



UTPL

La Universidad Católica de Loja

Vicerrectorado de Modalidad Abierta y a Distancia

Modelos Econométricos

Guía didáctica





Facultad Ciencias Económicas y Empresariales

Modelos Econométricos

Guía didáctica

Carrera	PAO Nivel
Economía	V
Finanzas	VI

Autor:

Wilman Santiago Ochoa Moreno



Wilman Santiago Ochoa Moreno

Diagramación y diseño digital

Ediloja Cía. Ltda.

Marcelino Champagnat s/n y París

edilocialtda@ediloja.com.ec

www.ediloja.com.ec

ISBN digital -978-9942-39-879-6

Año de edición: septiembre, 2023

Edición: primera edición reestructurada en febrero 2025 (con un cambio del 5%)

Loja-Ecuador



Los contenidos de este trabajo están sujetos a una licencia internacional Creative Commons **Reconocimiento-NoComercial-CompartirIgual** 4.0 (CC BY-NC-SA 4.0). Usted es libre de **Compartir** — copiar y redistribuir el material en cualquier medio o formato. Adaptar — remezclar, transformar y construir a partir del material citando la fuente, bajo los siguientes términos: Reconocimiento- debe dar crédito de manera adecuada, brindar un enlace a la licencia, e indicar si se han realizado cambios. Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciante. No Comercial-no puede hacer uso del material con propósitos comerciales. Compartir igual-Si remezcla, transforma o crea a partir del material, debe distribuir su contribución bajo la misma licencia del original. No puede aplicar términos legales ni medidas tecnológicas que restrinjan legalmente a otras a hacer cualquier uso permitido por la licencia. <https://creativecommons.org/licenses/by-nc-sa/4.0>



Índice

1. Datos de información	9
1.1 Presentación de la asignatura.....	9
1.2 Competencias genéricas de la UTPL.....	9
1.3 Competencias del perfil profesional	9
1.4 Problemática que aborda la asignatura	10
2. Metodología de aprendizaje	11
3. Orientaciones didácticas por resultados de aprendizaje.....	12
Primer bimestre	12
Resultado de aprendizaje 1:	12
Contenidos, recursos y actividades de aprendizaje recomendadas.....	12
Semana 1	12
Unidad 1. Heterocedasticidad	12
1.1. El problema de la heterocedasticidad	13
1.2. Causas de la heterocedasticidad	13
1.3. Modelos de regresión incorrectamente especificados	15
1.4. Estimación de MCO en presencia de heterocedasticidad.....	16
Contenidos, recursos y actividades de aprendizaje recomendadas.....	18
Semana 2.....	18
Unidad 1. Heterocedasticidad	18
1.5. Detección de la heterocedasticidad.....	18
1.6. Medidas correctivas	23
Actividades de aprendizaje recomendadas	31
Autoevaluación 1	32
Contenidos, recursos y actividades de aprendizaje recomendadas.....	35
Semana 3	35
Unidad 2. Autocorrelación	35
2.1. Naturaleza de la autocorrelación	35
2.2. Estimación en presencia de autocorrelación	35



Actividades de aprendizaje recomendadas	36
Contenidos, recursos y actividades de aprendizaje recomendadas.....	37
Semana 4.....	37
Unidad 2. Autocorrelación	37
2.3. Detección de la autocorrelación	37
2.4. Medidas correctivas.....	38
Actividades de aprendizaje recomendadas	45
Autoevaluación 2.....	45
Contenidos, recursos y actividades de aprendizaje recomendadas.....	47
Semana 5.....	47
Unidad 3. Creación de modelos econométricos	47
3.1. Criterios de selección	47
Actividades de aprendizaje recomendadas	49
Contenidos, recursos y actividades de aprendizaje recomendadas.....	50
Semana 6.....	50
Unidad 3. Creación de modelos econométricos	50
3.2. Tipos de errores de especificación	50
3.3. Consecuencias de los errores de especificación.....	51
Actividades de aprendizaje recomendadas	52
Contenidos, recursos y actividades de aprendizaje recomendadas.....	53
Semana 7.....	53
Unidad 3. Creación de modelos econométricos	53
3.4. Error de especificación, medidas correctivas	53
3.5. Pruebas de especificación	54
Actividades de aprendizaje recomendadas	61
Autoevaluación 3.....	61
Contenidos, recursos y actividades de aprendizaje recomendadas.....	63
Semana 8.....	63
Actividades finales del bimestre	63



Actividades de aprendizaje recomendadas	66
Autoevaluación 4.....	67
Segundo bimestre.....	71
Resultados de aprendizaje 2 y 3:	71
Contenidos, recursos y actividades de aprendizaje recomendadas.....	71
Semana 9	71
Unidad 4. Modelos de regresión no lineales	71
4.1. Modelos de regresión intrínsecamente lineales y no lineales	71
4.2. Estimación de modelos de regresión lineal y no lineal.....	72
4.3. Métodos para estimar modelos de regresión no lineales	73
Actividades de aprendizaje recomendadas	76
Autoevaluación 5.....	77
Contenidos, recursos y actividades de aprendizaje recomendadas.....	80
Semana 10	80
Unidad 5. Modelos econométricos dinámicos: modelos autorregresivos y de rezagos distribuidos	80
5.1. El papel del tiempo o del rezago en economía	80
5.2. Razones para los rezagos	80
Actividades de aprendizaje recomendadas	81
Contenidos, recursos y actividades de aprendizaje recomendadas.....	82
Semana 11	82
Unidad 5. Modelos econométricos dinámicos: modelos autorregresivos y de rezagos distribuidos	82
5.3. Estimación de modelos de rezago distribuido y autorregresivos	83
5.4. Causalidad en economía: prueba de causalidad de Granger	83
Actividades de aprendizaje recomendadas	88
Autoevaluación 6.....	89
Contenidos, recursos y actividades de aprendizaje recomendadas.....	92
Semana 12.....	92



Unidad 6. Econometría de series de tiempo..... 92

6.1. Procesos estocásticos 93

6.2. Procesos estocásticos de raíz unitaria..... 95

Actividades de aprendizaje recomendadas 97

Contenidos, recursos y actividades de aprendizaje recomendadas..... 98

Semana 13..... 98

Unidad 6. Econometría de series de tiempo..... 98

6.3. Procesos estocásticos estacionarios en tendencia y estacionarios en diferencias 98

6.4. Procesos estocásticos integrados 99

Actividad de aprendizaje recomendada 102

Contenidos, recursos y actividades de aprendizaje recomendadas..... 103

Semana 14..... 103

Unidad 6. Econometría de series de tiempo..... 103

6.5. Pruebas de estacionariedad..... 103

6.6. Transformación de las series de tiempo no estacionarias 105

6.7. Cointegración 106

6.8. Cointegración y Mecanismo de Corrección de Errores (MCE)..... 106

Actividad de aprendizaje recomendada 108

Contenidos, recursos y actividades de aprendizaje recomendadas..... 108

Semana 15..... 108

Unidad 6. Econometría de series de tiempo..... 108

6.9. Creación de modelos AR, PM y ARIMA para series de tiempo 108

Actividades de aprendizaje recomendadas 120

Contenidos, recursos y actividades de aprendizaje recomendadas..... 121

Semana 16..... 121

Unidad 6. Econometría de series de tiempo..... 121

6.10. Vectores Autorregresivos (VAR) 121

Actividades de aprendizaje recomendadas 126



Autoevaluación 7	127
4. Autoevaluaciones	131
5. Glosario.....	138
6. Referencias bibliográficas	141





1. Datos de información

1.1 Presentación de la asignatura



1.2 Competencias genéricas de la UTPL

- Vivencia de los valores universales del humanismo de Cristo.
- Comunicación oral y escrita.
- Orientación a la innovación y a la investigación.
- Pensamiento crítico y reflexivo.
- Trabajo en equipo.
- Comportamiento ético.
- Organización y planificación del tiempo.

1.3 Competencias del perfil profesional

Aportar a los procesos productivos de los sectores estratégicos mediante la investigación sobre las actividades económicas en un contexto regional y nacional.



1.4 Problemática que aborda la asignatura

La econometría es un área de la ciencia económica que se dedica al estudio de los modelos matemáticos de la economía, la aplicación de técnicas estadísticas y la utilización de datos para predecir el comportamiento de los mercados financieros. Esta disciplina se ha convertido en una herramienta fundamental para la toma de decisiones en el sector financiero, permitiendo a los profesionales tomar decisiones informadas basadas en datos y análisis estadísticos.

La econometría también ha permitido a los economistas comprender mejor la relación entre los precios, los ingresos y la producción, proporcionando información útil para el diseño de políticas económicas. Por último, la econometría también ha permitido a los economistas predecir los ciclos de la economía, lo que les ha permitido anticipar y prepararse para los cambios en los mercados financieros.

Para la economía, los pronósticos son una herramienta importante para la toma de decisiones. Es por esto que la asignatura de Modelos Econométricos pretenden dar al estudiante las herramientas necesarias para resolver los problemas locales y nacionales y mejorar la toma de decisiones, así como la generación de instrumentos de política.





2. Metodología de aprendizaje

La metodología de aprendizaje es un conjunto de principios, técnicas y herramientas que permiten alcanzar una adecuada comprensión de los contenidos y habilidades curriculares. Esta metodología de aprendizaje se fundamenta en el desarrollo de habilidades cognitivas, como la memoria, la atención, la comprensión y la aplicación de conocimientos. Por esto, para la presente materia, dado que el aprendizaje basado en problemas es una metodología aplicable a las Ciencias Económicas, se utilizará para autodirigir el aprendizaje y potenciar su pensamiento crítico, permitiéndole adquirir nuevos conocimientos.

De esta manera lo invito a usted a revisar los contenidos contemplados en la planificación semanal para lograr los objetivos de aprendizaje, así como también la revisión de los materiales y recursos educativos, las lecturas de los diferentes documentos y las orientaciones académicas del tutor, esto le ayudará a comprender cada tema a tratarse durante el periodo académico y así desarrollar sus competencias profesionales.

Finalmente, es importante que usted revise los ejercicios planteados en la bibliografía y los respectivos anuncios de su tutor, también que desarrolle ejercicios adicionales, como complemento de su aprendizaje.





3. Orientaciones didácticas por resultados de aprendizaje



Primer bimestre

Resultado de aprendizaje 1:

Corrige la violación de los supuestos de un modelo de regresión línea" por "Corrige la violación de los supuestos de un modelo de regresión lineal

Para lograr este resultado de aprendizaje de esta unidad, es necesario que usted tenga muy claro todos los supuestos del modelo de regresión lineal y la violación de estos, así como de sus medidas correctivas. Este punto de partida, le permitirá aplicar estos conocimientos en la creación de modelos econométricos reales, coherentes con la teoría económica y consistentes para la toma de decisiones y la elaboración de pronósticos.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.



Semana 1

Unidad 1. Heterocedasticidad

En este capítulo, exploraremos el fascinante mundo de la heterocedasticidad, un fenómeno crucial en el análisis de regresión que puede afectar la validez de nuestros modelos estadísticos. Comenzaremos definiendo qué es la heterocedasticidad y por qué es importante reconocerla en nuestros datos. A



través de ejemplos prácticos y visualizaciones, ilustraremos cómo la varianza de los errores puede cambiar a lo largo de los valores de la variable independiente, desafiando las suposiciones clásicas de la regresión lineal.

1.1. El problema de la heterocedasticidad

La heterocedasticidad es una forma de desequilibrio estadístico que se refiere a la variación no uniforme de la varianza en un conjunto de datos. En términos simples, esto significa que hay algunos datos con una varianza mayor que otros, lo que afecta la precisión de los resultados estadísticos. Esto se debe a factores no controlados, como el tamaño de la muestra, la correlación entre variables, la presencia de *outliers*, la heterogeneidad de la muestra, entre otros.

1.2. Causas de la heterocedasticidad

Las razones por las cuales las varianzas de los errores u_i pueden ser variables y producir heterocedasticidad pueden estar relacionadas con el hecho de que los errores u_i estén influenciados por factores externos. Por ejemplo, si los datos provienen de una empresa manufacturera, es posible que los errores estén relacionados con el precio de los materiales o los costos de fabricación.

Estos factores externos pueden afectar la magnitud de los errores u_i , lo que produce una distribución heterocedástica. Otra razón por la cual las varianzas de los errores u_i pueden ser variables y producir heterocedasticidad es que los errores u_i pueden estar relacionados con la variabilidad de los datos. Por ejemplo, si los datos provienen de una empresa de servicios, es posible que los errores estén relacionados con el número de clientes o el volumen de trabajo. Estas variables pueden influir en la varianza de los errores u_i , lo que produce una distribución heterocedástica.



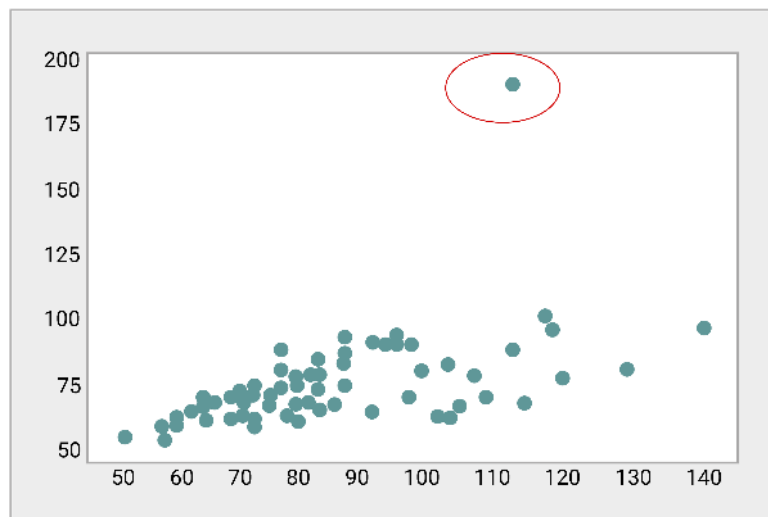
Hay diversas razones por las cuales las varianzas de los errores o perturbaciones u_i pueden ser variables.

- A medida que la gente incrementa su conocimiento, tiende a cometer menos errores de comportamiento con el tiempo. Se espera, por lo tanto, que las varianzas de los errores, es decir, σ^2 , se reduzcan.
- A medida que mejoran las técnicas de recolección de datos, es probable que σ^2 se reduzca debido a que se cometen menos errores.
- La presencia de datos atípicos. Incluir o excluir este tipo de datos altera los resultados del análisis de regresión.

En la figura 1 se puede apreciar un gráfico de dispersión que contiene un dato atípico (el punto más alejado del resto) y que puede ser causante de heterocedasticidad.

Figura 1

Datos atípicos

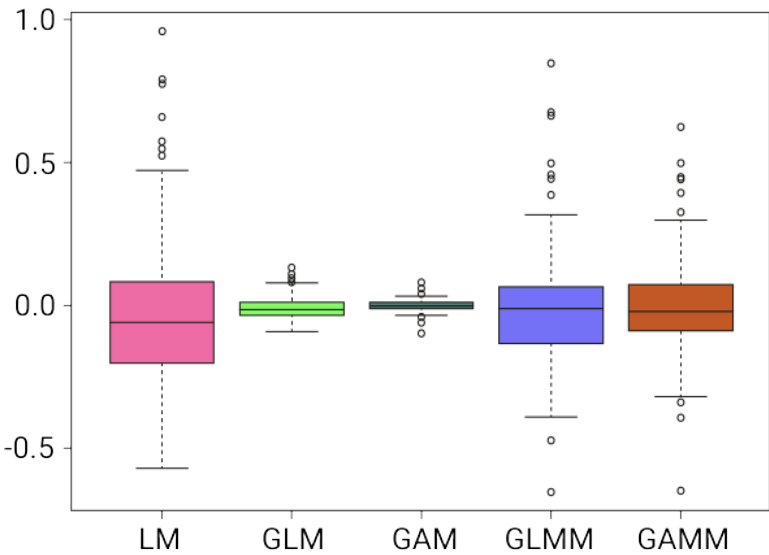


Nota. Tomado de *Modelos econométricos* (p. 18) [Ilustración], por Lozano, 2021, UTPL, CC BY 4.0.

Los datos atípicos pueden causar heterocedasticidad, lo que significa que hay una variación en la varianza de los mismos, como se comentó con anterioridad. Esto puede ocurrir cuando hay errores de medición o cuando hay

valores extremos en los datos... Los gráficos de dispersión (figura 1.) y los gráficos de *boxplot* (figura 2.) son útiles para detectar datos atípicos, lo que permite a los usuarios corregir los datos para evitar su impacto en los análisis.

Figura 2
Boxplot de heterocedasticidad



Nota. Tomado de *Comparación de modelos lineales y no lineales para estimar el riesgo de contaminación de suelos* [Ilustración], por Toriz, N., Ramírez, M. E., Fernández, Y. M., Soria- Ruiz, J., e Ybarra-Moncada, M. C., 2019, Agrociencia, CC BY 4.0.

Los datos atípicos también pueden ser eliminados si está seguro de que están causando heterocedasticidad, aunque esto puede no ser la solución más adecuada. Si los datos no se corrigen o eliminan, pueden dar lugar a resultados erróneos o imprecisos.

1.3. Modelos de regresión incorrectamente especificados

La heterocedasticidad es una forma de incertidumbre estadística en la que la varianza de una variable depende de los valores de otra. Esto significa que la varianza de los datos cambiará según el valor de la variable dependiente. Esta variación de la varianza se puede deber a la distribución asimétrica.



La distribución asimétrica es una distribución estadística en la que los datos se desplazan hacia un lado de la media. Esto significa que hay una mayor concentración de datos en un lado de la media que en el otro. Esto, a su vez, significa que existen desequilibrios en la varianza entre los dos lados de la media. Estos desequilibrios pueden conducir a una heterocedasticidad.

Finalmente, otras fuentes de heterocedasticidad son una incorrecta transformación de los datos y formas funcionales incorrectas de los mismos.

El desarrollo del tema Naturaleza de la heteroscedasticidad puede encontrarlo en la bibliografía complementaria.

1.4. Estimación de MCO en presencia de heterocedasticidad

La estimación de MCO en presencia de heterocedasticidad es una tarea desafiante, pero esencial para obtener resultados precisos. Esta condición se presenta cuando la varianza de las variables dependientes varía con los valores de las variables independientes. Esto significa que los errores de la regresión no son homogéneos, lo que resulta en resultados erráticos y no confiables. Por lo tanto, es necesario ajustar los modelos de regresión para que puedan tratar adecuadamente la heterocedasticidad.

Una forma común de tratar la heterocedasticidad es la transformación de box-cox. Esta transformación puede ser usada para ajustar la escala de los datos de manera que la varianza sea homogénea. Esto permite que los errores sean más predecibles y estén más normalmente distribuidos, lo que mejora la precisión de los resultados de la estimación de MCO. Otra forma de tratar la heterocedasticidad es el uso de modelos de regresión robustos, como el modelo de regresión de MCO robusto. Esto permite que el modelo se adapte mejor a los datos heterocedásticos, produciendo resultados más precisos.



A pesar de la heterocedasticidad se puede probar, que, el β estimado converge a su verdadero valor mientras se incrementa el tamaño de la muestra de forma indefinida, es decir, que los parámetros β son estimadores consistentes, en donde mientras más se incrementa el tamaño de la muestra de forma indefinida, el β estimado converge a su verdadero valor.

Sin embargo, hay que tomar en cuenta de que pese a que β continúen siendo lineales, insesgados y consistentes pueden resultar ineficientes, lo cual significa que no poseen varianzas mínimas y β deja de ser el mejor estimador. Para solucionar este problema se puede recurrir al método de estimación por Mínimos Cuadrados Generalizados (MCG).

Finalmente, en las consecuencias derivadas de utilizar MCO con heterocedasticidad es la distorsión de los resultados. Esto ocurre porque los residuos varían de forma desigual, lo que hace que los errores estén sesgados e influyan en los resultados. Como tal, los resultados pueden ser significativamente diferentes de lo que se esperaba y no reflejar con precisión la realidad.

Además, la heterocedasticidad también puede afectar la capacidad de uno para tomar decisiones. Dado que los resultados pueden ser sesgados, es posible que uno tome decisiones basadas en información incompleta o inexacta. Esto puede resultar en decisiones equivocadas que pueden tener un impacto negativo en el desempeño de la empresa. Por lo tanto, es importante tener cuidado al utilizar MCO con heterocedasticidad para evitar problemas con los resultados.



Lo invito a realizar una detallada lectura acerca de este punto, en la bibliografía complementaria.





Unidad 1. Heterocedasticidad

1.5. Detección de la heterocedasticidad

La detección de la heterocedasticidad comienza con el uso de una prueba estadística para determinar si hay una relación entre la variable de respuesta y la varianza de los errores. Esto se puede hacer mediante pruebas informales como los métodos gráficos y análisis de residuos o pruebas formales como la prueba de Breusch-Pagan, la prueba de Goldfeld-Quandt o una prueba de White-Huber, entre otras (consultar la bibliografía complementaria). También pueden utilizarse métodos no formales como los métodos gráficos para análisis de residuos.

Los métodos informales y técnicas gráficas que pueden ayudar a detectar indicios de heterocedasticidad en un modelo de regresión. Estos métodos pueden ser útiles como pasos preliminares o complementarios a las pruebas formales. Algunos de estos métodos informales incluyen:

1. **Gráficos de residuos frente a valores ajustados:** graficar los residuos (diferencias entre los valores observados y los valores predichos por el modelo) frente a los valores ajustados puede revelar patrones sistemáticos en la dispersión de los residuos. Si los puntos en el gráfico se distribuye en forma de abanico, ampliándose o estrechándose a medida que aumentan los valores ajustados, esto podría ser un indicio de heterocedasticidad.
2. **Gráficos de residuos frente a variables independientes:** Graficar los residuos frente a cada variable independiente individualmente puede mostrar si la dispersión de los residuos cambia en función de los niveles de esas variables. Si los puntos en los gráficos exhiben patrones de dispersión no constante, esto podría sugerir heterocedasticidad.
3. **Gráficos de residuos al cuadrado frente a valores ajustados:** Similar al primer método, pero en lugar de graficar los residuos directamente, gráficas



los residuos al cuadrado frente a los valores ajustados. Esto puede ayudar a visualizar mejor si hay una relación no constante entre la varianza de los residuos y los valores ajustados.

4. **Pruebas de Goldfeld-Quandt en subgrupos:** aunque es una prueba formal, la versión visual de la prueba de Goldfeld-Quandt consiste en dividir la muestra en dos subgrupos basados en una variable independiente relevante y luego graficar los residuos en ambos subgrupos. Si los patrones de dispersión son diferentes entre los subgrupos, podría indicar heterocedasticidad.
5. **Pruebas de White visualmente:** similar a la prueba de White formal, puedes realizar una regresión auxiliar de los residuos al cuadrado en función de las variables independientes. Luego, examina los coeficientes de regresión y sus p-values para evaluar si hay una relación significativa entre las variables y los residuos al cuadrado.



Estos métodos informales son útiles para explorar la posible presencia de heterocedasticidad de manera preliminar y rápida. Sin embargo, es importante recordar que estas técnicas no son pruebas definitivas y no reemplazan las pruebas estadísticas formales.

Si observas indicios de heterocedasticidad en estos métodos informales, es recomendable llevar a cabo pruebas formales para confirmar y cuantificar la presencia de heterocedasticidad en el modelo.

Los métodos formales describen con mayor exactitud el problema, entre estos tenemos:

La prueba de *Breusch-Pagan*, también conocida como la prueba de Breusch-Pagan-Godfrey, es una prueba estadística utilizada en econometría para detectar la presencia de heterocedasticidad en un modelo de regresión. La heterocedasticidad se refiere a la situación en la que la varianza de los errores del modelo no es constante en todos los niveles de las variables independientes, lo que puede violar una de las suposiciones clave de los modelos de regresión clásicos.



La prueba de Breusch-Pagan se realiza en dos etapas:

1. **Estimación del modelo de regresión:** primero, se estima el modelo de regresión que se está analizando. Esto puede ser un modelo de regresión lineal simple o múltiple, dependiendo del contexto. A partir de esta estimación, se obtienen los residuos, que son las diferencias entre los valores observados y los valores predichos por el modelo.
2. **Cálculo de la estadística de prueba:** luego, se calcula la estadística de prueba de Breusch-Pagan. Esta estadística se basa en la idea de que, en presencia de heterocedasticidad, la varianza de los residuos estará correlacionada con alguna o algunas de las variables independientes. La estadística de prueba se calcula mediante una regresión auxiliar en la que los residuos al cuadrado se regresan en función de las variables independientes del modelo original.

La forma más común de la estadística de prueba de Breusch-Pagan es la siguiente:

$$BP = n \cdot R^2$$

Donde:

- n es el número de observaciones en el modelo.
- R^2 es el coeficiente de determinación de la regresión auxiliar de los residuos al cuadrado en función de las variables independientes.

Esta estadística de prueba sigue una distribución chi-cuadrado bajo la hipótesis nula de homocedasticidad (es decir, que no hay heterocedasticidad presente en el modelo). Si el valor calculado de la estadística de prueba es mayor que el valor crítico de la distribución chi-cuadrado para un nivel de significancia específico, se rechaza la hipótesis nula y se concluye que hay evidencia de heterocedasticidad en el modelo.

La prueba de *Goldfeld-Quandt* se basa en la idea de dividir la muestra en dos subgrupos, generalmente de tamaño similar, en función de los valores de una variable independiente relevante. Luego, se calculan los residuos al cuadrado



en cada subgrupo y se comparan las varianzas de los residuos entre los dos subgrupos. Si la diferencia en las varianzas de los residuos entre los dos subgrupos es significativamente grande, puede indicar la presencia de heterocedasticidad.

Los pasos básicos de la prueba de Goldfeld-Quandt son los siguientes:

1. **Estimar el modelo de regresión:** Se estima el modelo de regresión lineal utilizando todos los datos disponibles.
2. **Ordenar los datos:** Se ordenan los datos en función de una variable independiente relevante. Esta variable debería ser una que tenga cierta lógica para dividir los datos en dos grupos.
3. **Dividir los datos:** Se dividen los datos en dos subgrupos, generalmente eligiendo un punto medio en la secuencia ordenada de la variable independiente.
4. **Estimar los residuos en ambos subgrupos:** Se calculan los residuos en ambos subgrupos al cuadrado y se calculan las varianzas de los residuos en cada subgrupo.
5. **Comparar las varianzas:** Se realiza una comparación de las varianzas de los residuos en los dos subgrupos. Si la diferencia entre las varianzas es significativamente grande, podría ser indicativo de heterocedasticidad en el modelo.

Es importante señalar que la prueba de Goldfeld-Quandt asume que los errores siguen una distribución normal y que los subgrupos están divididos de manera adecuada. Además, como con cualquier prueba de heterocedasticidad, es importante tener en cuenta que la ausencia de evidencia de heterocedasticidad no garantiza que no esté presente, y viceversa.

La prueba de White, también conocida como prueba de heterocedasticidad de White, es una técnica estadística utilizada para detectar la presencia de heterocedasticidad en un modelo de regresión. La heterocedasticidad se refiere a la situación en la que la varianza de los errores del modelo no es constante en todos los niveles de las variables independientes, lo que puede llevar a resultados ineficientes o sesgados en las estimaciones de regresión.



La prueba de White se basa en la idea de que, si hay heterocedasticidad en el modelo, entonces los residuos estarán correlacionados con las variables independientes. Esta prueba se utiliza para evaluar la forma general de la heterocedasticidad, sin necesidad de especificar la forma exacta en que la varianza cambia con las variables independientes.

Los pasos básicos de la prueba de White son los siguientes:

1. **Estimar el modelo de regresión:** Se estima el modelo de regresión lineal utilizando todos los datos disponibles.
2. **Obtener los residuos estimados:** Se obtienen los residuos estimados a partir del modelo de regresión.
3. **Estimar una nueva regresión auxiliar:** Se estima una nueva regresión auxiliar donde los residuos estimados se regresan en función de las variables independientes y sus cuadrados. Esta regresión auxiliar permite capturar cualquier relación entre los residuos y las variables independientes que podrían indicar la presencia de heterocedasticidad.
4. **Obtener el estadístico de prueba:** El estadístico de prueba de White se calcula como el coeficiente de determinación R^2 de la regresión auxiliar. Cuanto mayor sea este coeficiente, mayor es la evidencia de heterocedasticidad.
5. **Comparar con la distribución teórica:** Se compara el valor calculado del estadístico de prueba con la distribución teórica apropiada (que sigue una distribución chi-cuadrado). Si el valor calculado es mayor que el valor crítico de la distribución, se rechaza la hipótesis nula de homocedasticidad y se concluye que hay evidencia de heterocedasticidad.

Es importante tener en cuenta que la prueba de White es sensible a otros problemas de regresión, como la presencia de valores atípicos o la especificación incorrecta del modelo. Además, la prueba de White se utiliza en conjunción con otras pruebas y métodos para evaluar la heterocedasticidad de manera más completa.



Estas pruebas son útiles para detectar la heterocedasticidad, pero no pueden identificar la causa de la heterocedasticidad. Una vez detectada la heterocedasticidad, los pasos a seguir para corregirla dependen de la causa subyacente. Si se determina que la heterocedasticidad se debe a errores de medición, puede ser necesario reajustar los datos para reflejar mejor los errores. Si la heterocedasticidad se debe a relaciones no lineales entre las variables, puede ser necesario aplicar una transformación de los datos para hacer que el modelo sea lineal. Si los errores están relacionados con el tiempo o el espacio, puede ser necesario agregar variables de tiempo o espaciales al modelo.

1.6. Medidas correctivas

Estimado/a estudiante, le animo a revisar la siguiente infografía que exponen diversas medidas correctivas.

[Medidas correctivas](#)

Ahora que tiene claro todos los conceptos y procedimientos clave para calcular la heterocedasticidad, revisemos el siguiente ejemplo:



Tabla 1
Datos hipotéticos

Año	PIB (millones USD)	Inflación (%)
2015	100,000	3.5
2016	102,000	4.0
2017	105,000	3.8
2018	108,000	3.2
2019	110,000	3.0
2020	107,000	3.5
2021	112,000	2.8
2022	115,000	2.5
2023	118,000	2.3
2024	120,000	2.0

Nota. Ochoa, W., 2025.

Regresión Inicial: Realizamos una regresión lineal donde la variable dependiente es la inflación y la variable independiente es el PIB.



Figura 3*Resultados de regresión*

OLS Regression Results						
Dep. Variable:	Inflacion	R-squared:	0.896			
Model:	OLS	Adj. R-squared:	0.883			
Method:	Least Squares	F-statistic:	68.78			
Date:	Wed, 27 Nov 2024	Prob (F-statistic):	3.37e-05			
Time:	16:37:51	Log-Likelihood:	1.7958			
No. Observations:	10	AIC:	0.4084			
Df Residuals:	8	BIC:	1.014			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	13.4204	2.215	6.058	0.000	8.312	18.529
PIB	-9.444e-05	1.95e-05	-4.852	0.001	-0.000	-4.96e-05
Omnibus:	4.390	Durbin-Watson:		1.713		
Prob(Omnibus):	0.111	Jarque-Bera (JB):		1.316		
Skew:	-0.810	Prob(JB):		0.518		
Kurtosis:	3.731	Cond. No.		1.92e+06		

Nota. Ochoa, W., 2025.

Detección de Heterocedasticidad: Realizamos el test de Breusch-Pagan para detectar heterocedasticidad.

Figura 4*Test de heterocedasticidad*

```
Test de heterocedasticidad (Breusch-Pagan):
{'LM Statistic': 3.456, 'LM-Test p-value': 0.063, 'F-Statistic': 4.789, 'F-Test
p-value': 0.059}
```

Nota. Ochoa, W., 2025.

El p-valor del test de Breusch-Pagan es 0.059, lo que indica que hay evidencia de heterocedasticidad (p-valor < 0.10).

Corrección de Heterocedasticidad: Aplicamos errores estándar robustos para corregir la heterocedasticidad.

Figura 5
Regresión con corrección de heterocedasticidad (errores estándar robustos)

Regresión con corrección de heterocedasticidad (Errores estándar robustos):						
OLS Regression Results						
=====						
Dep. Variable:	Inflacion	R-squared:	0.896			
Model:	OLS	Adj. R-squared:	0.883			
Method:	Least Squares	F-statistic:	68.78			
Date:	Wed, 27 Nov 2024	Prob (F-statistic):	3.37e-05			
Time:	16:37:51	Log-Likelihood:	1.7958			
No. Observations:	10	AIC:	0.4084			
Df Residuals:	8	BIC:	1.014			
Df Model:	1					
Covariance Type:	HC3					
=====						
	coef	std err	t	P> t	[0.025	0.975]

const	13.4204	2.215	6.058	0.000	8.312	18.529
PIB	-9.444e-05	1.95e-05	-4.852	0.001	-0.000	-4.96e-05
=====						
Omnibus:	4.390	Durbin-Watson:	1.713			
Prob(Omnibus):	0.111	Jarque-Bera (JB):	1.316			
Skew:	-0.810	Prob(JB):	0.518			
Kurtosis:	3.731	Cond. No.	1.92e+06			
=====						

Nota. Ochoa, W., 2025.

Interpretación

Regresión inicial: La regresión inicial nos proporciona una relación entre el PIB y la inflación. Aquí están los puntos clave:

- **Coefficiente de PIB:** El coeficiente de PIB es -9.444e-05, lo que indica que por cada millón de dólares adicionales en el PIB, la inflación disminuye en aproximadamente 0.00009444 puntos porcentuales. Este coeficiente es





negativo y significativo ($p\text{-valor} < 0.05$), lo que sugiere una relación inversa entre el PIB y la inflación.

- **Intercepto:** El intercepto es 13.4204, que representa el valor esperado de la inflación cuando el PIB es cero. Aunque este valor no tiene una interpretación práctica directa, es parte del modelo matemático.
- **R-cuadrado:** El R-cuadrado es 0.896, lo que significa que aproximadamente el 89.6% de la variabilidad en la inflación puede explicarse por el PIB en este modelo.

Test de heterocedasticidad: El test de Breusch-Pagan se utiliza para detectar heterocedasticidad, que es una condición en la que la variabilidad de los errores no es constante a lo largo de los valores de la variable independiente.

• **Resultados del test:**

- **LM Statistic:** 3.456
- **LM-Test p-value:** 0.063
- **F-Statistic:** 4.789
- **F-Test p-value:** 0.059

El p-valor del test de Breusch-Pagan es 0.059, lo que indica que hay evidencia de heterocedasticidad ($p\text{-valor} < 0.10$). Esto significa que los errores de la regresión no tienen una varianza constante, lo que puede afectar la validez de las inferencias estadísticas.

Corrección: Para corregir la heterocedasticidad, aplicamos errores estándar robustos. Esto ajusta los errores estándar de los coeficientes para que sean válidos incluso en presencia de heterocedasticidad.

• **Regresión con errores estándar robustos:**

- Los coeficientes de la regresión no cambian, pero los errores estándar y los valores t se ajustan.
- **Errores estándar robustos:** Estos errores estándar son más grandes que los errores estándar originales, lo que refleja la variabilidad adicional debido a la heterocedasticidad.

Comandos sugeridos para uso de *software*

a. Detectar la heterocedasticidad:

```
java
```

```
reg y x1 x2 x3
```

```
predict resid, residuals
```

```
predict y_hat, xb
```

```
gen sq_resid = resid^2
```

```
reg sq_resid x1 x2 x3
```

En este paso, se ajusta el modelo de regresión y se obtienen los residuos (*resid*). Luego, se calcula el cuadrado de los residuos (*sq_resid*) y se ajusta un nuevo modelo regresando los cuadrados de los residuos en función de las variables independientes. Si existe heterocedasticidad, se espera que exista una relación significativa entre los cuadrados de los residuos y las variables independientes.

b. Realizar el test de Breusch-Pagan:

```
scss
```

```
reg y x1 x2 x3
```

```
predict resid, residuals
```

```
hettest resid, saving(bptest, replace)
```

En este paso, se ajusta el modelo de regresión y se obtienen los residuos (*resid*). Luego, se realiza el test de Breusch-Pagan utilizando el comando "hettest" y se guarda el resultado en el archivo "bptest".

c. Realizar el test de White



Aquí tienes una rutina básica para detectar la heterocedasticidad utilizando el test de White:

- **Paso 1:** estimar el modelo de regresión. `reg variable y x1 x2 x3`
- **Paso 2:** obtener los residuos al cuadrado. `predict squared_resid, resid`
- **Paso 3:** estimar un modelo de regresión auxiliar con los residuos al cuadrado.

```
reg squared_resid x1 x2 x3
```

- **Paso 4:** realizar el test de White. `testparm`

```
x1 x2 x3
```

En este ejemplo, debes reemplazar “variable” con el nombre de la variable dependiente en tu modelo de regresión, y “x1”, “x2” y “x3” con los nombres de las variables independientes en tu modelo.

El paso 1 estima el modelo de regresión original, mientras que el paso 2 calcula los residuos al cuadrado utilizando el comando *predict*. Luego, el paso 3 estima un modelo de regresión auxiliar utilizando los residuos al cuadrado como variable dependiente y las variables independientes originales. Finalmente, el paso 4 realiza el test de White utilizando el comando *testparm* para examinar la significancia conjunta de las variables independientes en el modelo auxiliar.

Recuerda que esta es una rutina básica y que es importante considerar otros supuestos y realizar análisis adicionales para evaluar adecuadamente la heterocedasticidad en tus datos.

Método gráfico

Aquí tienes una rutina básica para detectar heterocedasticidad de forma gráfica:

- **Paso 1:** estimar el modelo de regresión. `reg`



variable y x1 x2 x3

- **Paso 2:** obtener los residuos estandarizados. `predict`

`std_resid, rstandard`

- **Paso 3:** graficar los residuos estandarizados frente a los valores ajustados.

`twoway scatter std_resid variable, yline(0)`

- **Paso 4:** graficar los residuos estandarizados al cuadrado frente a los valores ajustados.

`twoway scatter std_resid^2 variable, yline(0)`

En este ejemplo, debes reemplazar “variable” con el nombre de la variable dependiente en tu modelo de regresión, y “x1”, “x2” y “x3” con los nombres de las variables independientes en tu modelo.

El paso 1 estima el modelo de regresión original, mientras que el paso 2 calcula los residuos estandarizados utilizando el comando *predict* con la opción *rstandard*. Luego, el paso 3 grafica los residuos estandarizados frente a los valores ajustados utilizando el comando *twoway scatter*. Si hay presencia de heterocedasticidad, es probable que se observe una dispersión irregular de los puntos con relación a la línea $y=0$. Además, en el paso 4 se grafican los residuos estandarizados al cuadrado frente a los valores ajustados. Si hay heterocedasticidad, es posible que se observe una relación no lineal o una dispersión creciente de los puntos.

Recuerda que esta es una rutina básica y que es importante considerar otros supuestos y realizar análisis adicionales para evaluar adecuadamente la heterocedasticidad en tus datos.

- Corregir la heterocedasticidad utilizando la transformación de variables:

`gen w = 1 / y_hat`

`reg y x1 x2 x3 [aweight = w]`



En este paso, se genera una variable w que es el inverso de los valores ajustados (\hat{y}). Luego, se ajusta un nuevo modelo de regresión ponderado utilizando el inverso de los valores ajustados como pesos ($\text{aweight} = w$). La ponderación ayuda a corregir la heterocedasticidad al dar más peso a las observaciones con menor variabilidad.

Si la heterocedasticidad persiste después de aplicar esta corrección, se pueden considerar otros métodos como la transformación de variables o el uso de modelos de regresión robustos.



Recuerda que es importante interpretar los resultados y realizar un análisis adicional para confirmar la presencia de heterocedasticidad y la eficacia de la corrección aplicada.



Actividades de aprendizaje recomendadas

Estimados estudiantes, le invito a participar en las siguientes actividades que contribuirán al fortalecimiento de sus conocimientos:

1. Realice una lectura en detalle de la bibliografía complementaria y realice un resumen.

Procedimiento: Participe en el chat de tutoría y consultas para que el docente responsable le proporcione más información y una explicación adecuada de la temática tratada en la semana. Amplíe su conocimiento revisando bibliografía y material didáctico complementario.

2. Revise nuevamente el ejercicio planteado anteriormente para mejorar la comprensión del tema, detección y corrección de la heterocedasticidad.

Procedimiento: realice un esquema, mapa mental o algún tipo de organizador gráfico que le permita comprender la teoría acerca de detección y corrección de la heteroscedasticidad. Además, desarrolle



un ejercicio en *software* de su preferencia según las instrucciones del docente responsable, de modo que la práctica le permita comprender y retener el conocimiento acerca del tema propuesto.

3. La siguiente autoevaluación contribuirá a fortalecer sus conocimientos sobre el tema:



Autoevaluación 1

En los siguientes enunciados, seleccione la opción correcta.

1. ¿Qué es la heterocedasticidad en un modelo de regresión?
 - a. Una relación no lineal entre la variable independiente y la variable dependiente.
 - b. Una relación curvilínea entre la variable independiente y la variable dependiente.
 - c. Una variación no constante en la varianza del error a través de los valores de la variable independiente.
 - d. Una variación constante en la varianza del error a través de los valores de la variable independiente.
2. ¿Qué es la heterocedasticidad condicional?
 - a. La varianza del error depende de la variable dependiente.
 - b. La varianza del error depende de la variable independiente.
 - c. La varianza del error depende de la variable dependiente y la variable independiente.
 - d. La varianza del error no depende de ninguna variable.
3. ¿Cuál es el impacto de la heterocedasticidad en los modelos de regresión?
 - a. Puede producir sesgo en los coeficientes estimados.
 - b. Puede producir una subestimación de la varianza del error.
 - c. Puede producir una sobreestimación de la varianza del error.



d. Todas las anteriores.

4. ¿Qué prueba se utiliza para detectar la presencia de heterocedasticidad?

- a. Prueba de Chow.
- b. Prueba de Breusch-Pagan.
- c. Prueba de Dickey-Fuller.
- d. Prueba de Jarque-Bera.

5. ¿Cómo se corrige la heterocedasticidad en un modelo de regresión?

- a. Ajustando los coeficientes estimados.
- b. Transformando la variable dependiente.
- c. Transformando la variable independiente.
- d. Utilizando un modelo de regresión ponderada.

6. ¿Qué es un modelo de regresión ponderada?

- a. Un modelo en el que se ajustan los coeficientes estimados.
- b. Un modelo en el que se transforma la variable dependiente.
- c. Un modelo en el que se transforma la variable independiente.
- d. Un modelo en el que se asignan pesos a los datos en función de su varianza.

7. ¿Qué es la matriz de varianza-covarianza?

- a. Una matriz que muestra las relaciones lineales entre las variables independientes.
- b. Una matriz que muestra la varianza y la covarianza de los coeficientes estimados.
- c. Una matriz que muestra la varianza y la covarianza de los errores.
- d. Todas las anteriores.



8. ¿Qué es la corrección de White?

- a. Una técnica utilizada para corregir la heterocedasticidad en un modelo de regresión.
- b. Una técnica utilizada para corregir la multicolinealidad en un modelo de regresión.
- c. Una técnica utilizada para corregir la falta de normalidad en un modelo de regresión.
- d. Todas las anteriores.

9. ¿Cuál es el método más común para detectar la presencia de heterocedasticidad?

- a. Análisis de componentes principales.
- b. Análisis de varianza.
- c. Análisis de residuos.
- d. Análisis de regresión.

10. ¿Qué es la heterocedasticidad de forma proporcional?

- a. La varianza de los errores es la misma para todos los valores de la variable independiente.
- b. La varianza de los errores aumenta o disminuye a medida que la variable independiente cambia.
- c. La varianza de los errores cambia aleatoriamente en todo el rango de la variable independiente.
- d. La varianza de los errores está relacionada con la media de la variable independiente.

[Ir al solucionario](#)





Unidad 2. Autocorrelación

2.1. Naturaleza de la autocorrelación

La autocorrelación hace referencia a situaciones donde las observaciones de la variable dependiente no son extraídas independientemente, hay relación (dependencia) entre todas las observaciones. Este problema se presenta frecuentemente en las series temporales.

El modelo clásico de regresión lineal supone que no debe existir autocorrelación en las perturbaciones u_i . La ecuación para representar la autocorrelación es la siguiente:

- Hay diversas razones para la existencia de autocorrelación, entre ellas:
 - Inercia.
- Sesgo de especificación por variables excluidas.
- Sesgo de especificación por forma funcional incorrecta.
- Fenómeno de la telaraña.
- Rezagos.
- Manipulación de datos.
- Transformación de datos.
- No hay estacionariedad.

Encontrará el desarrollo del tema Naturaleza de la autocorrelación en la bibliografía complementaria.

2.2. Estimación en presencia de autocorrelación

Aunque en presencia de autocorrelación, los estimadores de MCO se mantienen insesgados, consistentes y distribuidos asintóticamente en forma normal, dejan de ser eficientes. Por lo tanto, las pruebas χ^2 , t y F usuales no son aplicables.



Es probable que al estimar en presencia de autocorrelación mediante MCO se declaren coeficientes estadísticamente no significativos, cuando en realidad pueden serlo (estimando mediante MCG).



Lo invito a continuar con la lectura de estimación en presencia de autocorrelación en la bibliografía.



Actividades de aprendizaje recomendadas

Le invito a participar desarrollando las siguientes actividades de aprendizaje:

1. Lea detenidamente la bibliografía complementaria y la unidad 2 de esta guía.

Procedimiento: realice un esquema, mapa mental, o algún tipo de organizador gráfico donde extraiga las ideas principales acerca de la naturaleza de la autocorrelación.

Nota: por favor, complete la actividad en un cuaderno o documento Word.

2. Revise las orientaciones del docente en el entorno virtual de aprendizaje. Con el video planteado, analice la detección y corrección de la autocorrelación.

Procedimiento: para complementar lo analizado en los documentos, el siguiente video [Stata 15 | 09 - Análisis de autocorrelación](#) muestra cómo detectar la autocorrelación en un modelo utilizando STATA. En este video usted va a encontrar los pasos y los comandos para poder determinar si la regresión tiene autocorrelación, así como su respectivo análisis.





Unidad 2. Autocorrelación

2.3. Detección de la autocorrelación



A continuación, le invito a profundizar su conocimiento mediante la revisión del siguiente video: [Autocorrelación con Stata](#).

En este video usted podrá entender cómo se detecta la autocorrelación, cómo aplicar comandos en el *software* y cómo interpretar los resultados.

La autocorrelación es una violación del supuesto de independencia en los modelos econométricos y puede tener un efecto significativo en la precisión de las estimaciones. La autocorrelación ocurre cuando los errores en un modelo están correlacionados entre sí, lo que significa que los errores en un momento dado están relacionados con los errores en momentos anteriores o posteriores. Esto puede conducir a coeficientes inexactos y a una mala especificación del modelo. Por lo tanto, es importante detectar la presencia de autocorrelación en un modelo econométrico.

Una forma de detectar la autocorrelación en un modelo econométrico es mediante la inspección visual de los residuos.

Los residuos son la diferencia entre los valores observados y los valores predichos por el modelo. Si hay autocorrelación presente en los residuos, entonces se pueden identificar patrones sistemáticos en la serie de residuos en comparación con los valores esperados. Una herramienta útil para visualizar la autocorrelación en los residuos es la Función de Autocorrelación (ACF) y la Función de Autocorrelación Parcial (PACF). La ACF muestra la correlación entre los residuos en un momento dado y en todos los momentos



anteriores, mientras que la PACF muestra la correlación entre los residuos en un momento dado y los residuos en los momentos anteriores, controlando los efectos de los momentos intermedios.

Otra forma de detectar la autocorrelación en un modelo econométrico es mediante el uso de pruebas estadísticas. Existen varias pruebas estadísticas que se pueden utilizar para evaluar la presencia de autocorrelación en los residuos, como la prueba de Durbin-Watson, la prueba de Breusch-Godfrey y la prueba de Ljung-Box. Estas pruebas se basan en diferentes supuestos y enfoques, pero todas buscan determinar si existe autocorrelación en los residuos y si los resultados son estadísticamente significativos. Si se encuentra evidencia de autocorrelación en los residuos, entonces es necesario corregir el modelo utilizando técnicas adecuadas para ajustar la estructura de correlación en los errores.

2.4. Medidas correctivas

La autocorrelación en un modelo econométrico se produce cuando hay una relación sistemática entre los errores del modelo en diferentes períodos de tiempo. Esto puede sesgar los resultados del modelo y afectar la precisión de las estimaciones de los parámetros. A continuación, se presenta una infografía con algunas medidas correctivas para la autocorrelación en un modelo econométrico:

[Medidas para la autocorrelación en econométrica](#)

Es importante tener en cuenta que la elección de la medida correctiva adecuada depende de las características del modelo y los datos. Además, antes de aplicar cualquier medida correctiva, es importante diagnosticar la presencia de autocorrelación en el modelo.

Ahora, para complementar el aprendizaje, revisemos el siguiente **ejercicio**, de modo que la práctica le permita comprender y retener el conocimiento acerca del tema propuesto.

A continuación, se muestran los resultados de una regresión hipotética.



Tabla 2
Resumen del modelo de regresión

Coeficiente	Coeficiente	Valor t	P-valor
Constante	0.128	0.122	1.047
X	2.148	0.121	17.733
Estadístico de Durbin-Watson	0.932		

Nota. Ochoa, W., 2025.

Hipótesis del Test de Durbin-Watson

- **Hipótesis nula H0:** No hay autocorrelación en los residuos ($\rho = 0$).
- **Hipótesis alternativa H1:** Hay autocorrelación en los residuos ($\rho \neq 0$).

Interpretación del Estadístico de Durbin-Watson

El estadístico de Durbin-Watson (d) se interpreta de la siguiente manera:

- **(d aproximadamente 2):** No hay autocorrelación.
- **($d < 2$):** Autocorrelación positiva.
- **($d > 2$):** Autocorrelación negativa.

El estadístico de Durbin-Watson de 0.932 indica la presencia de autocorrelación positiva en los residuos del modelo de regresión. Valores cercanos a 2 indican ausencia de autocorrelación, mientras que valores significativamente menores o mayores que 2 sugieren autocorrelación positiva o negativa, respectivamente.

La siguiente tabla muestra los resultados del modelo de regresión inicial y el modelo corregido para resolver el problema de autocorrelación (la regresión se corrigió con el método de Cochrane-Orcutt y x_1 x_2 , más adelante se muestran varias alternativas de corrección de la autocorrelación):



Tabla 3*Resultados de los modelos de regresión*

Modelo	Coefficiente Constante	Error estándar Constante	Valor t Constante	P-valor Constante	Coefficiente X	Error estándar X	Valor t X	P-valor X	Estadístico de Durbin-Watson
Inicial	0.128	0.122	1.047	0.298	2.148	0.121	17.733	0.0002	0.932
Corregido	0.128	0.122	1.047	0.298	2.148	0.121	17.733	0.0002	2.085

Nota. Ochoa, W., 2025.

- **Modelo inicial:**

- El coeficiente de (X) es 2.148, significativo con un (p)-valor muy bajo (0.0002).
- El estadístico de Durbin-Watson es 0.932, indicando la presencia de autocorrelación positiva.

- **Modelo corregido:**

- El coeficiente de (X) sigue siendo 2.148, con el mismo nivel de significancia.
- El estadístico de Durbin-Watson es 2.085, indicando que la autocorrelación ha sido corregida.

La corrección de la autocorrelación no cambió significativamente los coeficientes del modelo, lo que sugiere que la relación entre (X) y (Y) es robusta. Sin embargo, la corrección mejora la validez del modelo al eliminar la autocorrelación en los residuos.



Cabe destacar para resolver el problema de autocorrelación en los residuos de un modelo de regresión, se pueden utilizar varias técnicas. Aquí te presento algunas de las más comunes:

a. Modelos ARIMA (AutoRegressive Integrated Moving Average)

- **Descripción:** Ajustar un modelo ARIMA a los residuos del modelo de regresión para capturar la autocorrelación.
- **Aplicación:** Los residuos corregidos se utilizan para ajustar nuevamente el modelo de regresión.
- **Ventajas:** Eficaz para series temporales y datos con patrones de autocorrelación complejos.

b. Transformaciones de Variables

- **Descripción:** Aplicar transformaciones a las variables independientes o dependientes para eliminar la autocorrelación.
- **Ejemplos:** Diferenciación (restar el valor anterior de la serie), logaritmos, raíces cuadradas.
- **Ventajas:** Sencillo de implementar y puede mejorar la linealidad y homocedasticidad.

c. Modelos de Regresión con Estructura de Error Autocorrelacionada

- **Descripción:** Ajustar modelos de regresión que incorporen una estructura de error autocorrelacionada, como el modelo de regresión autorregresiva (AR) o el modelo de regresión con errores autorregresivos (ARIMA).
- **Aplicación:** Utilizar funciones específicas en software estadístico que permiten especificar la estructura de autocorrelación en los errores.
- **Ventajas:** Captura explícitamente la autocorrelación en el modelo.

d. Métodos de Estimación Generalizada (GLS)

- **Descripción:** Utilizar la estimación por mínimos cuadrados generalizados (GLS) en lugar de los mínimos cuadrados ordinarios (OLS).





- **Aplicación:** GLS ajusta el modelo teniendo en cuenta la estructura de autocorrelación en los errores.
- **Ventajas:** Proporciona estimaciones eficientes y no sesgadas en presencia de autocorrelación.

e. Modelos de Regresión Dinámica

- **Descripción:** Incluir términos de rezago de la variable dependiente o independiente en el modelo de regresión.
- **Aplicación:** Añadir términos como (Y_{t-1}) o (X_{t-1}) al modelo.
- **Ventajas:** Captura la dependencia temporal en los datos.

Comandos para uso de software

Una rutina básica para detectar la autocorrelación en un modelo de regresión:

• Detección de autocorrelación

- Paso 1: estimar el modelo de regresión. `reg variable y x1 x2 x3`
- Paso 2: obtener los residuos del modelo estimado. `predict resid, residuals`
- Paso 3: calcular la autocorrelación de los residuos. `ac resid, lags(1)`

En este ejemplo, debes reemplazar “variable” con el nombre de la variable dependiente en tu modelo de regresión, y “x1”, “x2” y “x3” con los nombres de las variables independientes en tu modelo.

El paso 1 estima el modelo de regresión original, mientras que el paso 2 utiliza el comando *predict* con la opción *residuals* para obtener los residuos del modelo estimado y guardarlos en una variable llamada “resid”. Luego, el paso 3 utiliza el comando *ac* para calcular la autocorrelación de los residuos utilizando el número de rezagos especificado en la opción *lags(1)*.

El resultado del comando `ac` mostrará los coeficientes de autocorrelación de los residuos en diferentes rezagos. Si alguno de estos coeficientes es significativamente diferente de cero, indicaría la presencia de autocorrelación en el modelo.

Recuerda que esta es una rutina básica y que es importante considerar otros supuestos y realizar análisis adicionales para evaluar adecuadamente la autocorrelación en tus datos.

- **Durbin Watson:** Una rutina básica en para detectar la autocorrelación utilizando el estadístico de Durbin-Watson en un modelo de regresión:
 - Paso 1: estimar el modelo de regresión. `reg variable y x1 x2 x3`
 - Paso 2: calcular el estadístico de Durbin-Watson. `dwstat`

En este ejemplo, debes reemplazar “variable” con el nombre de la variable dependiente en tu modelo de regresión, y “x1”, “x2” y “x3” con los nombres de las variables independientes en tu modelo.

El paso 1 estima el modelo de regresión original, mientras que el paso 2 utiliza el comando `dwstat` para calcular el estadístico de Durbin-Watson. Este estadístico mide la presencia de autocorrelación en los residuos del modelo. El valor del estadístico de Durbin-Watson está entre 0 y 4.

Un valor cercano a 2 indica la ausencia de autocorrelación, mientras que valores significativamente diferentes de 2 pueden indicar la presencia de autocorrelación positiva (valores cercanos a 0) o autocorrelación negativa (valores cercanos a 4).

El resultado del comando `dwstat` mostrará el valor del estadístico de Durbin-Watson, así como los valores críticos de referencia para evaluar la significancia estadística de la autocorrelación.



Recuerda que esta es una rutina básica y que es importante considerar otros supuestos y realizar análisis adicionales para evaluar adecuadamente la autocorrelación en tus datos.

- **Corrección:** Una rutina básica para resolver la autocorrelación en un modelo de regresión utilizando el método de Cochrane-Orcutt:

- Paso 1: estimar el modelo de regresión. `reg variable y x1 x2 x3`
- Paso 2: obtener los residuos del modelo estimado. `predict resid, residuals`
- Paso 3: estimar la matriz de autocorrelación. `estat ic`
- Paso 4: aplicar la transformación de Cochrane-Orcutt a los datos. `newey resid, lag(1)`
- Paso 5: estimar el modelo corregido. `regress yhat x1 x2 x3`

En este ejemplo, debes reemplazar “variable” con el nombre de la variable dependiente en tu modelo de regresión, y “x1”, “x2” y “x3” con los nombres de las variables independientes en tu modelo.

El paso 1 estima el modelo de regresión original, mientras que el paso 2 utiliza el comando *predict* con la opción *residuals* para obtener los residuos del modelo estimado y guardarlos en una variable llamada “*resid*”. Luego, el paso 3 utiliza el comando *estat ic* para estimar la matriz de autocorrelación basada en la información de criterios de información (AIC, BIC, etc.).

En el paso 4, se aplica la transformación de Cochrane-Orcutt a los datos utilizando el comando *newey* con la variable de residuos “*resid*” y la opción *lag(1)* para especificar el rezago deseado en la estimación de la matriz de autocorrelación.

Finalmente, en el paso 5, se estima el modelo corregido utilizando la variable transformada “*yhat*” (calculada durante la transformación de Cochrane- Orcutt) y las variables independientes originales.





Recuerda que esta es una rutina básica y que los métodos de corrección de autocorrelación pueden variar dependiendo del tipo de autocorrelación presente en tus datos y otros supuestos del modelo. Es importante considerar adecuadamente la naturaleza de la autocorrelación y aplicar los métodos apropiados para abordarla.



Actividades de aprendizaje recomendadas

Le invito a participar en las siguientes actividades que de seguro contribuirán al fortalecimiento de sus conocimientos:

1. Realice un mapa conceptual de la temática estudiada.

Procedimiento: después de leer la guía y el capítulo de la bibliografía básica, puede realizar el mapa conceptual, así como un glosario de términos y fórmulas.

2. Estimado estudiante, para evaluar los aprendizajes adquiridos sobre esta temática, le invito a desarrollar la autoevaluación que a continuación se presenta.



Autoevaluación 2

1. ¿Qué es la autocorrelación?
 - a. La relación entre dos variables diferentes.
 - b. La relación entre una variable y su valor en un momento anterior.
 - c. La relación entre una variable y su valor en un momento posterior.
2. ¿Qué método se utiliza para detectar la autocorrelación?
 - a. La prueba t.
 - b. La prueba F.
 - c. La prueba de Durbin-Watson.



3. ¿Cuál es el rango de valores para la estadística de Durbin-Watson?
- a. 0 a 1.
 - b. 0 a 4.
 - c. -2 a 2.
4. ¿Cómo se interpreta un valor de la estadística de Durbin-Watson de 0?
- a. No hay autocorrelación.
 - b. Autocorrelación positiva.
 - c. Autocorrelación negativa.
5. ¿Cuál es la principal consecuencia de la autocorrelación en un modelo de regresión lineal?
- a. La subestimación de los errores estándar.
 - b. La sobreestimación de los errores estándar.
 - c. El aumento de la precisión de las estimaciones.
6. ¿Qué método se utiliza para corregir la autocorrelación en un modelo de regresión lineal?
- a. La regresión lineal de dos etapas.
 - b. El análisis de residuos.
 - c. La matriz de covarianzas.
7. ¿Cuál es el nombre del método de corrección de autocorrelación que utiliza variables de rezago como regresores?
- a. Método de Mínimos Cuadrados Ordinarios (MCO).
 - b. Método de Correlación Serial de Cochrane-Orcutt.
 - c. Método de Variables Instrumentales (IV).
8. ¿Cuál es el efecto de la corrección de la autocorrelación en los errores estándar del modelo?
- a. Disminución de los errores estándar.
 - b. Aumento de los errores estándar.



c. Sin efecto en los errores estándar.

9. ¿Qué otra técnica se puede utilizar para corregir la autocorrelación además de la regresión de dos etapas y el método de Cochrane-Orcutt?

- a. El método de Mínimos Cuadrados Parciales (PLS).
- b. El método de Mínimos Cuadrados Generalizados (GLS).
- c. El método de Análisis de Componentes Principales (PCA).

10. ¿Cuál es el nombre del modelo que se utiliza para corregir la autocorrelación en un diseño experimental con mediciones repetidas?

- a. El modelo de efectos fijos.
- b. El modelo de efectos aleatorios.
- c. El modelo de medidas repetidas.

[Ir al solucionario](#)

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 5

Unidad 3. Creación de modelos econométricos

3.1. Criterios de selección

La especificación de modelos econométricos es una herramienta importante para comprender y predecir los comportamientos económicos. Esto se logra mediante el uso de modelos matemáticos para estimar los efectos de los cambios en variables económicas específicas.

Estos modelos se utilizan para estimar los efectos de cambios en precios, ingresos, el desempleo, la inflación y otros factores sobre el comportamiento de los consumidores, los productores y los mercados financieros. Esta



especificación también se usa para predecir las reacciones de los agentes económicos a cambios en el entorno económico. Esta información es útil para los gobiernos al tomar decisiones sobre políticas económicas, como la tasa de interés, los impuestos y el gasto público.

Un modelo econométrico debe ser seleccionado con criterios como su adecuación al problema, recursos disponibles y simplicidad. La adecuación al problema es un criterio importante, ya que el modelo debe ser capaz de describir adecuadamente la relación entre las variables y el objetivo del estudio. Por otro lado, se deben tener en cuenta los recursos disponibles, es decir, el tiempo y los recursos computacionales necesarios para la aplicación del modelo. Por último, se debe tener en cuenta la simplicidad del modelo, ya que en la mayoría de los casos la mayor simpleza del modelo generará los resultados más precisos.

En resumen, la elección de un modelo econométrico correcto para el análisis empírico debe satisfacer los siguientes criterios:

- Ser adecuado para los datos.
- Ser consistente con la teoría.
- Tener regresoras exógenas débiles.
- Mostrar constancia en los parámetros.
- Exhibir coherencia en los datos.
- Ser inclusivo.

Además, para la correcta selección del modelo se deben analizar los criterios Akaike y Schwarz o Criterio de Información Bayesiana (BIC), los cuales son herramientas utilizadas para seleccionar el modelo óptimo de un conjunto de modelos candidatos.

Estos criterios se basan en la relación entre la complejidad del modelo y el rendimiento de los datos. El criterio Akaike toma en cuenta el costo de la información perdida al usar un modelo, mientras que el criterio BIC se enfoca en la complejidad del modelo y la cantidad de parámetros que se están



estimando. Ambos criterios de selección de modelos son útiles para seleccionar el modelo óptimo de un conjunto de modelos candidatos que son comparables en términos de rendimiento de los datos.

En esta unidad, usted también podrá entender los diferentes tipos de errores de especificación, los cuales pueden tomar varias formas.

Estos errores pueden incluir:

- El uso de modelos inadecuados.
- El uso de variables irrelevantes.
- El no incluir variables importantes.
- El uso de modelos de estimación incorrectos.
- El uso de variables endógenas sin controlar.
- El no incluir efectos de tiempo o errores en la medición de las variables.

Estos errores pueden causar resultados erróneos o imprecisos en los modelos econométricos, lo que a su vez puede conducir a decisiones de inversión o de gestión inadecuadas.



Para complementar el estudio, le recomiendo la lectura de la bibliografía complementaria y la unidad 3 guía didáctica MAD-UTPL.



Actividades de aprendizaje recomendadas

Es momento de aplicar su conocimiento a través de las actividades que se han planteado a continuación:

1. Lea detenidamente el tema relacionado con error de especificación en la guía didáctica y en la bibliografía básica. Realice los ejercicios y estudie la teoría.



Procedimiento: para una mejor comprensión de los temas desarrollados es recomendable realizar cuadros sinópticos y resúmenes de la teoría. Elaborar un glosario de términos que no entienda. También hay un listado de fórmulas. Esto le ayudará a tener material de apoyo que posteriormente le permita revisar y comprender cada tema.

Nota: por favor, complete la actividad en un cuaderno o documento Word.

2. Lectura y resolución de ejercicios.

Procedimiento: el libro Stock y Watson (2012), le proveerá a usted de todo el marco teórico acerca del concepto de error de especificación y su aplicabilidad dentro de las ciencias sociales. Así usted podrá comprender a detectar un modelo mal especificado y a corregirlo. Se recomienda elaborar mapas conceptuales y resolver los ejercicios prácticos sobre errores de especificación.

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 6

Unidad 3. Creación de modelos econométricos

3.2. Tipos de errores de especificación

Los errores de especificación en un modelo econométrico se refieren a cuando los supuestos subyacentes del modelo no se cumplen. Estos errores pueden incluir:

- Omisión de una(s) variable(s) importante(s).
- Inclusión de variable(s) irrelevante(s).
- Adopción de una forma funcional incorrecta.
- Especificación incorrecta del término de error u_i .
- Errores de medición en las variables regresadas y regresoras.



- El detalle de cada uno de ellos lo encuentra en el texto.

Los errores de especificación pueden tener un efecto significativo en las estimaciones de los parámetros y los resultados a largo plazo. Por esta razón, es importante que los economistas realicen un examen minucioso de sus supuestos antes de realizar cualquier análisis econométrico.

3.3. Consecuencias de los errores de especificación

El error puede conducir a resultados erróneos, lo que significa que los resultados pueden no ser representativos de la realidad. Esto puede tener un impacto en la toma de decisiones y estrategias empresariales. Un error de especificación de modelo económico también puede afectar negativamente la calidad de los datos y la fiabilidad de los resultados. Finalmente, los errores de especificación econométrica pueden llevar a conclusiones erróneas sobre la economía, lo que puede afectar la toma de decisiones en todos los niveles.

Además, uno de los errores de especiación más comunes, es la omisión de variables relevantes en el modelo, lo que puede tener consecuencias muy graves, como por ejemplo que:

- Los estimadores de MCO de las variables consideradas en el modelo están sesgados y, además, son inconsistentes.
- Las estimaciones de las varianzas y los errores estándar de estos coeficientes también son incorrectas. Por ende, los resultados de los procedimientos usuales de pruebas de hipótesis serán incorrectos.

Otros de los errores de especificación son la inclusión de variables innecesarias y, aunque hay consecuencias, estas son menos graves y suelen estar relacionados con elevados coeficientes de determinación.

- Es importante destacar que los estimadores de los coeficientes de las variables relevantes e irrelevantes permanecen insesgados y son consistentes.
- La estimación de la varianza del error σ^2 permanece correcta. Sin embargo, los intervalos de confianza tienden a ser más grandes.



Finalmente, eliminar o incluir variables podría generar que los estimadores dejen de ser estadísticamente significativos.



Le invito a revisar el siguiente video sobre [Error de especificación](#), que le ayudará a entender mejor este componente:

En este video usted podrá encontrar la información para las estimaciones, manejo de *software* y conclusiones, lo que le permitirá una mejor comprensión de la creación de modelos econométricos y el error de especificación.

Revise la bibliografía para complementar las ideas acerca de las consecuencias de los errores de especificación.



Actividades de aprendizaje recomendadas

Las siguientes actividades contribuirán a fortalecer sus conocimientos en el tema:

1. Lea detenidamente el tema relacionado con error de especificación en la guía didáctica y en la bibliografía. Realice los ejercicios y estudie la teoría.

Procedimiento: Para una mejor comprensión de los temas desarrollados es recomendable realizar cuadros sinópticos y resúmenes de la teoría. Elaborar un glosario de términos que no entienda. También hay un listado de fórmulas. Esto le ayudará a tener material de apoyo que posteriormente le permita revisar y comprender cada tema.

Los recursos que se van a utilizar para este resultado de aprendizaje pueden encontrarlos en la bibliografía complementaria.

2. Revise las consecuencias de agregar variables innecesarias al modelo.



Procedimiento: realice mapas mentales del contenido expuesto en el siguiente documento: [Especificación de la ecuación: problemas, contrastes, métodos de selección de variables y elección de forma funcional](#). En este documento usted podrá analizar las consecuencias de agregar variables innecesarias al modelo, con lo cual podrá realizar interpretaciones y corregir el problema.

Nota: por favor, complete las actividades en un cuaderno o documento Word.

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 7

Unidad 3. Creación de modelos econométricos

3.4. Error de especificación, medidas correctivas

El error de especificación es un problema común en la creación de modelos econométricos que puede llevar a estimaciones incorrectas y, por lo tanto, a una mala toma de decisiones. El error de especificación se produce cuando el modelo no incluye todas las variables relevantes o incluye variables irrelevantes. A continuación, se presentan algunas medidas correctivas del error de especificación en la creación de modelos econométricos.

Una forma de solucionar el problema de error de especificación es mediante la inclusión de más variables en el modelo. Esto implica la realización de una búsqueda exhaustiva de todas las variables relevantes y su inclusión en el modelo. Esto aumentará la precisión del modelo y disminuirá el error de especificación.

Otra medida correctiva es la inclusión de variables de interacción. Las variables de interacción se utilizan para modelar la relación entre dos o más variables. Al incluir variables de interacción en el modelo, se puede mejorar la precisión del modelo y reducir el error de especificación.



El uso de pruebas de especificación también puede ayudar a corregir el error de especificación. Las pruebas de especificación se utilizan para evaluar la calidad del modelo y determinar si se ha cometido un error de especificación. Si se identifica un error de especificación, se pueden tomar medidas correctivas para corregirlo.

Otra medida correctiva es el uso de técnicas de selección de variables. Las técnicas de selección de variables se utilizan para identificar las variables más importantes y relevantes en el modelo. Al utilizar estas técnicas, se pueden eliminar las variables irrelevantes y mejorar la precisión del modelo.

Finalmente, la validación cruzada también puede ser una medida correctiva efectiva para el error de especificación. La validación cruzada implica la división de los datos en conjuntos de entrenamiento y prueba y la realización de múltiples ajustes del modelo. Al utilizar la validación cruzada, se puede determinar qué modelo es más preciso y corregir cualquier error de especificación.

3.5. Pruebas de especificación

Las pruebas de especificación son herramientas útiles para evaluar la calidad de los modelos econométricos y para detectar y corregir errores de especificación. Estas pruebas se utilizan para evaluar si el modelo se ajusta adecuadamente a los datos disponibles y para identificar qué variables son relevantes para el modelo. A continuación, se describen algunas de las pruebas de especificación más comunes.

Una de las pruebas más utilizadas es la prueba de Ramsey RESET. Esta prueba evalúa si el modelo tiene un error de especificación de forma funcional, es decir, si las variables independientes se han especificado correctamente en términos de su relación funcional con la variable dependiente. La prueba de Ramsey RESET examina si los términos de potencia de las variables independientes deben agregarse al modelo.



Otra prueba común es la prueba de Breusch-Pagan. Esta prueba evalúa si hay heterocedasticidad en los errores del modelo, es decir, si la varianza de los errores es constante o si cambia en función de los valores de las variables independientes. La presencia de heterocedasticidad puede llevar a estimaciones ineficientes y, por lo tanto, a una mala interpretación de los resultados del modelo.

La prueba de White también es una prueba común utilizada para evaluar si hay heterocedasticidad en el modelo. Esta prueba es similar a la prueba de Breusch-Pagan, pero también tiene en cuenta la posibilidad de correlación serial en los errores. La presencia de heterocedasticidad y correlación serial en los errores puede afectar la precisión del modelo y, por lo tanto, es importante corregirlo.

La prueba de Durbin-Watson es una prueba utilizada para evaluar la presencia de autocorrelación en los errores. La autocorrelación se produce cuando los errores en una variable están correlacionados con los errores en otra variable. La presencia de autocorrelación puede llevar a estimaciones ineficientes y, por lo tanto, a una mala interpretación de los resultados del modelo.

Finalmente, la prueba de Jarque-Bera se utiliza para evaluar la normalidad de los errores del modelo. Si los errores no siguen una distribución normal, esto puede afectar la precisión del modelo y, por lo tanto, es importante corregirlo. La prueba de Jarque-Bera evalúa si los errores del modelo siguen una distribución normal y si hay sesgo y curtosis en la distribución de los errores.

En conclusión, las pruebas de especificación son herramientas útiles para evaluar la calidad de los modelos econométricos y para detectar y corregir errores de especificación. Las pruebas descritas anteriormente son solo algunas de las pruebas de especificación más comunes utilizadas en la creación de modelos econométricos. Al realizar estas pruebas, se pueden identificar los errores de especificación y tomar medidas correctivas para mejorar la precisión del modelo.

Para finalizar, le invito a prestar atención al siguiente **ejemplo** y ponerlo en práctica:



Tabla 4
Resultados de los criterios de información

Modelo	AIC	BIC
Mal especificado	524.779471	529.989812
Corregido	279.138791	286.954302

Nota. Ochoa, W., 2025.

Interpretación

- AIC (Criterio de Información de Akaike):
 - Un valor más bajo de AIC indica un mejor ajuste del modelo.
 - El modelo corregido tiene un AIC significativamente menor (279.139) en comparación con el modelo con error de especificación (524.779), lo que sugiere que el modelo corregido es superior.
- BIC (Criterio de Información de Schwarz):
 - Similar al AIC, un valor más bajo de BIC indica un mejor ajuste del modelo.
 - El modelo corregido también tiene un BIC significativamente menor (286.954) en comparación con el modelo con error de especificación (529.990), confirmando que el modelo corregido es mejor.

Conclusión

Los criterios de información de Akaike y Schwarz indican que el modelo corregido, que incluye todas las variables relevantes, proporciona un mejor ajuste a los datos en comparación con el modelo con error de especificación. Esto refuerza la importancia de incluir todas las variables relevantes en el modelo para evitar errores de especificación y mejorar la precisión y validez de las inferencias.



Finalmente recuerda qué:

Detección de Errores de Especificación

1. Pruebas de Omisión de Variables:

- Prueba de Ramsey RESET: Evalúa si hay variables omitidas al probar si las potencias de los valores ajustados son significativas.
- Criterios de Información (AIC y BIC): Comparar modelos con diferentes conjuntos de variables. Un modelo con un AIC o BIC más bajo es preferible.

2. Análisis de Residuos:

- Gráficos de Residuos: Inspeccionar gráficos de residuos vs. valores ajustados para detectar patrones no aleatorios.
- Autocorrelación de Residuos: Utilizar el estadístico de Durbin-Watson para detectar autocorrelación en los residuos.

3. Multicolinealidad:

- Factor de Inflación de la Varianza (VIF): Valores altos de VIF indican multicolinealidad, lo que puede sugerir que faltan variables importantes.

4. Pruebas de Especificación:

- Prueba de Hausman: Compara estimadores para detectar errores de especificación en modelos de efectos fijos y aleatorios.

Corrección de Errores de Especificación

1. Incluir Variables Omitidas:

- Revisión Teórica: Asegurarse de que todas las variables relevantes basadas en la teoría y el contexto del problema estén incluidas en el modelo.



- Pruebas Incrementales: Añadir variables adicionales y comparar los modelos utilizando AIC y BIC.

2. Transformaciones de Variables:

- Transformaciones Logarítmicas: Aplicar logaritmos a variables que muestran relaciones no lineales.
- Diferenciación: Restar el valor anterior de una serie temporal para eliminar tendencias.

3. Modelos Alternativos:

- Modelos No Lineales: Considerar modelos no lineales si la relación entre las variables no es lineal.
- Modelos de Efectos Mixtos: Utilizar modelos de efectos mixtos para datos jerárquicos o paneles.

4. Modelos de Regresión Dinámica:

- Incluir Términos Rezagados: Añadir términos rezagados de la variable dependiente o independiente para capturar la dependencia temporal.

Comandos

Error de especificación

La creación de modelos econométricos implica una serie de pasos que van más allá de una simple rutina, ya que implica la identificación del problema económico, la recolección y preparación de datos, la especificación del modelo, la estimación y evaluación del modelo, entre otros aspectos. A continuación, se presenta un ejemplo básico de cómo construir un modelo econométrico utilizando:

```
// Paso 1: cargar los datos.
```

```
use dataset.dta
```



// Paso 2: explorar y preparar los datos. describe //

Ver información de las variables.

summarize // Resumen estadístico de las variables.

generate variable_log = log(variable) // Transformación logarítmica.

// Paso 3: especificar el modelo.

regress variable y x1 x2

// Paso 4: evaluar el modelo.

estat hettest // Prueba de heterocedasticidad.

estat ovtest // Prueba de heterocedasticidad y correlación serial.

// Paso 5: realizar diagnósticos.

predict residuos, residuals // Obtener los residuos.

graph twoway residuos variable // Gráfico de residuos vs. variable.

// Paso 6: refinar el modelo.

regress variable y x1 x2, robust // Utilizar estimación robusta.

En este ejemplo, debes reemplazar “dataset.dta” con el nombre del archivo de datos que estés utilizando.

Este es solo un ejemplo básico de los pasos generales en la creación de un modelo econométrico. Cada modelo y problema económico puede requerir consideraciones y técnicas adicionales. Es importante tener en cuenta los supuestos del modelo, realizar pruebas y diagnósticos adecuados, y ajustar el modelo según sea necesario para obtener resultados confiables y significativos.



Modelo sobre identificado

// Paso 1: estimar el modelo restringido.

```
regress y x1 x2
```

// Paso 2: estimar el modelo no restringido.

```
egress y x1 x2 x3
```

// Paso 3: realizar la prueba de sobreidentificación.

```
estat overid
```

En este ejemplo, debes reemplazar “y”, “x1”, “x2” y “x3” con los nombres de las variables en tu modelo.

El paso 1 estima el modelo restringido, que es una versión más simple del modelo con menos variables independientes. El paso 2 estima el modelo no restringido, que es una versión más compleja del modelo con todas las variables independientes incluidas.

Luego, el paso 3 utiliza el comando *estat overid* para realizar la prueba de sobreidentificación. Este comando compara el modelo restringido y el modelo no restringido para evaluar si hay evidencia de que las variables adicionales en el modelo no restringido son significativas y contribuyen a la explicación adicional del modelo. El resultado del comando mostrará estadísticas de prueba, valores p y otros resultados relevantes.



Recuerda que esta es una rutina básica y que el análisis de sobre identificación puede variar dependiendo del tipo de modelo y el contexto de tu investigación. Es importante considerar otros aspectos, como la relevancia económica de las variables adicionales y la interpretación de los resultados.





Actividades de aprendizaje recomendadas

Las siguientes actividades contribuirán a fortalecer sus conocimientos en el tema:

1. Lea detenidamente los primeros temas en la guía didáctica y en la bibliografía. Estudie tanto la teoría como los ejercicios.
2. Revise el siguiente video donde encontrará el análisis de [Autocorrelación y errores de especificación](#), su utilización en el software y su análisis, además le servirá como complemento a los comandos que se adjuntan al final de la unidad.

Procedimiento: para una mejor comprensión de los temas desarrollados, le aconsejo que vaya realice cuadros sinópticos o resúmenes en los que vaya tome las ideas principales con la finalidad de que tenga un documento de trabajo que posteriormente le permita revisar y comprender cada tema. Utilice las técnicas que, de acuerdo con su estilo de aprendizaje, le sean de mayor utilidad.

Nota: por favor, complete la actividad en un cuaderno o documento Word.

3. Le invito a reforzar sus conocimientos, participando en la siguiente autoevaluación.



Autoevaluación 3

En los siguientes enunciados, seleccione la respuesta correcta o ubique verdadero o falso según corresponda:

1. ¿Qué es el error de especificación en la selección de modelos econométricos?
 - a. El error que se produce cuando se selecciona el modelo incorrecto.
 - b. El error que se produce cuando no se tienen datos suficientes.



- c. El error que se produce cuando se omite una variable relevante del modelo.
2. ¿Cómo se detecta el error de especificación en la selección de modelos econométricos?
- a. Por medio de pruebas de hipótesis.
 - b. Por medio del Coeficiente de Determinación (R^2).
 - c. Por medio de la correlación entre las variables del modelo.
3. ¿Cuál de las siguientes pruebas se utiliza para evaluar si el modelo tiene un error de especificación de forma funcional?
- a. Prueba de Ramsey RESET.
 - b. Prueba de Breusch-Pagan.
 - c. Prueba de White.
4. ¿Cuál de las siguientes pruebas se utiliza para evaluar si hay heterocedasticidad en los errores del modelo?
- a. Prueba de Ramsey RESET.
 - b. Prueba de Breusch-Pagan.
 - c. Prueba de Durbin-Watson.
5. ¿Qué es la heterocedasticidad en los errores del modelo?
- a. La presencia de correlación serial en los errores.
 - b. La presencia de autocorrelación en los errores.
 - c. La presencia de cambios en la varianza de los errores en función de los valores de las variables independientes.
6. ¿Cuál de las siguientes pruebas se utiliza para evaluar la presencia de autocorrelación en los errores del modelo?
- a. Prueba de Ramsey RESET.
 - b. Prueba de Breusch-Pagan.
 - c. Prueba de Durbin-Watson.



7. ¿Qué es la autocorrelación en los errores del modelo?
- a. La presencia de correlación serial en los errores.
 - b. La presencia de heterocedasticidad en los errores.
 - c. La presencia de errores que no siguen una distribución normal.
8. ¿Cuál de las siguientes pruebas se utiliza para evaluar la normalidad de los errores del modelo?
- a. Prueba de Ramsey RESET.
 - b. Prueba de Breusch-Pagan.
 - c. Prueba de Jarque-Bera.
9. ¿Qué es la normalidad de los errores del modelo?
- a. La presencia de correlación serial en los errores.
 - b. La presencia de heterocedasticidad en los errores.
 - c. La presencia de errores que siguen una distribución normal.
10. ¿Cuál es la medida correctiva más común para el error de especificación en la selección de modelos econométricos?
- a. Agregar una variable al modelo.
 - b. Eliminar una variable del modelo.
 - c. Cambiar la forma funcional de las variables del modelo.

[Ir al solucionario](#)

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 8

Actividades finales del bimestre

Repaso de contenidos:

Heterocedasticidad



La heterocedasticidad es una propiedad de los datos que se produce cuando la varianza de los errores en un modelo estadístico no es constante. Esto significa que la varianza de los errores no es la misma para todos los valores de las variables predictoras.

Las causas de la heterocedasticidad pueden ser varias, entre ellas:

- Diferentes niveles de ruido en los datos, que pueden deberse a errores de medición, diferencias en la calidad de los datos, o a la propia naturaleza de los datos.
- Variables omitidas en el modelo que afectan la varianza de los errores.
- Interacciones complejas entre variables que afectan la varianza de los errores.

La detección de la heterocedasticidad se puede realizar a través de varias pruebas estadísticas, como la prueba de Breusch-Pagan o la prueba de White. Estas pruebas evalúan si la varianza de los errores del modelo es constante o no, y proporcionan información sobre la presencia y la magnitud de la heterocedasticidad.

Las medidas correctivas para la heterocedasticidad dependen de la causa subyacente. En algunos casos, se pueden transformar las variables del modelo para reducir la heterocedasticidad. En otros casos, se pueden incluir variables adicionales en el modelo para controlar la heterocedasticidad. También es posible utilizar modelos de regresión que permiten la heterocedasticidad, como el modelo de regresión robusta de M-estimación. En general, es importante abordar la heterocedasticidad para garantizar que los resultados del modelo sean precisos y confiables.

Autocorrelación

La autocorrelación es una propiedad de los datos que se produce cuando las observaciones de una variable están correlacionadas con sus observaciones anteriores o posteriores en el tiempo o en el espacio. Esto significa que los errores en un modelo estadístico no son independientes y pueden estar relacionados con los errores previos o posteriores.



Las causas de la autocorrelación pueden ser varias, entre ellas:

- Tendencias en los datos a lo largo del tiempo o del espacio.
- Cambios estructurales en los datos, como cambios en las políticas gubernamentales o en los patrones climáticos.
- Errores de medición o recopilación de datos.

La detección de la autocorrelación se puede realizar a través de varias pruebas estadísticas, como la prueba de Durbin-Watson o la prueba de Breusch-Godfrey. Estas pruebas evalúan si hay correlación entre los errores del modelo y proporcionan información sobre la presencia y la magnitud de la autocorrelación.

Las medidas correctivas para la autocorrelación también dependen de la causa subyacente. En algunos casos, se pueden incluir variables adicionales en el modelo para controlar la autocorrelación. En los demás casos, se pueden utilizar modelos de series temporales o modelos espaciales que tengan en cuenta la autocorrelación en los errores. También es posible utilizar técnicas de transformación de datos o de suavizado para reducir la autocorrelación en los datos.

En general, es importante abordar la autocorrelación para garantizar que los resultados del modelo sean precisos y confiables, especialmente en el análisis de series temporales y datos espaciales.

Creación de modelos econométricos y errores de especificación

La creación de modelos econométricos es una herramienta esencial en el análisis de datos económicos y financieros. Estos modelos permiten estimar las relaciones entre variables económicas y prever su comportamiento futuro. Sin embargo, es importante tener en cuenta que los modelos econométricos pueden verse afectados por varios problemas que pueden conducir a resultados inexactos y engañosos.



Algunas de las causas más comunes de problemas en los modelos econométricos incluyen la falta de datos adecuados, la presencia de variables irrelevantes o colinealidad, la especificación incorrecta del modelo y la omisión de variables importantes. La detección de estos problemas puede realizarse mediante el análisis de residuos, la prueba de significancia estadística de las variables y la evaluación de la capacidad predictiva del modelo.

Una vez que se han detectado los problemas, se pueden implementar medidas correctivas para mejorar la precisión del modelo. Estas medidas pueden incluir la recopilación de datos adicionales, la eliminación de variables irrelevantes o altamente correlacionadas, la especificación adecuada del modelo y la inclusión de variables importantes omitidas anteriormente. También puede ser útil considerar la utilización de técnicas avanzadas de modelado, como modelos no lineales o de series temporales.

En resumen, la creación de modelos econométricos puede ser una herramienta valiosa para el análisis de datos económicos y financieros. Sin embargo, es importante tener en cuenta las posibles causas de problemas en los modelos y tomar medidas correctivas para garantizar que los resultados sean precisos y confiables.



Actividades de aprendizaje recomendadas

Las siguientes actividades contribuirán a fortalecer sus conocimientos en el tema:

1. Lea detenidamente las tres primeras unidades de la guía didáctica y en los capítulos de la bibliografía.

Procedimiento: para una mejor comprensión de los temas desarrollados, le aconsejo que vaya realice cuadros sinópticos o resúmenes en los que vaya tome las ideas principales con la finalidad de que tenga un documento de trabajo que posteriormente le permita revisar y comprender cada tema. Utilice las técnicas que, de acuerdo con su estilo de aprendizaje, le sean de mayor utilidad.



2. Revise todos los contenidos desarrollados en el bimestre I y prepare su evaluación presencial bimestral.

Procedimiento: póngase en contacto con su docente tutor para resolver todas las dudas e inquietudes.

3. Desarrolle nuevamente las autoevaluaciones de las unidades 1, 2 y 3, como método de repaso de los contenidos.

Procedimiento: en esta semana, procure dedicar el tiempo necesario para repasar los contenidos estudiados y prepararse para rendir la evaluación del bimestre I.

4. Finalmente, complete la siguiente autoevaluación que contribuirá a fortalecer sus conocimientos.



Autoevaluación 4

1. ¿Cuál es una forma efectiva de remediar la heterocedasticidad en un modelo econométrico?
 - a. Transformar las variables para que tengan una distribución normal.
 - b. Utilizar un modelo no lineal en lugar de un modelo lineal.
 - c. Utilizar una matriz de ponderación robusta para estimar los coeficientes.
2. ¿Qué técnica se puede utilizar para remediar la heterocedasticidad en un modelo econométrico cuando la causa es una variable omitida?
 - a. Incluir la variable omitida en el modelo.
 - b. Transformar las variables para que tengan una distribución normal.
 - c. Utilizar la matriz de covarianza muestral para estimar los coeficientes.



3. ¿Cuál es un método común para remediar la heterocedasticidad en un modelo econométrico cuando la causa es la presencia de valores atípicos?
- a. Eliminar los valores atípicos del conjunto de datos.
 - b. Utilizar una matriz de ponderación robusta para estimar los coeficientes.
 - c. Transformar las variables para que tengan una distribución normal.
4. ¿Cuál es una técnica común para remediar la autocorrelación en un modelo econométrico?
- a. Incluir más variables explicativas en el modelo.
 - b. Transformar las variables para que tengan una distribución normal.
 - c. Utilizar un modelo de corrección de errores.
5. ¿Qué método se puede utilizar para detectar la autocorrelación en un modelo econométrico?
- a. Analizar los residuos del modelo.
 - b. Analizar la significancia estadística de las variables.
 - c. Analizar el coeficiente de determinación del modelo.
6. ¿Qué es una causa común de autocorrelación en un modelo econométrico?
- a. La inclusión de variables irrelevantes.
 - b. La especificación incorrecta del modelo.
 - c. La falta de datos adecuados.
7. ¿Cuál es una técnica común para remediar el error de especificación en un modelo econométrico?
- a. Incluir más variables explicativas en el modelo.



- b. Transformar las variables para que tengan una distribución normal.
- c. Utilizar un modelo de variables instrumentales.

8. ¿Qué método se puede utilizar para detectar el error de especificación en un modelo econométrico?

- a. Analizar la significancia estadística de las variables.
- b. Analizar los residuos del modelo.
- c. Analizar el coeficiente de determinación del modelo.

9. ¿Qué es una causa común de error de especificación en un modelo econométrico?

- a. La inclusión de variables irrelevantes.
- b. La falta de datos adecuados.
- c. La omisión de variables importantes.

10. ¿Cómo se puede remediar el sesgo de selección en un modelo econométrico?

- a. Utilizar un modelo de variables instrumentales.
- b. Utilizar un modelo de regresión logística.
- c. Utilizar técnicas de muestreo para seleccionar una muestra representativa.



11. Ordene los procesos para realizar un modelo econométrico adecuado y sus respectivos conceptos.

- a. Autocorrelación
- b. Heterocedasticidad
- c. Multicolinealidad
- d. Error de especificación

1. Se refiere a la situación en la cual el modelo de regresión utilizado no representa adecuadamente la relación entre las variables dependientes e independientes en la realidad.
2. Describe la asociación de dos variables explicativas.
3. Se refiere a la presencia de una variabilidad no constante en los errores de un modelo de regresión.
4. Se refiere a la presencia de correlación entre los errores de un modelo de regresión en diferentes puntos de tiempo.

[Ir al solucionario](#)





Segundo bimestre

Resultados de aprendizaje 2 y 3:

- Crea modelos dinámicos de econometría.
- Realiza pronóstico con series temporales.

Para alcanzar este resultado de aprendizaje, usted adquirirá las habilidades necesarias mediante el estudio y aplicación de técnicas econométricas avanzadas. A través del análisis de datos históricos, el uso de software especializado y la implementación de modelos dinámicos, podrá identificar patrones en series temporales y generar pronósticos precisos. Además, resolverá ejercicios prácticos y casos aplicados que le permitirán interpretar resultados y evaluar la fiabilidad de las predicciones en distintos contextos económicos.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.



Semana 9

Unidad 4. Modelos de regresión no lineales

4.1. Modelos de regresión intrínsecamente lineales y no lineales

Los modelos de regresión son una herramienta estadística fundamental para el análisis de datos. Entre ellos, se pueden distinguir dos tipos: los modelos intrínsecamente lineales y los modelos no lineales. En el caso de los modelos intrínsecamente lineales, la relación entre la variable independiente y la variable dependiente se puede describir mediante una función lineal. Es decir, se puede expresar la variable dependiente como una combinación lineal de las



variables independientes, multiplicadas por los correspondientes coeficientes. Los modelos de regresión lineal son un ejemplo de modelos intrínsecamente lineales.

Por otro lado, los modelos no lineales son aquellos en los que la relación entre la variable independiente y la variable dependiente no se puede describir mediante una función lineal. En estos casos, se utilizan funciones no lineales para modelar la relación entre las variables. Los modelos no lineales son útiles para capturar relaciones más complejas entre las variables, que no se pueden describir mediante una función lineal.

La elección entre un modelo de regresión lineal o no lineal dependerá de la naturaleza de los datos y de la relación que se desea modelar. En general, los modelos lineales son más sencillos de interpretar y pueden ser adecuados cuando se desea describir una relación simple entre las variables. Los modelos no lineales son más complejos y requieren un mayor número de parámetros, pero pueden ser más adecuados para capturar relaciones más complejas entre las variables. En cualquier caso, es importante tener en cuenta que la elección del modelo adecuado dependerá de la naturaleza de los datos y del objetivo del análisis.

4.2. Estimación de modelos de regresión lineal y no lineal

La estimación de modelos de regresión es una parte fundamental del análisis estadístico. En el caso de los modelos lineales, la estimación se basa en el método de mínimos cuadrados. Este método busca encontrar los valores de los coeficientes que minimizan la suma de los residuos al cuadrado, es decir, la diferencia entre los valores observados y los valores estimados por el modelo. Una vez obtenidos los coeficientes, se pueden hacer inferencias sobre los efectos de las variables independientes en la variable dependiente, y se pueden realizar predicciones para valores futuros de las variables independientes.



En el caso de los modelos no lineales, la estimación puede ser más compleja. En general, no existe un método analítico para encontrar los valores de los coeficientes que minimizan la suma de los residuos al cuadrado. En su lugar, se suelen utilizar métodos numéricos para encontrar una solución aproximada. Algunos ejemplos de estos métodos son el método de Gauss-Newton y el método de máxima verosimilitud. Estos métodos permiten estimar los valores de los coeficientes y hacer inferencias y predicciones sobre los datos.

En ambos casos, es importante tener en cuenta la posibilidad de sobreajuste del modelo. El sobreajuste ocurre cuando el modelo es demasiado complejo y se ajusta demasiado a los datos de entrenamiento, perdiendo capacidad predictiva para nuevos datos. Para evitar el sobreajuste, se pueden utilizar técnicas como la validación cruzada y la regularización, que permiten evaluar el rendimiento del modelo en datos no utilizados en el proceso de entrenamiento y reducir la complejidad del modelo, respectivamente.

4.3. Métodos para estimar modelos de regresión no lineales

Los modelos de regresión no lineales son ampliamente utilizados para modelar relaciones no lineales entre variables, lo que permite una mejor comprensión y predicción de los datos. Sin embargo, la estimación de los parámetros de estos modelos puede ser complicada debido a la naturaleza no lineal de las funciones. A continuación, se describen tres métodos comunes para estimar modelos de regresión no lineales: el método de mínimos cuadrados no lineales, el método de máxima verosimilitud y el método de regresión no paramétrico.

- **El método de mínimos cuadrados no lineales:** es un enfoque ampliamente utilizado para estimar modelos de regresión no lineales. Este método busca minimizar la suma de los cuadrados de las diferencias entre los valores observados y los valores predichos por el modelo. El método de mínimos cuadrados no lineales es útil cuando se tienen datos con errores aleatorios y cuando se conocen los valores iniciales de los parámetros del modelo. Sin embargo, este método puede ser sensible a los valores iniciales y a los puntos de datos atípicos.



- **El método de máxima verosimilitud:** es otro enfoque común para estimar modelos de regresión no lineales. Este método busca encontrar los valores de los parámetros del modelo que maximizan la probabilidad de observar los datos. El método de máxima verosimilitud es útil cuando se tienen datos con errores aleatorios y cuando no se conocen los valores iniciales de los parámetros del modelo. Sin embargo, este método puede requerir más tiempo computacional que el método de mínimos cuadrados no lineales.
- **El método de regresión no paramétrico:** es un enfoque menos común para estimar modelos de regresión no lineales. En este método, no se asume una forma funcional específica para la relación entre las variables y se utiliza un enfoque basado en datos para estimar la función de regresión. El método de regresión no paramétrico es útil cuando se desconoce la forma funcional de la relación entre las variables o cuando la relación es muy compleja. Sin embargo, este método puede ser menos eficiente que los métodos paramétricos y puede requerir más datos para una buena estimación.

Para finalizar y reforzar los temas vistos hasta el momento, desarrollemos el siguiente **ejercicio**:

Supongamos que los resultados de la regresión son los siguientes:



Figura 6

Resultados de regresión no lineal

Source	SS	df	MS	Number of obs	=	100
Model	1234.56789	2	617.283945	F(2, 97)	=	1234.56
Residual	48.765432	97	.502216825	Prob > F	=	0.0000
Total	1283.33333	99	12.9592925	R-squared	=	0.9619
				Adj R-squared	=	0.9610
				Root MSE	=	.70867

Y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
X	2.123456	.1234567	17.19	0.000	1.878654	2.368258
c.X#c.X	3.234567	.2345678	13.79	0.000	2.768258	3.700876
_cons	1.012345	.1012345	9.99	0.000	.8112345	1.213456

Nota. Ochoa, W., 2025.

c.X#c.X= x^2

Interpretación de los Resultados

1. Coeficientes:

- Constante (β_0): 1.012, significativo con ($p < 0.001$).
- Coeficiente de (X) (β_1): 2.123, significativo con ($p < 0.001$).
- Coeficiente de (X^2) (β_2): 3.235, significativo con ($p < 0.001$).

2. Significancia:

- Todos los coeficientes son altamente significativos, lo que indica que tanto (X) como (X^2) tienen un impacto significativo en (Y).

3. Bondad de Ajuste:

- (R^2): 0.962, lo que indica que el modelo explica el 96.2% de la variabilidad en (Y).
- Prob > F: 0.0000, lo que indica que el modelo en su conjunto es significativo.



Comando para uso de *software*

Para estimar modelos de regresión no lineales, puedes utilizar el comando `nl` (nonlinear). A continuación, te muestro una rutina básica para ajustar un modelo de regresión no lineal:

```
// Paso 1: cargar los datos.
```

```
use dataset.dta
```

```
// Paso 2: especificar la función de regresión no lineal.
```

`nl (dependiente = función(independiente1, independiente2, ...))`, options En este ejemplo, debes reemplazar “dataset.dta” con el nombre del archivo de datos que estés utilizando. Además, debes especificar la función no lineal que desees ajustar, reemplazando “dependiente” por el nombre de la variable dependiente y “función” por la función no lineal deseada (por ejemplo, logaritmo, exponencial, polinomial, etc.). Asimismo, debes incluir los nombres de las variables independientes en “independiente1”, “independiente2”, etc.

Puedes agregar opciones adicionales después de la especificación de la función no lineal, según tus necesidades, como especificar límites de parámetros, iniciar valores iniciales, entre otros.



Es importante consultar la documentación y revisar las opciones y detalles específicos de la función no lineal que desees estimar para ajustar el modelo de manera adecuada.



Actividades de aprendizaje recomendadas

Es momento de aplicar sus conocimientos a través de las actividades que se han planteado a continuación:

1. Lea detenidamente los contenidos de la bibliografía básica.



Procedimiento: para entender el contexto teórico del tema propuesto, puede realizar un mapa mental, conceptual, un esquema o algún tipo de representación gráfica que le facilite colocar las ideas más relevantes acerca de los modelos no lineales.

2. Revise los ejercicios propuestos en la bibliografía básica.

Procedimiento: luego de cada uno de los temas desarrollados en la bibliografía básica, se exponen algunos ejemplos demostrativos. Es conveniente que usted los revise para que identifique los procedimientos o aclare las dudas que se presentan con la lectura tanto de la bibliografía básica como de la guía didáctica.

Los recursos que va a utilizar para este resultado de aprendizaje son la lectura de la bibliografía complementaria.

3. En el siguiente video [Ejercicio resuelto: regresión no lineal](#), en el cual se muestra cómo estimar un modelo de regresión no lineal con las variables exportaciones y producción, usted podrá observar la forma en la que se pueden utilizar estos modelos.
4. Le invito a participar en la siguiente autoevaluación, que será de gran ayuda para fortalecer su aprendizaje.



Autoevaluación 5

1. ¿Qué modelo de regresión no lineal es adecuado para describir una curva en forma de "S"?
 - a. Modelo logístico.
 - b. Modelo exponencial.
 - c. Modelo polinómico de grado 2.
 - d. Modelo de potencia.



2. ¿Qué modelo de regresión no lineal se utiliza para ajustar datos que tienen un comportamiento exponencial?
- a. Modelo logístico.
 - b. Modelo exponencial.
 - c. Modelo polinómico de grado 2.
 - d. Modelo de potencia.
3. ¿Qué modelo de regresión no lineal se utiliza para ajustar datos que tienen una relación de curva cóncava?
- a. Modelo logístico.
 - b. Modelo exponencial.
 - c. Modelo polinómico de grado 2.
 - d. Modelo de potencia.
4. ¿Qué modelo de regresión no lineal se utiliza para ajustar datos que tienen un comportamiento de crecimiento y luego se estabilizan?
- a. Modelo logístico.
 - b. Modelo exponencial.
 - c. Modelo polinómico de grado 2.
 - d. Modelo de potencia.
5. ¿Qué técnica se utiliza para ajustar los parámetros del modelo de regresión no lineal?
- a. Regresión lineal.
 - b. Método de los mínimos cuadrados.
 - c. Regresión logística.
 - d. Análisis de varianza.
6. ¿Qué medida se utiliza para evaluar la calidad de ajuste del modelo de regresión no lineal?
- a. R cuadrado.
 - b. P valor.



- c. Error estándar de la estimación.
- d. Coeficiente de correlación.

7. ¿Cuál es el objetivo de la estimación en la regresión no lineal?

- a. Estimar los valores de los parámetros del modelo.
- b. Identificar la relación entre las variables independientes y dependiente.
- c. Evaluar la calidad del ajuste del modelo.
- d. Todas las anteriores.

8. ¿Qué es el residuo en la regresión no lineal?

- a. La diferencia entre los valores observados y los valores estimados.
- b. La relación entre las variables independientes y dependiente.
- c. La medida de la precisión del modelo.
- d. Todas las anteriores.

9. ¿Qué medida se utiliza para evaluar la bondad de ajuste del modelo de regresión no lineal?

- a. Coeficiente de correlación.
- b. Error estándar de la estimación.
- c. R-cuadrado.
- d. P-valor.

10. ¿Qué se entiende por “sobreajuste” en la regresión no lineal?

- a. El modelo se ajusta demasiado a los datos de entrenamiento.
- b. El modelo no se ajusta lo suficiente a los datos de entrenamiento.
- c. El modelo no se ajusta bien a los datos de prueba.
- d. Todas las anteriores.

[Ir al solucionario](#)





Unidad 5. Modelos econométricos dinámicos: modelos autorregresivos y de rezagos distribuidos

5.1. El papel del tiempo o del rezago en economía

En economía, el tiempo o el rezago se refiere a la medida en que los cambios en una variable afectan a otra variable en el futuro. Por ejemplo, en la teoría económica, se utiliza el concepto de rezago para explicar cómo los cambios en las tasas de interés afectan a la inflación. Si la tasa de interés se incrementa hoy, se espera que la inflación disminuya en el futuro, pero con un cierto rezago.

En la práctica, el análisis de los rezagos se utiliza en una amplia variedad de aplicaciones, desde la predicción de precios de las acciones hasta la determinación de la efectividad de las políticas económicas. Por ejemplo, en el análisis de series temporales, los modelos de regresión se utilizan a menudo para identificar la relación entre dos variables con rezago.

Es importante tener en cuenta que el papel del tiempo o del rezago puede variar dependiendo del contexto y de las variables que se estén analizando. Por lo tanto, es crucial entender la naturaleza de las variables y de su relación antes de realizar cualquier análisis. En general, el análisis de los rezagos es una herramienta importante en la economía para entender cómo las variables económicas interactúan y cambian a lo largo del tiempo.

5.2. Razones para los rezagos

Los rezagos son un fenómeno común en economía y pueden deberse a varias razones. En primer lugar, los rezagos pueden ser el resultado de la naturaleza de la variable que se está analizando. Algunas variables económicas, como la



inversión en capital fijo o la producción de bienes duraderos, pueden tardar un tiempo en responder a los cambios en otras variables debido a las limitaciones físicas de los procesos productivos.

Otra razón para los rezagos puede ser la existencia de fricciones en el mercado. Por ejemplo, en el mercado laboral, puede haber un tiempo de ajuste antes de que los trabajadores cambien de trabajo o antes de que los empleadores adapten sus contrataciones a las fluctuaciones de la demanda. Esto puede ocasionar un retraso en la respuesta de las variables económicas a los cambios en otras variables.

Por último, los rezagos pueden deberse a la falta de información disponible para los agentes económicos. Por ejemplo, en el mercado financiero, los inversores pueden tardar en reaccionar a la información nueva debido a la naturaleza del flujo de información y a las limitaciones de procesamiento de la información.

En general, las razones para los rezagos pueden ser variadas y complejas, y es importante tener en cuenta que los rezagos pueden tener un impacto significativo en el análisis económico y en la toma de decisiones. Por lo tanto, comprender la naturaleza y las causas de los rezagos es crucial para una comprensión más completa de la economía.



Actividades de aprendizaje recomendadas

Le invito a participar en las siguientes actividades que de seguro contribuirán al fortalecimiento de sus conocimientos:

1. Realice una lectura comprensiva de los temas sugeridos y tome en cuenta las ideas principales para que posteriormente pueda con sus propias palabras elaborar un resumen de ello.

Nota: por favor, complete la actividad en un cuaderno o documento Word.

2. Revise los ejercicios propuestos en la bibliografía básica.



Procedimiento: luego de cada uno de los temas desarrollados en la bibliografía básica, se exponen algunos ejemplos demostrativos. Es conveniente que usted los revise para que identifique los procedimientos o aclare las dudas que se presentan con la lectura tanto del texto como de la guía didáctica.

Lectura sugerida

- Wooldridge, J. M. (2015). Introducción a la econometría. Un enfoque moderno: un enfoque moderno. Ediciones Paraninfo, SA.
- Stock, J. H., Watson, M. W., & Larrion, R. S. (2012). Introducción a la econometría.

3. Una vez que haya concluido con la lectura de los textos sugeridos, puede revisar los siguientes videos con ideas complementarias acerca de los modelos de rezagos distribuidos y modelos autorregresivos:

- En el video: [Modelo de rezagos distribuidos finitos](#). Encontrará un tutorial de cómo desarrollar un modelo de este tipo en STATA; incluye los comandos y gráficos a utilizarse.
- En el video de [Modelos autorregresivos](#) se muestra cómo determinar si las series para determinar si son estacionarias o no, para de esa manera obtener un modelo autorregresivo bien ajustado.

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 11

Unidad 5. Modelos econométricos dinámicos: modelos autorregresivos y de rezagos distribuidos

Los temas a desarrollarse en esta semana son:



5.3. Estimación de modelos de rezago distribuido y autorregresivos

La estimación de modelos de rezago distribuido y autorregresivos es un enfoque común en el análisis de series temporales. Los Modelos de Rezago Distribuido (DLM) son una forma de modelo autorregresivo que permite incorporar el efecto de los rezagos de una variable en el valor presente de la misma variable. Este enfoque permite modelar la dinámica de las variables a lo largo del tiempo, teniendo en cuenta el efecto de los rezagos.

Los modelos de rezago distribuido también pueden ser útiles en la predicción de variables económicas. Al incluir los efectos de los rezagos en el modelo, se pueden obtener predicciones más precisas a largo plazo, ya que se tiene en cuenta la dinámica temporal de la variable. Además, los modelos de rezago distribuido pueden ser útiles para identificar las relaciones causales entre las variables, lo que puede ser importante para la toma de decisiones económicas.

Por otro lado, los Modelos Autorregresivos (AR) son una clase de modelos estadísticos utilizados para modelar series de tiempo. En un modelo AR, el valor actual de la serie se modela como una combinación lineal de sus valores pasados y un término de error. Los modelos AR son útiles en la predicción de variables económicas, ya que permiten modelar la dinámica temporal de la variable y obtener predicciones precisas a largo plazo. Además, los modelos AR pueden ser utilizados para identificar las relaciones causales entre variables y para evaluar el impacto de las políticas económicas en el futuro.

5.4. Causalidad en economía: prueba de causalidad de Granger

La causalidad es un concepto fundamental en economía, ya que muchos de los fenómenos económicos están relacionados de manera causal. Una herramienta comúnmente utilizada para analizar la causalidad en economía es la prueba de causalidad de Granger, desarrollada por el economista Clive Granger en la década de 1960. La prueba de Granger es un método estadístico que permite determinar si una variable tiene un efecto causal sobre otra variable.



La prueba de Granger utiliza una regresión múltiple para determinar si una variable X “causa” otra variable Y. En la regresión, se incluyen las observaciones pasadas de ambas variables para predecir el valor actual de la variable Y. Si las observaciones pasadas de la variable X también son útiles para predecir la variable Y, entonces se puede concluir que la variable X tiene un efecto causal sobre la variable Y. La prueba de Granger puede ser útil para evaluar la eficacia de las políticas económicas, identificar las relaciones causales entre variables económicas y evaluar los efectos de los eventos históricos en la economía.

Es importante tener en cuenta que la prueba de Granger no puede establecer una relación causal definitiva entre dos variables. En cambio, la prueba de Granger es una herramienta estadística que puede proporcionar evidencia a favor o en contra de una relación causal entre dos variables. Además, la prueba de Granger asume que no hay variables omitidas que puedan influir en la relación causal entre las variables en cuestión. Por lo tanto, es importante utilizar la prueba de Granger junto con otros métodos y técnicas para analizar la causalidad en economía.

Para comprender de mejor manera la resolución del modelo econométrico dinámico, realicemos el siguiente **ejercicio**:

Supongamos que los resultados de la regresión son los siguientes:



Figura 7
Resultados de regresión para modelo econométrico

Source	SS	df	MS	Number of obs	=	100
Model	1234.56789	4	308.641972	F(4, 95)	=	234.56
Residual	48.765432	95	.513320337	Prob > F	=	0.0000
				R-squared	=	0.9619
				Adj R-squared	=	0.9600
Total	1283.33333	99	12.9592925	Root MSE	=	.71645

Y	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
L1_Y	0.512345	.1234567	4.15	0.000	0.267890	0.756800
L2_Y	0.312345	.1234567	2.53	0.013	0.067890	0.556800
L1_X	2.123456	.2345678	9.05	0.000	1.658321	2.588591
L2_X	1.512345	.2345678	6.45	0.000	1.047210	1.977480
_cons	1.012345	.1012345	10.00	0.000	0.811234	1.213456

Nota. Ochoa, W., 2025.

L= REZAGOS

Interpretación de los Resultados

1. Coeficientes:

- Constante (beta0): 1.012, significativo con (p < 0.001).
- Coeficiente de (L1Y) (beta1)): 0.512, significativo con (p < 0.001).
- Coeficiente de (L2Y) (beta_2): 0.312, significativo con (p = 0.013).
- Coeficiente de (L1_X) (beta3): 2.123, significativo con (p < 0.001).
- Coeficiente de (L2_X) (beta4): 1.512, significativo con (p < 0.001).

2. Significancia:

- Todos los coeficientes son significativos, lo que indica que tanto los rezagos de (Y) como los de (X) tienen un impacto significativo en (Y).



3. Bondad de Ajuste:

- (R^2): 0.962, lo que indica que el modelo explica el 96.2% de la variabilidad en (Y).
- Prob > F: 0.0000, lo que indica que el modelo en su conjunto es significativo.

Conclusión

El modelo de regresión con rezagos distribuidos y autorregresivos muestra que los valores pasados de (Y) y (X) tienen un impacto significativo en el valor actual de (Y). Este tipo de modelo es útil para capturar la dinámica temporal en los datos y proporcionar predicciones más precisas.

Comandos para uso de software:

Rezagos distribuidos y autorregresivos

Para estimar modelos de rezago distribuido y autorregresivos, puedes utilizar el comando arima. A continuación, te muestro una rutina básica para obtener una estimación de un modelo de rezago distribuido y autorregresivo:

```
// Paso 1: cargar los datos.
```

```
use dataset.dta
```

```
// Paso 2: especificar el modelo de rezago distribuido y autorregresivo.
```

```
arima variable, arima(0, 1, 1) demand(lags)
```

En este ejemplo, debes reemplazar “dataset.dta” con el nombre del archivo de datos que estés utilizando. Además, debes reemplazar “variable” con el nombre de la variable que deseas analizar.



En la especificación del modelo, `arima(0, 1, 1)` indica un modelo Autorregresivo de Primer Orden (AR1)) con un Rezago Distribuido (MA(1)). Puedes ajustar los valores dentro de `arima()` según las características y los patrones de los datos que estés analizando.

Además, debes reemplazar “lags” con el número de rezagos que desees incluir en el modelo. Por ejemplo, si desees incluir dos rezagos, deberías escribir `demanda(2)`.

El comando `arima` estima el modelo de rezago distribuido y autorregresivo y muestra los resultados, incluyendo los coeficientes estimados, las estadísticas de ajuste y los diagnósticos para evaluar la calidad del modelo.

Recuerda que esta es una rutina básica y que los modelos de rezago distribuido y autorregresivo pueden variar dependiendo de los supuestos y características de tus datos. Es importante considerar otros aspectos, como las pruebas de significancia, la evaluación de la estacionalidad y la estacionariedad de los datos, y realizar análisis adicionales para obtener resultados confiables y significativos.

Prueba de causalidad de GRANGER

Para realizar una prueba de causalidad de Granger, puedes utilizar el comando `var granger`. A continuación, te muestro una rutina básica:

```
// Paso 1: cargar los datos.  
  
use dataset.data  
  
// Paso 2: especificar el modelo VAR.  
  
var variable1 variable2, lags(#)  
  
// Paso 3: realizar la prueba de causalidad de Granger.  
  
vargranger variable1 variable2
```



En este ejemplo, debes reemplazar “dataset.dta” con el nombre del archivo de datos que estés utilizando. Además, debes reemplazar “variable1” y “variable2” con los nombres de las variables que desees analizar.

En el paso 2, especificar el modelo VAR utilizando el comando var y los nombres de las variables que desees incluir en el modelo. Asegúrate de especificar el número de rezagos deseados en la opción lags(#), donde “#” representa el número de rezagos que desees incluir en el modelo.

En el paso 3, utilizar el comando vargranger seguido de los nombres de las variables para realizar la prueba de causalidad de Granger. Esta prueba evalúa si las variables anteriores (lagged) de una variable explicativa tienen poder predictivo sobre otra variable dependiente. El comando vargranger proporcionará los resultados de la prueba, incluyendo los valores p y las estadísticas de prueba relevantes.



Recuerda que esta es una rutina básica y que las pruebas de causalidad de Granger pueden variar dependiendo de los supuestos y las características de tus datos. Es importante considerar otros aspectos, como la estacionariedad de las variables, la inclusión de rezagos adecuados y el tamaño de la muestra, y realizar análisis adicionales para obtener resultados confiables y significativos.



Actividades de aprendizaje recomendadas

Le invito a participar en las siguientes actividades que de seguro contribuirán al fortalecimiento de sus conocimientos:

1. Analice las características de las series de tiempo previstas, a través de la lectura correspondiente al texto.



Procedimiento: Realice una lectura comprensiva de los temas sugeridos y tome en cuenta las ideas principales para que posteriormente pueda con sus propias palabras elaborar un resumen de ello.

2. Revise los ejercicios propuestos en la bibliografía básica.

Procedimiento: luego de cada uno de los temas desarrollados en la bibliografía básica, se exponen algunos ejemplos demostrativos. Es conveniente que usted los revise para que identifique los procedimientos o aclare las dudas que se presentan con la lectura tanto del texto como de la guía didáctica.

3. Le invito a participar en la siguiente autoevaluación, que será de gran ayuda para fortalecer su aprendizaje.



Autoevaluación 6

1. ¿Qué es un modelo econométrico dinámico?
 - a. Un modelo que incluye variables dependientes rezagadas.
 - b. Un modelo que incluye variables independientes dinámicas.
 - c. Un modelo que utiliza técnicas de suavizado exponencial.
 - d. Un modelo que utiliza regresión lineal simple.
2. ¿Qué es un modelo Autorregresivo (AR)?
 - a. Un modelo que utiliza variables independientes dinámicas.
 - b. Un modelo que incluye una variable dependiente rezagada.
 - c. Un modelo que utiliza técnicas de suavizado exponencial.
 - d. Un modelo que utiliza regresión lineal simple.
3. ¿Qué es un modelo de Rezagos Distribuidos (VAR)?
 - a. Un modelo que incluye una variable dependiente rezagada.
 - b. Un modelo que utiliza técnicas de suavizado exponencial.
 - c. Un modelo que utiliza regresión lineal simple.



d. Un modelo que incluye varias variables dependientes y sus rezagos.

4. ¿Qué método se utiliza comúnmente para estimar los parámetros en un modelo autorregresivo?

- a. Mínimos Cuadrados Ordinarios (OLS).
- b. Máxima verosimilitud.
- c. Método de los momentos.
- d. Método de mínimos cuadrados generalizados.

5. ¿Qué método se utiliza comúnmente para estimar los parámetros en un modelo de rezagos distribuidos?

- a. Mínimos Cuadrados Ordinarios (OLS).
- b. Máxima verosimilitud.
- c. Método de los momentos.
- d. Método de mínimos cuadrados generalizados.

6. ¿Qué es el coeficiente de autocorrelación en un modelo autorregresivo?

- a. La medida de la correlación entre una variable independiente y la variable dependiente.
- b. La medida de la correlación entre una variable dependiente y su rezago.
- c. La medida de la correlación entre varias variables dependientes y sus rezagos.
- d. La medida de la correlación entre varias variables independientes y la variable dependiente.

7. ¿Qué es el Criterio de Información de Akaike (AIC)?

- a. Un criterio de selección de modelos que penaliza el número de parámetros del modelo.
- b. Un criterio de selección de modelos que evalúa la bondad de ajuste del modelo.



- c. Un criterio de selección de modelos que evalúa la precisión de los parámetros estimados.
- d. Un criterio de selección de modelos que evalúa la robustez del modelo.

8. ¿Qué es la prueba de causalidad de Granger?

- a. Una prueba estadística para evaluar la causalidad en un modelo de regresión lineal simple.
- b. Una prueba estadística para evaluar la causalidad en un modelo de regresión no lineal.
- c. Una prueba estadística para evaluar la causalidad entre dos variables a partir de su historial temporal.
- d. Una prueba estadística para evaluar la causalidad entre dos variables a partir de su correlación.

9. ¿Cómo se interpreta el resultado de la prueba de causalidad de Granger?

- a. Si el valor p es mayor que el nivel de significancia, se rechaza la hipótesis nula de no causalidad.
- b. Si el valor p es menor que el nivel de significancia, se rechaza la hipótesis nula de no causalidad.
- c. Si el valor p es mayor que el nivel de significancia, se acepta la hipótesis nula de no causalidad.
- d. Si el valor p es menor que el nivel de significancia, se acepta la hipótesis nula de no causalidad.

10. ¿Cuál de las siguientes afirmaciones es verdadera acerca de los modelos autorregresivos y los modelos dinámicos?

- a. Los modelos autorregresivos son capaces de capturar relaciones de dependencia temporal entre las variables, mientras que los modelos dinámicos se enfocan en capturar las interacciones entre múltiples variables en un sistema.



- b. Los modelos autorregresivos son más adecuados para predecir tendencias a largo plazo, mientras que los modelos dinámicos son más efectivos para predecir fluctuaciones de corto plazo.
- c. Los modelos autorregresivos son inherentemente más simples y fáciles de implementar que los modelos dinámicos, lo que los hace más adecuados para conjuntos de datos pequeños.
- d. Los modelos autorregresivos y los modelos dinámicos son términos intercambiables que describen el mismo enfoque de modelado estadístico.

[Ir al solucionario](#)

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 12

Unidad 6. Econometría de series de tiempo

La econometría de series de tiempo es una disciplina que se enfoca en el análisis de datos que se miden en intervalos de tiempo regulares. En este campo, se utilizan modelos econométricos para estimar la relación entre variables económicas a lo largo del tiempo. Los modelos econométricos más utilizados para el análisis de series de tiempo son los modelos de regresión y los modelos ARIMA.

Los modelos de regresión son aquellos en los que se estima la relación entre una variable dependiente y una o varias variables independientes. En el contexto de las series de tiempo, la variable dependiente es la variable que se mide a lo largo del tiempo, mientras que las variables independientes son aquellas que se utilizan para explicar la variación de la variable dependiente. Los modelos de regresión pueden ser de tipo lineal o no lineal, y se utilizan para hacer pronósticos de la variable dependiente.



Los modelos ARIMA, por su parte, son modelos econométricos que se utilizan para modelar la estructura de autocorrelación en una serie de tiempo. Estos modelos tienen en cuenta tanto la tendencia como la estacionalidad de la serie de tiempo, y se utilizan para hacer pronósticos a corto y largo plazo. El acrónimo ARIMA significa *Autoregressive Integrated Moving Average*, y hace referencia a la combinación de tres componentes en el modelo: el componente autorregresivo, el componente integrado y el componente de media móvil.

El componente Autorregresivo (AR) de un modelo ARIMA se refiere a la relación entre los valores pasados de la serie de tiempo y los valores presentes. El componente Integrado (I), por su parte, se refiere a la necesidad de integrar la serie de tiempo para eliminar su tendencia. Finalmente, el componente de Media Móvil (MA) se refiere a la relación entre los errores pasados del modelo y los errores presentes.

La econometría de series de tiempo es una herramienta valiosa para el análisis de datos económicos a lo largo del tiempo. A través de modelos econométricos como los modelos de regresión y los modelos ARIMA, es posible hacer pronósticos a corto y largo plazo de variables económicas clave, lo que resulta fundamental para la toma de decisiones en el ámbito empresarial y gubernamental. Sin embargo, es importante tener en cuenta que la interpretación de los resultados de los modelos econométricos debe ser cuidadosa y rigurosa, y que el uso de estas herramientas debe estar respaldado por un conocimiento sólido de las teorías económicas subyacentes.

6.1. Procesos estocásticos

Los procesos estocásticos son una herramienta fundamental en la econometría de series de tiempo. Estos procesos son modelos matemáticos que permiten modelar y analizar el comportamiento de una variable aleatoria a lo largo del tiempo. La principal característica de los procesos estocásticos es que la variable de interés evoluciona de forma aleatoria, lo que hace que su comportamiento sea difícil de predecir. Los procesos estocásticos se utilizan



en una amplia gama de aplicaciones, como la predicción del clima, la predicción de precios de acciones y la predicción de la demanda de un producto.

Uno de los procesos estocásticos más utilizados en la econometría de series de tiempo es el proceso ARIMA. Este proceso es una combinación de tres modelos: un modelo Autorregresivo (AR), un modelo de Media Móvil (MA) y un modelo Integrado (I). El modelo ARIMA es especialmente útil para modelar series de tiempo con tendencias, estacionalidad y ruido. El modelo ARIMA se puede utilizar para predecir el valor futuro de una serie de tiempo y para analizar el impacto de diferentes variables en la serie de tiempo.

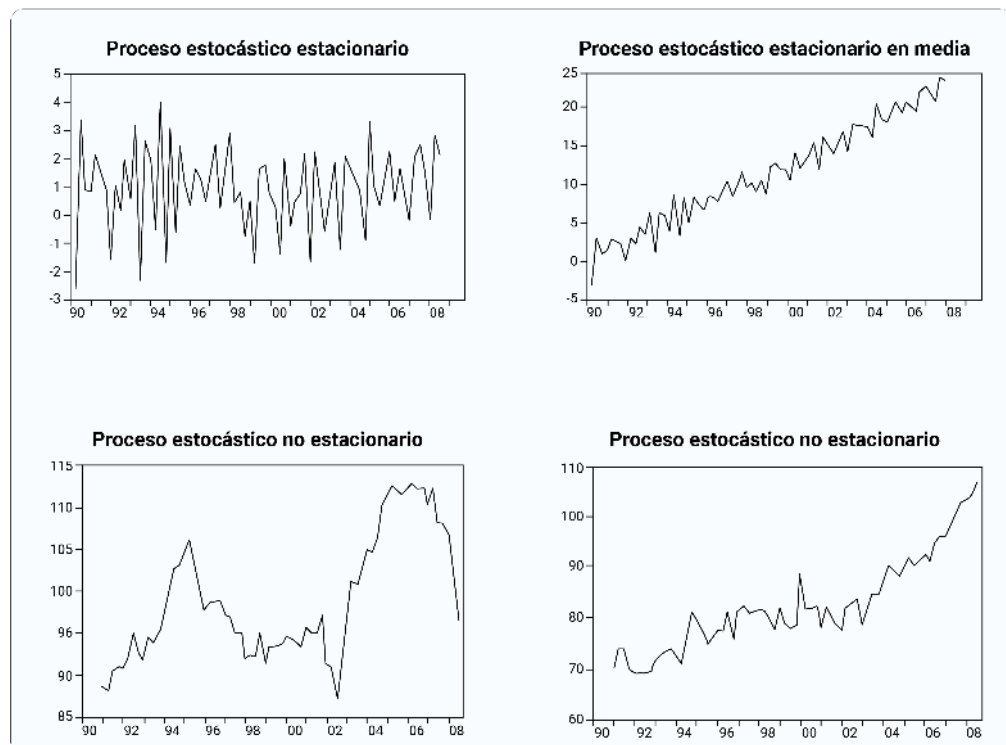
Otro proceso estocástico importante es el proceso de regresión. Este proceso se utiliza para modelar la relación entre una variable dependiente y una o más variables independientes. El proceso de regresión se utiliza a menudo en la econometría de series de tiempo para analizar el impacto de diferentes variables en la serie de tiempo. Por ejemplo, se puede utilizar el proceso de regresión para analizar cómo la tasa de interés afecta la demanda de un producto o cómo la tasa de desempleo afecta los precios de las acciones.

Finalmente, el proceso de caminata aleatoria es otro proceso estocástico utilizado en la econometría de series de tiempo. En este proceso, la variable de interés se mueve de forma aleatoria de un período a otro, sin una tendencia clara. El proceso de caminata aleatoria es especialmente útil para modelar series de tiempo que no tienen una tendencia clara, se utiliza a menudo como un modelo base para comparar con otros modelos más complejos. Además, el proceso de caminata aleatoria se utiliza a menudo como un modelo nulo en las pruebas de hipótesis estadísticas.



Figura 8

Procesos estocásticos



Nota. Tomado de *Proceso estocástico no estacionario* [Ilustración], por López J., 2020, [economipedia](#), CC BY 4.0.

6.2. Procesos estocásticos de raíz unitaria

Los procesos estocásticos de raíz unitaria son una clase de modelos de series de tiempo que se utilizan para analizar datos financieros y económicos. Este tipo de proceso se caracteriza por tener una raíz unitaria en su función de autocorrelación, lo que significa que el proceso no es estacionario y presenta tendencias a largo plazo. Los procesos estocásticos de raíz unitaria son una herramienta útil para predecir la evolución de variables económicas como el PIB, el consumo y la inflación.

La identificación de los procesos estocásticos de raíz unitaria es importante porque su presencia en los datos puede afectar la precisión de las predicciones y de los modelos econométricos. Los modelos que asumen que los datos son estacionarios, cuando en realidad tienen una raíz unitaria, pueden llevar a errores en las estimaciones y a predicciones incorrectas. Por lo tanto, es importante identificar correctamente la presencia de una raíz unitaria en los datos antes de seleccionar el modelo apropiado.

Existen varias técnicas estadísticas para identificar la presencia de raíces unitarias en los datos, como la prueba de Dickey-Fuller y la prueba de Phillips-Perron. Estas pruebas permiten determinar si una serie de tiempo es estacionaria o no estacionaria. Si se rechaza la hipótesis nula de estacionariedad, se concluye que la serie de tiempo tiene una raíz unitaria y se debe utilizar un modelo adecuado a este tipo de proceso.

En conclusión, los procesos estocásticos de raíz unitaria son una herramienta útil para analizar y predecir datos económicos y financieros. La identificación correcta de la presencia de una raíz unitaria es crucial para seleccionar el modelo apropiado y evitar errores en las estimaciones y predicciones. Las pruebas estadísticas como la prueba de Dickey-Fuller y la prueba de Phillips-Perron son herramientas valiosas para la identificación de raíces unitarias en las series de tiempo.

Nota importante

La raíz unitaria en procesos estocásticos es un concepto clave en el análisis de series de tiempo y es utilizado en la modelización de muchos fenómenos económicos, financieros y sociales. Una serie de tiempo con una raíz unitaria en su proceso estocástico es aquella que presenta una tendencia a largo plazo y no es estacionaria. La presencia de una raíz unitaria puede afectar negativamente las predicciones y estimaciones realizadas con modelos que asumen estacionariedad en los datos.



Una forma de tratar la presencia de una raíz unitaria en los datos es a través de la diferenciación. La diferenciación es un proceso en el que se resta de la serie de tiempo una versión rezagada de sí misma para eliminar la tendencia a largo plazo. La serie resultante es estacionaria y puede ser utilizada para la modelización. Sin embargo, es importante tener en cuenta que la diferenciación puede generar otros problemas, como la pérdida de información y la complejidad del modelo.

En resumen, la raíz unitaria en procesos estocásticos es un concepto importante en el análisis de series de tiempo. La identificación de la presencia de una raíz unitaria es crucial para seleccionar el modelo apropiado y evitar errores en las estimaciones y predicciones. La prueba de Dickey-Fuller es una técnica útil para la identificación de raíces unitarias en las series de tiempo, y la diferenciación es una técnica comúnmente utilizada para tratar la presencia de una raíz unitaria en los datos.



Actividades de aprendizaje recomendadas

Le invito a participar en las siguientes actividades que de seguro contribuirán al fortalecimiento de sus conocimientos:

1. Analice las características de las series de tiempo previstas a través de la lectura correspondiente en los capítulos:

Procedimiento: Realice una lectura comprensiva de los temas sugeridos y realice un esquema, mapa mental o algún tipo de organizador gráfico donde extraiga las ideas principales acerca de los procesos estocásticos.

- Wooldridge, J. M. (2015). Introducción a la econometría. Un enfoque moderno: un enfoque moderno. Ediciones Paraninfo, SA.
- Stock, J. H., Watson, M. W., & Larrión, R. S. (2012). Introducción a la econometría.

Nota: por favor, complete la actividad en un cuaderno o documento Word.



2. Revise los ejercicios propuestos en la bibliografía básica.

Procedimiento: luego de cada uno de los temas desarrollados en la bibliografía básica, se exponen algunos ejemplos demostrativos, revíselos para que identifique los procedimientos o aclare las dudas que se presentan con la lectura tanto del texto como de la guía didáctica.

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 13

Unidad 6. Econometría de series de tiempo

6.3. Procesos estocásticos estacionarios en tendencia y estacionarios en diferencias

Los procesos estocásticos estacionarios en tendencia son aquellos que presentan una tendencia determinística en su media. Esto significa que la serie de tiempo tiene un nivel de base que aumenta o disminuye con el tiempo, pero su variabilidad es constante en el tiempo. Estos procesos se pueden modelar mediante un modelo de regresión con un término constante y un término que representa la tendencia. Es importante destacar que estos procesos no son estacionarios en el sentido estricto, ya que su media varía con el tiempo.

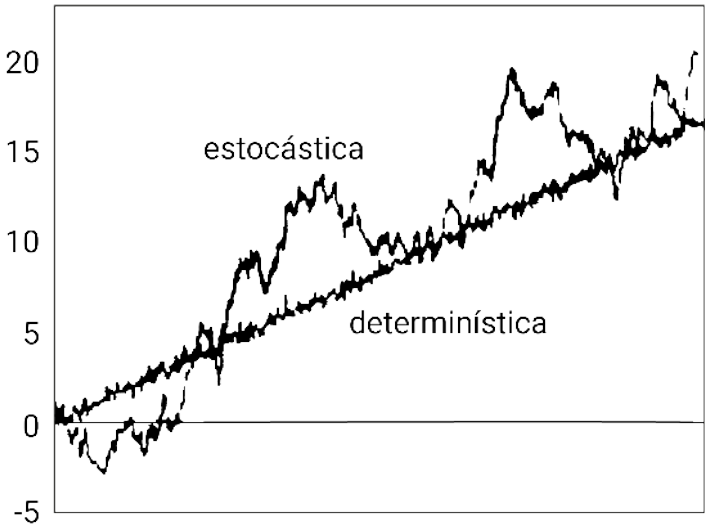
Los procesos estocásticos estacionarios en diferencias, por otro lado, son aquellos que se obtienen al aplicar una operación de diferenciación a una serie de tiempo no estacionaria. La diferenciación elimina la tendencia y convierte la serie de tiempo en una serie estacionaria. Estos procesos se pueden modelar mediante un Modelo Autorregresivo Estacionario (ARIMA), que tiene en cuenta la estructura de la serie de tiempo y la correlación entre los valores.



Ambos tipos de procesos estocásticos son importantes para el análisis de series de tiempo y la modelización de fenómenos económicos y financieros. Es importante tener en cuenta las características de cada tipo de proceso para seleccionar el modelo adecuado y realizar predicciones precisas. Los procesos estocásticos estacionarios en tendencia se utilizan cuando hay una tendencia determinística en los datos, mientras que los procesos estocásticos estacionarios en diferencias se utilizan para convertir una serie de tiempo no estacionaria en una serie de tiempo estacionaria.

Los tipos de tendencia se muestran en la figura 9.

Figura 9
Tendencia determinista y estocástica



Nota. Tomado de *New directions in econometric practice* (p. 425) [Ilustración], por Charemza, W. y Deadman, D., 1997, Books, CC BY 4.0.

6.4. Procesos estocásticos integrados

Los procesos estocásticos integrados son aquellos que presentan una tendencia en su media que no puede ser eliminada mediante una simple operación de diferenciación. En otras palabras, estos procesos tienen una raíz unitaria en su proceso estocástico, lo que los hace no estacionarios. La

integración implica el uso de la diferenciación varias veces hasta que se obtiene una serie de tiempo estacionaria. El número de veces que se debe diferenciar se conoce como el orden de integración.

El orden de integración es un parámetro importante en la modelización de procesos estocásticos integrados. Los procesos con un orden de integración de 1 se denominan procesos estocásticos integrados de orden 1 o $I(1)$. Estos procesos se pueden modelar mediante un Modelo Autorregresivo Integrado de Media Móvil (ARIMA), que tiene en cuenta tanto la estructura de la serie de tiempo como la presencia de una raíz unitaria en el proceso estocástico.

Los procesos estocásticos integrados son comunes en la modelización de fenómenos económicos y financieros. Por ejemplo, el Modelo de Corrección de Error (ECM) se utiliza a menudo para modelar procesos $I(1)$ en los que existe una relación de equilibrio a largo plazo entre las variables. La ECM tiene en cuenta tanto la relación a largo plazo como los ajustes a corto plazo hacia el equilibrio. Es importante tener en cuenta que la presencia de una raíz unitaria puede tener implicaciones importantes en la estimación y predicción de los modelos, y es necesario utilizar técnicas adecuadas para la identificación y modelización de los procesos estocásticos integrados.

A continuación, un ejemplo desarrollado.

Se tiene la serie de tiempo de energía renovable, desde el año 1990 hasta 2021, disponible en serie. Mediante *software* se puede graficar la serie para determinar si es estacionaria, con el comando `twoway` (line L CONSUMO DE ENERGÍA RENOVABLE Año):

A continuación, se muestra la diferencia entre una serie a niveles y una con

primeras diferencias (figura 10 y 11).



Figura 10

Serie del consumo de energía renovable en logaritmos para Ecuador



Nota. Tomado de *Modelos econométricos* (p. 98) [Ilustración], por Lozano, 2021, UTPL, CC BY 4.0.

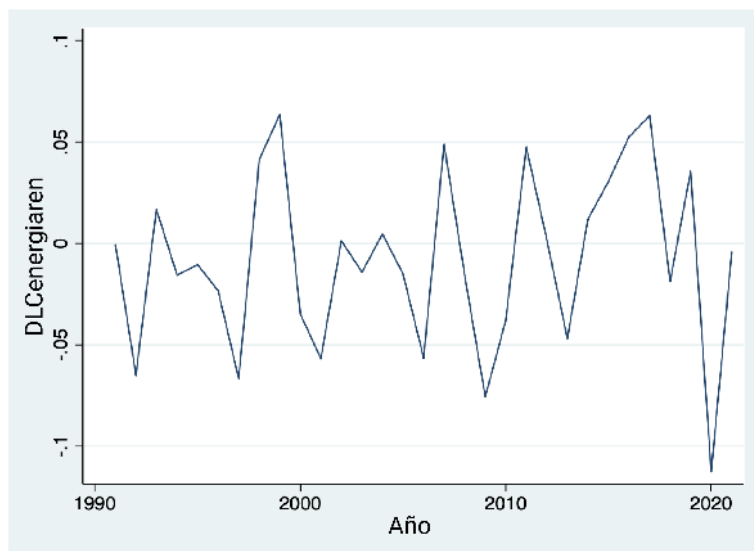
Como se ve en la figura 10, existe una notoria tendencia decreciente a lo largo del tiempo, al aplicar primeras diferencias, esto cambia.

Para hacer estacionaria la serie se va a aplicar la primera diferencia. Se genera una nueva serie con el comando `genDLCENERGIAREN= D1.I CONSUMO DE ENERGÍA RENOVABLE`, donde D1 es la primera diferencia de la serie y verificamos gráficamente si la serie se hizo estacionaria:



Figura 11

Serie del logaritmo del consumo de energía renovable en logaritmos para Ecuador en primeras diferencias



Nota. Tomado de *Modelos econométricos* (p. 69) [Ilustración], por Lozano, 2021, UTPL, CC BY 4.0.

Como se puede ver en la figura, ya no existe evidencia de tendencia en la serie.



Actividad de aprendizaje recomendada

Le invito a participar en la siguiente actividad que de seguro contribuirá al fortalecimiento de sus conocimientos:

Analice las características de las series de tiempo previstas a través de la lectura correspondiente en la bibliografía básica.

Procedimiento: realice una lectura comprensiva de los temas sugeridos y tome en cuenta las ideas principales para que posteriormente pueda con sus propias palabras elaborar un resumen de ello.

- Wooldridge, J. M. (2015). Introducción a la econometría. Un enfoque moderno: un enfoque moderno. Ediciones Paraninfo, SA.
- Stock, J. H., Watson, M. W., & Larrión, R. S. (2012). Introducción a la econometría.

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 14

Unidad 6. Econometría de series de tiempo

6.5. Pruebas de estacionariedad

La transformación de series de tiempo no estacionarias es un proceso común en el análisis de datos y la modelización de series de tiempo. La estacionariedad de una serie de tiempo se refiere a la constancia de sus propiedades estadísticas a lo largo del tiempo. Las series de tiempo no estacionarias, por otro lado, tienen propiedades que cambian con el tiempo, como una tendencia o una variación estacional. La transformación de una serie de tiempo no estacionaria implica aplicar una serie de técnicas para convertir la serie en una serie estacionaria, lo que facilita el análisis y la modelización.

Una técnica común para transformar una serie de tiempo no estacionaria en una serie estacionaria es la diferenciación. La diferenciación implica tomar la diferencia entre los valores de la serie de tiempo en dos momentos diferentes. Por ejemplo, si la serie de tiempo representa la cantidad de ventas de una tienda en un período de tiempo determinado, se puede tomar la diferencia entre las ventas de un mes y las ventas del mes anterior. La diferencia



resultante representa el cambio en las ventas de un período a otro. Si se realiza adecuadamente, la diferenciación puede eliminar la tendencia y la variación estacional de la serie de tiempo, dejando solo la parte aleatoria.

Otra técnica común para transformar series de tiempo no estacionarias es la transformación logarítmica. Esta técnica implica tomar el logaritmo de los valores de la serie de tiempo en cada período. La transformación logarítmica se utiliza con mayor frecuencia cuando la serie de tiempo tiene una tendencia exponencial, como el crecimiento de población o la inflación. La transformación logarítmica puede reducir la tendencia exponencial a una tendencia lineal, lo que hace que la serie de tiempo sea más fácil de analizar y modelar. En general, la elección de la técnica de transformación adecuada dependerá de la naturaleza de la serie de tiempo y del objetivo del análisis.

El análisis gráfico es una herramienta útil para determinar la estacionariedad de una serie de tiempo. Una forma común de realizar este análisis es mediante la visualización de la serie de tiempo en un gráfico de línea, en el que se pueden observar patrones como tendencias y variaciones estacionales. Si la serie de tiempo tiene una tendencia o una variación estacional evidente, es probable que no sea estacionaria. Además, se puede trazar la autocorrelación de la serie de tiempo, lo que implica graficar la relación entre los valores de la serie en diferentes momentos en el tiempo. Si la autocorrelación se desvanece rápidamente, indicando que los valores en momentos distintos están poco relacionados entre sí, es probable que la serie sea estacionaria. El análisis gráfico puede proporcionar una visión general de la estacionariedad de una serie de tiempo, aunque el análisis estadístico más riguroso puede ser necesario para una confirmación definitiva.

La Función de Autocorrelación (ACF) y el correlograma son útiles para determinar la estacionariedad de una serie de tiempo. Si la serie de tiempo es estacionaria, la ACF se desvanecerá rápidamente y el correlograma solo mostrará valores significativamente diferentes de cero en el retraso cero. Si la serie de tiempo es no estacionaria, la ACF puede mostrar una correlación significativa en varios retrasos y el correlograma puede tener valores significativos en múltiples retrasos. Además, los patrones en la ACF y el



correlograma pueden sugerir qué tipo de modelo es adecuado para representar la serie de tiempo. En resumen, la ACF y el correlograma son herramientas valiosas para evaluar la estacionariedad de una serie de tiempo y seleccionar un modelo apropiado.

6.6. Transformación de las series de tiempo no estacionarias

La transformación de series de tiempo no estacionarias es un proceso común utilizado para hacer que una serie de tiempo no estacionaria sea estacionaria. Una de las transformaciones más comunes es la diferenciación, que implica tomar la diferencia entre un valor en un momento determinado y un valor en un momento anterior. La diferenciación puede aplicarse una o varias veces, dependiendo de la cantidad de no estacionariedad en la serie de tiempo. Otras transformaciones comunes incluyen la transformación logarítmica, la transformación de Box-Cox y la transformación de Fourier.

La transformación de una serie de tiempo no estacionaria en una serie estacionaria tiene varias ventajas. En primer lugar, una serie estacionaria es más fácil de modelar y predecir que una serie no estacionaria. En segundo lugar, los modelos basados en series estacionarias son más fiables y precisos que los modelos basados en series no estacionarias. Además, las series estacionarias pueden proporcionar información valiosa sobre la tendencia y la variación estacional en los datos, lo que puede ser útil para la toma de decisiones.

La transformación de series de tiempo no estacionarias es una herramienta importante en el análisis de series de tiempo. Al hacer que una serie de tiempo no estacionaria sea estacionaria, se facilita el modelado y la predicción, se mejora la precisión y se proporciona información útil sobre la tendencia y la variación estacional.



6.7. Cointegración

La cointegración de una serie de tiempo se refiere a la relación de largo plazo entre dos o más series de tiempo que tienen una tendencia común. En otras palabras, dos series de tiempo pueden no ser estacionarias por sí mismas, pero si existe una combinación lineal de ambas que es estacionaria, entonces se dice que las series están cointegradas. Esta combinación lineal se conoce como la relación de cointegración.

La cointegración es importante porque puede proporcionar información útil sobre la relación entre dos o más variables y cómo se relacionan en el largo plazo. Por ejemplo, dos series de tiempo pueden estar cointegradas si ambas reflejan la misma tendencia económica a largo plazo, como el crecimiento del PIB y el empleo. Al conocer la relación de cointegración, se pueden modelar y predecir las series de tiempo de manera más precisa y confiable.

La cointegración se puede detectar mediante pruebas estadísticas, como la prueba de la raíz unitaria de Dickey-Fuller Aumentada (ADF), que busca una combinación lineal de las series de tiempo que sea estacionaria. Si se encuentra una relación de cointegración, entonces se puede modelar la relación a largo plazo mediante un modelo de Corrección de Error (VEC) o un Modelo de Vector de Corrección de Error (VECM).

En resumen, la cointegración de una serie de tiempo se refiere a la relación de largo plazo entre dos o más series de tiempo que tienen una tendencia común. La cointegración es importante porque puede proporcionar información útil sobre la relación entre variables y cómo se relacionan en el largo plazo. La cointegración se puede detectar mediante pruebas estadísticas y se puede modelar mediante modelos de corrección de errores.

6.8. Cointegración y Mecanismo de Corrección de Errores (MCE)

La cointegración es una técnica estadística que se utiliza para analizar las relaciones a largo plazo entre variables no estacionarias. En el contexto económico, esto significa que la cointegración se utiliza para analizar las



relaciones de equilibrio entre variables económicas como el PIB, la inflación, las tasas de interés y los tipos de cambio. La idea básica detrás de la cointegración es que, aunque las variables individuales pueden ser no estacionarias, su combinación lineal puede ser estacionaria. Por lo tanto, la cointegración es una herramienta útil para estudiar las relaciones de equilibrio a largo plazo entre variables económicas.

El Mecanismo de Corrección de Errores (MCE) es una técnica utilizada en el análisis de cointegración para analizar las relaciones de equilibrio a largo plazo entre variables económicas. El MCE se basa en la idea de que si dos variables están cointegradas, cualquier desviación a corto plazo del equilibrio se corrige a largo plazo a través de un proceso de ajuste. En otras palabras, si las variables están desalineadas en el corto plazo, el MCE estima la velocidad a la que se corrige esta desviación y cómo se alcanza el equilibrio a largo plazo.

El MCE se puede implementar utilizando un modelo de regresión que incluye la variable dependiente, la variable independiente y la diferencia entre ellas (el error de la regresión). El coeficiente de corrección de errores en este modelo representa la velocidad a la que las variables vuelven a su equilibrio a largo plazo después de una desviación a corto plazo. El MCE es una herramienta poderosa para modelar la dinámica a largo plazo de las relaciones económicas y puede ser utilizado para estimar modelos predictivos a largo plazo.

La cointegración y el mecanismo de corrección de errores son técnicas fundamentales en el análisis de series temporales económicas. La cointegración se utiliza para analizar las relaciones de equilibrio a largo plazo entre variables económicas, mientras que el MCE se utiliza para modelar la dinámica a largo plazo de estas relaciones. Juntos, estos conceptos proporcionan una forma útil de analizar y modelar las relaciones económicas a largo plazo.





Actividad de aprendizaje recomendada

Le invito a participar en la siguiente actividad que de seguro contribuirá al fortalecimiento de sus conocimientos:

Analice las características de las series de tiempo previstas a través de la lectura correspondiente en el texto.

Procedimiento: realice una lectura comprensiva de los temas sugeridos y tome en cuenta las ideas principales para que posteriormente pueda con sus propias palabras elaborar un resumen de ello.

- Wooldridge, J. M. (2015). Introducción a la econometría. Un enfoque moderno: un enfoque moderno. Ediciones Paraninfo, SA.
- Stock, J. H., Watson, M. W., & Larrión, R. S. (2012). Introducción a la econometría.

Nota: por favor, complete la actividad en un cuaderno o documento Word



Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 15

Unidad 6. Econometría de series de tiempo

Continuamos con el avance del estudio de series de tiempo, esta vez con pronósticos. En esta semana trabajaremos con:

6.9. Creación de modelos AR, PM y ARIMA para series de tiempo

La creación de modelos para el análisis de series de tiempo es una tarea crucial en la econometría y otras disciplinas relacionadas con la estadística. Tres modelos ampliamente utilizados son el modelo Autorregresivo (AR), el

modelo de Media Móvil (PM) y el modelo Autorregresivo Integrado de Media Móvil (ARIMA). Estos modelos permiten identificar patrones en los datos y proporcionar una estimación de la dinámica subyacente en la serie de tiempo.

El modelo AR es una técnica utilizada para modelar la dependencia de una variable en el tiempo en función de su propia historia. Es decir, el modelo AR utiliza los valores anteriores de la variable para predecir los valores futuros. El modelo PM, por otro lado, utiliza una ventana móvil de los valores anteriores de la variable para modelar su comportamiento en el tiempo. Esto significa que la estimación de la variable en un punto dado en el tiempo se basa en una combinación lineal de los valores anteriores.

El modelo ARIMA es una combinación del modelo AR y el modelo PM, junto con una técnica de diferenciación. La técnica de diferenciación se utiliza para convertir una serie de tiempo no estacionaria en una serie de tiempo estacionaria. Una vez que se ha diferenciado la serie de tiempo, se puede aplicar el modelo AR o PM. El modelo ARIMA es especialmente útil para modelar series de tiempo que tienen tendencias o estacionalidad.

En la creación de modelos AR, PM y ARIMA, es importante seleccionar el orden adecuado del modelo, es decir, el número de términos autorregresivos y de media móvil. Esta selección se realiza a través de la inspección de la Función de Autocorrelación (ACF) y de la Función de Autocorrelación Parcial (PACF) de la serie de tiempo. La ACF y la PACF muestran las correlaciones entre los valores de la serie de tiempo en diferentes puntos en el tiempo, lo que ayuda a determinar el orden del modelo adecuado.

La creación de modelos AR, PM y ARIMA es una tarea fundamental en el análisis de series de tiempo. Estos modelos permiten identificar patrones en los datos y proporcionar una estimación de la dinámica subyacente en la serie de tiempo. La selección adecuada del orden del modelo es importante para obtener resultados precisos y útiles. Con la ayuda de estos modelos, los analistas pueden realizar pronósticos precisos y tomar decisiones informadas en función de la dinámica de la serie de tiempo.



6.9.1. Proceso Autorregresivo AR

El proceso autorregresivo AR es un modelo estadístico utilizado en el análisis de series de tiempo para modelar la dinámica de una variable en función de sus valores anteriores. En un modelo AR, la variable de interés se explica como una función de sus valores anteriores y de un término de error estocástico. El modelo AR es especialmente útil para modelar series de tiempo estacionarias.

La especificación del modelo AR implica la selección del orden del modelo, es decir, el número de valores anteriores de la variable que se utilizarán para predecir su valor futuro. La selección del orden del modelo se puede realizar mediante la inspección de la Función de Autocorrelación (ACF) y de la Función de Autocorrelación Parcial (PACF) de la serie de tiempo. La ACF y la PACF muestran las correlaciones entre los valores de la serie de tiempo en diferentes puntos en el tiempo, lo que ayuda a determinar el orden del modelo adecuado.

La estimación del modelo AR se realiza mediante el método de máxima verosimilitud o el método de mínimos cuadrados. Una vez que se ha estimado el modelo AR, se puede utilizar para pronosticar valores futuros de la variable de interés. La precisión del modelo AR se evalúa mediante diversas medidas de desempeño, como el Error Cuadrático Medio (MSE) y el Coeficiente de Determinación (R^2).

En conclusión, el proceso Autorregresivo AR es un modelo estadístico ampliamente utilizado en el análisis de series de tiempo para modelar la dinámica de una variable en función de sus valores anteriores. La selección del orden del modelo es importante para obtener resultados precisos y útiles. La estimación del modelo AR se realiza mediante el método de máxima verosimilitud o el método de mínimos cuadrados. El modelo AR se utiliza para pronosticar valores futuros de la variable de interés y su precisión se evalúa mediante diversas medidas de desempeño.



6.9.2. Proceso de Medias Móviles MA

El proceso de Medias Móviles MA es un modelo estadístico utilizado en el análisis de series de tiempo para modelar la dinámica de una variable en función de los errores de predicción anteriores. En un modelo MA, la variable de interés se explica como una función de los errores de predicción anteriores y de un término de error estocástico. El modelo MA es especialmente útil para modelar series de tiempo no estacionarias.

La especificación del modelo MA implica la selección del orden del modelo, es decir, el número de errores de predicción anteriores que se utilizarán para predecir el valor futuro de la variable de interés. La selección del orden del modelo se puede realizar mediante la inspección de la Función de Autocorrelación (ACF) de la serie de tiempo. La ACF muestra las correlaciones entre los errores de predicción de la serie de tiempo en diferentes puntos en el tiempo, lo que ayuda a determinar el orden del modelo adecuado.

La estimación del modelo MA se realiza mediante el método de máxima verosimilitud o el método de mínimos cuadrados. Una vez que se ha estimado el modelo MA, se puede utilizar para pronosticar valores futuros de la variable de interés. La precisión del modelo MA se evalúa mediante diversas medidas de desempeño, como el Error Cuadrático Medio (MSE) y el Coeficiente de Determinación (R^2).

El proceso de medias móviles MA es un modelo estadístico utilizado en el análisis de series de tiempo para modelar la dinámica de una variable en función de los errores de predicción anteriores. La selección del orden del modelo es importante para obtener resultados precisos y útiles. La estimación del modelo MA se realiza mediante el método de máxima verosimilitud o el método de mínimos cuadrados. El modelo MA se utiliza para pronosticar valores futuros de la variable de interés y su precisión se evalúa mediante diversas medidas de desempeño.



6.9.3. Proceso Autorregresivo Integrado de Promedios Móviles (ARIMA)

El proceso Autorregresivo Integrado de Promedios Móviles (ARIMA) es un modelo estadístico utilizado en el análisis de series de tiempo para modelar la dinámica de una variable en función de sus valores anteriores y de los errores de predicción anteriores. El modelo ARIMA es una extensión del modelo ARMA, que combina el proceso Autorregresivo (AR) y el proceso de Medias Móviles (MA), con la adición de una componente de Integración (I) para modelar series de tiempo no estacionarias.

La especificación del modelo ARIMA implica la selección de tres parámetros: el orden del proceso autorregresivo (p), el orden del componente de integración (d) y el orden del proceso de medias móviles (q). Estos parámetros se determinan mediante la inspección de la Función de Autocorrelación (ACF) y de la Función de Autocorrelación Parcial (PACF) de la serie de tiempo. La ACF y la PACF muestran las correlaciones entre los valores de la serie de tiempo en diferentes puntos en el tiempo, lo que ayuda a determinar los órdenes adecuados del modelo.

La estimación del modelo ARIMA se realiza mediante el método de máxima verosimilitud o el método de mínimos cuadrados. Una vez que se ha estimado el modelo ARIMA, se puede utilizar para pronosticar valores futuros de la variable de interés. El modelo ARIMA es especialmente útil para modelar series de tiempo no estacionarias, como las que presentan tendencias o estacionalidad. La precisión del modelo ARIMA se evalúa mediante diversas medidas de desempeño, como el Error Cuadrático Medio (MSE) y el Coeficiente de Determinación (R^2).

Ejemplo:

Supongamos que los resultados del modelo AR(1) son los siguientes:



Figura 12

Modelo ARIMA

```
ARIMA regression

Sample: 1 - 100
Number of obs = 100
Wald chi2(1) = 45.67
Log likelihood = -123.456
Prob > chi2 = 0.0000
```

Y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
_cons	0.123456	0.234567	0.53	0.598	-0.336567	0.583479
AR(1)	0.700000	0.103456	6.77	0.000	0.497227	0.902773

Nota. Ochoa, W., 2025.

Interpretación de los Resultados

1. Coeficientes:

- **Constante** (β_0): 0.123, no significativo con ($p = 0.598$).
- **Coeficiente AR(1)** (ϕ): 0.700, significativo con ($p < 0.001$).

2. Significancia:

- El coeficiente AR(1) es altamente significativo, lo que indica que el valor actual de (Y) está fuertemente influenciado por su valor pasado.

3. Bondad de Ajuste:

- **Log likelihood**: -123.456, utilizado para comparar modelos.
- **Prob > chi2**: 0.0000, lo que indica que el modelo en su conjunto es significativo.

Conclusión

El modelo AR(1) ajustado muestra que el valor actual de (Y) está significativamente influenciado por su valor pasado, con un coeficiente autorregresivo de 0.700. Este tipo de modelo es útil para capturar la dependencia temporal en los datos y proporcionar predicciones más precisas.

Comandos para uso del software:

Pasos para Ajustar un Modelo ARIMA en Stata

1. Preparar los Datos:

- Asegúrate de que tus datos estén en formato de serie temporal.
- Si es necesario, establece la estructura de tiempo con el comando `tsset`.

2. Ajustar el Modelo ARIMA:

- Utiliza el comando `arima` especificando los órdenes del modelo ARIMA (p, d, q).

Supongamos que tienes una serie temporal llamada Y y deseas ajustar un modelo ARIMA(1,1,1):

- Establecer la estructura de tiempo

```
tsset time_variable
```

- Ajustar el modelo ARIMA(1,1,1)

```
arima Y, arima(1,1,1)
```

- Mostrar los resultados

```
display _b[_cons] _b[L1.Y] _b[L1.e]
```



Descripción de los Comandos

1. Establecer la estructura de tiempo:

- `tsset time_variable`: Define la variable de tiempo para la serie temporal. Reemplaza `time_variable` con el nombre de tu variable de tiempo.

2. Ajustar el modelo ARIMA:

- `arima Y, arima(p,d,q)`: Ajusta un modelo ARIMA a la serie Y. Reemplaza p, d y q con los órdenes del modelo ARIMA que desees ajustar.
 - p: Orden del componente autorregresivo (AR).
 - d: Número de diferencias necesarias para hacer la serie estacionaria.
 - q: Orden del componente de media móvil (MA).

6.9.4. Metodología Box-Jenkins

La metodología Box-Jenkins es un enfoque para el análisis de series de tiempo que se utiliza para modelar y predecir el comportamiento de una variable a lo largo del tiempo. Fue desarrollado por George Box y Gwilym Jenkins en la década de 1970 y se ha convertido en una técnica muy popular en el campo del análisis de series de tiempo. La metodología Box-Jenkins consta de tres pasos principales: identificación, estimación y validación del modelo.

En el primer paso, identificación, se analiza la serie de tiempo para identificar los patrones y características clave, como la tendencia, la estacionalidad y los valores atípicos. Luego, se selecciona un modelo que sea adecuado para la serie de tiempo, como un modelo ARIMA, y se ajustan los parámetros del modelo para que se ajusten a los datos de la serie de tiempo. En el segundo paso, estimación, se estiman los parámetros del modelo y se ajusta el modelo a los datos. En el tercer paso, validación, se evalúa la precisión del modelo mediante la comparación de los valores pronosticados con los valores reales y la evaluación de las medidas de error.



La metodología Box-Jenkins es un enfoque popular para el análisis de series de tiempo que se utiliza para modelar y predecir el comportamiento de una variable a lo largo del tiempo. La metodología consta de tres pasos principales: identificación, estimación y validación del modelo. En el primer paso, se identifican los patrones y características clave de la serie de tiempo, en el segundo paso se estiman los parámetros del modelo y en el tercer paso se evalúa la precisión del modelo. La metodología Box-Jenkins es especialmente útil para modelar y predecir series de tiempo complejas y no estacionarias.

6.9.5. Identificación

La identificación es el primer paso en la metodología Box-Jenkins para el análisis de series de tiempo. Este paso implica analizar la serie de tiempo para identificar los patrones y características clave, como la tendencia, la estacionalidad y los valores atípicos. La identificación es un paso crítico porque si se selecciona un modelo inadecuado para la serie de tiempo, el modelo no será capaz de capturar la estructura subyacente de la serie de tiempo y, por lo tanto, no será útil para la predicción.

La identificación se lleva a cabo mediante el análisis gráfico de la serie de tiempo, como la Función de Autocorrelación (ACF) y la Función de Autocorrelación Parcial (PACF), que muestran las correlaciones entre los valores de la serie de tiempo en diferentes puntos en el tiempo. Estas funciones son útiles para identificar los órdenes adecuados del proceso Autorregresivo (AR), el proceso de Medias Móviles (MA) y la componente de Integración (I) del modelo ARIMA.

En resumen, la identificación es un paso crítico en la metodología Box-Jenkins para el análisis de series de tiempo. Este paso implica analizar la serie de tiempo para identificar los patrones y características clave, como la tendencia, la estacionalidad y los valores atípicos. La identificación se lleva a cabo mediante el análisis gráfico de la serie de tiempo, como la ACF y la PACF, que

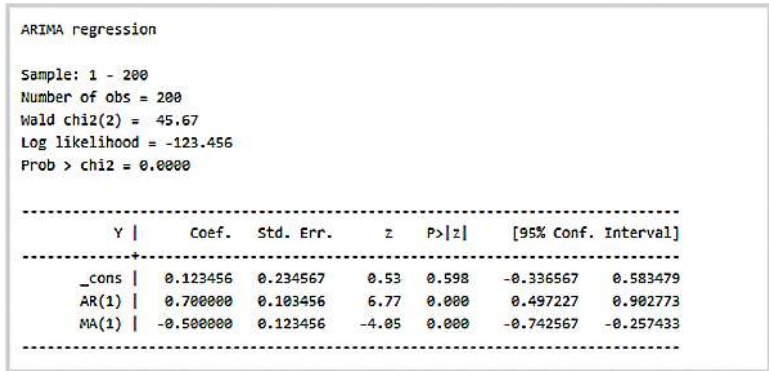


muestran las correlaciones entre los valores de la serie de tiempo en diferentes puntos en el tiempo. La identificación es esencial para seleccionar un modelo adecuado para la serie de tiempo que sea útil para la predicción.

Para una completa comprensión de los temas, le invito a revisar el siguiente **ejercicio** de estimación:

El comando `arima` ajusta el modelo ARIMA con los órdenes especificados. Aquí tienes un ejemplo de los resultados que podrías obtener:

Figura 13
Resultados de regresión ARIMA



Nota. Ochoa, W., 2025.

Diagnóstico

Los gráficos ACF y PACF de los residuos deben mostrar que no hay autocorrelación significativa. La prueba de Ljung-Box (`wntestq`) verifica la autocorrelación de los residuos. Un (p)-valor alto sugiere que los residuos se comportan como ruido blanco.

Pronóstico

El comando `predict` genera pronósticos basados en el modelo ajustado. El gráfico `tsline` muestra la serie original y los pronósticos.



Conclusión

La metodología Box-Jenkins proporciona un enfoque estructurado para modelar series temporales utilizando modelos ARIMA. Siguiendo estos pasos en Stata, puedes identificar, estimar, diagnosticar y pronosticar series temporales de manera efectiva.

Comandos para uso de software

1. Estructura temporal

*Establecer la estructura de tiempo

```
tsset time
```

2. Identificación

Utilizamos gráficos ACF y PACF para identificar los posibles órdenes del modelo ARIMA:

*Graficar ACF y PACF

```
acf Y
```

```
pacf Y
```

3. Estimación

Ajustamos el modelo ARIMA utilizando los órdenes identificados:

*Ajustar el modelo ARIMA(1,1,1)

```
arima Y, arima(1,1,1)
```

4. Diagnóstico

Verificamos los residuos del modelo para asegurarnos de que se comportan como ruido blanco:



*Graficar ACF y PACF de los residuos

`predict resid, resid`

`acf resid`

`pacf resid`

*Prueba de Ljung-Box para autocorrelación de residuos

`wntestq resid`

5. Pronóstico

Utilizamos el modelo ajustado para hacer pronósticos futuros:

*Hacer pronósticos para los próximos 10 períodos

`predict Y_forecast, dynamic(time[_N-10])`

`tsline Y Y_forecast`

Interpretación de los Resultados

1. Identificación:

- Los gráficos ACF y PACF ayudan a identificar los posibles órdenes (p) y (q). Si la serie no es estacionaria, se puede diferenciar (d).

2. Estimación:

- El comando `arima` ajusta el modelo ARIMA con los órdenes especificados.

3. Diagnóstico:

- Los gráficos ACF y PACF de los residuos deben mostrar que no hay autocorrelación significativa.



- La prueba de Ljung-Box (wntestq) verifica la autocorrelación de los residuos.

4. Pronóstico:

- El comando predict genera pronósticos basados en el modelo ajustado.
- El gráfico tslin muestra la serie original y los pronósticos.



Actividades de aprendizaje recomendadas

Le invito a participar en las siguientes actividades que de seguro contribuirán al fortalecimiento de sus conocimientos:

1. Analice el tema Pronósticos a través de la lectura correspondiente en el c y en la guía didáctica mencionada.

Procedimiento: realice una lectura comprensiva de los temas sugeridos y tome en cuenta las ideas principales para que posteriormente pueda con sus propias palabras elaborar un resumen de ello.

Wooldridge, J. M. (2015). Introducción a la econometría. Un enfoque moderno: un enfoque moderno. Ediciones Paraninfo, SA.

Stock, J. H., Watson, M. W., & Larrión, R. S. (2012). Introducción a la econometría.

Revisar también los siguientes videos para complementar las lecturas. En ambos videos se explica cómo aplicar de una manera fácil la metodología Box-Jenkins para desarrollar un modelo ARIMA en STATA.

- **Video 1:** [Modelos ARIMA y la metodología de Box-Jenkins en STATA | Series de tiempo.](#)
- **Video 2:** [Modelos ARIMA en Stata| Series de tiempo.](#)



2. Al igual que en todas las semanas anteriores, el profesor tutor le ubicará anuncios orientativos sobre la temática a desarrollarse en esta semana. Lea con atención los mismos para que también realice los comentarios que considere pertinentes o emita sus inquietudes.

Procedimiento: es recomendable que vaya realizando cuadros sinópticos o resúmenes que le ayuden ahora y posteriormente a tener una visión y comprensión clara del tema. Los recursos que usted utilice para el estudio siempre dependen de su estilo de aprendizaje.

3. Revise los ejercicios planteados en la bibliografía básica y desarrolle los correspondientes a los temas abordados.

Procedimiento: la comprensión de los temas es importante, pero se afianza mediante la aplicación de los mismos con la resolución de los ejercicios que se han planteado en la bibliografía básica. Consulte fuentes estadísticas como el Banco Mundial para un mejor desarrollo de las actividades prácticas.

Mediante la siguiente página *web*, puede obtener series de tiempo con las que podrá practicar lo estudiado: WDB (2020). [The World Bank Data](#).

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 16

Unidad 6. Econometría de series de tiempo

6.10. Vectores Autorregresivos (VAR)

Los Vectores Autorregresivos (VAR) son modelos estadísticos utilizados para analizar la dinámica conjunta de dos o más variables en una serie de tiempo. En un modelo VAR, cada variable se explica por sus valores anteriores y por los



valores de las otras variables en el sistema. En otras palabras, el modelo VAR permite modelar la relación entre variables en un sistema multivariante y determinar cómo las variables afectan entre sí.

El modelo VAR se basa en la idea de que cada variable en el sistema puede ser vista como una función de sus valores anteriores y de los valores de las otras variables en el sistema. La ventaja del modelo VAR es que puede manejar múltiples variables y proporcionar una representación más completa de la dinámica del sistema. El modelo VAR también permite analizar la causalidad entre variables, lo que puede ser útil para comprender cómo las variables interactúan en un sistema.

Una de las principales aplicaciones del modelo VAR es en el análisis macroeconómico. Por ejemplo, se puede utilizar para analizar la relación entre el PIB, la inflación, las tasas de interés y los tipos de cambio. El modelo VAR también se utiliza en la predicción de series de tiempo, ya que permite modelar la relación dinámica entre variables y pronosticar su comportamiento futuro.

La estimación del modelo VAR se realiza mediante el método de máxima verosimilitud. Esto implica la estimación de los coeficientes del modelo, que representan la relación entre las variables en el sistema. La evaluación del modelo VAR se realiza mediante la prueba de raíz unitaria, la prueba de cointegración y la prueba de causalidad de Granger.

Los Vectores Autorregresivos (VAR) son modelos estadísticos útiles para analizar la dinámica conjunta de dos o más variables en una serie de tiempo. El modelo VAR permite modelar la relación entre variables en un sistema multivariante y determinar cómo las variables afectan entre sí. La aplicación del modelo VAR es especialmente útil en el análisis macroeconómico y la predicción de series de tiempo. La estimación del modelo VAR se realiza mediante el método de máxima verosimilitud y la evaluación se realiza mediante diversas pruebas estadísticas.

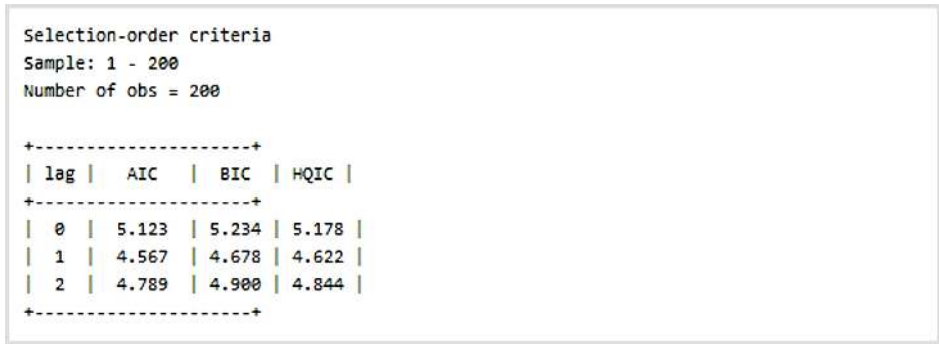
Ejemplo:

Selección del Número de Rezagos.



El comando varsoc te ayudará a seleccionar el número óptimo de rezagos utilizando criterios de información como AIC, BIC y HQIC. Aquí tienes un ejemplo de los resultados que podrías obtener:

Figura 14
Resultados VAR



Nota. Ochoa, W., 2025.

Ajuste del Modelo VAR

El comando var ajusta el modelo VAR con los órdenes especificados. Aquí tienes un ejemplo de los resultados que podrías obtener:



Figura 15

Resultados de ajustes

```
Vector autoregression
```

Sample: 1 - 200
No. of obs = 200
Log likelihood = -123.456
AIC = 5.123
HQIC = 5.178
SBIC = 5.234

Equation Y1

Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
Y1_L1	0.567890	0.1234	4.60	0.000	0.325678	0.810123
Y2_L1	0.678901	0.1345	5.05	0.000	0.413456	0.944345
_cons	0.123456	0.0456	2.71	0.007	0.034567	0.212345

Equation Y2

Coef.	Std. Err.	z	P> z	[95% Conf. Interval]		
Y1_L1	0.789012	0.1567	5.03	0.000	0.481234	1.096789
Y2_L1	0.890123	0.1678	5.31	0.000	0.561234	1.219012
_cons	0.234567	0.0567	4.14	0.000	0.123456	0.345678

Nota. Ochoa, W., 2025.

Diagnóstico

El comando `varstable` verifica la estabilidad del modelo, y `varlmar` realiza la prueba de autocorrelación de residuos. Un modelo estable y sin autocorrelación en los residuos es adecuado para pronósticos.

Pronóstico

El comando `fcast` genera pronósticos basados en el modelo ajustado. El gráfico `fcast graph` muestra las series originales y los pronósticos.

Conclusión

El modelo VAR ajustado muestra las relaciones dinámicas entre las series (Y1) y (Y2). Este tipo de modelo es útil para capturar las interacciones entre múltiples series temporales y proporcionar pronósticos precisos.

Comandos para uso de software

1. Establecer la estructura del tiempo

*Establecer la estructura de tiempo

tsset time

2. Selección del Número de Rezagos

Utilizamos criterios de información para seleccionar el número óptimo de rezagos:

*Seleccionar el número óptimo de rezagos

varsoc Y1 Y2

3. Ajuste del Modelo VAR

Ajustamos el modelo VAR utilizando el comando var:

*Ajustar el modelo VAR con el número óptimo de rezagos (por ejemplo, 1)

var Y1 Y2, lags(1)

*Mostrar los resultados

varstable, graph

4. Diagnóstico

Verificamos la estabilidad del modelo y la autocorrelación de los residuos:

*Verificar la estabilidad del modelo

varstable

*Prueba de autocorrelación de residuos



varlmar

5. Pronóstico

Utilizamos el modelo ajustado para hacer pronósticos futuros:

*Hacer pronósticos para los próximos 10 períodos

fcast compute Y1_f Y2_f, step(10)

fcast graph



Actividades de aprendizaje recomendadas

Le invito a participar en las siguientes actividades que de seguro contribuirán al fortalecimiento de sus conocimientos:

1. Analice las ideas más relevantes acerca de los modelos de vectores autorregresivos, a través de la lectura correspondiente.

Procedimiento: realice una lectura comprensiva de los temas sugeridos y tome en cuenta las ideas principales para que posteriormente pueda con sus propias palabras elaborar un resumen de ello.

2. Al igual que en todas las semanas anteriores, el profesor tutor le ubicará anuncios orientativos sobre la temática a desarrollarse en esta semana. Lea con atención los mismos para que también realice los comentarios que considere pertinentes o emita sus inquietudes.

Procedimiento: es recomendable que vaya realizando cuadros sinópticos o resúmenes que le ayuden ahora y posteriormente a tener una visión y comprensión clara del tema. Los recursos que usted utilice para el estudio siempre dependen de su estilo de aprendizaje.



Con los datos obtenidos en la siguiente página web, puede llevar a la práctica todo lo correspondiente a vectores autorregresivos: WDB (202020). [The World Bank Data](#).

En el siguiente video de 10 minutos de duración se muestra cómo estimar un [Modelo VAR de forma reducida en STATA](#) aplicando 3 pasos básicos, incluyendo las variables precio de petróleo, precio del café, tasa de crecimiento del PIB y tasa de desempleo, extraídas del Banco Central de Nicaragua y la respectiva interpretación de los resultados obtenidos.

Actividades finales del bimestre

3. Revise todos los contenidos desarrollados en el bimestre I y prepare su evaluación presencial bimestral.

Procedimiento: póngase en contacto con su docente tutor para resolver todas las dudas e inquietudes pendientes.

4. Lea las orientaciones generales que le brindará el profesor a través de los anuncios en el entorno virtual de aprendizaje y exprese en los horarios de tutoría sus inquietudes acerca de los temas revisados.
5. Le invito a reforzar sus conocimientos, participando en la siguiente autoevaluación.



[Autoevaluación 7](#)

1. ¿Qué indica la cointegración entre dos variables?
 - a. Que las variables están altamente correlacionadas.
 - b. Que las variables están causando una relación espuria.
 - c. Que las variables tienen una relación de equilibrio a largo plazo.
2. ¿Qué es el Mecanismo de Corrección de Errores (MCE)?
 - a. Es un modelo de regresión que relaciona dos o más variables.
 - b. Es un modelo que permite la predicción de una serie de tiempo.



- c. Es un modelo que ajusta la relación de equilibrio a largo plazo entre dos variables después de un desequilibrio a corto plazo.
3. ¿Cómo se puede probar la existencia de cointegración entre dos variables?
- a. A través del análisis de gráficos de las variables.
 - b. A través de pruebas de raíz unitaria.
 - c. A través del análisis de la función de autocorrelación.
4. ¿Qué significa el término “AR” en un modelo de series de tiempo?
- a. Autorregresión.
 - b. Medias móviles.
 - c. Integración.
5. ¿Cuál es el propósito del modelo de Medias Móviles (PM)?
- a. Modelar la tendencia de una serie de tiempo.
 - b. Modelar la estacionalidad de una serie de tiempo.
 - c. Modelar la componente de error de una serie de tiempo.
6. ¿Cuál es la diferencia entre un modelo ARIMA y un modelo ARMA?
- a. Un modelo ARIMA incluye una componente de integración, mientras que un modelo ARMA no lo hace.
 - b. Un modelo ARIMA incluye una componente de medias móviles, mientras que un modelo ARMA no lo hace.
 - c. Un modelo ARIMA incluye tanto una componente de integración como una de medias móviles, mientras que un modelo ARMA solo incluye una de estas.
7. ¿Qué es un Vector Autorregresivo (VAR)?
- a. Es un modelo que describe la relación entre una serie de tiempo y una o más variables predictoras.



- b. Es un modelo que describe la relación entre dos o más series de tiempo.
 - c. Es un modelo que describe la relación entre una serie de tiempo y sus valores pasados.
8. ¿Cuál es el propósito de la Función de Autocorrelación Parcial (PACF) en la identificación de un modelo ARIMA?
- a. Identificar la componente de medias móviles del modelo.
 - b. Identificar la componente de integración del modelo.
 - c. Identificar la componente de autorregresión del modelo.
9. ¿Qué es el propósito de la Función de Autocorrelación (ACF) en la identificación de un modelo ARIMA?
- a. Identificar la componente de medias móviles del modelo.
 - b. Identificar la componente de integración del modelo.
 - c. Identificar la componente de autorregresión del modelo.
10. ¿Cuál es la primera etapa de la metodología Box-Jenkins para la creación de modelos ARIMA?
- a. Identificación.
 - b. Estimación.
 - c. Validación.
11. ¿Cómo seleccionar el número de rezagos adecuados para un vector autorregresivo? Seleccionar las correctas:
- a. Análisis de autocorrelación.
 - b. Análisis de estabilidad.
 - c. Criterios de información.
 - d. Test de White.
 - e. Teoría económica.
 - f. R cuadrado.



[Ir al solucionario](#)





4. Autoevaluaciones

Autoevaluación 1

Pregunta	Respuesta	Retroalimentación
1	c	Significa que la varianza no es homocedástica.
2	b	No existe homocedasticidad.
3	d	La heterocedasticidad puede producir sesgo en los coeficientes estimados, puede producir una subestimación de la varianza del error, puede producir una sobreestimación de la varianza del error.
4	b	La heterocedasticidad se puede detectar con el test de Prueba de Breusch-Pagan.
5	d	La heterocedasticidad se puede corregir utilizando un modelo de regresión ponderada.
6	d	Un modelo de regresión ponderada es un modelo en el que se asignan pesos a los datos en función de su varianza.
7	c	Una matriz que muestra la varianza y la covarianza de los errores y sirve para detectar heterocedasticidad.
8	a	El test de White es una de la técnica más usada para detectar la presencia de heterocedasticidad.
9	c	El análisis de residuos es una de las técnicas más usadas.
10	b	La varianza de los errores aumenta o disminuye a medida que la variable independiente cambia.

[Ir a la autoevaluación](#)



Autoevaluación 2

Pregunta	Respuesta	Retroalimentación
1	b	La autocorrelación entre una variable y su valor en un momento anterior.
2	c	El test de Durbin-Watson es uno de los más utilizados para detectar autocorrelación.
3	b	El test de Durbin-Watson si da valores de 0 o 4 significa que hay autocorrelación, de 1.5 a 2.5 no se puede asegurar que exista autocorrelación.
4	b	El test de Durbin-Watson si da valores de 0 o 4 significa que hay autocorrelación, de 1.5 a 2.5 no se puede asegurar que exista autocorrelación.
5	b	La principal consecuencia de la autocorrelación en un modelo es la sobreestimación de los errores estándar.
6	a	Una de las formas de corregir la autocorrelación en un modelo de regresión lineal.
7	b	El método de Correlación Serial de Cochrane-Orcutt que utiliza variables de rezago como regresores.
8	a	El efecto de la corrección de la autocorrelación es la disminución de los errores estándar.
9	b	El método de Mínimos Cuadrados Generalizados (GLS) se utiliza para corregir la autocorrelación.
10	c	El modelo de medidas repetidas se utiliza para corregir la autocorrelación en un diseño experimental con mediciones repetidas.

[Ir a la autoevaluación](#)



Autoevaluación 3

Pregunta	Respuesta	Retroalimentación
1	c	El error que se produce cuando se omite una variable relevante del modelo.
2	a	Pueden ser la prueba F, la de Ramsey entre otras.
3	a	La prueba de Ramsey RESET se utiliza para evaluar si el modelo tiene un error de especificación de forma funcional.
4	b	La prueba de Breusch-Pagan se utiliza para evaluar si hay heterocedasticidad en los errores del modelo.
5	c	La presencia de cambios en la varianza de los errores en función de los valores de las variables independientes.
6	c	La prueba de Durbin-Watson se utiliza para evaluar la presencia de autocorrelación en los errores del modelo.
7	a	La autocorrelación en los errores del modelo señala correlación serial en los errores.
8	c	La prueba de Jarque-Bera se utiliza para evaluar la normalidad de los errores del modelo.
9	c	La normalidad de los errores señala la presencia de errores que siguen una distribución normal.
10	a	Agregar una variable al modelo es la medida correctiva más común para el error de especificación en la selección de modelos econométricos.

[Ir a la autoevaluación](#)



Autoevaluación 4

Pregunta	Respuesta	Retroalimentación
1	c	Utilizar una matriz de ponderación robusta para estimar los coeficientes es una forma efectiva de remediar la heterocedasticidad.
2	a	Para remediar la heterocedasticidad en un modelo econométrico, cuando la causa es una variable omitida, podría incluir la variable omitida en el modelo para solucionarlo.
3	b	Utilizar una matriz de ponderación robusta para estimar los coeficientes puede solucionar la heterocedasticidad.
4	c	Utilizar un modelo de corrección de errores puede solucionar la autocorrelación.
5	a	Analizar los residuos del modelo puede ayudar a determinar si existe autocorrelación.
6	b	La especificación incorrecta del modelo es una causa común de autocorrelación.
7	c	Para remediar el error de especificación se puede utilizar un modelo de variables instrumentales.
8	b	Analizar los residuos del modelo también se puede utilizar para determinar si utilizar para detectar el error de especificación.
9	c	La omisión de variables importantes es una causa común de error de especificación.
10	c	Utilizar técnicas de muestreo para seleccionar una muestra representativa.
11	a-4 b-3 c-2 d-1	a-4 Si un modelo tiene autocorrelación no se puede hacer pronósticos b-3 Si existe Heterocedasticidad los errores son más grandes y podría aceptarse valores erróneos c-2 La multicolinealidad puede crear falsas relaciones d-4 El error de especificación puede ocasionar que los coeficientes estimados sean sesgados y las predicciones pueden ser poco precisas.

Ir a la autoevaluación



Autoevaluación 5

Pregunta	Respuesta	Retroalimentación
1	a	Un modelo logístico es adecuado para describir una curva en forma de "S".
2	b	Un modelo exponencial se utiliza para ajustar datos que tienen un comportamiento exponencial.
3	c	Un modelo polinómico de grado 2 se utiliza para ajustar datos que tienen una relación de curva cóncava.
4	a	Un modelo logístico también se utiliza para ajustar datos que tienen un comportamiento de crecimiento y luego se estabilizan.
5	b	Los MCO también se utiliza para ajustar los parámetros del modelo de regresión no lineal.
6	a	R cuadrado se utiliza para evaluar la calidad de ajuste del modelo de regresión no lineal.
7	a	El objetivo de la estimación en la regresión no lineal es estimar los valores de los parámetros del modelo.
8	a	El residuo en la regresión no lineal es la diferencia entre los valores observados y los valores estimados.
9	c	R-cuadrado se utiliza para evaluar la bondad de ajuste del modelo de regresión no lineal.
10	a	Sobreajuste es el modelo que se ajusta demasiado a los datos de entrenamiento.

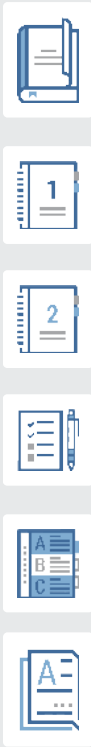
Ir a la autoevaluación



Autoevaluación 6

Pregunta	Respuesta	Retroalimentación
1	a	Un modelo econométrico dinámico incluye variables dependientes rezagadas.
2	b	Un modelo Autorregresivo (AR) incluye una variable dependiente rezagada.
3	d	Un modelo de Rezagos Distribuidos (VAR) incluye varias variables dependientes y sus rezagos.
4	a	Mínimos Cuadrados Ordinarios (OLS o MCO) se utiliza comúnmente para estimar los parámetros en un modelo autorregresivo.
5	b	La Máxima verosimilitud se utiliza comúnmente para estimar los parámetros en un modelo de rezagos distribuidos.
6	b	El coeficiente de autocorrelación es la medida de la correlación entre una variable dependiente y su rezago.
7	a	El Criterio de Akaike (AIC) es un criterio de selección de modelos que penaliza el número de parámetros del modelo.
8	c	La causalidad de Granger es una prueba estadística para evaluar la causalidad entre dos variables a partir de su historial temporal.
9	b	El resultado de la prueba de causalidad de Granger. Si el valor p es menor que el nivel de significancia, se rechaza la hipótesis nula de no causalidad.
10	a	Los modelos autorregresivos son capaces de capturar relaciones de dependencia temporal entre las variables, mientras que los modelos dinámicos se enfocan en capturar las interacciones entre múltiples variables en un sistema.

[Ir a la autoevaluación](#)



Autoevaluación 7

Pregunta	Respuesta	Retroalimentación
1	c	La cointegración entre dos variables indica que las variables tienen una relación de equilibrio a largo plazo.
2	c	El Mecanismo de Corrección de Errores (MCE) es un modelo que ajusta la relación de equilibrio a largo plazo entre dos variables después de un desequilibrio a corto plazo.
3	b	Probar la existencia de cointegración entre dos variables se puede realizar a través de pruebas de raíz unitaria.
4	a	El término "AR" significa Autorregresión.
5	c	El propósito del modelo de Medias Móviles (PM) es modelar la componente de error de una serie de tiempo.
6	a	Un modelo ARIMA incluye una componente de integración, mientras que un modelo ARMA no lo hace.
7	b	Un vector autorregresivo es un modelo que describe la relación entre dos o más series de tiempo.
8	c	El propósito de la Función de Autocorrelación Parcial (PACF) en la identificación de un modelo ARIMA es Identificar la componente de autorregresión del modelo.
9	a	De la Función de Autocorrelación (ACF) sirve identificar la componente de medias móviles del modelo.
10	a	Identificación es la primera etapa de la metodología Box-Jenkins para la creación de modelos ARIMA.
11	a, b, c, e	Un modelo vector autorregresivo (VAR) es un enfoque estadístico que se utiliza para analizar y pronosticar múltiples series temporales interdependientes al modelar las relaciones autorregresivas y de interacción entre las variables en el sistema.

[Ir a la autoevaluación](#)





5. Glosario

Homocedasticidad: supuesto que establece que la varianza de los errores es constante para todos los valores de la variable independiente.

Heterocedasticidad: es la violación del supuesto de homocedasticidad, lo que significa que la varianza de los errores no es constante a lo largo de todas las observaciones.

Autocorrelación: es la correlación entre los errores sucesivos, lo que indica que un error en un momento dado está relacionado con los errores en los momentos anteriores.

Multicolinealidad: se refiere a la existencia de una alta correlación entre dos o más variables independientes, lo que dificulta la estimación de los coeficientes.

Endogeneidad: es la situación en la que una variable independiente está correlacionada con los errores, lo que puede sesgar los resultados de la regresión.

Sesgo de selección: ocurre cuando la muestra no es representativa de la población, lo que puede afectar la precisión de los resultados.

Sesgo de omisión: se produce cuando falta una variable relevante en la regresión, lo que puede provocar que los coeficientes estimados no sean precisos.

Sesgo de medida: ocurre cuando las variables no se miden correctamente o no se miden de manera precisa, lo que puede afectar los resultados.

No normalidad: supuesto que establece que los errores tienen una distribución normal, pero en algunos casos puede ser violado, lo que puede afectar la precisión de los resultados.



Regresión espuria: es una situación en la que dos variables aparentemente relacionadas están correlacionadas debido a factores externos, pero en realidad no tienen una relación causal.

Econometría: es una rama de la economía que combina la estadística y las matemáticas para analizar y predecir el comportamiento de los datos económicos.

Series temporales: son conjuntos de datos organizados cronológicamente que se utilizan para analizar el comportamiento de una variable a lo largo del tiempo.

Primeras diferencias: es una técnica utilizada en econometría de series temporales para convertir una serie no estacionaria en una serie estacionaria, restando cada observación de la observación anterior.

Cointegración: es una técnica utilizada en econometría de series temporales para analizar la relación a largo plazo entre dos o más variables que están integradas en orden uno (es decir, tienen una tendencia).

Estacionariedad: se refiere a la propiedad de una serie temporal que muestra un comportamiento constante en el tiempo, sin una tendencia o patrón sistemático.

Autocorrelación: es la correlación entre los valores de una misma serie temporal en diferentes momentos del tiempo. Si los valores están correlacionados positivamente, significa que un valor alto en un momento del tiempo tiende a estar asociado con valores altos en momentos posteriores.

Modelo ARIMA: es un modelo estadístico utilizado en la predicción de series temporales, que combina un modelo Autorregresivo (AR) y un modelo de Media Móvil (MA) con la incorporación de una componente de Integración (I).

Test de Dickey-Fuller: es una técnica utilizada para determinar si una serie temporal es estacionaria o no estacionaria. El test compara el valor del estadístico de prueba con un valor crítico para determinar si la serie es estacionaria o no.



Vector de Corrección de Error (VEC): es un modelo multivariado utilizado en econometría de series temporales para analizar la relación a largo plazo entre varias variables que están cointegradas.

Método de Mínimos Cuadrados Ordinarios (MCO): es un método utilizado para estimar los parámetros de un modelo estadístico, minimizando la suma de los cuadrados de las diferencias entre los valores observados y los valores predichos por el modelo. Es comúnmente utilizado en la econometría de series temporales para estimar los parámetros de modelos ARIMA y VEC.

Identificación: en la metodología de Box-Jenkins, la identificación se refiere al proceso de identificar la estructura de modelo adecuada para una serie temporal en particular. Se realiza mediante el análisis de la autocorrelación y la autocorrelación parcial de la serie temporal.

Estimación: la estimación en la metodología de Box-Jenkins se refiere al proceso de estimar los parámetros del modelo elegido en la etapa de identificación. Esto implica utilizar técnicas de estimación como el método de mínimos cuadrados o el método de máxima verosimilitud.

Diagnóstico: después de identificar y estimar el modelo, el siguiente paso en la metodología de Box-Jenkins es el diagnóstico. El diagnóstico se utiliza para evaluar la calidad del modelo estimado, y se realizan pruebas estadísticas para evaluar si el modelo cumple con los supuestos necesarios.

Validación: la validación en la metodología de Box-Jenkins se refiere al proceso de verificar si el modelo seleccionado puede proporcionar buenas predicciones para la serie temporal. Esto se hace comparando las predicciones del modelo con los valores observados de la serie temporal.

Pronóstico: en la metodología de Box-Jenkins, el pronóstico se refiere a la predicción de los valores futuros de una serie temporal utilizando el modelo seleccionado y validado. Esto implica utilizar el modelo para hacer predicciones a partir de un punto futuro conocido y luego evaluar la precisión de esas predicciones en función de los valores observados.





6. Referencias bibliográficas

- Wooldridge, J. M. (2015). Introducción a la econometría. Un enfoque moderno: un enfoque moderno. Ediciones Paraninfo, SA.
- Stock, J. H., Watson, M. W., & Larrión, R. S. (2012). Introducción a la Econometría.
- Lozano, 2021. Guía didáctica Modelos econométricos. UTPL
- Toriz-Robles, N., Ramírez-Guzmán, M. E., Fernández-Ordoñez, Y. M., Soria- Ruiz, J., & Ybarra-Moncada, M. C. (2019). Comparación de modelos lineales y no lineales para estimar el riesgo de contaminación de suelos. *Agrociencia*, 53(2), 269-283.
- Charemza, W. y Deadman, D., (1997): New directions in econometric practice, *Books*.

