedio de 4.71. Por el contrario, en el año 2022 se alcanzó el rendimiento más alto como curso, con un promedio de 5.85, destacando como un año excepcional en términos de desempeño académico en esta asignatura.

3.3.1. Objetivos específicos:

- Analizar la evolución del rendimiento estudiantil en la asignatura de Física Técnica I, dentro del marco del curso “1er año Mc.El.”, impartido en la Escuela de Ingeniería.
- Elaborar un modelo adecuado que se ajuste con mayor precisión al desempeño de los estudiantes en la asignatura de Física Técnica I, perteneciente al curso “1er año Mc.El.”, dictado en la Escuela de Ingeniería.

3.3.2. Preguntas de investigación

- ¿Se identifica alguna tendencia en el rendimiento de los estudiantes en la asignatura de Física Técnica I, dentro del curso “1er año Mc.El.”, impartido en la Escuela de Ingeniería?
- ¿Cuál es el modelo más adecuado para describir el desempeño de los estudiantes en la asignatura de Física Técnica I, dentro del curso “1er año Mc.El.”, dictado en la Escuela de Ingeniería?

3.3.3. Metodología

La metodología empleada en este estudio abarcó una serie de pasos fundamentales para el análisis exhaustivo de la serie temporal en cuestión. En primer lugar, se realizó un análisis descriptivo detallado de los datos, con el fin de comprender su comportamiento general, identificar posibles tendencias y visualizar patrones significativos. Se incluyó un gráfico de línea para representar visualmente la evolución de los datos, añadiendo una línea de tendencia para resaltar cualquier dirección en el conjunto de datos.

Posteriormente, se procedió a generar gráficos de las funciones de autocorrelación ACF y autocorrelación parcial PACF. Estas visualizaciones jugaron un papel crucial en la evaluación preliminar de la estacionariedad de los datos y en la identificación de posibles estructuras de autocorrelación. Estos correlogramas proporcionaron una comprensión visual inicial de la relación entre las observaciones en diferentes puntos temporales.

Con base en las observaciones obtenidas de los correlogramas, se llevó a cabo una evaluación más rigurosa de la estacionariedad de la serie temporal. Para ello, se realizaron pruebas de estacionariedad utilizando tanto el test ADF (Augmented Dickey-Fuller) como el test KPSS (Kwiatkowski, Phillips, Schmidt, Shin). Estas pruebas estadísticas sirvieron para reafirmar o refutar las hipótesis extraídas de los correlogramas, proporcionando una evaluación más sólida de la estacionariedad de la serie temporal.

Finalmente, una vez confirmada la estacionariedad de los datos, se procedió a la selección del modelo ARIMA más adecuado para el conjunto de datos analizados. Esto implicó la identificación

de los parámetros óptimos del modelo ARIMA, aprovechando tanto la información derivada de los correlogramas como de las pruebas estadísticas de estacionariedad.

En resumen, la metodología empleada abarcó desde la exploración inicial de los datos y la representación gráfica, pasando por el análisis de la estacionariedad con correlogramas y pruebas estadísticas, hasta la selección y estimación de un modelo ARIMA apropiado. Este enfoque integral garantizó un análisis completo y riguroso de la serie temporal, proporcionando una base sólida para su modelado y predicción.

3.3.4. Estadística Descriptiva de las notas promedios semestrales de los estudiantes

Para realizar un estudio efectivo de la asignatura en estudio, se inició reduciendo la base de datos para incluir únicamente las columnas relevantes (Figura 21). Posteriormente, se llevó a cabo un análisis exploratorio de datos centrado en el rendimiento promedio de los estudiantes en la asignatura de Física Técnica I a lo largo de 14 años consecutivos. Este análisis se realizó principalmente utilizando Python y técnicas de análisis de series de tiempo.

Durante este proceso, se realizaron cálculos para obtener estadísticas descriptivas iniciales relacionadas con el rendimiento promedio de los estudiantes de Física Técnica I. Luego, se procedió a la creación de gráficos de series de tiempo, los cuales permitieron visualizar la evolución del rendimiento promedio a lo largo de los años. Además, se realizó un análisis de tendencias para identificar patrones o cambios significativos en el rendimiento a lo largo del tiempo. Este enfoque proporcionó una comprensión más profunda del comportamiento de los estudiantes en la asignatura a lo largo del período de estudio.

FISICA TECNICA	
AÑO	
2010	4.853824
2011	5.039487
2012	5.005641
2013	4.711053
2014	5.453077
2015	5.022045
2016	4.826744
2017	5.158000
2018	5.440577
2019	5.691786
2020	5.254314
2021	5.761404
2022	5.849167
2023	5.540870

Figura 21: Base de datos de Física Técnica I.

Se observó que la nota mínima promedio anual de los estudiantes en la asignatura de Física Técnica I fue de 4.71 en el año 2013. Por otro lado, se registró una nota promedio semestral máxima de 5.85 en el año 2022. En cuanto al desempeño promedio, este se situó en 5.26, con una

Datos	14
Promedio	5.257713
Desviación estándar	0.368790
Mínimo	4.711053
25 %	5.009742
50 %	5.206157
75 %	5.518921
Máximo	5.849167

Cuadro 31: Desempeño de los estudiantes de Física Técnica I.

desviación estándar de 0.37.

La Figura 22 muestra las notas promedio del curso por semestre a lo largo del periodo. La ecuación de la tendencia, $y = 0,07x + 4,61$, describe una relación lineal entre el año (x) y el promedio de las calificaciones en la asignatura de Física Técnica I (y). En esta instancia, el coeficiente 0,07 representa la pendiente de la línea de tendencia, lo que sugiere un aumento lineal en el promedio de las notas a medida que pasa el tiempo. Por otro lado, el término independiente (4,61) representa el valor esperado del promedio de las notas en el año cero del estudio (en este caso, 2010 ajustado a $x = 0$).

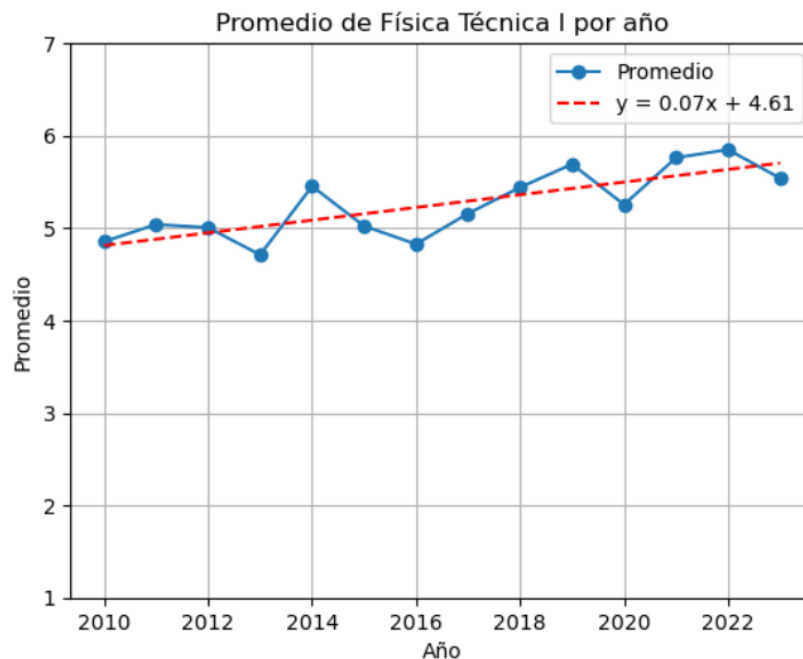


Figura 22: Serie de tiempo de las notas promedios semestrales de los estudiantes en Física Técnica I.

Al interpretar la ecuación, podemos concluir que la tendencia general a lo largo de los años muestra un incremento de aproximadamente 0.07 unidades en el promedio de las notas de la asignatura de Física Técnica I por semestre. Es decir, en promedio, se espera que el rendimiento de los estudiantes en esta asignatura aumente en 0.07 puntos cada semestre. Sin embargo, es importante tener en cuenta que esta es una simplificación y que pueden existir otros factores no considerados en la ecuación que influyan en el rendimiento académico.

Es necesario analizar la naturaleza de los datos para seleccionar el modelo correcto y mostrar si se trata de datos cíclicos sin tendencia o con actividad estacional.

3.3.5. Pruebas de estacionariedad

Se examinaron los gráficos de la función de autocorrelación (ACF) y de la función de autocorrelación parcial (PACF) de las notas promedio semestrales de los estudiantes en Física Técnica I. Estos gráficos, mostrados en la Figura 23, proporcionaron información inicial sobre la estacionariedad de la serie temporal.

En los gráficos ACF y PACF, no se observaron picos significativos en ningún retraso, y además tenemos correlaciones positivas y negativas, lo que sugiere que la serie temporal podría ser estacionaria o que la autocorrelación ha disminuido. Esto indica una posible estacionariedad en la serie de datos, aunque se requieren pruebas estadísticas adicionales para confirmar esta observación.

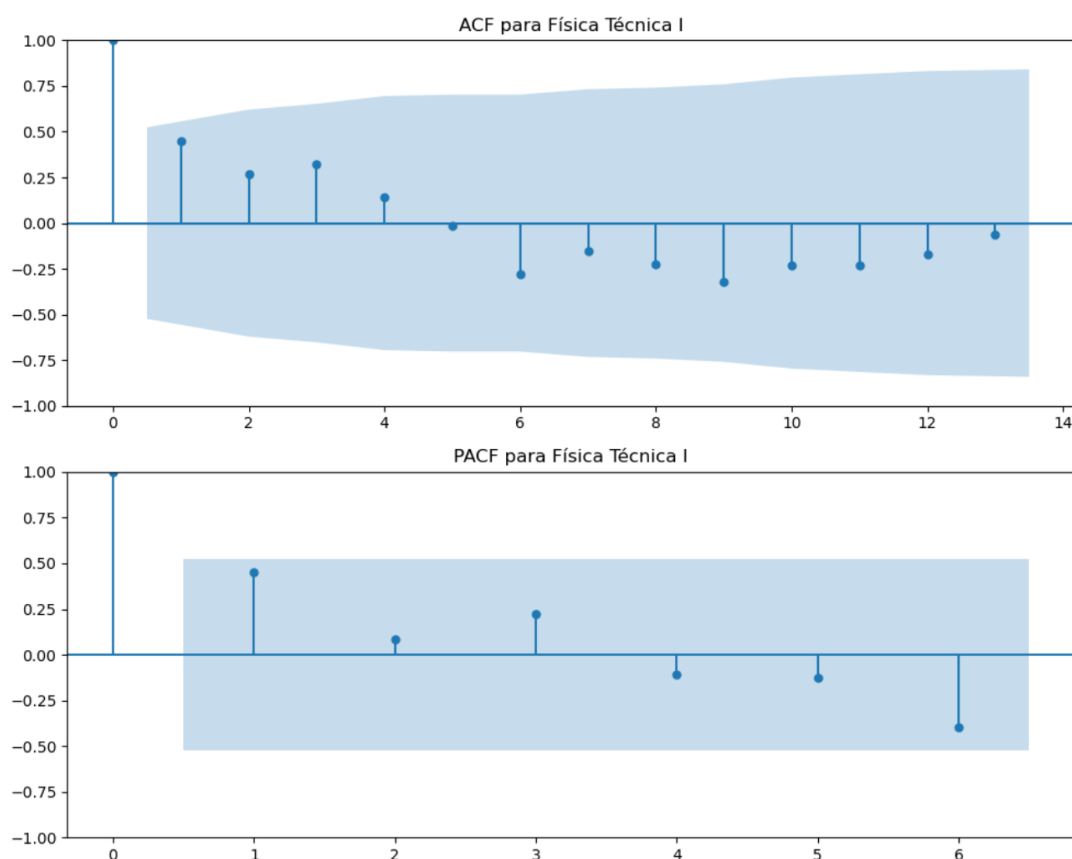


Figura 23: Gráfico ACF y PACF de las notas promedio semestrales de los estudiantes en Física Técnica I.

Con estas observaciones preliminares, se procedió a realizar las pruebas de estacionariedad ADF y KPSS para verificar la presencia de estacionariedad en la serie temporal.

Un ciclo estacionario tiene una media y una varianza constante en el tiempo y no hay patrones en el ciclo. Es necesario que la serie sea estacionaria antes de continuar con la estimación de un modelo ARIMA, ya que en este análisis se utilizó la prueba Augmented Dickey-Fuller (ADF) y la prueba de Kwiatkowski, Phillips, Schmidt y Shin (KPSS) para verificar la estacionariedad en el rendimiento académico de los estudiantes en métodos cuantitativos.

-
- Probamos la hipótesis para la prueba ADF de que;

- H0: La secuencia no es estacionaria.
- H1: La secuencia estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es estacionaria.

Parámetro	Valor
Estadística ADF	-2.04122
Valor-p	0.26880
Lags utilizados	0
Valor critico (1 %)	-4.06885
Valor critico (5 %)	-3.12715
Valor critico (10 %)	-2.70173

Cuadro 32: Resultado de la prueba augmented Dickey-Fuller.

- Probamos la hipótesis para la prueba KPSS de que;

- H0: La secuencia es estacionaria.
- H1: La secuencia no es estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es no estacionaria.

Parámetro	Valor
Estadística KPSS	0.62470
Valor-p	0.02039
Lags utilizados	1
Valor critico (1 %)	0.739
Valor critico (5 %)	0.463
Valor critico (10 %)	0.347

Cuadro 33: Resultado de la prueba KPSS.

Basándonos en los resultados de las pruebas de estacionariedad ADF y KPSS mostrados en el Cuadro 32 y Cuadro 33, podemos tomar decisiones sobre la estacionariedad de la secuencia.

Para la prueba ADF, la hipótesis nula (H0) establece que la secuencia no es estacionaria, mientras que la hipótesis alternativa (H1) sugiere lo contrario. Con un nivel de significancia del 5 %, el valor-p asociado a la prueba ADF es 0.26880, que es superior al umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H0) y concluimos que la secuencia es no estacionaria.

En contraste, para la prueba KPSS, la H0 postula que la secuencia es estacionaria, mientras que la H1 sugiere que no lo es. Aquí, con un nivel de significancia del 5 %, el valor-p asociado a la prueba KPSS es 0.02039, que es inferior al umbral crítico de 0.05. Por lo tanto, rechazamos la hipótesis nula (H0) y concluimos que la secuencia es no estacionaria.

Los resultados obtenidos sugieren una consistencia en la no estacionariedad de la secuencia de datos. Esto se evidencia por la confirmación de la no estacionariedad a través de la prueba ADF,

junto con la falta de evidencia suficiente para rechazar la no estacionariedad según la prueba KPSS. En consecuencia, se concluye que la serie es no estacionaria.

3.3.6. Selección de un modelo

Ahora procederemos a seleccionar el modelo ARIMA más apropiado para nuestra serie temporal. Utilizaremos las capacidades de Python para explorar y evaluar diversos modelos, con el fin de identificar aquel que mejor se ajuste a nuestras necesidades. Este análisis nos proporcionará la base necesaria para tomar decisiones informadas respecto a la elección del modelo más adecuado.

MODELO	AIC	BIC	HQ
ARIMA (1,0,1)	14.525	17.081	14.288
ARIMA (0,0,0)	14.762	16.040	14.644
ARIMA (1,0,0)	13.299	15.216	13.121
ARIMA (0,1,0)	12.453	13.018	12.337

Cuadro 34: Identificación del modelo.

El modelo con los mínimos Criterios de Información de Akaike (AIC), Criterio de Información Bayesiano (BIC) y Hannan-Quinn (HQ) es el más adecuado para la secuencia. Por lo tanto, al probar todos los modelos competidores en el Cuadro 34, el modelo ARIMA (0, 1, 0) tiene los valores más bajos y, por lo tanto, el mejor modelo predictivo.

3.3.7. Estimación de parámetros

El Cuadro 35 a continuación muestra estimaciones de los parámetros del sistema ARIMA (0, 1, 0).

Tipo	Coefficiente	Error estándar	Valor Z	Valor-p
sigma2 (σ^2)	0.1308	0.067	1.947	0.052

Cuadro 35: Muestra de la estimación de parámetros.

El modelo ARIMA(0, 1, 0) indica que estamos utilizando un modelo autorregresivo integrado de media móvil con cero componentes autorregresivas (AR), una diferencia (I) de orden 1 y cero componentes de media móvil (MA). Esto sugiere que hemos aplicado una diferenciación de primer orden para estabilizar la serie temporal.

En la estimación de parámetros, el parámetro σ^2 tiene un valor de 0,1308 con un error estándar de 0,067. El valor Z es 1,947, y aunque no es muy alto, sugiere cierto grado de significancia. El valor-p asociado es 0,052, lo que indica una tendencia hacia la significancia estadística, aunque no alcanza un nivel de confianza convencional del 95 %.

En resumen, con un modelo ARIMA(0, 1, 0), hemos aplicado una diferenciación de primer orden para estabilizar la serie de tiempo de notas promedio semestrales en Física Técnica I. Aunque la significancia estadística del parámetro σ^2 no es extremadamente alta, el modelo ajustado puede proporcionar predicciones útiles para la serie de tiempo, aunque con cierta cautela debido a la marginal significancia estadística de los parámetros estimados.

3.3.8. Resumen y conclusión

Los resultados del análisis revelaron una variabilidad significativa en el rendimiento de los estudiantes en la asignatura de Física Técnica I a lo largo de los 14 años de estudio. En promedio,

los alumnos finalizaron la asignatura con una calificación de 5.26, pero se observaron fluctuaciones destacadas en diferentes períodos. Por ejemplo, el año 2013 registró el rendimiento más bajo con un promedio de 4.71. Por el contrario, el año 2022 destacó como el de mejor desempeño, con un promedio de 5.85.

Los objetivos específicos de este estudio incluyeron analizar la evolución del rendimiento estudiantil en la asignatura y elaborar un modelo adecuado que se ajustara con precisión a dicho desempeño. Para lograr estos objetivos, se implementó una metodología integral que comenzó con un análisis descriptivo detallado de los datos, seguido de gráficos de las funciones de autocorrelación ACF y PACF para identificar estacionariedad. Además, se utilizó el test ADF y el test KPSS para respaldar las hipótesis obtenidas del análisis de los gráficos de autocorrelación.

Basándonos en los resultados de las pruebas estadísticas, se seleccionó el modelo ARIMA más apropiado utilizando los criterios de información AIC, BIC y HQ. Se encontró que el modelo ARIMA(0,1,0) era el más adecuado para la serie de tiempo estudiada. Este modelo implica una diferenciación de primer orden para estabilizar la serie de tiempo y una estimación del parámetro σ^2 con un valor de 0,1308 y un error estándar de 0,067. Aunque el valor Z asociado no es muy alto (1,947), sugiere cierto grado de significancia. El valor-p asociado (0,052) indica una tendencia hacia la significancia estadística, aunque no alcanza un nivel de confianza convencional del 95 %.

En resumen, este estudio proporcionó una visión completa del rendimiento estudiantil en la asignatura Física Técnica I, desde un análisis descriptivo inicial hasta la selección y estimación de un modelo ARIMA adecuado. Los resultados obtenidos contribuyen a una mejor comprensión de los factores que influyen en el desempeño académico de los estudiantes en esta área específica y pueden servir como base para futuras investigaciones y acciones de mejora educativa.

3.4. Matemática Técnica I

Durante el análisis de la asignatura Matemática Técnica I, se examinaron los promedios obtenidos por semestre a lo largo de los 14 años de estudio. En promedio, los alumnos finalizaban la asignatura con una calificación de 5.70. Sin embargo, se observaron variaciones significativas en el rendimiento a lo largo de los años. En particular, en el año 2020 se registró el rendimiento más bajo como curso, con un promedio de 5.02. Por el contrario, en el año 2022 se alcanzó el rendimiento más alto como curso, con un promedio de 6.10, destacando como un año excepcional en términos de desempeño académico en esta asignatura.

3.4.1. Objetivos específicos:

- Analizar la evolución del rendimiento estudiantil en la asignatura Matemática Técnica I, dentro del marco del curso “1er año Mc.El.”, impartido en la Escuela de Ingeniería.
- Elaborar un modelo adecuado que se ajuste con mayor precisión al desempeño de los estudiantes en la asignatura Matemática Técnica I, perteneciente al curso “1er año Mc.El.”, dictado en la Escuela de Ingeniería.

3.4.2. Preguntas de investigación

- ¿Se identifica alguna tendencia en el rendimiento de los estudiantes en la asignatura Matemática Técnica I, dentro del curso “1er año Mc.El.”, impartido en la Escuela de Ingeniería?
- ¿Cuál es el modelo más adecuado para describir el desempeño de los estudiantes en la asignatura Matemática Técnica I, dentro del curso “1er año Mc.El.”, dictado en la Escuela de Ingeniería?

3.4.3. Metodología

La metodología empleada en este estudio abarcó una serie de pasos fundamentales para el análisis exhaustivo de la serie temporal en cuestión. En primer lugar, se realizó un análisis descriptivo detallado de los datos, con el fin de comprender su comportamiento general, identificar posibles tendencias y visualizar patrones significativos. Se incluyó un gráfico de línea para representar visualmente la evolución de los datos, añadiendo una línea de tendencia para resaltar cualquier dirección en el conjunto de datos.

Posteriormente, se procedió a generar gráficos de las funciones de autocorrelación ACF y autocorrelación parcial PACF. Estas visualizaciones jugaron un papel crucial en la evaluación preliminar de la estacionariedad de los datos y en la identificación de posibles estructuras de autocorrelación. Estos correlogramas proporcionaron una comprensión visual inicial de la relación entre las observaciones en diferentes puntos temporales.

Con base en las observaciones obtenidas de los correlogramas, se llevó a cabo una evaluación más rigurosa de la estacionariedad de la serie temporal. Para ello, se realizaron pruebas de estacionariedad utilizando tanto el test ADF (Augmented Dickey-Fuller) como el test KPSS (Kwiatkowski, Phillips, Schmidt, Shin). Estas pruebas estadísticas sirvieron para reafirmar o refutar las hipótesis extraídas de los correlogramas, proporcionando una evaluación más sólida de la estacionariedad de la serie temporal.

Finalmente, una vez confirmada la estacionariedad de los datos, se procedió a la selección del modelo ARIMA más adecuado para el conjunto de datos analizados. Esto implicó la identificación

de los parámetros óptimos del modelo ARIMA, aprovechando tanto la información derivada de los correlogramas como de las pruebas estadísticas de estacionariedad.

En resumen, la metodología empleada abarcó desde la exploración inicial de los datos y la representación gráfica, pasando por el análisis de la estacionariedad con correlogramas y pruebas estadísticas, hasta la selección y estimación de un modelo ARIMA apropiado. Este enfoque integral garantizó un análisis completo y riguroso de la serie temporal, proporcionando una base sólida para su modelado y predicción.

3.4.4. Estadística Descriptiva de las notas promedios semestrales de los estudiantes

Para realizar un estudio efectivo de la asignatura en estudio, se inició reduciendo la base de datos para incluir únicamente las columnas relevantes (Figura 24). Posteriormente, se llevó a cabo un análisis exploratorio de datos centrado en el rendimiento promedio de los estudiantes en la asignatura Matemática Técnica I a lo largo de 14 años consecutivos. Este análisis se realizó principalmente utilizando Python y técnicas de análisis de series de tiempo.

Durante este proceso, se realizaron cálculos para obtener estadísticas descriptivas iniciales relacionadas con el rendimiento promedio de los estudiantes de Matemática Técnica I. Luego, se procedió a la creación de gráficos de series de tiempo, los cuales permitieron visualizar la evolución del rendimiento promedio a lo largo de los años. Además, se realizó un análisis de tendencias para identificar patrones o cambios significativos en el rendimiento a lo largo del tiempo. Este enfoque proporcionó una comprensión más profunda del comportamiento de los estudiantes en la asignatura a lo largo del período de estudio.

MATEMATICA TECNICA I	
AÑO	
2010	6.083824
2011	5.742051
2012	5.833333
2013	5.696842
2014	5.676667
2015	5.553636
2016	5.690698
2017	6.006000
2018	5.592692
2019	5.487321
2020	5.019412
2021	5.818070
2022	6.102083
2023	5.455652

Figura 24: Base de datos de Matemática Técnica I.

Se observó que la nota mínima promedio anual de los estudiantes en la asignatura de Matemática Técnica I fue de 5.02 en el año 2020. Por otro lado, se registró una nota promedio semestral máxima

Datos	14
Promedio	5.697020
Desviación estándar	0.282531
Mínimo	5.019412
25 %	5.563400
50 %	5.693770
75 %	5.829518
Máximo	6.102083

Cuadro 36: Desempeño de los estudiantes de Matemática Técnica I.

de 6.10 en el año 2022. En cuanto al desempeño promedio, este se situó en 5.70, con una desviación estándar de 0.28.

La Figura 25 muestra las notas promedio del curso por semestre a lo largo del periodo. La ecuación de la tendencia, $y = -0,02x + 5,90$, describe una relación lineal entre el año (x) y el promedio de las calificaciones en la asignatura de Matemática Técnica I (y). En este caso, el coeficiente de $-0,02$ indica la pendiente de la recta de tendencia, lo que sugiere una disminución lineal en el promedio de las notas a medida que pasa el tiempo. Por otro lado, el término independiente ($5,90$) representa el valor esperado del promedio de las notas en el año cero del estudio (en este caso, 2010 ajustado a $x = 0$).

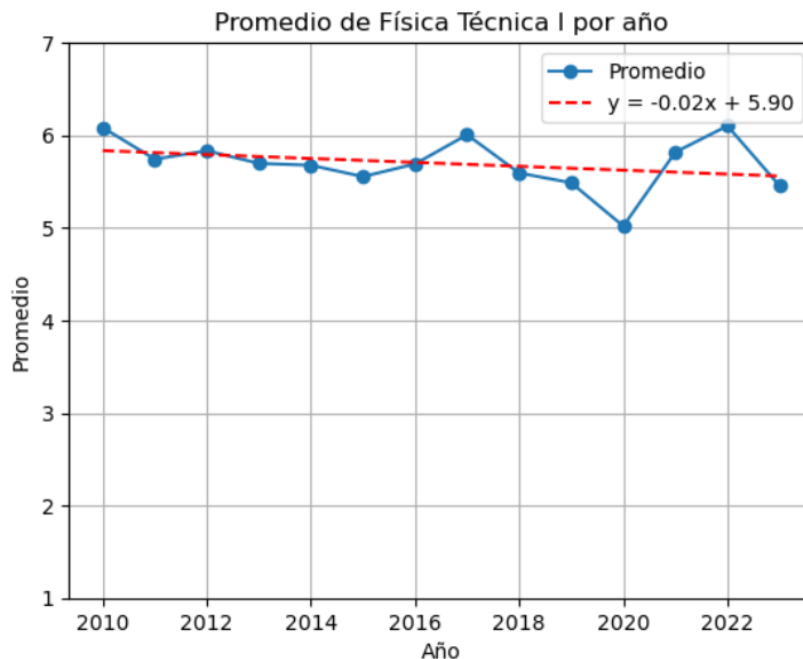


Figura 25: Serie de tiempo de las notas promedios semestrales de los estudiantes en Matemática Técnica I.

Al interpretar la ecuación, podemos concluir que la tendencia general a lo largo de los años muestra una disminución de aproximadamente 0.02 unidades en el promedio de las notas de la asignatura de Matemática Técnica I por año. Es decir, en promedio, se espera que el rendimiento de los estudiantes en esta asignatura disminuya en 0.02 puntos cada semestre. Sin embargo, es importante tener en cuenta que esta es una simplificación y que pueden existir otros factores no

considerados en la ecuación que influyan en el rendimiento académico.

Es necesario analizar la naturaleza de los datos para seleccionar el modelo correcto y mostrar si se trata de datos cíclicos sin tendencia o con actividad estacional.

3.4.5. Pruebas de estacionariedad

Se examinaron los gráficos de la función de autocorrelación (ACF) y de la función de autocorrelación parcial (PACF) de las notas promedio semestrales de los estudiantes en Matemática Técnica I. Estos gráficos, mostrados en la Figura 26, proporcionaron información inicial sobre la estacionariedad de la serie temporal.

En los gráficos ACF y PACF, no se observaron picos significativos en ningún retraso, y además tenemos correlaciones positivas y negativas, lo que sugiere que la serie temporal podría ser estacionaria o que la autocorrelación ha disminuido. Esto indica una posible estacionariedad en la serie de datos, aunque se requieren pruebas estadísticas adicionales para confirmar esta observación.

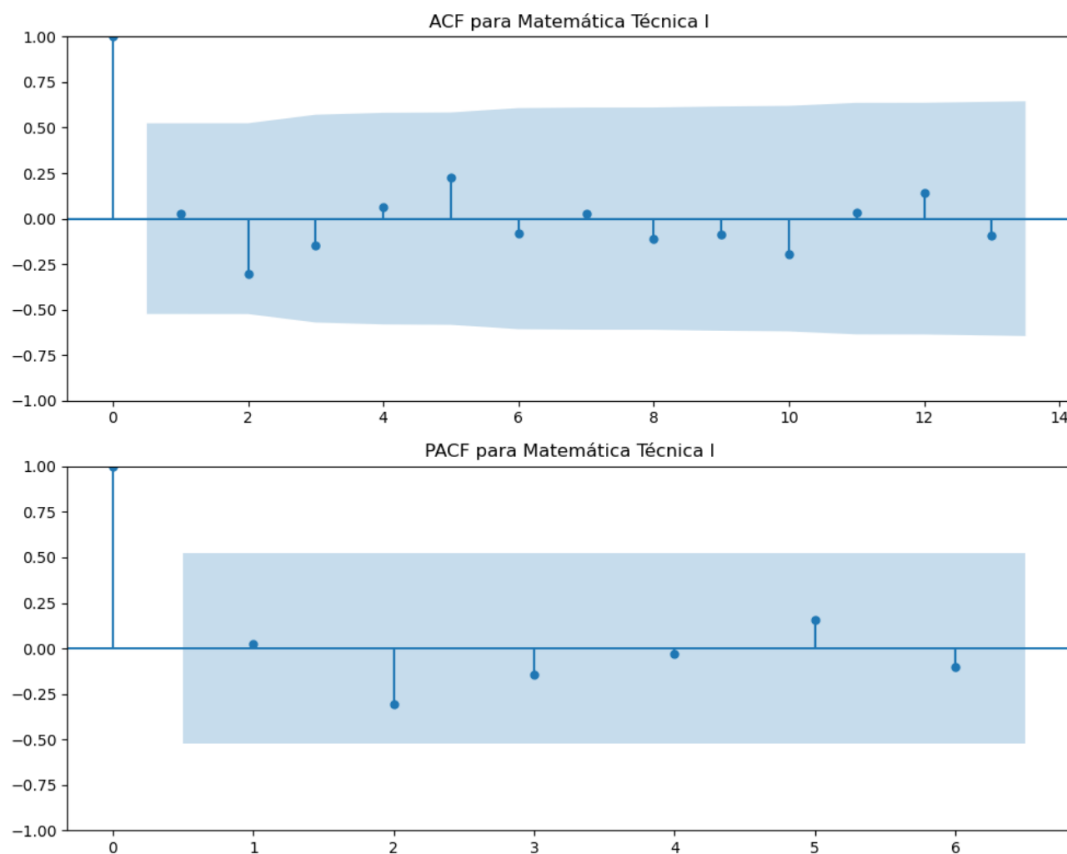


Figura 26: Gráfico ACF y PACF de las notas promedio semestrales de los estudiantes en Matemática Técnica I.

Con estas observaciones preliminares, se procedió a realizar las pruebas de estacionariedad ADF y KPSS para verificar la presencia de estacionariedad en la serie temporal.

Un ciclo estacionario tiene una media y una varianza constante en el tiempo y no hay patrones en el ciclo. Es necesario que la serie sea estacionaria antes de continuar con la estimación de un modelo ARIMA, ya que en este análisis se utilizó la prueba Augmented Dickey-Fuller (ADF) y la prueba de Kwiatkowski, Phillips, Schmidt y Shin (KPSS) para verificar la estacionariedad en el

rendimiento académico de los estudiantes en métodos cuantitativos.

- Probamos la hipótesis para la prueba ADF de que;
 - H0: La secuencia no es estacionaria.
 - H1: La secuencia estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es estacionaria.

Parámetro	Valor
Estadística ADF	1.94944
Valor-p	0.99860
Lags utilizados	5
Valor critico (1 %)	-4.66519
Valor critico (5 %)	-3.36719
Valor critico (10 %)	-2.80296

Cuadro 37: Resultado de la prueba augmented Dickey-Fuller.

- Probamos la hipótesis para la prueba KPSS de que;
 - H0: La secuencia es estacionaria.
 - H1: La secuencia no es estacionaria.

Con un nivel de significancia del 95 %, un valor-p inferior a 0,05 significa un rechazo de H0, lo que significa que la secuencia es no estacionaria.

Parámetro	Valor
Estadística KPSS	0.19179
Valor-p	0.1
Lags utilizados	0
Valor critico (1 %)	0.739
Valor critico (5 %)	0.463
Valor critico (10 %)	0.347

Cuadro 38: Resultado de la prueba KPSS.

Basándonos en los resultados de las pruebas de estacionariedad ADF y KPSS mostrados en el Cuadro 37 y Cuadro 38, podemos tomar decisiones sobre la estacionariedad de la secuencia.

Para la prueba ADF, la hipótesis nula (H0) establece que la secuencia no es estacionaria, mientras que la hipótesis alternativa (H1) sugiere lo contrario. Con un nivel de significancia del 5 %, el valor-p asociado a la prueba ADF es 0.99860, que es superior al umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H0) y concluimos que la secuencia es no estacionaria.

En contraste, para la prueba KPSS, la H0 postula que la secuencia es estacionaria, mientras que la H1 sugiere que no lo es. Aquí, con un nivel de significancia del 5 %, el valor-p asociado a la prueba KPSS es 0.1, que excede el umbral crítico de 0.05. Por lo tanto, no rechazamos la hipótesis nula (H0) y concluimos que la secuencia es estacionaria.

Los resultados obtenidos a través de las pruebas de estacionariedad presentan una discrepancia significativa en cuanto a la naturaleza de la serie temporal analizada. Mientras que la prueba ADF sugiere consistentemente la estacionariedad de la secuencia, indicada por la no rechazo de la hipótesis nula, la prueba KPSS no proporciona evidencia suficiente para refutar la estacionariedad. Esta discrepancia plantea interrogantes sobre la verdadera naturaleza de la serie temporal. Esto se puede deber a que no poseemos suficientes datos para obtener una conclusión más clara.

A la luz de los resultados contradictorios de las pruebas de estacionariedad ADF y KPSS, se ha optado por asumir la estacionariedad de la serie temporal analizada. Aunque la prueba ADF no proporcionó suficiente evidencia para rechazar la hipótesis nula de no estacionariedad, los gráficos de autocorrelación ACF y PACF mostraron indicios de posible estacionariedad, lo que respalda nuestra decisión. Reconocemos que esta asunción puede ser discutible y que hay incertidumbre en la verdadera naturaleza de la serie temporal. Sin embargo, dadas las limitaciones de los datos y la falta de consenso entre las pruebas estadísticas, consideramos que es prudente proceder bajo la suposición de estacionariedad para los fines de este análisis.

3.4.6. Selección de un modelo

Ahora procederemos a seleccionar el modelo ARIMA más apropiado para nuestra serie temporal. Utilizaremos las capacidades de Python para explorar y evaluar diversos modelos, con el fin de identificar aquel que mejor se ajuste a nuestras necesidades. Este análisis nos proporcionará la base necesaria para tomar decisiones informadas respecto a la elección del modelo más adecuado.

MODELO	AIC	BIC	HQ
ARIMA (1,0,1)	9.926	12.482	9.690
ARIMA (0,0,0)	7.302	8.580	7.183
ARIMA (1,0,0)	9.291	11.208	9.114

Cuadro 39: Identificación del modelo.

El modelo con los mínimos Criterios de Información de Akaike (AIC), Criterio de Información Bayesiano (BIC) y Hannan-Quinn (HQ) es el más adecuado para la secuencia. Por lo tanto, al probar todos los modelos competidores en el Cuadro 39, el modelo ARIMA (0, 0, 0) tiene los valores más bajos y, por lo tanto, el mejor modelo predictivo.

3.4.7. Estimación de parámetros

El Cuadro 40 a continuación muestra estimaciones de los parámetros del sistema ARIMA (0, 0, 0).

Tipo	Coefficiente	Error estándar	Valor Z	Valor-p
Constante	5.6970	0.079	71.866	0.000
sigma2 (σ^2)	0.0741	0.027	2.760	0.006

Cuadro 40: Muestra de la estimación de parámetros.

El modelo ARIMA(0,0,0) indica que estamos utilizando un modelo autorregresivo integrado de media móvil con cero componentes autorregresivas (AR), cero componentes de integración (I) y cero componentes de media móvil (MA). Esto significa que no hay dependencia de los valores pasados, ni tendencia ni error en los datos.

En cuanto a la estimación de parámetros, la constante tiene un valor de 5,6970 con un error estándar de 0,079. El valor Z es alto (71,866), lo que indica que la constante es estadísticamente significativa. Además, el valor-p asociado es muy bajo (0,000), lo que sugiere una alta significancia estadística.

En resumen, con un modelo ARIMA(0,0,0), se está utilizando una estimación constante para predecir la serie temporal de notas promedio semestrales en Matemática Técnica I. La constante estimada parece ser estadísticamente significativa, lo que sugiere que este modelo puede ser útil para predecir las notas con un nivel aceptable de precisión.

3.4.8. Resumen y conclusión

Los resultados del análisis revelaron una variabilidad significativa en el rendimiento de los estudiantes en la asignatura de Matemática Técnica I a lo largo de los 14 años de estudio. En promedio, los alumnos finalizaron la asignatura con una calificación de 5.70, pero se observaron fluctuaciones destacadas en diferentes periodos. Por ejemplo, el año 2020 registró el rendimiento más bajo con un promedio de 5.02. Por el contrario, el año 2022 destacó como el de mejor desempeño, con un promedio de 6.10.

Los objetivos específicos de este estudio incluyeron analizar la evolución del rendimiento estudiantil en la asignatura y elaborar un modelo adecuado que se ajustara con precisión a dicho desempeño. Para lograr estos objetivos, se implementó una metodología integral que comenzó con un análisis descriptivo detallado de los datos, seguido de gráficos de las funciones de autocorrelación ACF y PACF para identificar estacionariedad. Además, se utilizó el test ADF y el test KPSS para respaldar las hipótesis obtenidas del análisis de los gráficos de autocorrelación.

Basándonos en los resultados de las pruebas estadísticas, se seleccionó el modelo ARIMA más apropiado utilizando los criterios de información AIC, BIC y HQ. Se encontró que el modelo ARIMA(0,0,0) era el más adecuado para la serie temporal estudiada. Este modelo implica una constante de 5.6970, que resultó ser estadísticamente significativa. Si bien este modelo simple no captura la dinámica compleja de la serie temporal, proporciona una base sólida para la predicción del rendimiento académico en la asignatura de Matemática Técnica I.

En resumen, este estudio proporcionó una visión completa del rendimiento estudiantil en la asignatura Matemática Técnica I, desde un análisis descriptivo inicial hasta la selección y estimación de un modelo ARIMA adecuado. Los resultados obtenidos contribuyen a una mejor comprensión de los factores que influyen en el desempeño académico de los estudiantes en esta área específica y pueden servir como base para futuras investigaciones y acciones de mejora educativa.

Referencias

- [1] Abreh, Autor. "Time Series Modeling of Students' Performance in Mathematics among the Selected Four Senior High Schools in Bongo District, Ghana". *The International Journal of Humanities & Social Studies*, Vol. 8(3), Pp. 284, Marzo de 2020. Disponible en: <https://www.internationaljournalcorner.com/index.php/theijhss/article/view/151960/105821>.
- [2] Vanitha, S., & Jayashree, R. "A Prediction on Educational Time Series Data Using Statistical Machine Learning Model - An Experimental Analysis". *Journal of Theoretical and Applied Information Technology*, Vol. 100(14), Pp. 5189, 31 de Julio de 2022. Disponible en: <https://www.jatit.org/volumes/Vol100No14/12Vol100No14.pdf>.
- [3] Banerjee, D. "Forecasting of Indian stock market using time-series ARIMA model". *IEEE*, Pp. 31, 2241 Full Text Views.
- [4] Mahia, F., Dey, A. R., Masud, M. A., & Mahmud, M. S. "Forecasting Electricity Consumption using ARIMA Model". *IEEE*, Pp. 19, 857 Full Text Views.
- [5] Estadística y Data Science. "Curso de Series Temporales con R". GitHub, 2022. Disponible en: <https://github.com/ecabestadistica/curso-series-temporales>.
- [6] Academia Politécnica Naval de Chile. "Historia y Misión de la Academia". Sitio web oficial de APOLINAV. Disponible en: <https://www.apolinav.cl/>.