UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE ECONOMÍA

ECONOMETRÍA Y SERIES DE TIEMPO, ELEMENTOS DE JUICIO PARA PRONÓSTICO

TESINA

QUE PARA OBTENER EL TITULO DE:
LICENCIADO EN ECONOMÍA
P R E S E N T A:
JOSÉ ALBERTO REYES DE LA ROSA

DIRECTORA DE TESINA: M.I. GENOVEVA BARRERA GODINEZ

MÉXICO D.F. JUNIO 2006

AGRADECIMIENTOS

Es para mí motivo de gran satisfacción personal como profesional, agradecer a todos y cada uno de los personajes que considero participaron de manera directa e indirecta en el logro de la conclusión de este trabajo.

A mi Universidad UNAM mi máxima casa de estudios institución que me brindó la oportunidad de liberar una profesión.

A mis padres
Josefina y Crescencio. Con mi eterno reconocimiento, por haber sido los mejores maestros de economía y de la vida que jamás halla tenido, quien siempre están y estarán presentes por todas sus enseñanzas, las cuales me han permitido superar todos los obstáculos que se me han presentado y por haberme dado su mayor riqueza, su amor.

A mis hermanos
María Eugenia, Enrique, Jorge, Alma Delia, Marco Antonio, Elizabeth, por el apoyo y la confianza que depositaron en mí.

A mis sobrinos
Flavio Felipe, María Teresa, Jesús Alejandro, María del Carmen, Diana Laura, Rita Isabel, Alicia, Ana, Jorge, Daniel, Nallely, para que pronto lleguen a esta etapa.

A mis amigos
María del Carmen Ramírez Rubio mi admiración y respeto por ser madre y padre a la vez de dos pequeños que pronto le darán grandes satisfacciones (Hugo y Ernesto)
Patricio Sixtos Coss por ser de una sola pieza y ser leal a sus ideales.
Silvia Mancia por ser una gran amiga y por su gran valor al ganarle la batalla al cáncer.
Eduardo Hayashi mi primer profesor de computo, quien me permitió ser su ayudante y aprender.
Patricia Martínez siempre tiene una gran disposición de apoyarnos en todo lo que se necesite gracias PATY.
Ofelia Moreno, gracias por tan buena persona conmigo.
Inés Flores, como olvidar el apoyo que siempre me brindas gracias, y solo pocas personas tiene la dicha de ayudar a traer una vida a este mundo felicidades abuelita partera.

A mis compañeros de clase
Esther, Alejandra Ballesteros, Gerardo (caches), Francisco (Paco), Horacio, Humberto, Emma, Lucio Cañas.

A mis nuevos compañero y amigos
María del Socorro Pérez, Ángel Martínez, Elizabeth Martínez, Sinai, Joel Rojas, Carlos Omar Álvarez, Armando León., Cynthia González, Judit, Alan Fabrizzio,.Manuel Haro Zepeda, Alejandro Zamora.

A quien siempre estuvo al pendiente para la realización este trabajo
Sra. Carmen Guerrero, Sr. Evencio Peregrina, Ekhaterina y Paloma.

A Beatriz Miroslava, Rosario Gutierrez, Esperanza Rojas, Verónica Leticia Hernández, Perla Vega, Adriana Cruz, Karina Álvarez, Kathia Pérez.

A mis profesores

- Luciano Lara Flores, quien me enseñó que la economía no es solo cuestión de charlar y arreglar el mundo en un café, sino hay que trabajar para formarse como economista y ser mejor cada día y así contribuir en la transformación de nuestro país, cito su lema que en clase nos decía “El trabajo del hombre no solo ha transformado a la naturaleza y a la sociedad, sino al hombre mismo”, se refería a que cada día nos presenta la oportunidad de aprender y de enseñar y de ser mejores siempre.
- Enrique López Santiago, quien siempre tuvo la disponibilidad de orientarme desde que llegue a esta facultad hasta la fecha, gracias por el apoyo incondicional.
- Dra. María Antonieta Barrón, mi agradecimiento por dejarme trabajar en sus proyecto y aprender que la realidad dista mucho de lo que se hace en un escritorio, con esa exigencia que le pide a cada uno de sus colaboradores entendía que las cosas deben hacerse bien y a la primera, gracias doctora.
- Lic. Alfonso Gómez Navarro, a quien recuerdo que en sus clases siempre nos decía que el conocimiento no es gratis, que para acceder a él se debe trabajar desde muy temprano si es que se quiere convivir y ser amigo del mismo.
- Mtra. Alejandra Patiño, siempre brindando apoyo y mi profesora del módulo II del diplomado de estadística, gracias Ale que también a igual que el profesor Valbuena se preocupaban para que este trabajo ya se concluyera.
- Mtra. Hortensia Martínez, que me dio la oportunidad de estar en clase como su ayudante.

AGRADECIMIENTOS

Dr. Genarao Sánchez Barajas. Un gran líder y muy perseverante, gracias doctor por el apoyo brindado durante muchos años, tanto en el aula de clase como en el compartir sus conocimientos.

Lic. Rubén Valbuena Álvarez. Por el apoyo otorgado y que siempre se preocupo que este día llegará, agradezco la oportunidad que me dio de ser profesor titular y sobre todo el de aprender a lado de un “viejo lobo de mar”.

MI. Genoveva Barrrera Godinez. A la maestra con cariño, que desde que fuimos compañeros en el diplomado de econometría se gesto esta ilusión de elaborar un trabajo que sirviera de apoyo a los alumnos, hoy se termina una fase del mismo. Agradezco la confianza que me ha brindado para ser su ayudante durante muchos años y sobre todo el aprender en cada clase algo nuevo.

Lic. Arturo Mérida Monroy. Gracias por la confianza que me tuviste, sin conocer del todo el trabajo me dijiste que llenaba las expectativas para poder ser utilizado como apoyo para los alumnos, gracias por ese gesto.

Lic. Rocío Cruz Sánchez. Una de mis grandes amigas y confidentes, gracias por preocuparte para que este trabajo saliera lo antes posible.

ÍNDICE

INTRODUCCIÓN

Entre las técnicas de pronóstico cuantitativo, las de mayor aplicación son las técnicas basadas en el análisis de regresión. Ello se debe a una serie de factores tales como su facilidad de aplicación, su respaldo estadístico y la confianza que genera el método, ya que los principales resultados son familiares para cualquier persona que cuenta con conocimientos básicos de inferencia estadística.

Pero ante todo, ha influido que, por una parte, las técnicas de regresión permiten tratar muchos casos que están fuera del alcance de las técnicas de series de tiempo y, por otra parte, que los responsables sientan que sus pronósticos están mejor fundados, por la manera que se construyen los modelos.

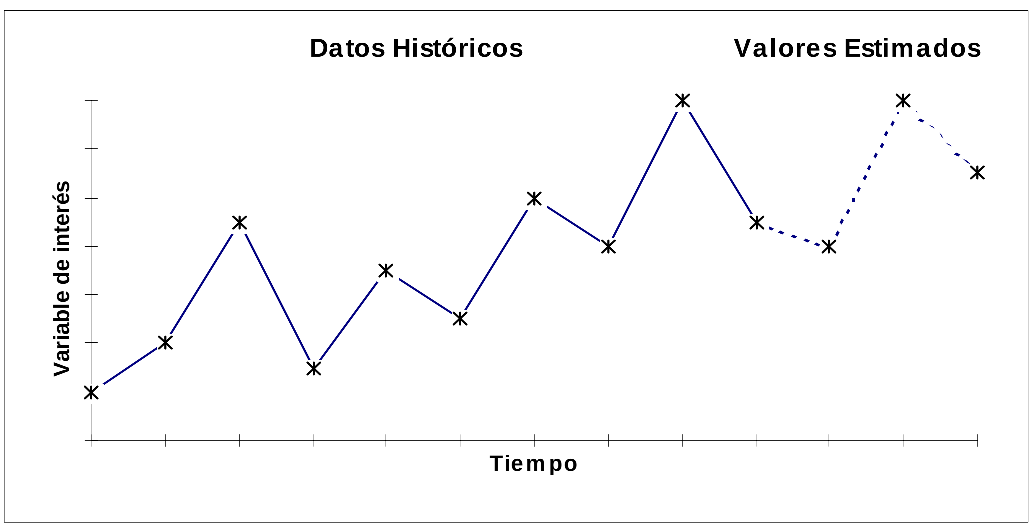
En las series de tiempo se busca conocer el valor de las variables de interés hurgando en su pasado, para de esta manera identificar el patrón de comportamiento y generar un modelo de la forma

$$Y_t = f(x_t, x_{t-1}, x_{t-2}, \ldots, x_{t-n}) \tag{1}$$

donde ***f*** es una función que permite calcular el valor futuro de $\boldsymbol{Y_t}$ a partir de un conjunto de datos históricos, tomando en cuenta factores de variación cíclica, tendencias, estacionalidad, autocorrelaciones, etc.

De manera implícita se asume que el conocimiento del futuro está en el pasado, pues la historia se repite o bien, sigue un curso regular en el tiempo (gráfica l), a esta clase de técnicas se les conoce como método de extrapolación.

Gráfica 1
Pronóstico por medio de una serie de tiempo



Fuente: Datos hipotéticos con fines ilustrativos.

En el análisis de regresión se busca conocer el futuro ***Y*** a partir de otras variables relacionadas ***X's***, haciendo uso de un modelo que expresa dicha interdependencia funcional

$$Y = f(x_1, x_2, x_3, \ldots, x_n) \tag{2}$$

donde:

Y es la variable ha estimar, conocida como variable dependiente, endógena o explicada.
X_i es la i-ésima variable relacionada, conocida como independiente, exógena o explicativa.
f( ) es la función que indica la relación que guarda ***Y*** con las ***X's***.

Así, por ejemplo, podemos estimar la demanda futura de energía a partir de los elementos que configuran el consumo: (X_1) tamaño de población; (X_2) producción industrial; (X_3) número de establecimientos comerciales, etc.

El papel de las técnicas de regresión consiste en identificar cuál es la función lineal o de otro tipo (cuadrática, exponencial, logarítmica etc.) que mejor representa a un conjunto de datos, como estimar el valor de los parámetros asociados. De hecho, es parte del objetivo de este trabajo, así como la corrección de los problemas que presentan las ecuaciones que se estiman, se presenta el sustento teórico-estadístico y después se elabora un ejercicio en el que se aplican las técnicas y métodos para continuar con la estimación de las ecuaciones o modelos, esto es posible realizar de una manera más rápida gracias a que existe software que nos facilita los cálculos; por tanto el uso de programas especializados en estadística y econometría (Excel, Econometric Views y SPSS) serán la herramienta para desarrollar los ejercicios.

Para lograr esto y con la idea de introducir los conceptos básicos, se inicia (capítulo 1) con la explicación de, qué es un problema de regresión lineal simple, la dependencia estadística y casual que existen entre los datos, el ajuste y correlación que se presentan en la elaboración de una regresión con el método de mínimos cuadrados ordinarios, se revisan los supuestos del modelo de regresión lineal clásico así como las propiedades de los estimadores, posteriormente se describe cómo se elabora un intervalo de confianza y cómo probar las hipótesis de los coeficientes, finalmente se estudia el diagrama de dispersión que nos permite ver el comportamiento que tendrán los datos para concluir con el modelo de regresión en forma matricial.

En el capítulo 2 se desarrolla el modelo de regresión múltiple con sus correspondientes conceptos generalizados en capítulo l, se explica en que consiste la violación de los supuestos del modelo de regresión lineal clásico, ¿cómo detectar que presentan un problema?, él ¿cómo se corrige? y lo más importante el desarrollo de un ejercicio en que se desglosa paso a paso la detección y la corrección, para después continuar con la estimación, que será la interpretación de los resultados obtenidos.

En el capítulo 3 se ven los componentes de las series de tiempo y algunos métodos de estimación, que nos concedan las herramientas para decidir si la ecuación estimada es la correcta o no, y de esta manera aplicar las metodologías necesarias para realizar un análisis exahustivo que nos de el mejor modelo. Una vez conocido el mejor modelo se aplican los métodos (Box-Jenkins, Autorregresivos) de pronóstico que nos permitan conocer o tener una idea del comportamiento de algunas variables que sea de nuestro interés el estudiarlas.

En los anexos se incluirán los datos que se utilizan a partir del capítulo dos, por si se desea realizar el ejercicio y comparar los resultados.

Es necesario comentar que el trabajo no pretende ser un manual o un libro de cómputo, solo se verá lo necesario para que los usuarios de este software puedan familiarizarse y desarrollar sus ejercicios de econometría de la mejor manera posible.

Se pretende que con el presente trabajo los alumnos interesados en estos temas, puedan realizar sus trabajos y/o practicas con fluidez y sobre todo entender los métodos empleados para la estimación así como la utilización del software que se emplea.

1 CAPITULO

1. REGRESIÓN LINEAL SIMPLE

1.1. Definición de regresión

La palabra regresión significa "acción de volver hacia atrás"[1]. Fue Francis Galton (1822-1911) quien dio forma a la expresión eugenesia, que es la aplicación de las leyes biológicas de la herencia al perfeccionamiento de la especie humana; "halló que la estatura de los hijos de padres altos, no seria tan altos como sus padres y lo mismo ocurría con padres de estatura baja que sus hijos no serian tan bajos, esto es que el comportamiento de los hijos tanto de padres altos como bajos, tendía a comportarse o a regresar a la estatura promedio de la población".

Desde el punto de vista geométrico, una curva de regresión es simplemente el lugar geométrico de las medias condicionales o expectativas de la variable dependiente para los valores fijos de las variables explicatorias.

Hay que recordar que el objetivo de la investigación es la especificación de una relación funcional entre dos variables como los ***Y= f(x)***, la ***Y*** o variable dependiente esta en función de ***X*** que es la variable independiente; no es posible esperar una explicación perfecta, entre ambas, y por ello se rescribe ***Y= f(x)+*** μ_i, donde μ_i es una variable aleatoria que recibe el nombre de perturbación estocástica.

El término de perturbación estocástica denotado por μ_i reemplaza a todas aquellas variables que han sido excluidas del modelo, pero que conjuntamente afectan a ***Y***.

1.2. Aspectos generales

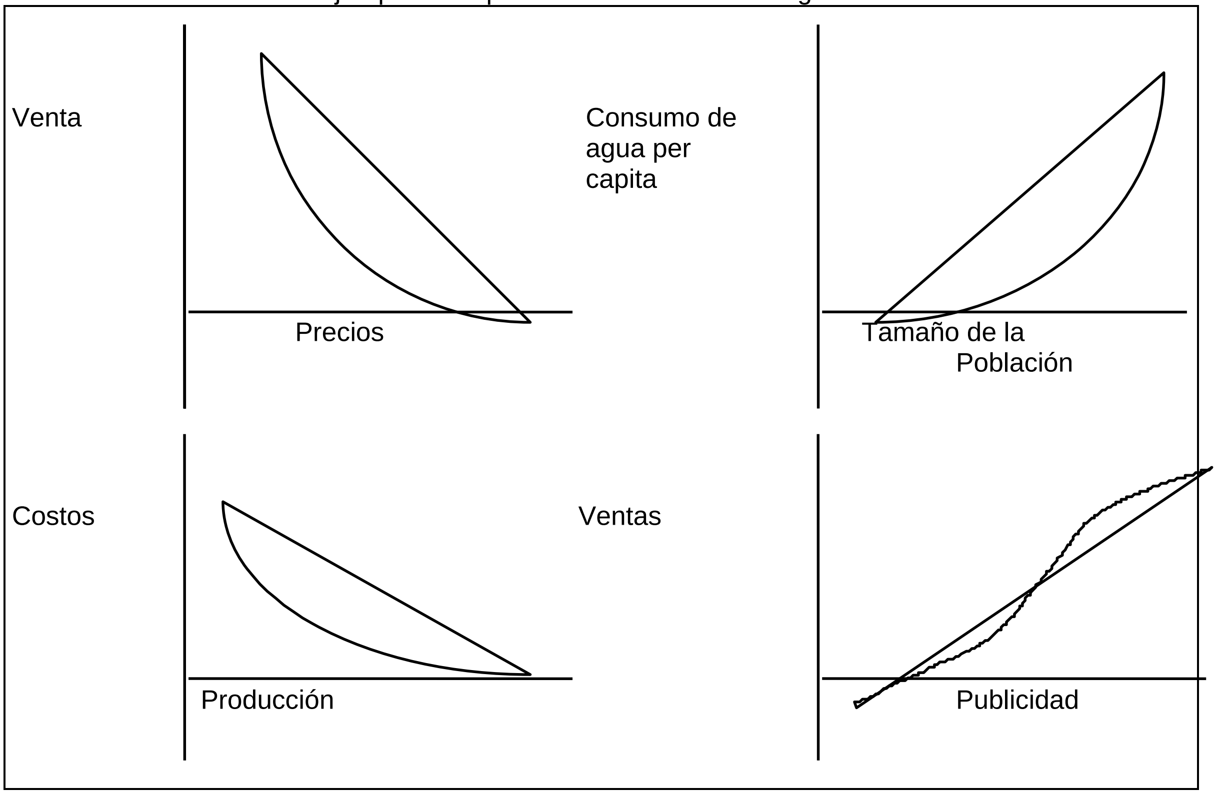
El análisis de regresión es una técnica estadística para investigar y modelar la relación funcional entre variables, siendo numerosas las aplicaciones en el pronóstico.

La idea básica consiste en identificar cuál es la curva que mejor se ajusta a un conjunto de ***N*** pares de datos (X_i, Y_i) y con ello se establece una ecuación que permita estimar el valor de ***Y*** dado que se conoce el valor de la variable independiente (***X***).

Por ejemplo, se puede estimar el consumo de gasolina dado un incremento en los precios, el consumo de agua per capita en función del tamaño de una población, los costos esperados según el nivel de producción, las ventas futuras de acuerdo al esfuerzo publicitario, (figura 1) etc. Desde luego, suponiendo que se cuenta con la información necesaria para cada caso.

[1] Enciclopedia Hispánica 1998, Volumen 7 pág. 2.

FIGURA 1
Ejemplos de aplicación del análisis de regresión



Los datos pueden provenir de observaciones de ***X*** y ***Y*** a lo largo del tiempo (anual, mensual, semanal, diario, etc.), sin que exista el requisito de que sean secuenciales; también pueden corresponder a observaciones sobre distintas unidades, equipos, poblaciones, regiones, etc., obtenidas o no en un mismo momento.

Cabe añadir que no solo se busca una expresión matemática que diga de qué manera están relacionadas las variables, sino también definir con qué precisión se puede hacer una predicción, lo que en buena medida depende del grado de dispersión de los datos y de la bondad de la curva seleccionada, correspondiendo al análisis de correlación establecer el grado de asociación y la calidad del ajuste.

2. Dependencia estadística y dependencia funcional

Se puede decir que el análisis de regresión se ocupa de lo que se conoce como dependencia estadística entre variables y no de la dependencia funcional o determinista. En la dependencia estadística se manejan esencialmente variables aleatorias o estocásticas es decir, variables que tienen distribuciones de probabilidad. La dependencia funcional, en cambio, se ocupa de variables que no son aleatorias ni estocásticas.

Para tener claridad en estos conceptos [Barrera G. 1997] en la clase de series de tiempo expone que “la dependencia de una cosecha de la temperatura ambiente, lluvias, el sol y los fertilizantes, por ejemplo, es de naturaleza estadística en el sentido de que las variables explicatorias, aunque ciertamente son importantes, no permiten al decisor predecir el producto de la cosecha con seguridad por errores en la medición de tales variables o porque otros muchos factores (variables) afectan la cosecha, los cuáles son difíciles de identificar individualmente. Por tanto, existe alguna variabilidad intrínseca o aleatoria en la variable dependiente producto de la cosecha, que no puede ser completamente explicado, cualquiera que sea el número de variables explicatorias”

2.1. Regresión y causalidad

Aunque los modelos de regresión lineal se ocupan de la dependencia de una variable en función de otras, no implica necesariamente causalidad. En el ejemplo de la cosecha, citado previamente, no existe una razón para suponer que las lluvias no dependen estadísticamente del producto de cosecha. El hecho de que tratemos a la cosecha como dependiente de las lluvias (entre otras causas) se debe a consideraciones no estadísticas. El sentido común sugiere que la relación no puede ser al contrario pues no es posible controlar las lluvias mediante cambios en el producto la cosecha.

El punto esencial es, entonces que**,** una relación estadística *per se*, no puede implicar lógicamente la causalidad. Para aducir causalidad debe apelar a consideraciones teóricas o apriorísticas.

2.2. Asociación y causalidad

Entre las ventajas de los modelos elaborados con las técnicas regresión (modelos explicativos o causales) es que contribuyen a un mejor entendimiento del problema bajo consideración, lo que a su vez permite explorar distintas alternativas y diseñar políticas de intervención[2].

En otros términos, dado que ***Y*** está en función de ***X*,** se puede buscar cómo actuar sobre ***X*** para producir un efecto deseado en ***Y***; a diferencia de los métodos de extrapolación en los que se limita a predecir el valor ***Y***, para luego ver cómo se adapta a la nueva circunstancia.

Sin embargo, el identificar una función que se ajusta bien a un conjunto de datos no es razón suficiente para inferir que un cambio en una variable va a influir en que la otra también cambie.

Para ilustrar lo anterior, pensemos en que se ha obtenido una muestra de la altura de ***N*** mujeres y la de sus respectivos esposos, cuyos resultados se presentan en cuadro 1 y en la gráfica 2.

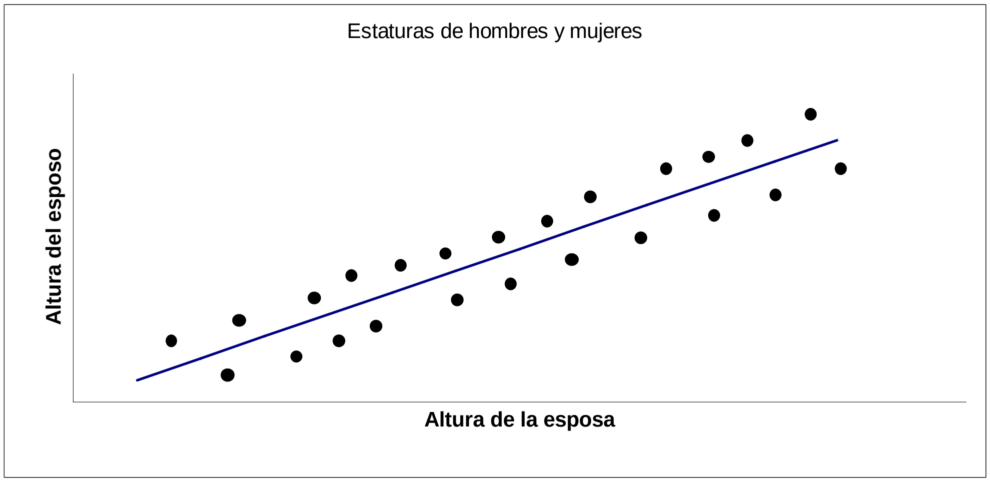
Cuadro 1
Estaturas de hombre y mujeres

Hombres	Mujeres
1.65	1.61
1.66	1.62
1.70	1.62
1.69	1.63
1.70	1.65
1.71	1.64
1.72	1.66
1.74	1.67
1.75	1.69
1.76	1.70
1.76	1.70
1.78	1.71
1.79	1.72
1.79	1.72

Fuente: Datos hipotéticos con fines ilustrativos.

[2] El análisis de intervención permite hacer una prueba formal del cambio en la media de una serie de tiempo. Para ver más sobre el tema ver Ender, W. 1990.

Gráfica 2
Puntos de la altura de N parejas



Fuente: Datos hipotéticos con fines ilustrativos.

Al observar la gráfica, se nota la existencia de una relación funcional y una alta correlación entre ambas variables, de suerte que al conocer la altura de la esposa podemos estimar con cierta precisión la altura del esposo. No obstante, a nadie se le ocurriría pensar qué esposo va a reflejar algún efecto si la esposa ingiere la "vitamina de crecimiento". Más que una relación causa-efecto se encuentra una relación de asociación, la cual es el resultado de la influencia de factores como las preferencias personales, patrones culturales, clases económicas etc.

El porqué de una relación está fuera del dominio de la estadística el único medio que se dispone para hablar de relaciones causa-efecto es el conocimiento acerca del fenómeno y del mecanismo bajo el que opera, donde la regresión es sólo un instrumento de apoyo para el análisis de los datos.

Todo lo anterior no quiere decir que exista algún impedimento para utilizar las relaciones de asociación en el pronóstico, de hecho algunos trabajos son realizados sobre estas bases; generalmente porque se cuenta con mayor información de la variable asociada o porque es más fácil el acceso a ella, además de que se conocen sus tendencias.

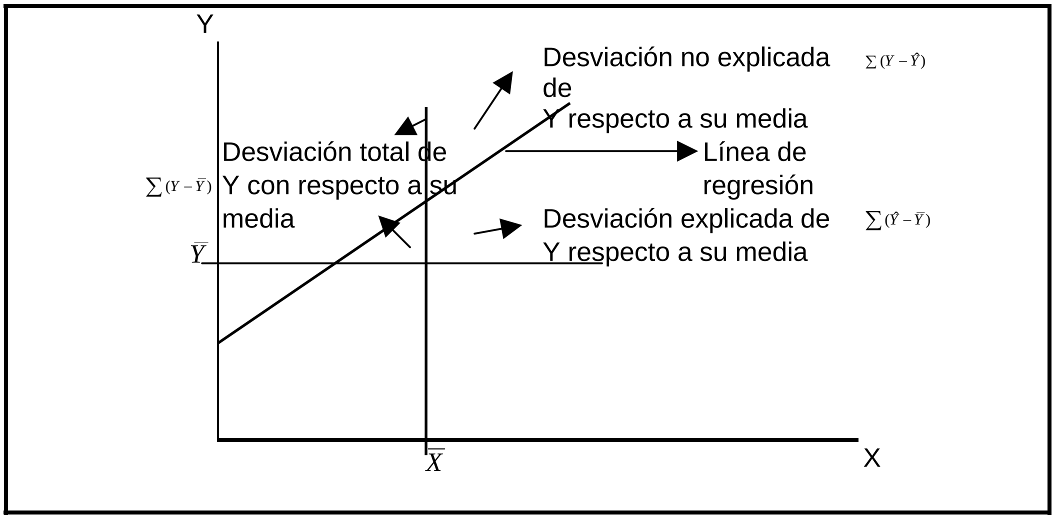
3. Coeficiente de determinación

Se define como el que mide la relación que existe entre dos variables, se obtiene de la relación de dos tipos de variación, la variación de los valores de ***Y*** en el conjunto de datos alrededor de la línea de regresión ajustada y de su propia media, podemos verlo en la figura 2.

$$R^2 = 1 - \frac{\sum\left(Y - \hat{Y}\right)^2}{\sum\left(Y - \overline{Y}\right)^2} \; ; \; R^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} \qquad (3)$$

Toma valores que oscilan de $0 \leq R^2 \leq 1$

Figura 2
Coeficiente de determinación



4. Coeficiente de correlación

El coeficiente de correlación indica la fuerza o grado de asociación que existe entre dos variables, o sea, describe la eficacia con que una variable es explicada por otra, a continuación se describen algunas propiedades.

- Puede ser negativo o positivo
- Tiene como límites $-1 \leq r \leq 1$
- Es simétrico, una correlación igual a cero no implica necesariamente independencia.
- Cuando la pendiente de la ecuación estimada es positiva, r es la raíz cuadrada positiva, pero si es de pendiente negativa, r es la raíz cuadrada negativa.
- El signo de r indica la dirección de la relación entre las dos variables.

Se presenta el desarrollo para llegar al coeficiente de correlación:

Se sabe que $r_{xy} = \dfrac{\operatorname{cov} \quad xy}{S_x \quad S_y}$

Cov_{xy} es el estimador de la covarianza de ***X*** y ***Y*** dada por $\hat{COV}_{XY} = \dfrac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{N}$

S_x, S_y Son los estimadores de la desviación estándar de ***X*** y ***Y***, respectivamente

$$S_x = \sqrt{\frac{\sum(X_i - \bar{X})^2}{N}}\ ; S_y = \sqrt{\frac{\sum(Y_i - \bar{Y})^2}{N}}$$

Por lo tanto

$$r_{XY} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{NS_xS_Y} = \frac{\dfrac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{N}}{N * \sqrt{\dfrac{\sum(X_i - \bar{X})^2}{N}} * \sqrt{\dfrac{\sum(Y_i - \bar{Y})^2}{N}}} \qquad (4)$$

Agrupando términos da como resultado

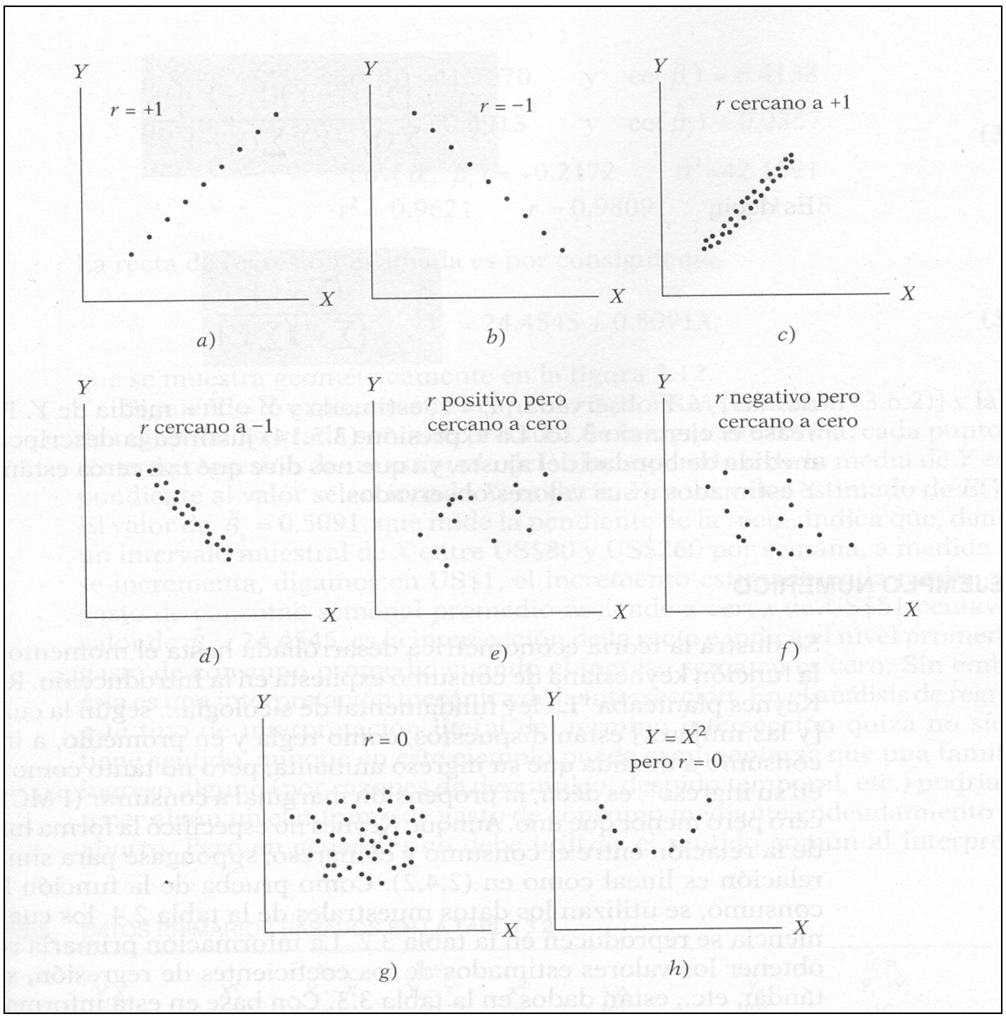
$$r_{XY} = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{N * \sqrt{\sum (X_i - \overline{X})^2} * \sqrt{\sum (Y_i - \overline{Y})^2}} \quad (5)$$

En términos de desviaciones se expresa: $r_{XY} = \frac{\sum x_i y_i}{\sqrt{\left(\sum x_i^2\right)\left(\sum y_i^2\right)}}$; o bien (6)

$$r = \pm\sqrt{R^2} \; ; \; r = \frac{N\sum X_i Y_i - \left(\sum X_i\right)\left(\sum Y_i\right)}{\sqrt{\left[N\sum X_i^2 - \left(\sum X_i\right)^2\right]\left[N\sum Y_i^2 - \left(\sum Y_i\right)^2\right]}} \quad (7)$$

Se puede obtener como la raíz cuadrada del coeficiente de determinación ver [Sánchez,G.,2006].

Figura 3
Representación gráfica del coeficiente de correlación.



Fuente: Tomado de Damodar Gujaratí p.83. 4ª edición,2003.

5. Método de mínimos cuadrados

Supongamos que se esta interesado en una relación entre dos variables ***X*** y ***Y***. Con el objeto de describir esta relación, estadísticamente necesitamos una colección de observaciones para cada variable y una hipótesis sobre la forma matemática de la relación.

Primeramente se prueba una relación lineal, se trata de especificar la regla por la cual se determina la mejor línea recta.

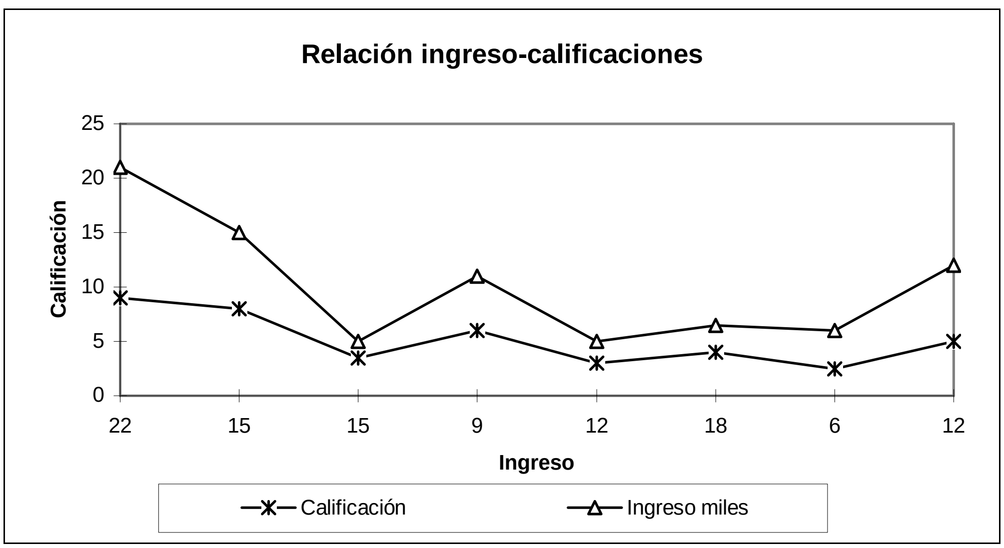
Se desea probar la hipótesis que las calificaciones promedio de los estudiantes, puede ser explicado por el ingreso percibido en un mes por los padres de los estudiantes, estos datos se observan en el cuadro 2 y en la gráfica 3.

Cuadro 2
Relación de calificaciones e ingresos

Y Calificación	X Ingreso miles de pesos
9.0	21.0
8.0	15.0
3.5	5.0
6.0	11.0
3.0	5.0
4.0	6.5
2.5	6.0
5.0	12.0

Fuente: Datos hipotéticos con fines ilustrativos.

Gráfica 3
Relación entre dos variables



Fuente: Datos hipotéticos con fines ilustrativos.

Muchas líneas rectas se pueden ajustar a través de los puntos ¿cuál línea recta?; suma de distancias verticales igual a cero (positivos y negativos). Desgraciadamente este procedimiento tiene la propiedad indeseable de que las desviaciones de igual tamaño pero de signo contrario se cancelan.

5.1. Desviación de mínimos cuadrados

Se supone un ajuste lineal, ***Y= a + bX*** , ***Y*** es la variable dependiente, ***X*** es la variable independiente. Ya que se desea explicar o predecir movimientos en ***Y*** nuestro objetivo es minimizar la suma vertical de los cuadrados de las desviaciones de la línea ajustada, es decir

Minimizar $\sum_{i=1}^{N} (Y_i - \hat{Y}_i)^2$ (8)

donde : ***Yᵢ***.- Valor observado de la i-ésima observación.

$\hat{Y}_i$.- Valor ajustado o estimado de la i-ésima observación

e_i.- Residuales

$$\hat{Y}_i = a + bX_i$$
$$Y_i = a + bX_i + e_i$$
$$e_i = Y_i - \hat{Y}_i \quad \text{por lo tanto } e_i^2 = (Y_i - \hat{Y}_i)^2 \qquad (9)$$

Es el valor ajustado de ***Y*** correspondiente a una observación particular de X_i

Para encontrar los estimadores de **a** y **b** que minimicen la suma de los errores al cuadrado, derivemos parcialmente, igualamos a cero y resolvemos las dos ecuaciones simultáneamente.

$$\frac{\partial}{\partial a}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = 0$$
$$\frac{\partial}{\partial b}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 = 0$$

Sustituyendo $\hat{Y}$

$$\frac{\partial}{\partial a}\sum_{i=1}^{n}(Y_i - a - bX_i)^2 = \sum_{i=1}^{n}\frac{\partial}{\partial a}(Y_i - a - bX_i)^2$$

derivando con respecto ***a***

$$= \sum_{i=1}^{n} 2(-1)(Y_i - a - bX_i)$$
$$= -2\sum_{i=1}^{n}(Y_i - a - bX_i) \text{ ....} \mathbf{a}$$

derivando con respecto a ***b***

$$\frac{\partial}{\partial b}\sum_{i=1}^{n}(Y_i - a - bX_i)^2 = \sum_{i=1}^{n}\frac{\partial}{\partial b}(Y_i - a - bX_i)^2$$

$$= \sum_{i=1}^{n} 2(-1 * X_i)(Y_i - a - bX_i)$$
$$= -2\sum_{i=1}^{n} X_i(Y_i - a - bX_i) \text{ ....} \mathbf{b}$$

Igualando **a** y **b** a cero, da un sistema de ecuaciones con dos incógnitas "**a**" y "**b**"

$$= -2\sum_{i=1}^{n}(Y_i - a - bX_i) = \sum_{i=1}^{n}(Y_i - a - bX_i) = 0$$

$$= -2\sum_{i=1}^{n} X_i(Y_i - a - bX_i) = \sum_{i=1}^{n} X_i(Y_i - a - bX_i) = 0$$

Introduciendo el signo de sumatoria, además que 2? 0 por lo que,

$$\sum_{i=1}^{n}(Y_i - a - bX_i) = \sum_{i=1}^{n} Y_i - na - b\sum_{i=1}^{n} X_i = 0 \qquad \text{Ecuación 1}$$

$$\sum_{i=1}^{n} X_i(Y_i - a - bX_i) = \sum_{i=1}^{n}(X_iY_i - aX_i - bX_i^2) = 0$$

$$\sum_{i=1}^{n} X_iY_i - a\sum_{i=1}^{n} X_i - b\sum X_i^2 = 0 \qquad \text{Ecuación 2}$$

El sistema de ecuaciones 1 y 2 se pueden expresar en la siguiente forma:

$$na + b\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i \qquad \text{Ecuación 3}$$

$$a\sum_{i=1}^{n} X_i + b\sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} X_iY_i \qquad \text{Ecuación 4}$$

Multiplicando la ecuación 1 por $\sum_{i=1}^{n} X_i$ y la ecuación 2 por n

$$na\sum_{i=1}^{n} X_i + b\left(\sum_{i=1}^{n} X_i\right)^2 = \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i \qquad \text{Ecuación 5}$$

$$na\sum_{i=1}^{n} X_i + nb\sum_{i=1}^{n} X_i^2 = n\sum_{i=1}^{n} X_iY_i \qquad \text{Ecuación 6}$$

Restando la ecuación 5 de la ecuación 6

$$b\left(\sum_{i=1}^{n} X_i\right)^2 - nb\sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i - n\sum_{i=1}^{n} X_iY_i \qquad \text{Ecuación 7}$$

Factorizando "b"

$$b\left[\left(\sum_{i=1}^{n} X_i\right)^2 - n\sum_{i=1}^{n} X_i^2\right] = \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i - n\sum_{i=1}^{n} X_iY_i \qquad \text{Ecuación 8}$$

Por lo tanto

$$b = \frac{\sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i - n\sum_{i=1}^{n} X_i Y_i}{\left(\sum_{i=1}^{n} X_i\right)^2 - n\sum_{i=1}^{n} X_i^2} \qquad \text{Ecuación 9}$$

Si multiplicamos por (-1) el numerador y el denominador

$$b = \frac{n\sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} Y_i \sum_{i=1}^{n} X_i}{n\sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2} \qquad \text{Ecuación 10}$$

Despejando "**a**" de la ecuación 3

$$na = \sum_{i=1}^{n} Y_i - b\sum_{i=1}^{n} X_i$$

$$a = \frac{1}{n}\sum_{i=1}^{n} Y_i - b\frac{1}{n}\sum_{i=1}^{n} X_i$$

$$a = \overline{Y} - b\overline{X} \text{ dado que } \frac{\sum Y_i}{n} = \overline{Y} \text{ y } \frac{\sum X_i}{n} = \overline{X} \qquad \text{Ecuación11}$$

De esta manera "**a**" queda en función del valor conocido de "**b**" y de las medias $\overline{Y}$ y $\overline{X}$

Otra forma de escribir la ecuación es con sus desviaciones

$$Y_i = a + bX_i + e_i$$

$$\overline{Y} = a + b\overline{X} + \overline{e}_i$$

$$\overline{e}_i = \sum_{i=1}^{n} e_i = 0$$

$$Y_i - \overline{Y} = a - a + b(X_i - \overline{X}) + e_i$$

Denotando con minúsculas las desviaciones

$$y_i = Y_i - \overline{Y} \qquad x_i = X_i - \overline{X}$$

$$y_i = bx_i + e_i$$
$$\hat{y}_i = bx_i$$
$$e_i^2 = (y_i - \hat{y}_i)^2$$

Nuevamente el procedimiento es minimizar la sumatoria de los cuadrados de los residuales

Minimizar $\sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

Sustituyendo $\hat{y}_i$ con $\hat{y}_i = bx_i$

Minimizar $\sum_{i=1}^{n} (y_i - bx_i)^2$ con respecto al parámetro “**b**”

$$\frac{\partial}{\partial b}\sum_{i=1}^{n}(y_i - bx_i)^2 = 0$$

$$\sum_{i=1}^{n} - 2x_i(y_i - bx_i) = 0$$

$$-2\sum_{i=1}^{n} x_i(y_i - bx_i) = 0$$

$$\sum_{i=1}^{n} x_i(y_i - bx_i) = 0$$

$$\sum_{i=1}^{n} x_i y_i - b\sum_{i=1}^{n} x_i^2 = 0$$

$$b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i$$

$$b = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} \tag{10}$$

Para demostrar que es un mínimo se obtiene la segunda derivada con respecto a “**b**” y se demuestra que $\frac{\partial}{\partial b} = \sum_{i=1}^{n} - 2x_i(y_i - bx_i) > 0$

$$\frac{\partial}{\partial b} = \left(-2\sum_{i=1}^{n} x_i y_i - 2\sum_{i=1}^{n} - bx_i^2\right)$$

$$\frac{\partial}{\partial b} = 2\sum_{i=1}^{n} x_i^2$$

$$= 2\sum_{i=1}^{n} x_i^2 > 0$$

La estimación de “**b**” se puede reexpresar de la siguiente manera, dado que $\sum x_i \overline{Y} = 0$:

$$b = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} x_i (Y_i - \overline{Y})}{\sum_{i=1}^{n} x_i^2} = b = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2} - \frac{\overline{Y} \sum x_i}{\sum_{i=1}^{n} x_i^2}$$ Ecuación 12

Por lo tanto

$$b = \frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2} \tag{11}$$

Es posible expresar a "**b**" como funciones lineales de las observaciones de **Y**

Si $w_i = \frac{x_i}{\sum_{i=1}^{n} x_i^2}$ por lo tanto $b = \sum_{i=1}^{n} w_i Y_i$

w_i son constantes fijas con las siguientes características

$\sum_{i=1}^{n} w_i = 0$ ya que $\sum_{i=1}^{n} w_i = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2}$

$\sum_{i=1}^{n} w_i^2 = \frac{1}{\sum_{i=1}^{n} x_i^2}$ ya que $w_i^2 = \frac{x_i^2}{\left(\sum_{i=1}^{n} x_i^2\right)^2}$ y $\sum_{i=1}^{n} w_i^2 = \frac{\sum_{i=1}^{n} x_i^2}{\left(\sum_{i=1}^{n} x_i^2\right)^2} = \frac{1}{\sum_{i=1}^{n} x_i^2}$

$\sum_{i=1}^{n} w_i x_i = \sum_{i=1}^{n} w_i X_i = 1$ ya que $\sum_{i=1}^{n} w_i (X_i - \overline{X}) = \sum_{i=1}^{n} w_i X_i - \overline{X} \sum_{i=1}^{n} w_i$

$w_i x_i = \frac{x_i^2}{\sum_{i=1}^{n} x_i^2}$ por lo tanto $\sum_{i=1}^{n} w_i x_i = \frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^2} = 1$

De manera similar se puede expresar "**a**" como

$a = \sum_{i=1}^{n} \left(\frac{1}{n} - \overline{X} w_i \right) Y_i$ ya que $a = \overline{Y} - b\overline{X}$

sustituyendo **b**

$$a = \overline{Y} - \overline{X} \sum_{i=1}^{n} w_i Y_i$$ Ecuación 13

$$a = \frac{1}{n}\sum_{i=1}^{n} Y_i - \overline{X}\sum_{i=1}^{n} w_i Y_i \qquad \text{Ecuación 14}$$

Factorizando Y_i

$$a = \sum_{i=1}^{n}\left(\frac{1}{n} - \overline{X}w_i\right)Y_i \text{ es decir } a = \overline{Y} - b\overline{X} \qquad (12)$$

Es decir también se puede expresar "**a**" como una combinación lineal de $\boldsymbol{Y_i}$ ya que $\left(\frac{1}{n} - \overline{X}w_i\right)$.

Una técnica para estimar la ecuación de regresión es mediante el método de mínimos cuadrados, que consiste en ajustar la línea recta óptima a la muestra de las observaciones, esto es minimizar la suma de los cuadrados de las desviaciones entre los puntos y la línea.

La función de regresión poblacional denota cómo el valor promedio poblacional de ***Y*** varía con las ***X's.*** Es importante señalar que es difícil pensar la forma que toma la función ***f(x)***, porque no se dispone de la totalidad de la población para efectuar el análisis. La función de regresión poblacional es una función lineal[3] de ***X*** de la siguiente forma:

$$E(Y/X) = \beta_1 + \beta_2 X \qquad \text{o bien} \qquad Y = \beta_1 + \beta_2 X_2 + \mu_i \qquad (13)$$

en el que β_1 y β_2 son parámetros desconocidos pero fijos que denominaremos de aquí en adelante coeficientes de regresión, también reciben el nombre β_1 de intersección, ordenada al origen o constante y β_2 pendiente o propensión marginal. El análisis de regresión el objetivo es estimar una función de regresión poblacional

La función de regresión muestral, ¿por qué estimar esta función?, porque en nuestra práctica profesional no contamos con toda la información suficiente, (la población), entonces debemos hechar mano de muestras para que la estimación sea lo más certera posible, entonces se debe estimar la función de regresión poblacional con base en información muestral; ¿qué es esto? es hacer inferencia estadística acerca de los verdaderos valores de β_1 y β_2. La cual se escribe:

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X \qquad \text{o bien} \qquad \hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + e_i \qquad (14)$$

6. Supuestos del modelo de regresión lineal clásico

Una vez de acuerdo, que para estimar el comportamiento de una población es necesario hacer inferencia sobre los verdaderos valores, es importante conocer qué tan cerca estamos de los mismos, para continuar se debe saber como se generan $\boldsymbol{X_2}$ y $\boldsymbol{\mu_i}$ de $Y = \beta_1 + \beta_2 X_2 + \mu_i$, y que tanto depende ***Y*** de estos; si no se sabe no se puede hacer inferencia; por tanto los supuestos que vamos a plantear para $\boldsymbol{X_2}$ y $\boldsymbol{\mu_i}$ son "fundamentales para llevar a cabo una interpretación válida de las estimaciones de la regresión".

"Supuesto 1: El modelo de regresión lineal es lineal en los parámetros $Y = \beta_1 + \beta_2 X_2 + \mu_i$

[3] El termino lineal "significa una regresión lineal en los parámetros o las β's, (es decir los parámetros, se elevan únicamente a la primera potencia), pudiendo o no ser lineales en las variables explicativas, las X"

Supuesto 2: Los valores de *X* son fijos en muestreo repetido, se supone no estocástica.

Supuesto 3: El valor medio o promedio de μ_i es igual a cero $E(\mu_i | X) = 0$

El valor promedio de μ_i, dado un valor de *X* es igual a cero

Supuesto 4: Homocedasticidad o igual varianza para $\boldsymbol{\mu_i}$

$$var(\mu_i | X) = E[\mu_i - E(\mu_i)]^2$$
$$= E(\mu_i^2) \text{ por el supuesto 3}$$
$$= \sigma^2$$

la varianza de $\boldsymbol{\mu_i}$ para cada ***X*** es constante.

Supuesto 5: No existe autocorrelación entre la $\boldsymbol{\mu_i}$

$$\text{cov}(\mu_i, \mu_j) = E[\mu_i - E(\mu_i)][\mu_j - E(\mu_j)]$$
$$= E(\mu_i \mu_j) \quad \text{por el supuesto 3}$$
$$= 0 \quad i \neq ? j$$

Las perturbaciones de $\boldsymbol{\mu_i}$ y $\boldsymbol{\mu_j}$ no están correlacionadas

Supuesto 6: Cero covarianza entre $\boldsymbol{\mu_i}$ y ***X***

$$\text{cov}(\mu_i, X) = E[\mu_i - E(\mu_i)][X - E(X)]$$
$$= E[\mu_i (X - E(X)] \text{ puesto que } E(\mu_i) = 0$$

$=E(\boldsymbol{\mu_i}$ X) - E(X)$E(\boldsymbol{\mu_i})$, ya que E(X) es una constante

$=E(\boldsymbol{\mu_i}$ X), puesto que $E(\mu_i) = 0$

=0, de acuerdo con el supuesto

Supuesto 7: El número de observaciones de ***n*** debe ser mayor que el número de parámetros por estimar.

Supuesto 8: Variabilidad en los valores de ***X***. No todos los valores de ***X*** en una muestra dada deben ser iguales. ***var(X)*** debe ser un número positivo finito.

Supuesto 9: El modelo de regresión esta correctamente especificado. No existen sesgos ni errores de especificación.

Supuesto 10: No hay multicolinealidad perfecta. No hay relaciones perfectamente lineales entre las variables explicativas" [Gujarati D. 2003].

7. Propiedades de los estimadores

Un estimador es una regla que establece cómo calcular una estimación basada en las mediciones contenidas en una muestra. Generalmente un estimador se expresa mediante una fórmula, se pueden tener varios estimadores para un sólo parámetro. Por ejemplo si cada uno de 10 economistas fueran asignados para estimar el costo de un gran proyecto, obtendrían casi seguramente distintas estimaciones del costo total. Los economistas serían los estimadores, que utilizarían sus conocimientos para hacer la estimación, cada uno representa una sola regla humana para obtener la estimación. De aquí surge la pregunta ¿cuáles son los mejores estimadores?.

La estimación puntual es similar al proceso de disparar con una pistola a un blanco. El estimador que genera estimaciones es la pistola, y una estimación en particular es la bala, y el parámetro de interés es el blanco. Al hacer muestreo y estimar el valor de un parámetro es equivalente a disparar un solo tiro al blanco.

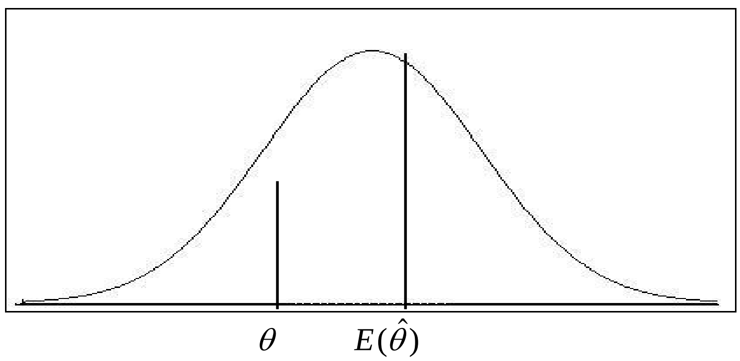
Como lo menciona el doctor [Guerrero, V. 2001] “La evaluación de un estimador se hace construyendo una distribución de frecuencias de las estimaciones obtenidas en un muestreo repetitivo y se observa como se agrupa la distribución alrededor del parámetro en estudio”.

Se escribe $\hat{\theta}$ para indicar que este es un estimador del parámetro θ
Se espera que un buen estimador tenga como media o esperanza al parámetro estimado.

Un estimador se dice que es el mejor estimador linealmente insesgado si cumple las siguientes condiciones:

- Sesgo de un estimador, un estimador $\hat{\theta}$ es insesgado si $E(\hat{\theta}) = \theta$, de lo contrario se dice que es sesgado. Si $E(\hat{\theta}) > \theta$ entonces lo que esta ocurriendo es:

Figura 4
El sesgo de B de un estimador puntual $\hat{\theta}$ está dado por B= $E(\hat{\theta}) - \theta$



- Se desea que Var$(\hat{\theta})$ sea mínima. Si se tiene dos estimadores insesgados para un mismo parámetro se prefiere al de menor varianza.

Para evaluar ambas cosas sesgo y varianza se puede utilizar $E\left[(\hat{\theta} - \theta)^2\right]$. La Media del Cuadrado del Error del estimador $\hat{\theta}$, se denota como MCE$(\hat{\theta})$

MCE $(\theta) = Var(\hat{\theta}) + B^2$

Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son estimadores insesgados de θ, se dice que $\hat{\theta}_1$ es más eficiente que $\hat{\theta}_2$ si $Var(\hat{\theta}_2) > Var(\hat{\theta}_1)$.

- Consistencia, si se lanzara una moneda *n* veces que tiene probabilidad ***p*** de ser águila, entonces ***Y***, el número de águilas en los ***n*** lanzamientos, tiene una distribución binomial. Si ***p*** es desconocido se puede estimar con ***Y/n***. ¿Qué pasa con esta proporción muestral si aumenta el número de lanzamientos n? intuitivamente se pensaría que ***Y/n*** debería estar más cerca de ***p***. Esto en términos de probabilidad se escribe así $P\left[\ \left|(y / n) - p\right| \leq \xi\ \right]$

Esta probabilidad debería ser cercana a la unidad para valores grandes de ***n***. Si la probabilidad de $P\left[\ \left|(y / n) - p\right| \leq \xi\ \right]$ tiende a uno cuando $n \rightarrow \infty$ entonces ***Y/n*** es un estimador consistente de ***p***.

- Suficiencia, al calcular un estimador con los valores de la muestra se “resume” la información en un número, este procedimiento de resumir los datos de la muestra mantiene toda la información con respecto al parámetro de interés, o se ha perdido u ocultado algo de la información en el proceso de sintetizar datos.

Por ejemplo consideremos los resultados de n pruebas de un experimento bernulli ***X_1,...X_n***. Sea $Y = \sum X_1$ el número de éxitos en ***n*** pruebas. Si se conoce el valor de ***Y***, es posible obtener más información con respecto a ***p*** al considerar otras funciones de ***X_1,...X_n***. dado ***Y***.

$$P(X_1 = X_1,..., X_n = X_n \mid Y) \qquad Y = \frac{P^y (1-p)^{n-y}}{\binom{n}{y} P^y (1-p)^{n-y}} = \frac{1}{\binom{n}{y}} \tag{15}$$

Una vez conocido ***Y***, ninguna otra función de ***X_1...X_n*** suministra más información con respecto a ***P***. Por lo tanto ***Y*** es un estadístico suficiente para ***P***. Es decir, el que se aproxima al parámetro (μ) que va a estimar, al aumentar la muestra.

8. Intervalos de confianza

Un estimador por intervalo es una regla que utiliza las mediciones de la muestra para obtener dos números que forman los extremos del intervalo. Sería conveniente que el intervalo tuviera dos propiedades, primero que contuviera al parámetro de interés y segundo que fuera estrecho.

A los estimadores por intervalo se les conoce generalmente como intervalos de confianza. Los extremos del intervalo se les conoce como límites de confianza inferior y superior. La probabilidad de que un intervalo contenga a θ se conoce como coeficiente de confianza. El coeficiente de confianza indica la fracción de veces, en un muestreo repetitivo, que los intervalos construidos contendrán al parámetro θ.

Si el límite inferior y el límite superior son los límites de confianza inferior y superior, respectivamente, para un parámetro θ. Entonces, si $P(L_1 < \theta < L_s) = 1 - \alpha$ la probabilidad $(1 - \alpha)$ es el coeficiente de confianza. En el libro de [Sánchez G, 2006] se ilustra la construcción de los intervalos de confianza.

9. Pruebas de hipótesis

Es importante tener clara la diferencia entre pruebas de hipótesis y estimación, ambas incluyen el uso de distribuciones muestrales, la diferencia sustancial es saber si se ha establecido o no una noción predeterminada del estado de la población que se intenta muestrear.

Si no se sabe nada acerca de la población, se usa una estimación para proporcionar estimaciones puntuales o por intervalo para los parámetros de la población de interés, si la información sobre la población se asegura o se sospecha se usa una prueba de hipótesis para determinar la factibilidad de esta información.

Al intentar alcanzar una decisión es útil hacer una hipótesis o conjeturas sobre la población implicada, tales hipótesis que pueden ser o no ser ciertas se llaman hipótesis estadísticas por lo general son enunciados acerca de las distribuciones de probabilidad de las poblaciones.

El procedimiento para probar una hipótesis sobre el estado de una población se describirá para facilitar la elaboración de las hipótesis a probar.

1. Establecer la hipótesis nula y alternativa
2. La hipótesis nula establece que la diferencia entre un estadístico muestral y el parámetro poblacional correspondiente supuesto, se debe a la variación propia del muestreo.
3. La hipótesis alternativa es la afirmación opuesta sobre la población, que debe ser cierta si la hipótesis nula es falsa.
4. Suponer que la hipótesis nula es cierta y determinar la distribución muestral apropiada para esta suposición. La hipótesis nula esta en tela de juicio, se supone que la hipótesis nula es cierta a menos que se produzca una evidencia muestral que demuestre lo contrario.
5. Bajo la suposición que la hipótesis nula es cierta, se desarrolla una distribución muestral a partir de la cual se obtendrá el estadístico calculado para la prueba♣.

Obtener una muestra aleatoria a partir de la población bajo estudio y con los datos recogidos calcular el estadístico apropiado para la prueba, este estadístico casi siempre es la media muestral o la proporción muestral.

Evaluar la posibilidad de seleccionar determinado valor estadístico para la prueba a partir de la distribución muestral supuesta.

Si parece poco probable que el estadístico de prueba calculado se pueda obtener de la distribución muestral, se rechaza la hipótesis nula, por otra parte, si parece poco probable que el estadístico de prueba se obtenga de la distribución muestral supuesta, la hipótesis nula no se rechaza.

Un aspecto importante de la estimación es determinar la posibilidad de error. La evaluación del error es una parte integral de las características poblacionales que se pretende estimar usando la evidencia muestral, al probar una hipótesis se debe evaluar la posibilidad de que la conclusión a la que se llego sea incorrecta.

"Error tipo I y tipo II en las pruebas de hipótesis siempre es posible uno de los dos errores. La muestra (aunque se selecciona de manera aleatoria) puede llevar a un rechazo incorrecto de la hipótesis nula (error tipo I) o a una aceptación incorrecta de la misma (error tipo II)" [Sánchez G. 2006].

9.1. Verificación de hipótesis e intervalos de confianza para los parámetros

Las estimaciones de los coeficientes obtenidas mediante el análisis de la regresión deben ser estadísticamente aceptables de esta forma es posible considerar su sentido económico. Por lo tanto, previamente a la interpretación de los resultados, se requiere aplicar una serie de tests, pruebas o contrastes estadísticos. Para cada muestra se obtendrán estimaciones diferentes, es decir, valores distintos de $\hat{\beta}_i$ lo que implica que en el muestreo, $\hat{\beta}_i$ es una variable aleatoria con su correspondiente función de distribución. Ello permite la construcción de estadísticos, que se utilizan para verificar hipótesis sobre los verdaderos valores de los parámetros.

Para construir dichos estadísticos se precisa conocer el comportamiento de cada variable aleatoria $\hat{\beta}_i$. Bajo la hipótesis de que cada $\boldsymbol{\mu_i}$ se distribuye como una variable aleatoria normal, con E($\boldsymbol{\mu_i}$) = 0 y Var($\boldsymbol{\mu_i}$) = s^2, se deduce,

$$\frac{\hat{\beta}_i - \beta_i}{S_{\hat{\beta}_i}} \approx t_{n-k} \qquad (16)$$

♣ La distribución muestral específica será diferente de una prueba a otra.

Es decir, la expresión anterior se distribuye con una t de Student con n - k grados de libertad. Mientras que $S_{\hat{\beta}_i}$ es lo que se denomina error estándar del estimador o del coeficiente. Con esta expresión se verifica cualquier hipótesis, que denominaremos nula (H_0), sobre el valor del parámetro β_i. La hipótesis más usual es la que establece $H_0:$ β_i=0, frente a una hipótesis alternativa $H_1:$ $\beta_i \neq 0$. En realidad, se trata de verificar que las variables ***X's*** son explicativas, ya que si se acepta H_0, diremos que el parámetro, con nuestros datos, pueden ser cero con una probabilidad de (1 – α), donde α es el nivel de significación elegido.

Este valor empírico t* (recibe también el nombre de calculado o experimental) se obtiene a partir de los datos muestrales (coeficiente/ error estándar del coeficiente) y debe ser comparado con el valor teórico que aparece en las tablas de la misma distribución. Este valor teórico señala el valor critico a partir del cual se rechaza la hipótesis nula, es decir, delimita las regiones de aceptación y rechazo.

Por lo tanto la regla de decisión será la siguiente: Aceptar la hipótesis nula si la t* empírica es menor que la t de tablas, de lo contrario, no aceptar la hipótesis.

$$\begin{bmatrix} t^* < t_{tablas} \quad aceptar \quad H_o \\ \\ t^* < t_{tablas} \quad no \;\; aceptar \;\; H_o \end{bmatrix}$$

Asimismo, estos resultados permiten construir intervalos de confianza para los coeficientes:

$$\hat{\beta}_i \pm t_{\alpha/2} S_{\hat{\beta}_i}$$

(17)

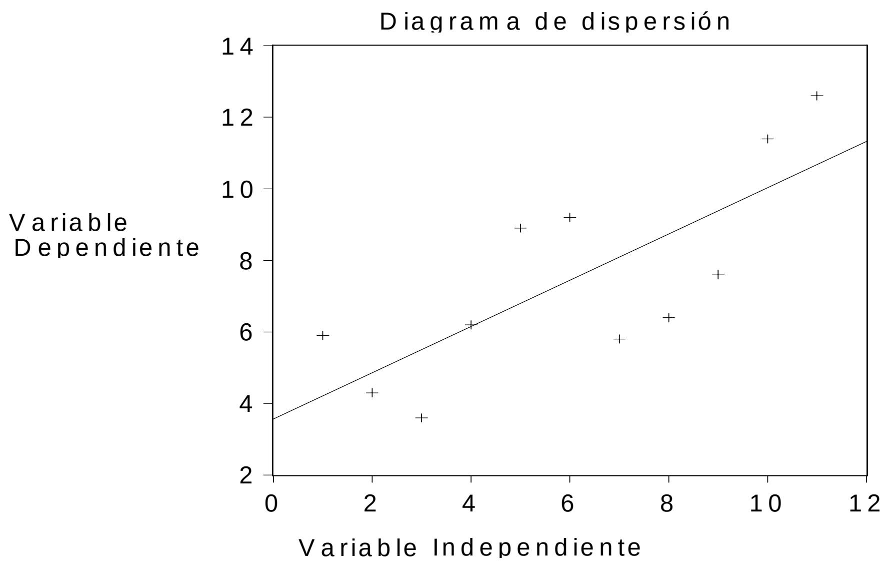
donde $t_{\alpha/2}$ es el valor de distribución teórica para un nivel de significación de α y n-k grados de libertad. El significado de dicho intervalo consiste en que con una confianza del (1-α) el verdadero valor del parámetro se encontrara dentro de dos puntos. Dicho de otro modo de los infinitos intervalos que se pudieran construir, en 1-α en ellos se encontraría el verdadero valor poblacional del parámetro.

10. Diagrama de dispersión

El punto de partida de la regresión es un conjunto de ***N*** pares de datos (***X$_i$, Y$_i$***), los cuales pueden provenir de registros históricos, observaciones de campo, experimentos, etc. Datos que se representa en un diagrama de puntos, llamado también diagrama de dispersión ver figura 5.

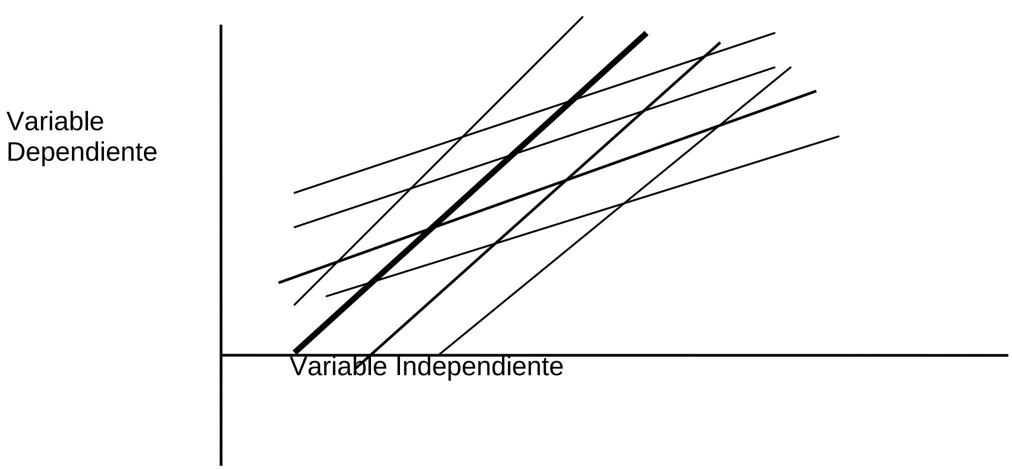
El diagrama de dispersión, nos brinda dos tipos de información, visualmente podemos buscar los patrones que indican que las variables están relacionadas; y si es que existe una relación entre ellas.

Figura 5
Diagrama de dispersión



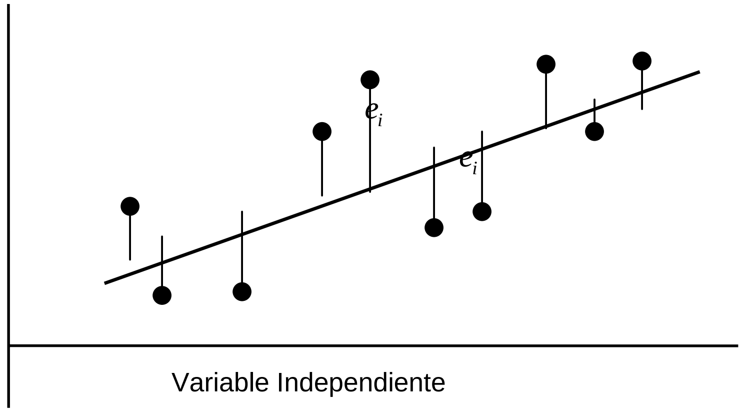
Con base en el conocimiento del fenómeno y con el apoyo del diagrama de dispersión, se propondrá el tipo de función que mejor represente la relación entre ambas variables, que por ahora supondremos de carácter lineal ***Y = a + b X*** con lo que queda una ecuación que representa una familia de rectas (ver figura 6); el paso siguiente consiste en fijar el valor de los coeficientes de regresión, la constante "**a** o β_1" y la pendiente "**b** o β_2", para identificar la recta más apropiada.

Figura 6
Líneas de regresión posibles



El criterio comúnmente empleado para este fin es el de mínimos cuadrados, que como ya se mencionó antes, busca (ver el apartado 1.5) minimizar la suma de los cuadrados de las desviaciones entre el valor real de ***Y*** y el valor estimado haciendo uso de la recta de regresión (ver figura 7).

Figura 7
Desviaciones verticales respecto a la línea de regresión



Este criterio no es el único posible, también podría pensarse en minimizar la suma del valor absoluto de las desviaciones, calcular alguna desviación perpendicular a la recta, seleccionar la recta que pase por el mayor número de puntos, ajustar a simple vista, etc.

11. Modelo matricial de regresión lineal

La siguiente ecuación es una expresión abreviada de un conjunto de **N** ecuaciones simultáneas

$$Y_1 = \beta_1 + \beta_2 X_{21} + \beta_3 X_{31} + \ldots + \beta_k X_{k1} + \mu_1$$
$$Y_2 = \beta_1 + \beta_2 X_{22} + \beta_3 X_{32} + \ldots + \beta_k X_{k2} + \mu_2$$

....

$$Y_N = \beta_1 + \beta_2 X_{2N} + \beta_3 X_{3N} + \ldots + \beta_{kN} X_{kN1} + \mu_N$$

Para un desarrollo más detallado de este método los profesores [Pindyck R. y Rubinfeld D.,1998] explican a detalle los desarrollos matemáticos para llegar al modelo generalizado.

Reescribiendo el sistema en forma alterna tenemos:

y X β μ

$$\begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_N \end{bmatrix} = \begin{bmatrix} 1 & X_{21} & X_{31} & \ldots X_{k1} \\ 1 & X_{22} & X_{32} & \ldots X_{k2} \\ . & . & . & .. \\ . & . & . & .. \\ . & . & . & .. \\ 1 & X_{2N} & X_{3N} & \ldots X_{kN} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ . \\ \beta_k \end{bmatrix} + \begin{bmatrix} \mu_1 \\ \mu_2 \\ . \\ . \\ . \\ \mu_N \end{bmatrix}$$

y X ? u

N * 1 ***N * k*** ***k * 1*** ***N * 1***

Donde:

y = un vector columna ***N * 1*** de observaciones de la variable dependiente ***Y***.
X = es una matriz ***N * k*** se obtienen ***N*** observaciones de los ***K-1*** variables $\boldsymbol{X_2}$ a $\boldsymbol{X_k}$. La columna de ***1's*** representa el intercepto. (Esta matriz se conoce también como la matriz de observaciones).
$\boldsymbol{\beta}$ = un vector columna ***k * 1*** de los parámetros desconocidos β_1, β_2, ..., β_k.
μ = un vector columna ***N * 1*** de las ***N*** perturbaciones $\boldsymbol{\mu_i}$.

El sistema se conoce como la representación matricial del modelo de regresión lineal general de (k variables).

"Cuando la primera variable es una constante, es decir, cuando hay un término independiente en el modelo, entonces la primera fila y la primera columna se convierten en la diagonal principal de la matriz.

¿Bajo que condiciones existirá un único estimador MCO? El sistema de ecuaciones normales esta formado por tantas ecuaciones como parámetros se pretende estimar, por lo que existirá una única solución si $|X'X| \neq 0$ y existirán infinitas soluciones si $|X'X| = 0$".[Novales A, 1988]

Se puede escribir de forma más condensada como:

$$\underset{N*1}{\boldsymbol{y}} = \underset{N*k}{\boldsymbol{X}} \; \underset{k*1}{\boldsymbol{\beta}} + \underset{N*1}{\mu}$$

O se escribe simplemente

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\mu} \qquad (18)$$

Se muestra una forma simple de expresar un modelo en forma matricial, pero el profesor [Intriligator M., 1990] presenta en su obra un enfoque y un desarrollo a detalle por medio del uso del álgebra matricial.

2 CAPITULO

1. MULTICOLINEALIDAD[4]

La interpretación del modelo de regresión depende implícitamente del supuesto de que las variables explicativas o independientes no están fuertemente relacionada entre si.

La multicolinealidad se refiere a la existencia de más de una relación lineal exacta.

Antes de comenzar a discutir las consecuencias de la multicolinealidad sobre los modelos regresión o econométricos es necesario aclarar lo siguiente:

- "La multicolinealidad es una cuestión de grado y no de clase. La distinción significativa no es entre la presencia o ausencia de este fenómeno en un modelo, sino entre sus varios grados.
- Como la multicolinealidad se refiere a una condición sobre las variables explicativas o independientes que se supone no estocásticas, entonces es una característica de la muestra y no de la población bajo estudio".[Luis O, 1992]

Esto implica que la multicolinealidad es una condición de interdependencia entre $X_1, X_2, \ldots X_k$ que puede existir independientemente de la dependencia entre las ***X's*** y la variable ***Y***, esto es que la multicolinealidad es una propiedad del conjunto de variables explicativas únicamente "cuando ocurre multicolinealidad, como si los miembros del subconjunto de variables explicativas afectadas actuaran al unísono. Como resultado de ello, los datos carecen de suficiente variación independiente para permitirnos determinar el efecto separado de cada variable explicativa. Entonces el problema de la multicolinealidad no es un problema de método de estimación, ni de la población, sino de los datos; un problema básico de la forma de generación de las observaciones un fenómeno muy frecuente en econometría donde se trabaja con datos de ciencias como la economía y las ciencias sociales en general". [Kmenta, 1979]

1.1. Como detectar multicolinealidad

Es necesario señalar que la multicolinealidad afecta la confiabilidad de los estimadores minimocuadráticos de un modelo por las grandes varianzas que ocasiona, no afecta la insesgadez de los mismos. El principal problema radica en que no permite determinar en forma separada o individual, la influencia o impacto de las variables explicativas afectadas sobre la variable dependiente.

Para detectar este problema estadístico resulta importante mencionar sus principales síntomas.

1. Un coeficiente de determinación R^2 elevado, esto por el hecho de que la colinealidad representa una doble, triple o múltiple contabilidad del poder explicativo de un modelo, por cada variable multicolineal incluida en el mismo, según sean dos, tres o más las que estén encadenadas entre sí.

2. Las t calculadas o empíricas de cada estimador son muy pequeñas o no significativas.

3. Las varianzas y covarianzas de los estimadores minimocuadráticos, son inusualmente grandes.

4. Por el punto anterior se producen intervalos de confianza muy grandes para los parámetros estimados mediante MCO.

[4] Se conoce en la literatura técnica como problema de "datos colineales", "colinealidad", "interconexión" o "multicolinealidad" se utiliza con más frecuencia esta última manera de nombrar el problema, por ser la más conocida.

5. Existe sensibilidad en los estimadores minimocuadráticos y sus errores estándar frente a cambios pequeños o marginales en los datos.

6. El valor del determinante de la matriz de correlaciones entre los regresores próximos a cero.

7. La Medida de multicolinealidad de Belsey, Kuck y Welsch, esta dado por:

$$k(x) = \frac{\sqrt{\lambda_{máximo}}}{\sqrt{\lambda_{mínimo}}} \tag{19}$$

8. Regla de Farrar y Glauber[5], análisis de las correlaciones parciales. Si satisface la condición de que $\frac{R_i}{R_j} \geq 1$ entonces la multicolinealidad existe en el modelo.

   Donde R_j, representa el coeficiente de correlación múltiple (raíz cuadrada del coeficiente de determinación simple $\sqrt{R^2}$) y R_i el coeficiente de correlación múltiple obtenida efectuando una regresión de ***X*** respecto a las otras variables independientes en el modelo (es la raíz cuadrada del coeficiente de determinación múltiple respectivo $\sqrt{R_i^2}$).

   Otra forma de trabajar con el contraste de Farrar-Glauber es el siguiente:

   H_0 : $|R_{xx}|=1$ Ortogonalidad[6] entre los regresores (ausencia de multicolinealidad).

$$G_{\exp} = -\left[T - 1 - \frac{1}{6}(2k+5)\right]\ln|R_{XX}| \tag{20}$$

$$G_{\exp} \sim \chi_n^2$$

   donde $n = \frac{1}{2}k(k-1)$ y T es el tamaño muestral.

   $In|R_{xx}| =$ El logaritmo natural del determinante de la matriz.

9. Regla de Klein[7] , está basada en regresiones auxiliares en las que se estima cada una de las variables explicativas en función de las demás. De esta forma es posible conocer, por una parte, el grado en que cada uno de los regresores está explicado por el resto, y por otra, tener una aproximación a la redundancia de información que se produce cuando todos participan conjuntamente en la explicación de la variable dependiente. Para ello se establece una comparación entre el coeficiente de determinación del modelo original, que incluye todos los regresores $\left(R^2_{y,x_1,\ldots,x_k}\right)$, y los coeficientes de determinación obtenidos en cada uno de las regresiones auxiliares $\left(R_j^2\right)$, en las que se habrá estimado el regresor X_j como función de todas las demás variables explicativas.

   Si $R^2_{y,x_1,\ldots,x_k} < R_j^2$ puede considerarse que la variable X_j contiene información ya incluida en el resto de los regresores, por lo que dicha variable es la causante de la multicolinealidad, para [Luis O, 1992] "a partir de estas regresiones auxiliares es posible obtener el factor de inflación de la varianza que adopta la siguiente expresión":

---

[5] Farrar, D.E. y R.R. Glauber, *"Multicollinearity in Regression Anallysis: The Problem Revisited"*, Review of Economics and Statistics, vol 49, 1967, pp. 92-107E.

[6] Una matriz Q de nxn se llama ortogonal si Q es invertible y $Q^{-1}=Q^t$, si esto se cumple entonces $Q^tQ=I$.

[7] Klein, L.R. *An Introduction to Econometrics*, Prentice-Hall, Englewood Cliffs, N:J. 1962, pp 64-101.

$$\frac{1}{\left(1-R_j^2\right)}=\frac{\sigma_{b_i}^2}{\sigma_{b_{ortogonal}}^2} \qquad (21)$$

Donde:

R_j^2 :Representa el coeficiente de determinación de las regresiones auxiliares.

$\sigma_{b_i}^2$:Es la varianza del coeficiente en el supuesto de multicolinealidad.

$\sigma_{b_{ortogonalidad}}^2$:Es la varianza del coeficiente en el supuesto de ortogonalidad o ausencia de multicolinealidad.

10. Medida de Theil, esta medida pueda calcularse a partir de la siguiente expresión:

$$m=R^2-\sum_{h=1}^{k}\left(R^2-R_h^2\right) \qquad (22)$$

Este estadístico que Theil denomina contribución marginal o incremental mide el efecto de contribución del regresor h, en el coeficiente de determinación R^2, por lo que es considerado como una de las medidas de la aportación del regresor en la explicación del regresando.

Siendo:

R_h^2 :El coeficiente de determinación excluido el regresor h.

R^2 :El coeficiente de determinación del modelo inicial.

En el caso de ortogonalidad el valor estadístico será próximo a cero.

1.2. Medidas remediales

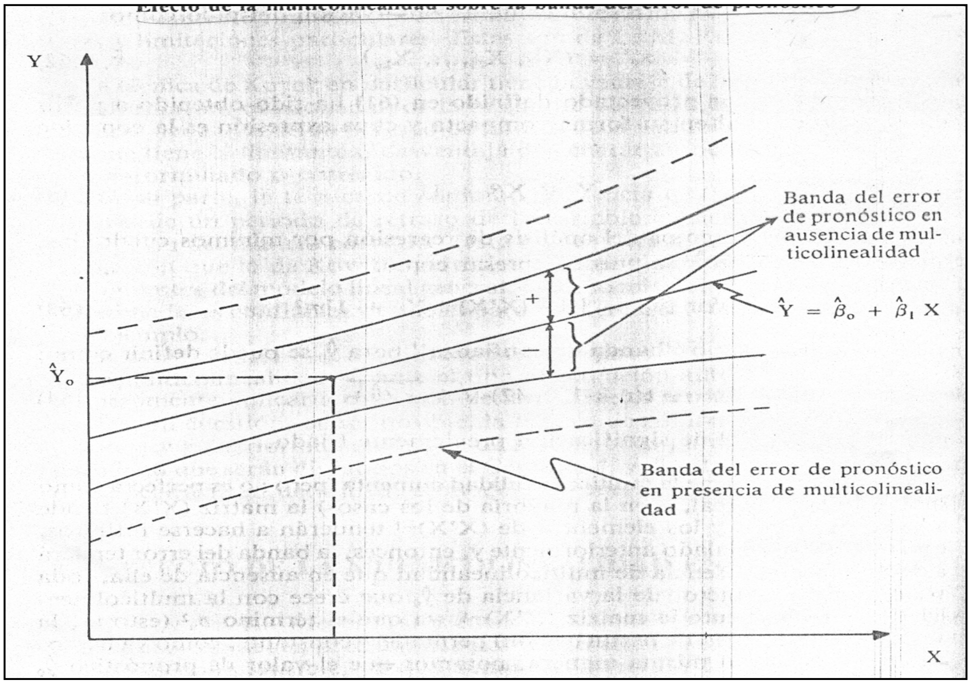
La técnica específica o procedimiento a aplicar, en cada caso para corregir o eliminar la multicolinealidad en el modelo dependerá de una serie de factores (que se mencionan a continuación) y de la naturaleza del problema en cuestión.

1. Aumentar el número de la muestra, esto es posible si se dispone de suficiente información. Este procedimiento es una buena alternativa, en particular si se hace que los grados de libertad (n-k) sean mayores a 30, con lo cual se estará reduciendo las varianzas de nuestros estimadores y como consecuencia aumentando la precisión de los mismos.

2. Incorporar alguna restricción (información *a priori*) sobre los parámetros en el modelo, de tal forma que los estimadores la satisfagan exactamente.

3. Conviene comentar que a pesar de todas las inconveniencias ocasionadas por la multicolinealidad en un modelo, en cuanto a la precisión y confiabilidad de los estimadores minimocuadráticos, si el objetivo del modelo es la predicción, este problema estadístico no representa una dificultad seria. Para quedar claro en esta idea la gráfica 4 muestra un problema en presencia y ausencia de multicolinealidad

Gráfica 4
Efecto de la multicolinealidad en la predicción



Fuente: Adoptado de Chiswick, B.R. y Chiswick, S.J. *Statistics and Econometrics: A problem soloing Text*. University Parc, 1975, p.168.

1.3. Soluciones

Con la intención de ilustrar la explicación del problema que nos representa la multicolinealidad, se citarán dos ejemplos, un ejercicio del curso: Econometría Aplicada impartido en la Facultad de Economía, en coordinación con el Centro de Educación Continua, dentro del programa de verano, actualización para profesores, el curso fue dictado por el Dr. Gerardo Esquivel y el segundo es un ejercicio que muestra la información sobre el sector minero.

La central camionera del sur desea estimar la demanda de boletos (DM_t) para el año 2000, en función del precio de los mismos (P_t) y de la calidad del servicio, evaluada a través de los gastos de la empresa realizada para la mejora del mismo (G_t), para ello se dispone de datos de los últimos 50 trimestres que aparecen en el anexo 1

Se dispone de otra estimación alternativa del parámetro correspondiente a P_t con valor de -0.5, obtenida a partir de otra muestra diferente a la tratada.

El modelo a estimar es el siguiente:

$DM_t = \beta_o + \beta_1 P_t + \beta_2 G_t + \mu_t$

Los resultados de la estimación por MCO es:

Cuadro 3
Regresión demanda de boletos en función de los precios y de la calidad en el servicio

LS // Dependent Variable is DM
Sample: 1986:2 1998:3
Included observations: 50 after adjusting endpoints

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	1310.944	365.4159	3.587540	**0.0008**
P	0.065659	2.519629	0.026059	**0.9793**
G	2.569635	28.03740	0.091650	**0.9274**

R-squared	0.813165	Mean dependent var	6424.275
Adjusted R-squared	0.805215	S.D. dependent var	1021.570
S.E. of regression	450.8645	Akaike info criterion	12.28046
Sum squared resid	9554102.	Schwartz criterion	12.39518
Log likelihood	-374.9584	F-statistic	**102.2797**
Durbin-Watson stat	2.205668	Prob(F-statistic)	0.000000

Donde se observa que ninguno de los coeficientes correspondientes a las variables explicativas (salvo el término constante), es significativo[8]. Una vez que se sabe si las variables independientes explican a la dependiente, es necesario interpretar el valor del coeficiente de las variables, así, por ejemplo el valor de 0.065659 lo entendemos como la propensión marginal y su interpretación es la siguiente, si se incrementa en una unidad P, se crece en 0.65 DM y así para la siguientes variables. Sin embargo el valor de F= 102.2797 lo que expresa que la significación estadística es alta, lo que quiere decir que la estimación de la regresión en general esta correcta o bien que al menos una variable independiente es diferente de las demás. Esto nos indica la posible existencia de multicolinealidad no exacta pero elevada entre las variables explicativas del modelo.

Cuadro 4
Regresión del precio en función de la calidad del servicio

LS // Dependent Variable is P
Sample: 1986:2 1998:3
Included observations: 50

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	1.528812	1.868187	0.818340	0.4172
G	0.089864	0.000107	841.7983	0.0000

R-squared	**0.999932**	Mean dependent var	1549.700
Adjusted R-squared	0.999931	S.D. dependent var	279.1337
S.E. of regression	2.321067	Akaike info criterion	1.723232
Sum squared resid	258.5928	Schwartz criterion	1.799713
Log likelihood	-112.0277	F-statistic	**708624.4**
Durbin-Watson stat	1.955028	Prob(F-statistic)	0.000000

[8] Quiere decir que existe una relación lineal entre la variable independiente con respecto a la dependiente o bien que la variable dependiente es explicada por la variable independiente.

También es posible realizar una regresión de las variables explicativas G_t y la constante para con P_t y comprobar si el R^2 es elevado. Al realizar la regresión de P_t sobre G_t se observa que su coeficiente es significativo y el R^2 tiene un valor de **0.9999**.
Al disponer de una estimación del coeficiente de una de las variables que causan el problema de multicolinealidad P_t a partir de otra muestra diferente a la tratada, se puede incorporar esta información y llevar a cabo la estimación del modelo restringido a dicha información adicional.
Es importante no perder de vista el valor del coeficiente de determinación y de la F estadística.

El modelo a estimar cuando incorporamos la información sería:

$DM_t - \beta_2 P_t = \beta_o + \beta_1 M_t + \mu_t$
con $\beta_2 = -0.5$

El estimador del vector de parámetros del modelo sujeto a la restricción $\beta_2 = -0.5$ es:

Se puede ver que en el cuadro 5 tanto el coeficiente de determinación como la F estadística ya no son tan grandes en comparación con el cuadro 4.

Recordar que si el R^2 tiende a uno o es uno la F estadística se va hacia infinito, por el contrario si el R^2 tiende a cero o es cero, entonces el modelo queda indeterminado.

Cuadro 5
Regresión de información adicional en función de la calidad en el servicio

LS // Dependent Variable is DA
Sample: 1986:2 1998:3
Included observations: 50

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	1314.1111	359.11380	3.6593165	0.0012
G	0.2516435	0.0205205	12.263033	**0.0000**

R-squared	**0.758043**	Mean dependent var	5649.425
Adjusted R-squared	0.753002	S.D. dependent var	897.7438
S.E. of regression	446.1690	Akaike info criterion	1.921212
Sum squared resid	85.45208	Schwartz criterion	1.671412
Log likelihood	-374.9613	F-statistic	**150.3820**
Durbin-Watson stat	2.201672	Prob(F-statistic)	0.000000

Donde $DA_t = DM_t - \beta_2 P_t$. Puede observarse que las estimaciones obtenidas para el coeficiente de G_t por esta vía, difieren de las obtenidas cuando se estima por MCO, siendo los parámetros estimados significativos

Otras soluciones alternas para tratar de solucionar el problema de multicolinealidad, es que si sólo se va a hacer previsión, una posible forma de solucionar el problema de multicolinealidad sería omitir una de las variables explicativas, ya que la información adicional que aporta a la explicación de la variable dependiente no es muy elevada.

De esta forma las estimaciones que se obtendrán presentarán los problemas propios de omitir una variable explicativa (estimación sesgada de los parámetros). Si se realiza la regresión omitiendo la variable P_t se tendría que realizar la estimación del cuadro 6.

Cuadro 6
Regresión de demanda de billetes en función de la calidad en el servicio

LS // Dependent Variable is DM
Sample: 1986:2 1998:3
Included observations: 50

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	1314.873	359.1252	3.661321	0.0006
G	0.296576	0.020521	14.45219	**0.0000**

R-squared	**0.813132**	Mean dependent var	6424.275
Adjusted R-squared	0.809239	S.D. dependent var	1021.570
S.E. of regression	446.1831	Akaike info criterion	12.24064
Sum squared resid	9555810.	Schwartz criterion	12.31712
Log likelihood	-374.9628	F-statistic	208.8658
Durbin-Watson stat	2.200697	Prob(F-statistic)	0.000000

Se observa que la estimación del parámetro β_1 es significativa y el coeficiente de determinación tiene un valor de 0.8131 o del 81.31%

Una razón que puede explicar la alta multicolinealidad entre el precio de los billetes y la calidad del servicio radica en que ambas tienden a moverse en el tiempo en la misma dirección. Para aminorar este problema, se puede estimar el modelo en diferencias[9], y se estimaría, por lo tanto el siguiente modelo:

$D_t - D_{t-1} = \beta_1 (G_t - G_{t-1}) + \beta_2 (P_t - P_{t-1}) + \mu_t$

donde $\mu_t = n_t - n_{t-1}$ $u_t = n_t - n_{t-1}$

$t-1$ representa un rezago o retardo de datos

Los resultados de la estimación se presentan en el cuadro 7:

Cuadro 7
Regresión de diferencias en función de la calidad en el servicio y de los precios

LS // Dependent Variable is D(DM)
Sample: 1986:3 1998:3
Included observations: 49 after adjusting endpoints

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	-167.0019	107.8887	-1.547909	**0.1285**
D(G)	2.547062	2.498905	1.019271	**0.3134**
D(P)	-15.82967	27.60757	-0.573382	**0.5692**

R-squared	0.254362	Mean dependent var	76.09512
Adjusted R-squared	0.221943	S.D. dependent var	703.1093
S.E. of regression	620.1954	Akaike info criterion	12.91934

[9] El estimar un modelo en diferencias es que la(s) variable(s) están siendo desfasadas un periodo de tiempo, con la intención de eliminar la relación estrecha que existe entre ellas.

Sum squared resid	17693545	Schwartz criterion	13.03516
Log likelihood	-383.0518	F-statistic	**7.846076**
Durbin-Watson stat	2.459251	Prob(F-statistic)	0.001170

Se observa entonces que ninguno de los coeficientes estimados es significativo, y el estadístico F es relativamente bajo.

Por lo tanto esta técnica de diferencias no es conveniente (para este ejemplo) porque no corrige el problema de multicolinealidad, por que ninguno de los coeficientes es significativo, el R^2 es de .25, el error estándar de los coeficientes es grande y la distribución F aunque es significativa es muy pequeña. La mejor estimación se presentó en el cuadro 4, 5 y 6 que son regresiones de una sola variable, es importante recordar que la multicolinealidad es una cuestión de grado y no de clase y por lo tanto se puede trabajar con multocolinealidad, tratando de que sea la menor posible.

Recordando además que la multicolinealidad siempre va a existir, porque no es posible separar la relación existente entre las variables, al menos en economía, porque una variable depende de la otra y esta a su vez de otra y así sucesivamente, hay que tener presente que el problema de multicolinealidad es un problema de muestra o sea de datos y los estimadores siguen siendo los mejores linealmente insesgados.

Es muy importante dar la interpretación del coeficiente, por ejemplo, en el cuadro 4 el resultado de la variable G (gasto) es de 0.089864 entonces se dice que a medida que se avanza una unidad en el eje de las x’s se incrementa en un 0.08 en el eje y, o bien la propensión marginal del gasto es de 0.08.

El segundo ejercicio cuyo objetivo es el análisis de la presencia de multicolinealidad cuya especificación es:

INGRESO = $\beta_1 + \beta_2$ CONSUMO + β_3 GASTOS + β_4 GEX + μ_i
Donde:
Ingreso = Al Ingreso de explotación de la industria minera
Consumo = El Consumo en bienes
Gasto = Al gasto de personal
GEX = Al gasto en explotación al sector minero

La utilización de estos regresores puede indicar, la presencia de multicolinealidad en el modelo. Este problema se detectará por distintas vías alternativas y finalmente se propondrá una solución para su corrección. La base de datos aparece en el anexo 2

a) Significancia global y significancia individual de los regresores
b) Análisis de la matriz de correlaciones de los regresores
c) Medida de Belsey, Kuck y Welsch
d) Contraste de Farrar-Glauber
e) Medida de Theil

Cuadro 8
Regresión del ingreso en función del consumo, del gasto y del gasto en exportaciones

LS // Dependent Variable is INGRESO
Sample: 1 17
Included observations: 17

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	3.729609	3.240869	1.150805	**0.2705**
CONSUMO	0.392848	0.276641	1.420067	**0.1791**
GASTO	0.663397	0.415687	1.595907	**0.1345**
GEX	0.648393	0.218018	2.974033	**0.0108**

R-squared	0.998971	Mean dependent var	235.0170
Adjusted R-squared	0.998734	S.D. dependent var	279.5921
S.E. of regression	9.949564	Akaike info criterion	7.635259
Sum squared resid	1286.920	Schwartz criterion	7.831309
Log likelihood	-60.89970	F-statistic	4207.203
Durbin-Watson stat	1.589402	Prob(F-statistic)	0.000000

Detección de la multicolinealidad

a) Significación global y significación individual de los regresores.
Se trata de un modelo globalmente bien estimado, el estadístico F nos indica que las variables conjuntamente son significativas. Por otra parte el valor del coeficiente de determinación R^2 es 0.9989 y al analizar la significación individual de las variables explicativas se puede comprobar que las variables consumo y gasto no son significativas en la explicación de los ingresos.

b) Análisis de la matriz de correlaciones de los regresores
A partir de la matriz de correlaciones de los regresores (R_{xx}) se puede profundizar en el proceso de detección de multicolinealidad en el modelo. Si los regresores fueran ortogonales, esto es, en ausencia de multicolinealidad, el determinante de esta matriz tomaría el valor de 1. A continuación se obtiene la matriz de correlaciones de las variables explicativas que se han empleado en esta estimación.

Cuadro 9
Matriz de correlación

	CONSUMO	GASTO	GEX
CONSUMO	1	0.990379	0.997769
GASTO	0.990379	1	0.995340
GEX	0.997769	0.995340	1

Se calcula el determinante que es 0.00033924, el valor que toma el determinante es muy próximo a cero, lo que también es indicativo de la posible existencia de multicolinealidad en el modelo. Se observa que la matriz de correlaciones refleja una fuerte correlación entre las variables que se han considerado en el modelo.

Los autovalores de la matriz (R_{xx})son:

λ =0.009850
λ =0.001152
λ =2.988990
A partir de los autovalores obtenidos se calcula el número de condición:

$$k(x) = \frac{\sqrt{\lambda_{máximo}}}{\sqrt{\lambda_{mínimo}}} = \sqrt{\frac{2.988990}{0.001152}} = 2594.609375 = 50.9373$$

Con el valor de k(x) es distinto de 1 entonces puede considerarse que existe multicolinealidad. Además, el número de condición k(x)=50.9373, toma un valor para el que puede asumirse que existe multicolinealidad “fuerte”.

d) Contraste de Farrar-Glauber

Este contraste plantea $H_0: |R_{xx}| = 1$ ortogonalidad de los regresores.

El estadístico calculado o experimental es: $G = -\left[T - 1 - \frac{1}{6}(2k+5\right]\ln|R_{xx}|$

$G \sim \chi_n^2$

donde: $n = k * \frac{k-1}{2}$

Como son conocidos todos los valores, puede calcularse sustituyendo;

$$G = -\left[17 - 1 - \frac{1}{6}\left(\left(2*4\right)+5\right)\right] * In\left(0.000033924\right) = 142.364198$$

Se calcula la probabilidad que queda a la derecha de este punto bajo una distribución χ^2 con 6 grados de libertad y 95% de nivel de confianza(valor crítico 12.5916), por lo tanto, no se acepta la hipótesis nula, lo que confirma que existe ortogonalidad entre los regresores.

e) Factor de inflación de la varianza

Para desarrollar esta metodología se requiere, previamente, realizar ciertas regresiones auxiliares, en concreto, se considera cada uno de los regresores como función de los demás. De cada una de estas regresiones se necesita el valor de R^2 y, a partir de ellos, pueden calcularse los aumentos de las varianzas causadas por la multicolinealidad midiendo el factor de inflación de la varianza, según la siguiente expresión.

$$\frac{\sigma_{b_i}^2}{\sigma_{b_{ortogonal}}^2} = \frac{1}{1 - R_j^2}$$

Cuadro 10
Regresión del consumo en función del gasto y del gasto en exportaciones

LS // Dependent Variable is CONSUMO
Sample: 1 17
Included observations: 17

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	2.079209	3.081281	0.674787	0.5108
GASTO	-0.639877	0.363360	-1.760998	0.1001
GEX	0.709049	0.091938	7.712272	0.0000

R-squared	**0.996352**	Mean dependent var	125.4016
Adjusted R-squared	0.995831	S.D. dependent var	148.8619
S.E. of regression	9.612225	Akaike info criterion	7.522734
Sum squared resid	1293.528	Schwartz criterion	7.669771
Log likelihood	-60.94324	F-statistic	1911.709
Durbin-Watson stat	1.937222	Prob(F-statistic)	0.000000

Para calcular el factor de inflación de la varianza :
1-/(1-0.996352) = 274.122

Ahora se estima la segunda regresión auxiliar.

Cuadro 11
Regresión del Gasto en función del Consumo y del gasto en exportaciones

LS // Dependent Variable is GASTO
Sample: 1 17
Included observations: 17

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	-0.182041	2.083114	-0.087389	0.9316
CONSUMO	-0.283398	0.160930	-1.760998	0.1001
GEX	0.407116	0.088371	4.606894	0.0004

R-squared	**0.992389**	Mean dependent var	55.46576
Adjusted R-squared	0.991302	S.D. dependent var	68.59110
S.E. of regression	6.396967	Akaike info criterion	6.708310
Sum squared resid	572.8966	Schwartz criterion	6.855348
Log likelihood	-54.02064	F-statistic	912.7659
Durbin-Watson stat	1.812016	Prob(F-statistic)	0.000000

Se procede al cálculo del factor de inflación de la varianza:
1/(1-0.992389) = 131.3887

Se estima la tercer regresión auxiliar

Cuadro 12
Regresión del gasto en exportaciones función del consumo y del gasto

LS // Dependent Variable is GEX
Sample: 1 17
Included observations: 17

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	-1.270813	3.958333	-0.321048	0.7529
CONSUMO	1.141627	0.148027	7.712272	0.1791
GASTO	1.480014	0.321261	4.606894	0.0004

R-squared	**0.998229**	Mean dependent var	223.9812
Adjusted R-squared	0.997976	S.D. dependent var	271.0886
S.E. of regression	12.19685	Akaike info criterion	7.999018
Sum squared resid	2082.685	Schwartz criterion	8.146056
Log likelihood	-64.99165	F-statistic	3945.000
Durbin-Watson stat	2.131833	Prob(F-statistic)	0.000000

Se procede al cálculo del factor de inflación de la varianza:
1/(1-0.998229) = 564.6527

Los factores multiplicativos alcanzan valores elevados en los tres casos, por lo que puede considerarse que la varianza de los coeficientes del modelo original se encuentra muy inflada, por lo que se toman como no significativos regresores que si podrían serlo.

A partir de estas mismas regresiones se podría calcular la medida de Klein, que compara los R_i^2 con el R^2 de la regresión original.

$$R^2_{i(Consumo)} = 0.9963\,;\ R^2_{i(Gasto)} = 0.9923\,;\ R^2_{i(Gex)} = 0.9982$$

Como los valores que toman estos coeficientes de determinación son inferiores, pero muy próximos, al de la regresión inicial (0.998971), puede considerarse que, aunque no a un nivel preocupante, sí existe multicolinealidad entre los regresores.

f) Medida de Theil
Esta medida refleja la diferencia entre la variablilidad explicada por el modelo completo y la suma de cada una de las aportaciones de los regresores al modelo.

$$m = R^2 - \sum(R^2 - R_h^2)$$

donde:

R_h^2 : Es el coeficiente de determinación de una regresión auxiliar en la que se excluye el regresor X_h.

R^2 : Es el coeficiente de determinación del modelo completo.

Debe realizarse en primer lugar las regresiones auxiliares necesarias para calcular esta medida, excluyendo en cada uno de los casos un regresor y calculando su aportación al modelo completo.

Cuadro 13
Regresión del ingreso en función del gasto y del gasto en exportaciones

LS // Dependent Variable is INGRESO
Sample: 1 17
Included observations: 17

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	4.546423	3.303186	1.376375	0.1903
GASTO	0.412023	0.389529	1.057746	0.3081
GEX	0.926941	0.098559	9.404948	0.0000

R-squared	**0.998811**	Mean dependent var	235.0170
Adjusted R-squared	0.998642	S.D. dependent var	279.5921
S.E. of regression	10.30447	Akaike info criterion	7.661818
Sum squared resid	1486.549	Schwartz criterion	7.808855
Log likelihood	-62.12545	F-statistic	5882.638
Durbin-Watson stat	1.501813	Prob(F-statistic)	0.000000

La variable consumo aporta 0.998971 – 0.998811 = 0.00016 un valor muy próximo a cero.

Cuadro 14
Regresión del ingreso en función del consumo y del gasto en exportaciones

LS // Dependent Variable is INGRESO
Sample: 1 17
Included observations: 17

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	3.608844	3.414296	1.056980	0.3084
CONSUMO	0.204843	0.263770	0.776595	0.4503
GEX	0.918472	0.144843	6.341147	0.0000

R-squared	**0.998769**	Mean dependent var	235.0170
Adjusted R-squared	0.998594	S.D. dependent var	279.5921
S.E. of regression	10.48485	Akaike info criterion	7.696525
Sum squared resid	1539.049	Schwartz criterion	7.843562
Log likelihood	-62.42046	F-statistic	5681.733
Durbin-Watson stat	1.505478	Prob(F-statistic)	0.000000

La variable Gasto aporta 0.998971 – 0.998769 = 0.000202 también muy próximo a cero.

Cuadro 15
Regresión del ingreso en función del consumo y del gasto

LS // Dependent Variable is INGRESO
Sample: 1 17
Included observations: 17

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	2.905623	4.033475	0.720377	0.4831
CONSUMO	1.133071	0.150837	7.511871	0.0000
GASTO	1.623027	0.327359	4.957940	0.0002

R-squared	**0.998271**	Mean dependent var	235.0170
Adjusted R-squared	0.998024	S.D. dependent var	279.5921
S.E. of regression	12.42839	Akaike info criterion	8.036629
Sum squared resid	2162.507	Schwartz criterion	8.183666
Log likelihood	-65.31134	F-statistic	4041.651
Durbin-Watson stat	2.027039	Prob(F-statistic)	0.000000

La variable Gex aporta 0.998971 – 0.998271 = 0.0007 un valor muy próximo a cero.

Calculando la medida de Theil se obtiene que este valor no dista mucho del coeficiente de determinación de la regresión completa, por lo que puede admitirse que existe multicolinealidad (si los regresores fuesen ortogonales la medida *m* sería cero).

Para calcular la medida de Theil, se necesita hacer lo siguiente:
0.998971-(0.00016+0.000202+0.0007) = 0.997909

1.4. Resumen de Multicolinealidad[10]

Es la existencia de una relación lineal perfecta o exacta entre algunas o todas las variables explicativas de un modelo de regresión.

Causas
Muestras pequeñas n<k
Restricción en el modelo o en la población objeto de muestreo
Incorrecta especificación del modelo
Modelo sobreestimado
Unidades de medición (método de recolección utilizado)
Estimación del modelo en diferencias

Síntomas
R^2 elevada o alta
t’s no significativas
σ elevada o grande

[10] Se le atribuye este término a Ragnar Frisch, Statistical Confluence Análisis by Jeans of Complete Regresión System, Institute of Economics, Oslo University, núm. 5, 1934.

Altas correlaciones entre parejas de regresores
Matriz de correlación $\longrightarrow$ 0

Pruebas para detectar
Valor del determinante de correlaciones entre los regresores próximos a cero
Regla de Farrar y Glauber o Pruebas auxiliares

$\frac{R_i}{R_j} \geq 1$ Existe multicolinealidad en el modelo

o bien

$$F_i = \frac{R_i^2/(k-1)}{(1-R_i^2)/(n-k+1)}$$

Regla de Klein o Factor de inflación de varianza

Si $\frac{1}{(1-R_j^2)} < 10$ No hay multicolinealidad

Si $\frac{1}{(1-R_j^2)} > 10$ Existe multicolinealidad

Prueba de Theil

$$m = R^2 - \sum_{h=1}^{k}(R^2 - R_h^2)$$

Si Theil $\longrightarrow$ 0 No hay multicolinealidad
Si Theil $\longrightarrow$ 1 Existe multicolinealidad
Matriz de correlación

$\lambda X_i X_j < R^2 \longrightarrow$ No hay multicolinelidad

$\lambda X_i X_j > R^2 \longrightarrow$ Existe multicolinealidad

Consecuencias
Varianzas y covarianzas grandes.
Intervalos de confianza amplios, aceptación fácil de la H_0.
Razones t's no significativas.
Los estimadores MCO y sus errores estándar son sensibles a pequeños cambio den la información.

Corrección
No hacer nada
Información a priori
Combinación de información de corte transversal y de series de tiempo
Eliminación de una(s) variable(s) y el sesgo de especificación
Transformación de variables
Introducir datos nuevos o adicionales
Reducción de colinealidad en las regresiones polinomiales

2. HETEROSCEDASTICIDAD

Uno de los supuestos del modelo de regresión lineal clásico [Gujarti, 2004] dice "es la varianza de cada término de perturbación μ_i, condicional a los valores seleccionados de las variables explicativas, es algún número constante igual a δ^2. Éste es el supuesto de homoscedasticidad, o igual (homo) *dispersión* (cedasticidad), es decir, *igual varianza".*
Si la varianza de las perturbaciones no es constante a lo largo de las observaciones, la regresión es heteroscedástica, esta circunstancia provoca en principio, una dificultad respecto a la estimación del modelo.

2.1. Como detectar heteroscerdasticidad

Generalmente se supone, por el tipo de datos utilizados, o por el tipo de análisis efectuado, que el modelo con el que se trabaja puede presentar heteroscedasticidad.

Lo que se debe hacer es estimar el modelo por mínimos cuadrados, y con los residuos obtenidos (como estimación de las verdaderas $\boldsymbol{\mu_i}$ desconocidas) se procede a un análisis gráfico, para el caso en que la varianza de los mismos es función de ***X***, de forma que al aumentar la variable independiente se incrementa también la varianza de los residuos, la representación gráfica será parecida a la figura 8; en el caso de que la relación entre la varianza y ***X*** sea inversa, aparecería una figura como la número 9; después se aplicará un contraste de hipótesis estadístico, con la finalidad de complementar el primer análisis de tipo indicativo.

Figura 8

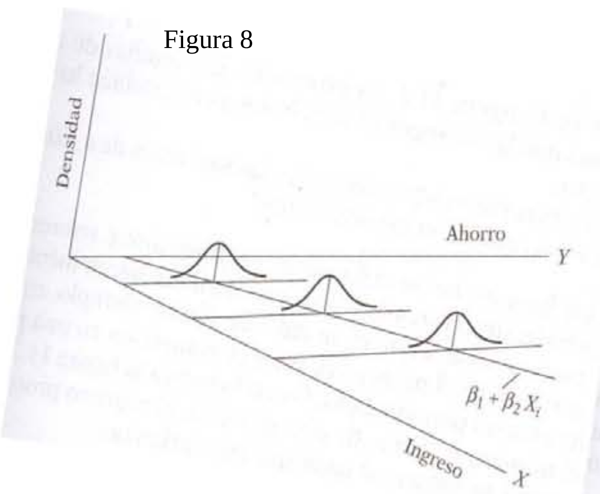
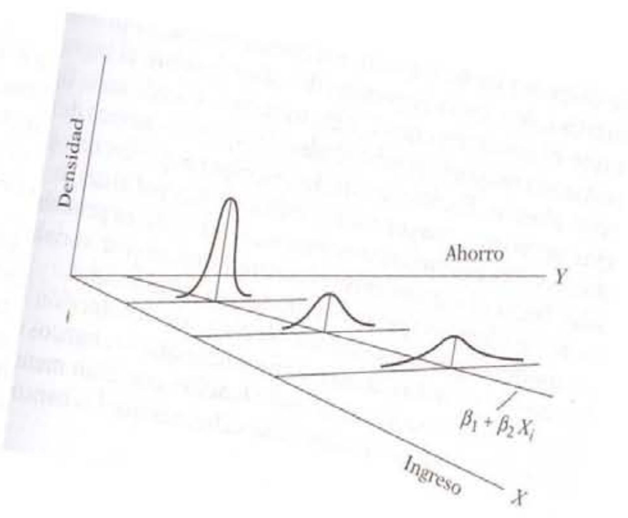


Figura 9



El análisis gráfico se basa en la representación de los residuos mínimo cuadráticos en función de la variable endógena. En caso de que se posea una información *a priori* sobre la posibilidad de que alguna de las variables explicativas del modelo provoque el problema de la heterocedasticidad, entonces se representará los residuos en función de dichas variables explicativas.

Los contrastes que se pueden utilizar para detectar heteroscedasticidad son de dos tipos:

- Los que se basan en las hipótesis establecidas, que se les denomina parámetricos. Se examinaran los contrastes Goldfeld - Quandt, Breusch-Pagan, Glejser y de White correspondientes a este tipo.

- Los no parámetricos, que no tienen en cuenta el supuesto de normalidad de las perturbaciones. A este tipo pertenecen los contrastes de picos y de rangos que se presentarán concepto y fórmulas pero no se desarrollara ejercicio para estos dos métodos. Si se desea profundizar en estos métodos revisar el texto de [Uriel E. y Contreras D, 1993].

2.2. Contraste de Goldfeld Quandt

Goldfeld Quandt suponen que a lo largo de la muestra, las perturbaciones tienen varianzas que varían de acuerdo con los valores que tome una determinada variable explicativa o una combinación de estas. Las varianzas de las perturbaciones serán pequeñas para los valores más pequeños de la variable explicativa, y grandes para los valores altos.

Los pasos a seguir en la aplicación del contraste de Goldfeld Quandt

1. Se ordenan las observaciones en forma decreciente de la variable explicativa que está ligada a la heteroscedasticidad.
2. Se omiten un número C de observaciones centradas y se dividen las N-C observaciones restantes en dos grupos de (T-C)/2 observaciones cada uno.
3. Se realizan regresiones separadas para cada uno de los grupos creados y se obtienen las sumas de cuadrados de residuos de las dos regresiones, SRC_1 que corresponden a la regresión sobre los valores más pequeños, y SRC_2 que corresponde a la de los valores más grandes. Los residuos de los grupos tienen (N-C-2k)/2 grados de libertad.

Donde:
C es el número de observaciones centrales eliminadas y
k el número de regresores del modelo.
T es el número de observaciones.
N número total de observaciones de cada grupo.

4. Con los residuos de las dos regresiones se calcula el estadístico F que esta dado por

$$F_{(T-C-2k)/2,(T-C-2k)/2} = \frac{SRC_2/(N-C-2k)/2}{SRC_1/(N-C-2k)/2} \qquad (23)$$

Si se mantiene la hipótesis de normalidad, y bajo la hipótesis nula de que las perturbaciones son homocedásticas, entonces la fórmula (16) se distribuye como una F de Snedecor con (T-C-2k)/2 grados de libertad para el numerador y denominador.
La hipótesis nula H_0, (*es la que se desea aceptar*) que se contrasta es

$H_o\!:\sigma_1^2 = \sigma_2^2 = ... = \sigma_N^2$ Existe homoscedasticidad
$H_1 : \sigma_1^2 \neq \sigma_2^2 \neq ... \neq \sigma_N^2$ No existe homoscedasticidad

El hecho de efectuar dos regresiones separadas para llevar a cabo la prueba se debe principalmente a dos razones:

a) En tratar de que los valores grandes y pequeños de la(s) variable(s) a la que presumiblemente esta asociada la heteroscedasticidad estén lo más separadas posible, a efectos de captar comportamientos diferentes de las varianzas de las perturbaciones.

b) En que una restricción estadística necesaria para poder aplicar la F ya que para esta expresión tenga efectivamente una distribución F, es que las formas cuadráticas que aparecen en el numerador y en el denominador sean independientes entre sí, esta independencia esta asegurada si se efectúan dos regresiones separadas.

NOTA: Se omite un número C de observaciones. En la práctica profesional se aconseja que sea $\frac{1}{3} < C < \frac{1}{5}$, la elección de C esta basada en un compromiso, ya que, cuanto mayor sea, más se acentuarán las diferencias entre los dos grupos, pero al mismo tiempo disminuirá la potencia del contraste.

2.3. Contrastes de Glejser

Los pasos a seguir para explicar este contraste son:

1. Se realiza la estimación de mínimos cuadrados ordinarios y se calculan los residuos.

2. Una vez evaluado para cada observación, se realiza una regresión en la que los valores absolutos de los residuos aparecen como variable dependiente, y como independiente la variable que se cree causante de la heterocedasticidad.

3. El análisis de la significación estadística de los estimadores de los parámetros puede ser considerado como un test de heterocedasticidad. Si los parámetros resultan ser significativos entonces existirá heteroscedasticidad.

2.4. Contraste de White

Este contraste no depende, en principio, de la hipótesis que se haya efectuado sobre la naturaleza de la heteroscedasticidad y en este sentido puede calificarse de robusto. Los pasos a seguir para efectuar la prueba de White:

1. Se aplica la regresión por el método de mínimos cuadrados al modelo original, calculándose los correspondientes residuos.
2. Se calcula la regresión de los cuadrados de los residuos del modelo original sobre los cuadrados y sobre todos los productos cruzados de los regresores del modelo, o sea:

$$\hat{\mu}_t = \beta_0 + \sum_{i=1}^{k} \sum_{j \geq i} \beta_{ij} \psi_{ijt} + \mu_{ii}$$

donde

$\hat{\mu}_t =$ residuos obtenidos al aplicar mínimos cuadrados al modelo original

k = número total de regresores del modelo original

$\psi_{ijt} = X_{it} \,.\, X_{jt}$, siendo X_{it} y X_{jt} regresores del modelo original

En la regresión auxiliar $\hat{\mu}_t = \beta_0 + \sum_{i=1}^{k} \sum_{j \geq i} \beta_{ij} \psi_{ijt} + \mu_{ii}$ se incluye el término independiente, aun cuando el modelo original no lo tenga; por ello se incluye de forma expresa un término independiente.

Bajo la hipótesis nula de homoscedasticidad se verifica que

N $R^2 \sim \chi^2_m$

Donde:

R^2 = Coeficiente de determinación en la regresión auxiliar
N = Tamaño de la muestra
m = Número de parámetros, excluyendo el término independiente de la regresión auxiliar.
Si k es el número de parámetros del modelo original que incluye término independiente y no existen otras redundancias, entonces resultará que: $m = \frac{k(k+1)}{2} - 1$

Si NR^2 es mayor que χ^2 de las tablas para un nivel de significación determinado, se rechaza la hipótesis nula de homoscedasticidad.
Al aplicar esta prueba se debe tener en cuenta que la distribución dada N $\overline{R}^2 \sim \chi^2_m$ solamente es válida para muestras grandes.

Cuando las perturbaciones son heteroscedásticas y se conoce el esquema que se sigue, la varianza de las perturbaciones, la solución consiste en aplicar Mínimos Cuadrados Generalizados, en este procedimiento se transforma el modelo para que las perturbaciones sean homoscedásticas y se apliquen mínimos cuadrados ordinarios al modelo transformado.
Sea el modelo

$$Y_i = \beta_1 + \beta_2 X_{2t} + ... + \beta_k X_{kt} + \mu_t$$

donde

$$E(\mu_t^2) = \sigma_t^2 = \lambda X_{jt}$$

siendo λ un factor de proporcionalidad.

El modelo se transforma dividiendo ambos miembros por la raíz cuadrada de X_{jt} o sea:

$$\frac{Y_t}{\sqrt{X_{jt}}} = \frac{\beta_1}{\sqrt{X_{jt}}} + \beta_2 \frac{X_{2t}}{\sqrt{X_{jt}}} + ... + \beta_k \frac{X_{kt}}{\sqrt{X_{jt}}} + \frac{\mu_t}{\sqrt{X_{jt}}}$$

El término de perturbación del modelo transformado es homoscedástico ya que

$$E\left(\frac{\mu_t}{\sqrt{X_{jt}}}\right)^2 = \frac{1}{X_{jt}} E(\mu_t^2) = \frac{1}{X_{jt}} \lambda X_{jt} = \lambda$$

En la practica el problema que se presenta es que no se conoce exactamente el esquema de la heteroscedasticidad; en $E(\mu_t^2) = \sigma_t^2 = \lambda X_{jt}$ se ha supuesto que la varianza de las perturbaciones es proporcional al valor de X_{jt}, pero también había podido suponerse que dicha varianza es proporcional a χ^2_{jt}, ***lnX$_t$***. Para resolver el problema, Glejser[11] sugiere que se calculen distintas regresiones en las que se toman como variable endógena los residuos del modelo original y como variable explicativa distintas formas funcionales de la variable a la que está asociado el esquema de heterocedasticidad, tales como:

$$|\hat{\mu}_t| = \beta_1 + \beta_2 X_{jt}^{1/2} + \mu_I$$

$$|\hat{\mu}_t| = \beta_1 + \beta_2 LnX_{jt} + \mu_I$$

[11] Glejser H.R. (1969) *A new Test for Heteroscedasticity*, Journal of the American Statistical Association 64.

$$\left|\hat{\mu}_t\right| = \beta_1 + \beta_2 \frac{1}{X_{jt}} + \mu_I$$

Se toma como base del esquema de heteroscedasticidad aquella regresión en la que se obtenga un R^2 más elevado. En caso de que en la regresión el coeficiente β_2 no fuera significativo, sería indicativo de que no existe heteroscedasticidad asociada a la variable X_{jt} .

Dado que $\left|\hat{\mu}_i\right|$ varía de forma aproximadamente proporcional a σ_t el modelo original se transforma dividiendo por los valores ajustados de la ecuación que se ha seleccionado. En caso de que el término independiente no fuera significativo, la transformación consistiría en dividir el modelo por el regresor. Por ejemplo si la regresión seleccionada fuera $\left|\hat{\mu}_t\right| = \beta_1 + \beta_2 \ln X_{jt} + \mu_i$ y el coeficiente β_1 no fuera significativo, se dividiría el modelo original por ln X_{jt}

2.5. Soluciones

El siguiente ejemplo brinda la oportunidad de trabajar los contrastes que anteriormente se explicaron, el banco pretende elaborar un modelo de regresión lineal con el cual explicar CM_t, el nivel de consumo mensual medio por familia en términos nominales a través de su YM_t ingreso medio e IP_t el índice de precios al consumidor

Se desea estimar el modelo $CM_t = \beta_o + \beta_1 YM_t + \beta_2 IP_t + \mu_t$

Se dispone de 50 observaciones generadas desde el mes de enero de 1995 a febrero de 1999. Se pueden consultar en el anexo 3.

Se sabe que el término de error del modelo esta relacionado con la variable YM_t por lo que pueden surgir problemas de heteroscedasticidad, para detectarla es necesario aplicar el contraste de Goldfeld y Quandt, Breusch-Pagan, Glejser y White.

Si se sospecha que la varianza del término de error depende positivamente de la variable YM_t $\sigma^2\mu_t = f(YM_t)$, donde los valores mayores de σ_t^2 ocurren en observaciones en la que YM_t es grande YM_t entonces el contraste de Goldfeld y Quandt, consistirá en lo siguiente:

Cuadro 16
Regresión de consumo mensual en función del ingreso y del índice de precios
(para probar el contraste Goldfeld y Quandt)

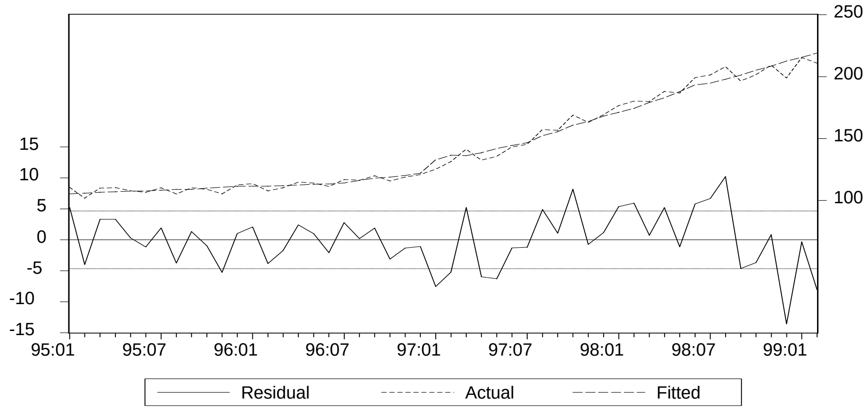
Estimada la regresión con las 50 observaciones
LS // Dependent Variable is CM_t
Sample: 1995:01 1999:02
Included observations: 50

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	51.13592	29.70301	1.721574	0.0917
YM_t	0.722600	0.069800	10.35238	0.0000
IP_t	-0.334157	0.364018	-0.917969	0.3633

R-squared	0.985286	Mean dependent var	144.0565
Adjusted R-squared	0.984660	S.D. dependent var	37.54941
S.E. of regression	4.650676	Akaike info criterion	3.132150
Sum squared resid	1016.553	Schwartz criterion	3.246871
Log likelihood	-146.2507	F-statistic	1573.630

Durbin-Watson stat	1.725384	Prob(F-statistic)	0.000000

Gráfica 5
Gráfica de residuales del cuadro 8



En la gráfica 5 se observa que los residuos tienen el comportamiento similar de diciembre de 1998, para 1999 se observa un movimiento descendente que provoca que los residuos salgan de las bandas y esto trae como consecuencia que las varianzas crezcan, por lo que se sospecha, que es un indicio de que existe heteroscedasticidad.

Ahora con la matriz de varianzas y covarianzas se observan que éstas no son tan grandes, pero no son relativamente pequeñas, por lo que se pueden considerar que existe presencia de heteroscedasticidad.

Cuadro 17
Matriz de covarianzas

	C	INGRESO	IPC
C	882.268614804	2.0126702559	-10.7816374673
INGRESO	2.0126702559	0.00487209381706	-0.0250456106484
IPC	-10.7816374673	-0.0250456106484	0.132509052901

A continuación se realizan el paso 2 y 3 del contraste, en esta regresión se trabaja con una muestra de 20 observaciones obteniendo los siguientes resultados.

Cuadro 18
Regresión para el período 1995.01 al 1996.08, nivel de consumo en función del ingreso y del índice de precios, se llamará a esta primera regresión (muestra 1)

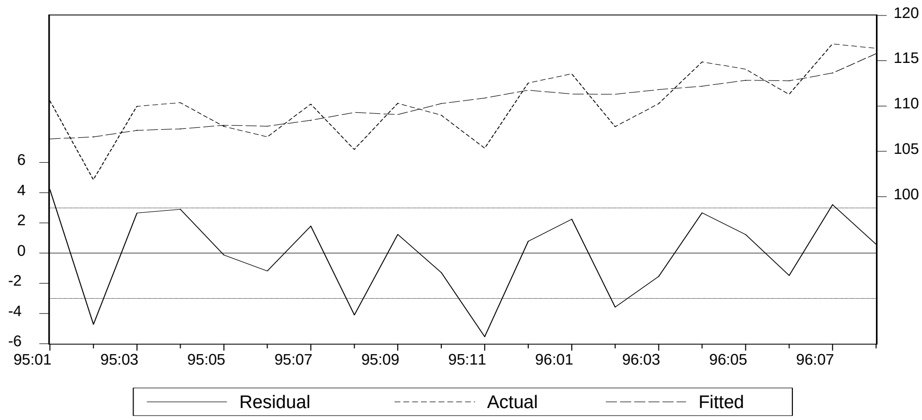
LS // Dependent Variable is CM_t
Sample**: 1995:01 1996:08**
Included observations: 20

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	83.54437	90.28387	0.925352	0.3677
YM_t	0.954588	1.178316	0.810129	0.4291
IP_t	-0.929724	2.298486	-0.404494	0.6909

R-squared	0.456946	Mean dependent var	110.2375
Adjusted R-squared	0.393057	S.D. dependent var	3.848347

S.E. of regression	2.998114	Akaike info criterion	2.333448
Sum squared resid	**152.8076**	Schwartz criterion	2.482807
Log likelihood	-48.71325	F-statistic	7.152218
Durbin-Watson stat	2.464310	Prob(F-statistic)	0.005574

Gráfica 6 (muestra 1)



Cuadro 19

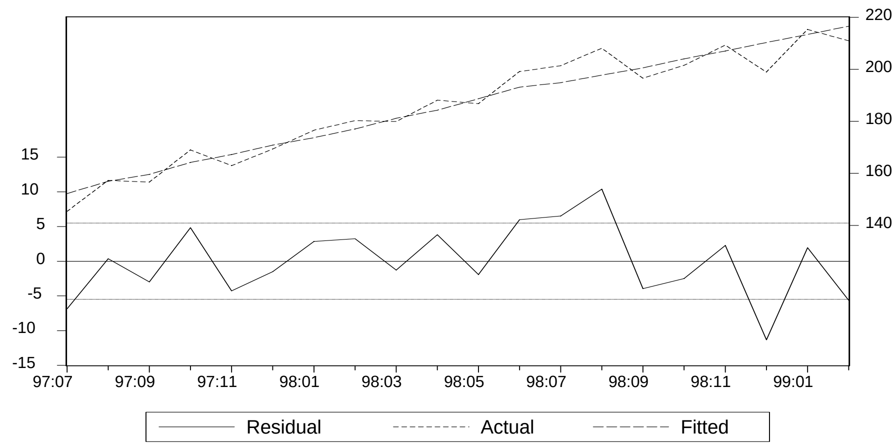
Regresión para el período 1997.07 al 1999.02, nivel de consumo en función del ingreso y del índice de precios, a esta segunda regresión se le llamará (muestra 2)

LS // Dependent Variable is CM_t
Sample: **1997:07 1999:02**
Included observations: 20

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	23.58319	235.5272	0.100129	0.9214
YM_t	0.569763	0.404417	1.408851	0.1769
IP_t	0.183243	2.617527	0.070006	0.9450

R-squared	0.935826	Mean dependent var	185.6604
Adjusted R-squared	0.928276	S.D. dependent var	20.56537
S.E. of regression	5.507667	Akaike info criterion	3.549763
Sum squared resid	**515.6848**	Schwartz criterion	3.699123
Log likelihood	-60.87640	F-statistic	123.9529
Durbin-Watson stat	2.016514	Prob(F-statistic)	0.000000

Gráfica 7 (muestra 2)



Ahora, dadas las sumas residuales de ambas regresiones, SRC_1 y SRC_2, entonces bajo la hipótesis nula de homoscedasticidad, el cociente es:

$$F = \frac{SRC_2}{SRC_1} = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2}$$

Sigue una distribución F, en nuestro problema se obtiene que $F = \frac{515.6848}{152.8076} = 3.374733$

Los valores de tablas al 95% y 99% respectivamente son:
$F=_{17,17}$ =2.27
$F=_{17,17}$ = 3.26

Por lo tanto no se acepta la hipótesis nula de que hay homoscedasticidad (rechazamos la hipótesis nula de que la varianza del término de error no depende positivamente de la variable YM_t).

Para realizar el contraste de Breusch-Pagan, se debe estimar por MCO el modelo original y obtener los errores correspondientes.

Cuadro 20
Serie de residuos del modelo original

1986.01	5.208694	-5.266702	1.912264	-1.210043	-1.132792
1986.02	-3.981123	1.020547	-3.112331	4.890235	5.796758
1986.03	3.320239	2.069169	-1.342693	1.065004	6.656564
1986.04	3.324172	-3.826185	-1.066957	8.169602	10.205958
1986.05	0.301454	-1.729209	-7.542756	-0.737035	-4.627312
1986.06	-1.147069	2.415285	-5.229079	1.175368	-3.678796
1986.07	1.932835	0.983452	5.234702	5.350779	0.831388
1986.08	-3.743546	-2.074574	-5.978817	5.917311	-13.583885
1986.09	1.335281	2.767031	-6.281586	0.719636	-0.330302
1986.10	-0.983800	0.181836	-1.306597	5.203771	-8.076147

Ahora se estima la regresión del cuadrado de los residuos anteriores sobre las variables de las cuales se piensa que depende la varianza del término de error, bajo la hipótesis nula de homoscedasticidad, el estadístico $\mu = TR^2$ donde R^2 de la última regresión se distribuye como una χ^2, siguiendo **p** el número de variables de las cuales se cree que depende la varianza del término de error (en nuestro caso P=1)

Los resultados de la regresión del cuadrado de los residuos calculados son:

Cuadro 21
Regresión de residuos al cuadrado (de la regresión original, cuadro 17) en función del ingreso

LS // Dependent Variable is **Residuo^2**
Sample: 1995:01 1999:02
Included observations: **50**

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	-25.67315	13.95684	-1.839468	0.0720
YM_t	0.252579	0.073250	3.448164	0.0012

R-squared	**0.198528**	Mean dependent var	20.33106
Adjusted R-squared	0.181831	S.D. dependent var	32.03578
S.E. of regression	28.97723	Akaike info criterion	6.772199
Sum squared resid	40304.63	Schwartz criterion	6.848680
Log likelihood	-238.2519	F-statistic	11.88983
Durbin-Watson stat	2.588052	Prob(F-statistic)	0.001184

Por lo que el estadístico μ toma el valor de :
μ =50*0.198528=9.9264

Dado que los valores de las tablas para 95% y 99% son: respectivamente:
$\chi_1^2 = 3.84$
$\chi_1^2 = 6.63$
Se concluye, la no aceptación de la hipótesis nula de homoscedasticidad

Dado que la pregunta del problema muestra que μ_t varía con YM_t, entonces es lógico pensar que σ_u^2 varía con YM_t^2 . Por este motivo será conveniente realizar también el contraste de Breusch-Pagan relacionando $\hat{u}_t^2$ con una constante y YM_t^2. El resultado de la regresión es el siguiente:

Cuadro 22
Regresión de residuos al cuadrado al cuadrado en función del ingreso al cuadrado

LS // Dependent Variable is **Residuo^2**
Sample: 1995:01 1999:02

Included observations: **50**				
Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	-3.082717	7.755143	-0.397506	0.6928
YM^2	0.000645	0.000182	3.548245	0.0009

R-squared	**0.207791**	Mean dependent var	20.33106
Adjusted R-squared	0.191286	S.D. dependent var	32.03578
S.E. of regression	28.80930	Akaike info criterion	6.760575
Sum squared resid	39838.85	Schwartz criterion	6.837056
Log likelihood	-237.9613	F-statistic	12.59004
Durbin-Watson stat	2.618150	Prob(F-statistic)	0.000879

El estadístico μ tomó en esta caso el valor:

μ = 50 * 0.20779= 10.3895

Dado que los valores de las tablas para 95% y 99% son respectivamente:

$\chi_1^2 = 3.84$

$\chi_1^2 = 6.63$ por lo que seguimos no aceptando la hipótesis nula de homoscedasticidad

Para realizar el contraste de Glejser se sigue un proceso similar al de Breusch-Pagan.

Se estima por MCO el modelo original obteniéndose los residuos correspondientes. Continuando con el cálculo de la regresión del valor absoluto de dichos residuos sobre la variable Y_t^α, se toman como valores del exponente α= -1, -½, 1, ½. Habrá que elegir aquel α que proporciones una mejor regresión (coeficientes significativos y suma residual pequeña).

Cuadro 23

Regresión de los residuos en valor absoluto (del cuadro 8) en función del ingreso para α=-1

LS // Dependent Variable is **RESIDUOS VALOR ABSOLUTO** Sample: 1995:01 1999:02 Included observations: 50				
Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	7.428527	1.418926	5.235315	0.0000
YM^{-1}	**-653.1169**	228.5984	-2.857050	0.0063

R-squared	0.145341	Mean dependent var	3.519573
Adjusted R-squared	0.127535	S.D. dependent var	2.847064
S.E. of regression	2.659324	Akaike info criterion	1.995322
Sum squared resid	339.4562	Schwartz criterion	2.071803
Log likelihood	-118.8300	F-statistic	8.162732
Durbin-Watson stat	2.481272	Prob(F-statistic)	0.006303

Cuadro 24

Regresión de los residuos en valor absoluto en función del ingreso para α=-1/2

LS // Dependent Variable is **RESIDUOS VALOR ABSOLUTO**					
Sample: 1995:01 1999:02					
Included observations: 50					
Variable		Coefficient	Std. Error	T-Statistic	Prob.
C		11.13064	2.652728	4.195923	0.0001
$YM^{-1/2}$		**-99.38050**	34.28924	-2.898300	0.0056
R-squared	0.148938		Mean dependent var		3.519573
Adjusted R-squared	0.131208		S.D. dependent var		2.847064
S.E. of regression	2.653721		Akaike info criterion		1.991104
Sum squared resid	338.0273		Schwartz criterion		2.067584
Log likelihood	-118.7245		F-statistic		8.400141
Durbin-Watson stat	2.491538		Prob(F-statistic)		0.005640

Cuadro 25

Regresión de los residuos en valor absoluto en función del ingreso para α=1/2

LS // Dependent Variable is **RESIDUOS VALOR ABSOLUTO**					
Sample: 1995:01 1999:02					
Included observations: 50					
Variable		Coefficient	Std. Error	T-Statistic	Prob.
C		-3.801716	2.502359	-1.519253	0.1353
$YM^{1/2}$		**0.548651**	0.185417	2.959012	0.0048
R-squared	0.154271		Mean dependent var		3.519573
Adjusted R-squared	0.136651		S.D. dependent var		2.847064
S.E. of regression	2.645394		Akaike info criterion		1.984818
Sum squared resid	335.9093		Schwartz criterion		2.061299
Log likelihood	-118.5674		F-statistic		8.755750
Durbin-Watson stat	2.507126		Prob(F-statistic)		0.004781

Cuadro 26

Regresión de los residuos en valor absoluto en función del ingreso para α=1

LS // Dependent Variable is **RESIDUOS VALOR ABSOLUTO**					
Sample: 1995:01 1999:02					
Included observations: 50					
Variable		Coefficient	Std. Error	T-Statistic	Prob.
C		-0.103196	1.272939	-0.081069	0.9357
YM		**0.019890**	0.006681	2.977217	0.0045
R-squared	0.155878		Mean dependent var		3.519573
Adjusted R-squared	0.138292		S.D. dependent var		2.847064
S.E. of regression	2.642879		Akaike info criterion		1.982916
Sum squared resid	335.2709		Schwartz criterion		2.059397
Log likelihood	-118.5198		F-statistic		8.863821
Durbin-Watson stat	2.511973		Prob(F-statistic)		0.004548

Como puede observarse en las cuatro regresiones realizadas, el coeficiente que multiplica a Y_t^{α} es significativo. Además todas ellas obtienen una suma residual casi idéntica, por lo que resulta difícil decidir a partir de estos resultados cuál es la forma funcional exacta que relaciona $\left|\hat{u}_t\right|$ con YM_t; lo que si parece claro es que esta relación existe.

Para realizar el contraste de White, es necesario estimar el modelo original por MCO, ignorando la posible heteroscedasticidad. Después, estimar la regresión auxiliar del cuadrado de los residuos anteriores sobre una constante, los regresores del modelo original (YM_t y IP_t), sus cuadrados (YM_t^2 y I_t^2) y sus productos cruzados de segundo orden (YM_t * IP_t).

El resultado de la regresión es:

Cuadro 27
Regresión de los residuos al cuadrado (cuadro 17) en función de el ingreso del índice de precios y sus cuadrados, además el producto de estos dos.

LS // Dependent Variable is **RESIDUO^2**
Sample: 1995:01 1999:02
Included observations: **50**

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	7141.080	6916.061	1.032536	0.3075
YM	70.43124	37.18205	1.894227	0.0648
IP	-220.0554	174.3414	-1.262209	0.2135
YM^2	0.103170	0.047333	2.179664	0.0347
IP^2	1.571322	1.097370	1.431898	0.1592
YMIP	-0.897650	0.462800	-1.939605	0.0589

R-squared	**0.352462**	Mean dependent var	20.33106
Adjusted R-squared	0.278878	S.D. dependent var	32.03578
S.E. of regression	27.20444	Akaike info criterion	6.718927
Sum squared resid	32563.59	Schwartz criterion	6.948370
Log likelihood	-232.9201	F-statistic	4.789927
Durbin-Watson stat	2.588298	Prob(F-statistic)	0.001397

Bajo el supuesto de homoscedasticidad, el estadístico μ =TR^2 , se distribuye como una χ_j^2 donde j es el número de regresores de la regresión exceptuando la constante.

Para este ejemplo μ =50*0.352456= 17.6228, mientras que los valores en tablas al 95% y 99% son: respectivamente:

$\chi_5^2 = 11.1$

$\chi_5^2 = 15.1$

Por lo que no se acepta la hipótesis nula de homoscedasticidad al 95% y 99%

2.6. Contraste de picos

Los pasos para realizar esta prueba son los siguientes:

1. Se ordenan las observaciones según la variable a la que se sospecha está asociado el esquema de heterocedasticidad de las perturbaciones.

2. Se estima el modelo por mínimos cuadrados ordinarios.

3. Se calcula el número de picos, se entiende que existe un pico si: $|\hat{\mu}_t| > |\hat{\mu}_{t'}|$

Es decir, cuando un residuo es mayor en valor absoluto que todos los que le preceden. El primer residuo, al que no se puede aplicar la regla anterior, no constituye un pico.

4. Se realiza el contraste usando las tablas de picos, estas proporcionan la distribución de probabilidad acumulada del número de picos bajo la hipótesis nula de homoscedasticidad.

2.7. Contraste de rangos

Las etapas a seguir son las siguientes:
1. Se aplican mínimos cuadrados ordinarios al modelo original.
2. Se calculan los rangos, o números de orden, de los residuos (en valor absoluto) y de las observaciones de la variable que se cree provoca la heterocedasticidad.
3. Se calcula el coeficiente de correlación de rango de Spearman, que viene dado por:

$$r_s = 1 - \frac{6\sum D_i^2}{N(N^2-1)} \qquad (24)$$

Donde D_i es la diferencia entre los rangos de la variable explicativa y $|\hat{\mu}_i|$. El coeficiente de correlación de rangos de Spearman se obtiene calculando el coeficiente de correlación ordinario en el caso específico en que los valores de las variables son rangos.

4. Se contrasta la hipótesis nula

$$H_0 = \rho_s = 0$$
$$H_1 = \rho_s \neq 0$$

mediante el estadístico

$$\frac{r_s\sqrt{N-2}}{\sqrt{1-r_s^2}} \qquad (25)$$

que sigue aproximadamente, bajo la hipótesis nula, una distribución t de student con N-2 grados de libertad.
Si se acepta H_0 se aceptará la existencia de homocedasticidad, mientras que si se rechaza H_0 se supondrá que existe heterocedasticidad.

2.8. Resumen de Heteroscedasticidad

La varianza de las perturbaciones no es constante a lo largo de las observaciones

Causas
Modelos de aprendizaje y error
Variación en las técnicas de investigación en los datos
Valores atípicos
Cambios estructurales
Incorrecta transformación de los datos
Forma funcional incorrecta

Contagio de estratos
Datos de corte transversal

Síntomas
Errores estándar de las variables independientes altos
Error estándar de la regresión grande
Prueba global F baja y varianza grande
Varianzas y covarianzas grandes

Pruebas para detectar
Naturaleza del problema
Método gráfico (diagrama de dispersión con abanicos)
Prueba Park
Prueba Glejser
Prueba de correlación por grado de Spearman
Prueba Goldfeld-Quandt
Prueba Breusch-Pagan-Godfrey
Prueba de heteroscedasticidad de white datos cruzados y no cruzados
Prueba Koenker-Basset
Prueba de homogeneidad de varianza de Bartlett
Cusum Q cuadrado

Consecuencias
Los estimadores dejan de ser eficientes
En algunos casos los estimadores dejan de ser insesgados.

Corrección
Método de mínimos cuadros ponderados.
Varianzas y errores estándar consistentes con heteroscedasticidad de White.
Estimación del modelo en logaritmos, inversos, tasas, deflactor.
Utilizando variables dummy.
Mínimos cuadrados ponderados.
Promedios móviles.

3. AUTOCORRELACIÓN[12]

Para hablar de la existencia o violación de este supuesto, [Martín G. 1997] define la autocorrelación como "la correlación serial existente entre los miembros de una serie de observaciones ordenadas en el tiempo o en el espacio", es decir, los errores correspondientes a un período de tiempo están directamente correlacionados con los errores correspondientes al período siguiente. Aunque es posible que la correlación serial sea tanto positiva como negativa, el interés primordial se basa, en el caso de que la autocorrelación sea positiva, o sea, que es el caso de que los errores correspondientes a un periodo de tiempo estén correlacionados en sentido positivo con los errores correspondientes al periodo siguiente.

La hipótesis de que los errores correspondientes a diferentes observaciones están intercorrelacionados falla a menudo cuando se trata con series de tiempo, cuando los términos de error correspondientes a diferentes períodos de tiempo, diferentes observaciones de corte transversal están correlacionados, entonces el término de error esta autocorrelacionado.

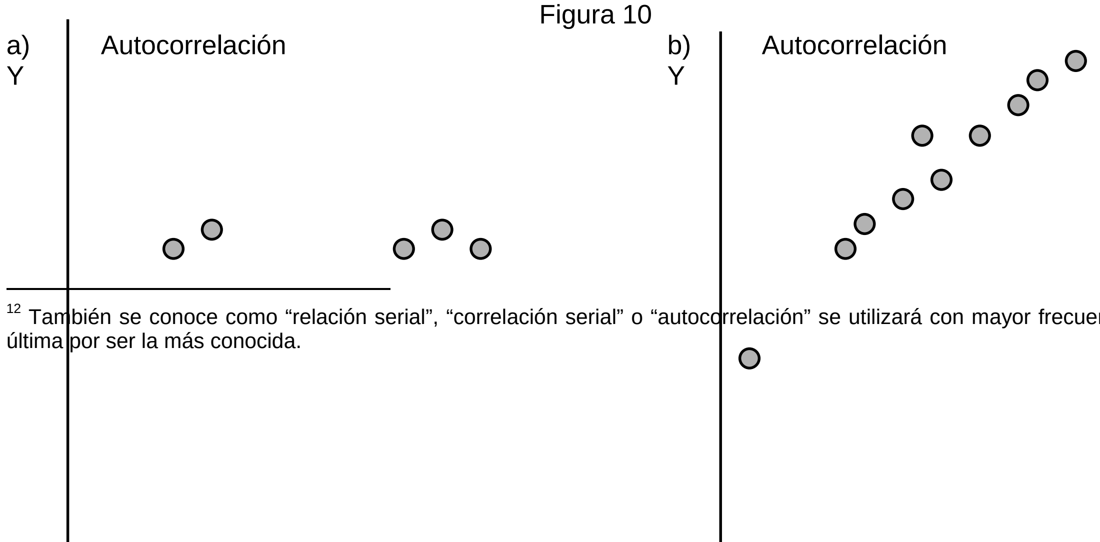
En los estudios de series de tiempo se da la correlación serial cuando los errores asociados con las observaciones correspondientes a un período de tiempo dado se trasladan a periodos de tiempo futuros.

En el ámbito de la regresión, el modelo de regresión lineal clásico supone que tal autocorrelación no existe en las perturbaciones μ_i (ver supuesto 5 página 18).

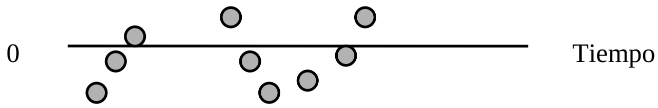
En la figura 10 se visualiza alguno de los patrones de autocorrelación y no autocorrelación

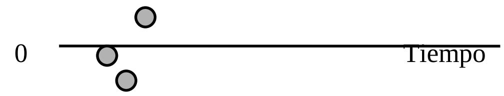
Por regla general, "la presencia de autocorrelación no afectará al insesgamiento o a la consistencia de los estimadores minimocuadráticos, pero la eficiencia si se verá afectada". [Spanos A, 1999].

Figura 10

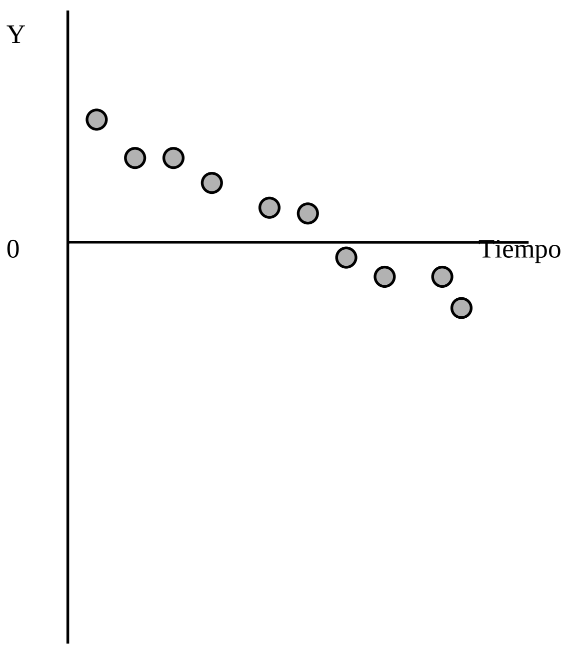


[12] También se conoce como "relación serial", "correlación serial" o "autocorrelación" se utilizará con mayor frecuencia esta última por ser la más conocida.

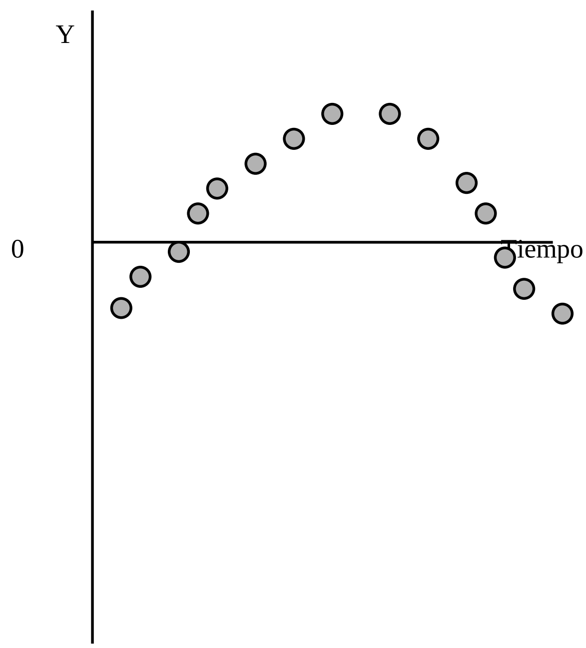




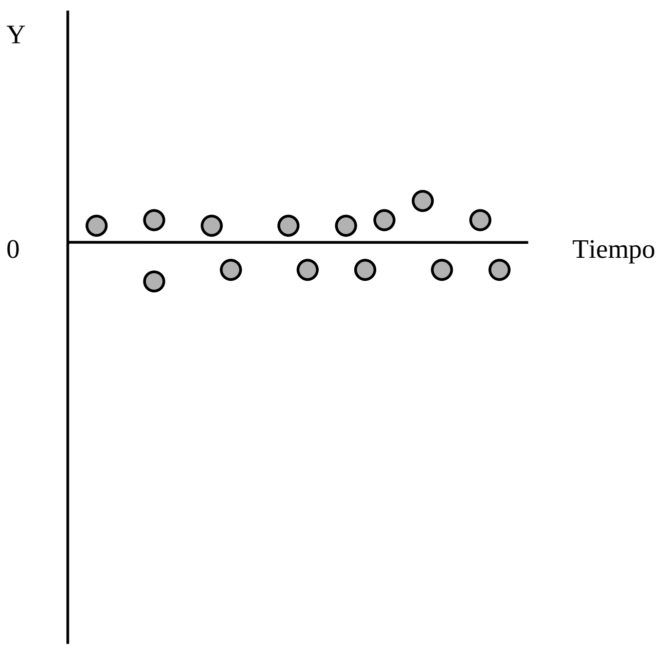
c) Autocorrelación



d) Autocorrelación



e) No autocorrelación



Una de las causas de la aparición de la autocorrelación reside en la propia naturaleza del término de perturbación, que recoge aquellas variables que se mueven conjuntamente a lo largo del tiempo, pero de hecho no son individualmente relevantes para la explicación de la variable dependiente. La presencia de autocorrelación se puede explicar por razones generales tales como el crecimiento económico y los movimientos cíclicos de la economía. También puede ser motivada por errores de especificación del modelo, ya sea por omisión de alguna variable o por una especificación funcional incorrecta, cualquiera de estas razones afectará al término de error.
La presencia de autocorrelación en las perturbaciones considera la posibilidad que exista una dependencia temporal entre los valores de la perturbación, de manera que, si la influencia sobre el instante ***t*** es sólo del instante anterior, la perturbación seguiría la siguiente expresión:

$$u_t = \rho \bullet u_{t-1} + \varepsilon_t$$

donde ε_t es una variable aleatoria "Ruido Blanco"[13]

A este tipo de dependencia se le denomina proceso autorregresivo de orden 1 o AR(1). En general, si la dependencia temporal se establece hasta ρ retardos, el proceso será un proceso autorregresivo de orden ρ, AR(ρ)

$$u_t = \rho \bullet u_{t-1} + \rho \bullet u_{t-2} + ... + \rho \bullet u_{t-\rho} + \varepsilon_t$$

donde ε_t es una variable aleatoria "ruido blanco"

Otra posibilidad consiste en considerar que la autocorrelación viene generada por una combinación de variables aleatorias ruido blanco, contemporáneas y retardadas, es decir, un proceso de medias móviles (MA). El orden de este proceso vendrá dado por el retardo (q) hasta el que depende la perturbación.

El proceso más general será :

$$u_t = \varepsilon_t + \theta_1 \bullet \varepsilon_{t-1} + \theta_2 \bullet \varepsilon_{t-2} + ... + \theta_q \bullet \varepsilon_{t-q}$$

donde ε_t es una variable aleatoria "ruido blanco"

3.1. Contraste Durbin Watson

El estadístico tiene la siguiente expresión:

$$d = DW = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} \qquad (26)$$

"el cual es la razón de las sumas de las diferencias al cuadrado de residuos sucesivos, con respecto a la suma de residuos al cuadrado. Obsérvese que en el numerador el número de observaciones es *n-1* debido a que se pierde una de ellas al tomar las diferencias consecutivas" [Martín G. y Labeaga A, 1997] .

Una ventaja del estadístico *d* consiste en que se basa en los residuos estimados que se calculan en el análisis de regresión.

---

[13] Un proceso de ruido blanco representa una variable que: oscila en torno a una media constante, con una volatilidad constante y cuyo pasado no contiene información útil para predecir valores futuros.

Los supuestos en los que se basa el estadístico ***d*** son:

1. El modelo de regresión incluye el término de intersección.
2. Las variables explicativas, las *X*, son no estocásticas o fijas para muestreo repetidos.
3. Las perturbaciones u_t se generan a través de un esquema autorregresivo de primer orden: $u_t = \rho \bullet u_{t-1} + \varepsilon_t$
4. El modelo de regresión no incluye el valor o los valores rezagados de la variable dependiente como una de las variables explicativas, por lo cual la prueba no es aplicable a modelos del siguiente tipo:
   $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + ... + \beta_k X_{kt} + \gamma Y_{t-1} + \mu_t$
   donde Y_{t-1} es el valor de *Y*, rezagado un periodo.
5. El mayor problema lo presenta la existencia de la zona de indecisión, en la que el contraste no es concluyente. La opción que se suele adoptar es la de considerar que sí existe autocorrelación e intentar ratificarla con la utilización de otros contraste. "[Gujarati D. 2003]

El contraste de Durbin Watson parte de la siguiente hipótesis nula:

$H_0: \rho = 0$	No existe autocorrelación
$H_1: 0 < \lvert\rho\rvert < 1$	Existe autocorrelación AR(1)

El estadístico *d* toma valores entre 0 y 4 de manera que:

Si $d \approx 0$ Existe autocorrelación positiva
Si $d \approx 2$ No existe autocorrelación
Si $d \approx 4$ Existe autocorrelación negativa

El estadístico ***d*** depende de las siguientes características

1. Tamaño de la muestra (n)
2. Número de regresores (k)
3. Forma correcta de la matriz de regresores (X).

Para realizar este contraste, Savin-White[14] establecieron las colas inferior (d_L) y superior (d_U) para los valores críticos del estadístico *d* en función de las características antes mencionadas. Si el valor muestral de ***d*** es inferior a 2, la verificación o rechazo de la hipótesis se realiza de la siguiente forma:

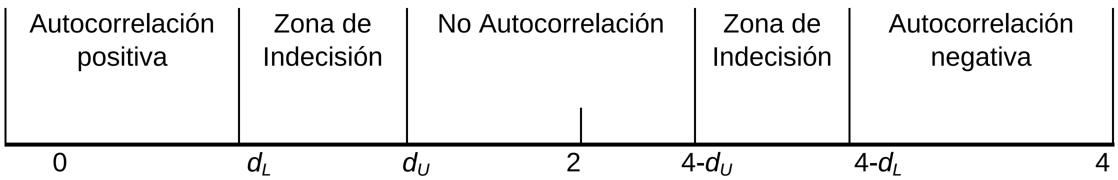
$d < d_L$ Se rechaza la H_0, existe autocorrelación positiva AR(1)
$d_L < d < d_U$ El contraste no es concluyente.
$d_U < d < 2$ No se rechaza la H_0.
Si el valor muestral de ***d*** es, superior a 2, los valores críticos para realizar el contraste serán:

$2 < d < 4 - d_U$	No se rechaza la hipótesis nula
$4 - d_U < d < 4 - d_L$	El contraste no es concluyente
$4 - d_L < d < 4$	Se rechaza la H_0, existe autocorrelación negativa.

En forma gráfica la representación de las regiones quedaría

[14] Savin, E and K. White. *The Durbin-Watson Test for Serial Correlation With Extreme Sample Sizes or Many Regressors*. Econometrica, 45, 1997.

Gráfica 8
Estadístico Durbin-Watson



3.2. Contraste h de Durbin

Considere la forma:

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + ... + \beta_r Y_{t-r} + \beta_{r+1} X_{1t} + \beta_{r+2} X_{2t} + ... + \beta_s X_{st} + \mu_t$$

con

$$u_t = \rho \bullet u_{t-1} + \varepsilon_t$$

donde ε_t es una variable aleatoria “Ruido Blanco”

La formulación de la hipótesis será:

H_0: ρ=0 No existe autocorrelación

H_1: $0 < |\rho| < 1$ Existe autocorrelación AR(1)

Y si la hipótesis nula es cierta, el estadístico:

$$h = r \bullet \sqrt{\frac{n}{1 - n \bullet Var(b_1)}} \to N(0,1) \qquad (27)$$

donde:

n Es el tamaño de la muestra

$Var(b_1)$ Varianza muestral estimada del coeficiente de Y_{t-1} en la regresión del modelo original.

r es el coeficiente de autocorrelación muestral, o sea:

$$r = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n} e_{t-1}^2} \qquad (28)$$

El coeficiente de correlación puede obtenerse por medio de una regresión auxiliar, estimando el parámetro por MCO, en un modelo como:

$$e_t = \rho \bullet e_{t-1} + \mu_t$$

Una forma alternativa de obtener este valor es a través del estadístico de Durbin-Watson desarrollando las sumatorias que forman el estadístico. De esta manera se obtiene que:

$$r = 1 - \frac{d}{2} \quad (29)$$

3.3. Contraste de Breusch-Godfrey

Este contraste permite construir la hipótesis alternativa de forma más general, buscando el orden *p* del proceso autorregresivo o el orden *q* de un proceso de medias móviles. El modelo autorregresivo general será:

$$u_t = \rho \bullet u_{t-1} + \rho \bullet u_{t-2} + ... + \rho \bullet u_{t-\rho} + \varepsilon_t$$

donde ε_t es una variable aleatoria "ruido blanco"

Mientras que el modelo de medias móviles será:

$$u_t = \varepsilon_t + \theta_1 \bullet \varepsilon_{t-1} + \theta_2 \bullet \varepsilon_{t-2} + ... + \theta_q \bullet \varepsilon_{t-q}$$

donde ε_t es una variable aleatoria "ruido blanco"

El contraste se realizará con las siguientes fases:

1. Estimar el modelo por MCO y obtener la serie de los residuos e_t
2. Estimar una regresión auxiliar de los residuos e_t sobre ***p*** retardos y sobre las variables explicativas *X's* del modelo, pudiendo introducir retardos de la variable dependiente sin que cambien las propiedades del contraste.
$$e_t = \gamma_1 e_{t-1} + \gamma_2 e_{t-2} + ... + \gamma_p e_{t-p} + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + ... + \beta_r Y_{t-r} + \beta_{r+1} X_{1t} + \beta_{r+2} X_{2t} + ... + \beta_{X_{st}}$$
3. Obtener el coeficiente de determinación R^2 de la regresión auxiliar estimada en el punto anterior.
4. Realización del contraste:

H_0: No existe autocorrelación

H_1: Existe autocorrelación de orden *p*

Si la hipótesis nula es cierta, el estadístico:

$$\chi^2_{calculada} = n \bullet R^2 \rightarrow \chi^2_p$$

donde ***p*** es el número de retardos de los residuos que se introducen en la regresión auxiliar.

3.4. Correlograma y función de autocorrelación parcial simple (FACP)

"El correlograma es la función gráfica de la función de autocorrelación, es decir, es un gráfico de los coeficientes de autocorrelación en función del orden del retardo. Por otra parte, la interpretación conjunta de las representaciones de los coeficientes de autocorrelación parcial simple y del correlograma, puede permitir, si la serie es estacionaria, identificar el tipo de proceso de autocorrelación, (ya que cada proceso tiene una representación distinta y única)" para conocer más procesos se puede ver el texto del profesor [Green W, 1999].

Es necesario recordar que los procesos autorregresivos presentan un correlograma decreciente exponencial, sinusoidal o con alternancia de signos, mientras que en la FACP sólo existen *p* valores distintos de cero, orden del proceso.

En cuanto a los procesos de medias móviles, el comportamiento de ambas representaciones es al revés, o sea, un proceso MA(*q*) tiene *q* valores distintos de cero en el correlograma y un decrecimeinto exponencial, sinusoidal o con alternancia de signos en la FACP.

La última posibilidad sería la combinación de ambos procesos, construyendo un proceso mixto ARMA (*p,q*), es decir

$$u_t = \rho \bullet u_{t-1} + \rho \bullet u_{t-2} + ... + \rho_p \bullet u_{t-\rho} + \varepsilon_t + \theta_1 \bullet \varepsilon_{t-1} + ... + \theta_q \bullet \varepsilon_{t-q}$$

La representación gráfica de ambas funciones tiene tantos comportamientos de proceso autorregresivo como del de medias móviles.

3.5. Corrección de autocorrelación

3.6. Método de Cochrane-Orcutt

El procedimiento de estimación de Cochrane-orcutt (C-O) se basa en un proceso iterativo, que se basa en la estimación de los residuos e_t para obtener información sobre el valor desconocido *p*.
A partir de la información, probar el posible incumplimiento de la hipótesis clásica de no autocorrelación por medio de:

a) Contraste gráficos.
b) Contraste de Durbin-Watson.
c) Correlograma.
d) Contraste de Breusch-Godfrey.
e) Con esta misma información, se propone especificar una nueva forma estableciendo una relación lineal de la variable independiente y de un rezago de la variable dependiente como factores explicativos. Bajo esta formulación, se contrasta la presencia de autocorrelación en las perturbaciones.
f) En caso de detectar el problema de autocorrelación, obtener una estimación MCO para los parámetros del modelo.

3.7. Soluciones

Para el análisis de la autocorrelación de los residuos del modelo de MCO, en primer lugar es necesario estimar el modelo original por este criterio. Los datos pueden ser consultados en el anexo 4.

El modelo es:

$$CP_t = \beta_1 + \beta_b PIB_t + \mu_t$$

Donde:
CP es el consumo total (incluye al de gobierno y al privado).
PIB es el Producto Interno Bruto.

Los resultados de la estimación son observados en el cuadro 27, a partir de esta estimación se generan los residuos mínimo cuadráticos correspondientes a este modelo, que se pueden ver en la gráfica 8,

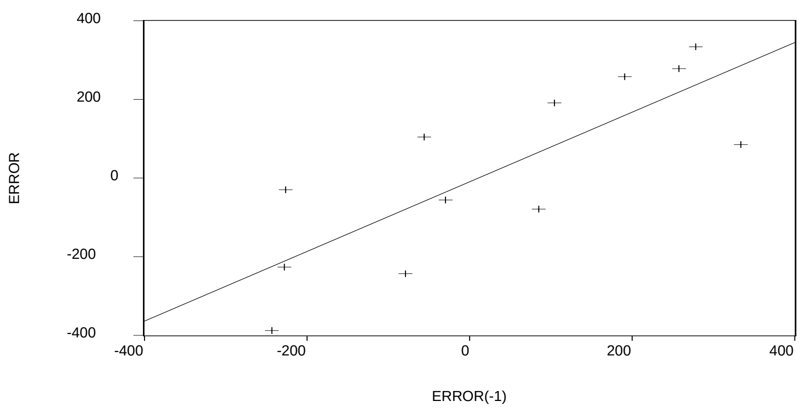
Cuadro 28
Regresión del consumo en función del producto interno bruto

LS // Dependent Variable is CP
Sample: 1987 1999
Included observations: 13

Variable	Coefficient	Std. Error	T-Statistic	Prob.

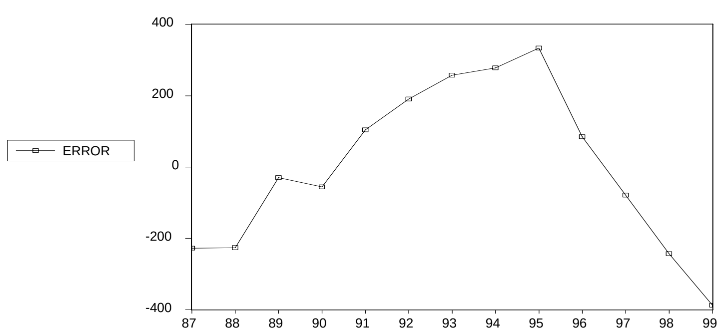
C	3079.706	237.8295	12.94922	0.0000
PIB	0.058374	0.004282	13.63260	0.0000
R-squared	0.944119	Mean dependent var		6192.553
Adjusted R-squared	0.939039	S.D. dependent var		971.3511
S.E. of regression	239.8289	Akaike info criterion		11.10049
Sum squared resid	632696.8	Schwartz criterion		11.18740
Log likelihood	-88.59938	F-statistic		185.8479
Durbin-Watson stat	0.343122	Prob(F-statistic)		0.000000

Gráfica 9
Diagrama de dispersión de los residuos vs. residuos con un rezago



a) En el diagrama de dispersión (contraste de gráficos) se observa que los puntos se colocan en el diagrama en el primer y tercer cuadrante, por lo que puede ser indicativo de presencia de autocorrelación positiva.

Gráfica 10
Residuales o errores en el tiempo



Esta representación de residuales acusa un comportamiento sistemático, posiblemente cíclico que sigue indicando la posible presencia de problemas de autocorrelación.

b) El contraste Durbin-Watson en el cuadro 27, su valor es de 0.343122, en las tablas de Savin-White o tablas del estadístico ***d*** de Durbin Watson, se obtienen los límites inferior y superior, al

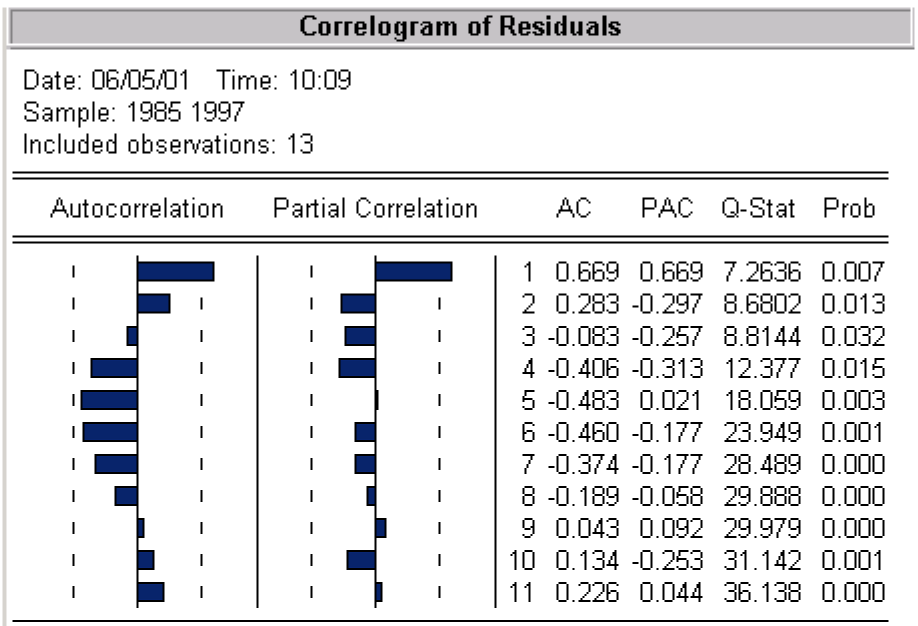
95% de nivel de confianza, con un tamaño de muestra igual a 13 y con k' igual a 1, los resultados son:

Para $d_L = 1.010$ y $d_U = 1.340$

Ambos valores de la tabla son mayores al estadístico ***d*** , entonces tenemos que:
$d < d_L$ se rechaza la hipótesis nula, es decir, existe autocorrelación positiva.

c) Ahora, al ver el correlograma, el correlograma de estos residuos cuadro 28, muestra un decrecimiento sinusoidal, mientras que en la FACP se observa una banda distinta de cero. Este comportamiento se asocia a un proceso autorregresivo de orden 1, es decir, un AR(1), porque la observación 1 sale de las bandas de confianza, sería de orden 2 si la observación número 2 sale de la banda y así sucesivamente.

Cuadro 29
Correlograma de residuales



Correlogram of Residuals

Date: 06/05/01 Time: 10:09
Sample: 1985 1997
Included observations: 13

	AC	PAC	Q-Stat	Prob
1	0.669	0.669	7.2636	0.007
2	0.283	-0.297	8.6802	0.013
3	-0.083	-0.257	8.8144	0.032
4	-0.406	-0.313	12.377	0.015
5	-0.483	0.021	18.059	0.003
6	-0.460	-0.177	23.949	0.001
7	-0.374	-0.177	28.489	0.000
8	-0.189	-0.058	29.888	0.000
9	0.043	0.092	29.979	0.000
10	0.134	-0.253	31.142	0.001
11	0.226	0.044	36.138	0.000

d) El contraste de Breusch-Godfrey se calcula con ayuda del programa Eviews. Este contraste solicita como información el número de retardos o rezagos de los residuos que debe introducir en el contraste. Hasta que no se detalle algún otro procedimiento para su justificación, se irán seleccionando progresivamente 1, 2, …, retardos, hasta que el contraste de significación individual del último retardo se sitúe en región de no rechazo de la hipótesis nula.

Cuadro 30
Regresión con un rezago en la regresión auxiliar.

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	**18.71771**	**Probability**	**0.001498**
Obs*R-squared	8.473176	Probability	0.003604
Test Equation:			

LS // Dependent Variable is RESID				
Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	149.1337	151.1754	0.986495	0.3472
PIB	-0.003342	0.002760	-1.210686	0.2539
RESID(-1)	0.974886	0.225334	4.326396	0.0015

R-squared	0.651783	Mean dependent var	5.07E-13
Adjusted R-squared	0.582139	S.D. dependent var	229.6187
S.E. of regression	148.4304	Akaike info criterion	10.19941
Sum squared resid	220315.9	Schwartz criterion	10.32978
Log likelihood	-81.74235	F-statistic	9.358853
Durbin-Watson stat	1.278439	Prob(F-statistic)	0.005120

El Eviews ofrece una salida de resultado final como la salida al complemento de la regresión auxiliar de los residuos de MCO en función de los regresores del modelo y de los rezagos elegidos de los mismos.

En el cuadro 29 se observa que el estadístico experimental o calculado del contraste es: $n*R^2$= (13*0.651783)= 8.473176, con una probabilidad asociada de 0.003604 < 0,05 que hace situarse a dicho estadístico en la región crítica y por tanto, se rechaza la hipótesis de no autocorrelación a un 95% de nivel de confianza.:

$$n \bullet R^2 > \chi^2$$

así también el valor de su probabilidad asociado a F es de 0.001498 comparado con el α=0.05 se acepta la H_0 y se observa que la autocorrelación persiste en el modelo.

Por otra parte la probabilidad asociada al estadístico ***t*** calculado en la regresión auxiliar, correspondiente al rezago de los residuos, es 0.0015, que por ser menor a 0.05, se sitúa al estadístico calculado en región de rechazo de la hipótesis de significación individual, es decir, el primer retardo de los residuos es significativo en la explicación de los residuos, al 95% de nivel de confianza.

e) En el cuadro 30 se muestra la regresión correspondiente a la introducción de dos rezagos de los residuos en la regresión auxiliar.

Cuadro 31
Regresión con dos rezagos en la regresión auxiliar.

Breusch-Godfrey Serial Correlation LM Test:			
F-statistic	10.11549	Probability	0.004987
Obs*R-squared	8.997396	Probability	0.011123

Test Equation:
LS // Dependent Variable is RESID

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	68.27062	167.3323	0.407994	0.6928
PIB	-0.001571	0.003186	-0.493088	0.6338
RESID(-1)	1.216267	0.315142	3.859431	0.0039
RESID(-2)	-0.428241	0.394440	-1.085692	0.3058

R-squared	0.692107	Mean dependent var	5.07E-13
Adjusted R-squared	0.589477	S.D. dependent var	229.6187

S.E. of regression	147.1215	Akaike info criterion	10.23018
Sum squared resid	194802.7	Schwartz criterion	10.40401
Log likelihood	-80.94236	F-statistic	6.743657
Durbin-Watson stat	1.956964	Prob(F-statistic)	0.011152

Al igual que el cuadro 31, el estadístico calculado del contraste de Breusch-Godfrey tiene una probabilidad asociada de 0.004987, menor que 0.05, y sigue situando al estadístico en región de rechazo de la hipótesis nula, es decir, existe autocorrelación. El orden de esta dependencia se contrasta en la regresión auxiliar, mediante el contraste de significancia individual de segundo rezago de los residuos.

Como la probabilidad asociada al t calculado es de 0.3058, mayor que 0.05, no se puede rechazar la hipótesis nula; por lo tanto, este segundo rezago ya no es significativo en el comportamiento de los residuos mínimo-cuadráticos.
La conclusión del contraste de Breusch-Godfrey es que existe autocorrelación en las perturbaciones y que es de orden uno.
La especificación alternativa propuesta se formaliza mediante la siguiente ecuación:

$$CP_t = \beta_1 + \beta_b PIB_t + \beta_3 CP_{t-1} + \mu_t$$

Al presentar este modelo una especificación dinámica con la matriz de regresores estocástica, el contraste de Durbin-Watson no se puede utilizar, por lo que se debe recurrir al contraste de ***h*** de Durbin y al contraste de Breusch-Godfrey.
El primer paso para la realización del contraste es la obtención de la regresión por MCO del modelo especificado:

Cuadro 32
Regresión del consumo en función del PIB y del consumo con un rezago

LS // Dependent Variable is CP
Sample: 1988 1999
Included observations: 12 after adjusting endpoints

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	780.6843	466.1902	1.674605	0.1283
PIB	-0.002714	0.010880	-0.249441	0.8086
CP(-1)	0.934178	0.169258	5.519250	0.0004

R-squared	0.984426	Mean dependent var	6333.763
Adjusted R-squared	0.980965	S.D. dependent var	864.0075
S.E. of regression	119.2059	Akaike info criterion	9.774022
Sum squared resid	127890.3	Schwartz criterion	9.895248
Log likelihood	-72.67139	F-statistic	284.4366
Durbin-Watson stat	1.618970	Prob(F-statistic)	0.000000

Se debe contrastar el posible incumplimiento de la hipótesis de no autocorrelación antes de comprobar si las variables son significativas o no.

Contraste de la ***h*** de Durbin (ver página 59).

Bajo la hipótesis nula el estadístico (ver fórmula 20).

La obtención de los valores necesarios para construir el estadístico se realiza con el cuadro 31 (salida de la regresión) del modelo estimado por MCO.

$$r = 1 - \frac{d}{2} = 1 - \frac{1.618970}{2} = 0.190515$$

$$Var(b_3) = (0.16925)^2 = 0.028648$$

$$h_{calculada} = 0.19051 \bullet \sqrt{\frac{12}{1-(12*0.028648)}} = 0.80836$$

El estadístico teórico sigue una distribución normal (0, 1), que al 95% de nivel de confianza, en contraste de una sola cola, es:

$$\boldsymbol{h_t} = 1.645$$

por lo tanto, como:

$$\boldsymbol{h_{calculada}} = 0.808436 < \boldsymbol{h_t} = 1.645$$

No se rechaza la hipótesis nula.

Es decir, con un 95% de confianza las perturbaciones de esta especificación son esféricas[15].

Cuadro 33

Regresión con un rezago en el consumo y en la regresión auxiliar

Breusch-Godfrey Serial Correlation LM Test:				
F-statistic	0.107581	Probability	0.751334	
Obs*R-squared	0.159231	Probability	0.689866	
Test Equation:				
LS // Dependent Variable is RESID				
Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	98.45569	575.6385	0.171037	0.8684
PIB	0.002360	0.013534	0.174365	0.8659
CP(-1)	-0.037549	0.211913	-0.177189	0.8638
RESID(-1)	0.136888	0.417347	0.327996	0.7513
R-squared	0.013269	Mean dependent var		-1.00E-13
Adjusted R-squared	-0.356755	S.D. dependent var		107.8258
S.E. of regression	125.5952	Akaike info criterion		9.927330
Sum squared resid	126193.3	Schwartz criterion		10.08897
Log likelihood	-72.59124	F-statistic		0.035860
Durbin-Watson stat	1.891146	Prob(F-statistic)		0.990206

Podemos observar que el estadístico nR^2 como la $\boldsymbol{t_{calculada}}$ del rezago de los residuos en la regresión auxiliar, tiene probabilidades asociadas mayores que 0.05, por lo que se sitúan en región de aceptación, no pudiendo rechazar la hipótesis nula.

Entonces el contraste de ***h*** de Durbin y el de Breusch-Godfrey confirman la ausencia de autocorrelación en las perturbaciones de esta formulación del modelo.

[15] Las perturbaciones que satisfacen los supuestos de homocedasticidad y no autocorrelación son, conocidas como perturbaciones esféricas.

f) Por último se calcula la regresión por el método de Cochrane-Orcutt (C-O), recordando que el modelo que presentaba problemas de autocorrelación respondía a la formulación inicial del consumo en función del PIB; sobre ese modelo será sobre el que se trate de obtener los estimadores en forma eficiente.

Como ya se menciono, el método de Cochrane-Orcutt, es un criterio de estimación iterativa para obtener estimadores lineales, insesgados y óptimos, bajo la especificación del modelo de mínimos cuadrados generalizados.

Cuadro 34
Proceso C-O

LS // Dependent Variable is CP
Sample: 1988 1999
Included observations: 12 after adjusting endpoints
Covergence achieved after 22 iterations

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	6295.619	3386.924	1.858801	0.0960
PIB	0.017790	0.032744	0.543310	0.6001
AR(1)	0.859675	0.099318	8.655807	0.0000

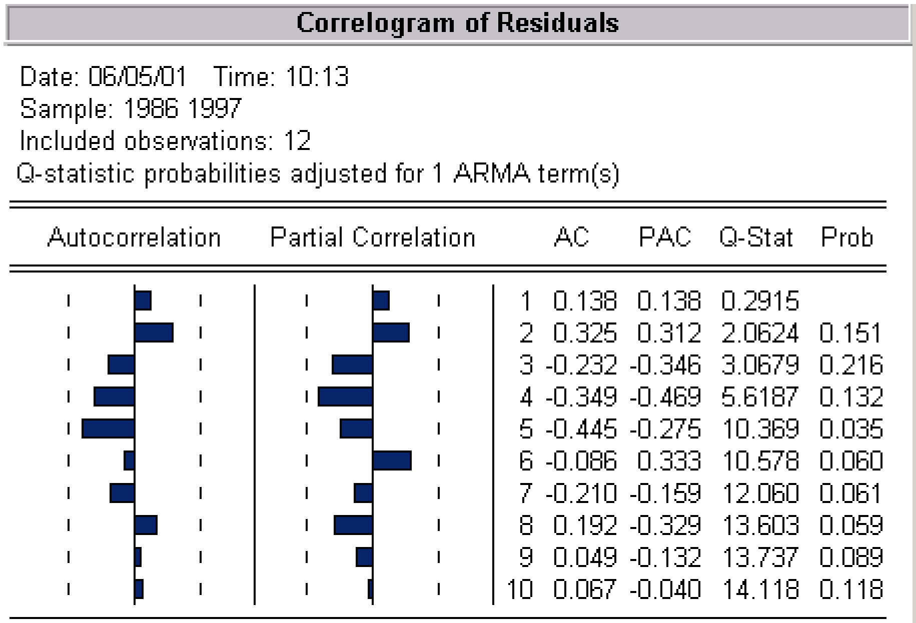
R-squared	0.984927	Mean dependent var	6333.763
Adjusted R-squared	0.981578	S.D. dependent var	864.0075
S.E. of regression	117.2705	Akaike info criterion	9.741284
Sum squared resid	123771.2	Schwartz criterion	9.862510
Log likelihood	-72.47496	F-statistic	294.0523
Durbin-Watson stat	1.518258	Prob(F-statistic)	0.000000

Inverted AR Roots .86

El cuadro 33, muestra los resultados de la repetición 22 veces el proceso de Cochrane-Orcutt para conseguir la convergencia. Cabe destacar que la variable PIB pasa a ser no significativa, mientras que el proceso autorregresivo, con una probabilidad asociada de 0.0000.

Se obtiene el correlograma de los residuos, (en la figura 11) y con esa estimación se podrá corroborar el diagnóstico de buen comportamiento de la perturbación.

Figura 11
Estadístico Q de Ljung-Box



Correlogram of Residuals

Date: 06/05/01 Time: 10:13
Sample: 1986 1997
Included observations: 12
Q-statistic probabilities adjusted for 1 ARMA term(s)

	AC	PAC	Q-Stat	Prob
1	0.138	0.138	0.2915	
2	0.325	0.312	2.0624	0.151
3	-0.232	-0.346	3.0679	0.216
4	-0.349	-0.469	5.6187	0.132
5	-0.445	-0.275	10.369	0.035
6	-0.086	0.333	10.578	0.060
7	-0.210	-0.159	12.060	0.061
8	0.192	-0.329	13.603	0.059
9	0.049	-0.132	13.737	0.089
10	0.067	-0.040	14.118	0.118

El estadístico Q de Ljung-Box permite verificar la hipótesis nula de:
H_o: $p_1 = p_2 = \ldots = p_p = 0$
Cuando se acepta esta hipótesis se está indicando que los residuos no están autocorrelacionados. En el presente caso, dado que la probabilidad es superior a 0.05, indica que se ha corregido la autocorrelación, ya que los residuos del nuevo no la presentan.

3.8. Resumen de Autocorrelación

La correlación existente entre los miembros de una serie de observaciones ordenadas en el tiempo o en el espacio

Causas
Datos de corte transversal (autocorrelación espacial)
Inercia
Sesgo de especificación (variables excluidas)
Sesgo de especificación (forma funcional incorrecta)
Fenómeno de la telaraña
Rezagos
Manipulación de datos
Transformación de datos
Estacionariedad[16]

[16] Un proceso estocástico es estacionario si para todo entero m>0 los conjuntos de variables $\{y_{t1}, y_{t2} \ldots, y_{tm}\}$ tienen la misma distribución, independientemente del valor de t. Es decir, que un conjunto determinado de m variables aleatorias cualesquiera del proceso estocástico tiene la misma distribución multivariable que cualquier otro conjunto de m variables aleatorias extraídas del mismo proceso.
¿Qué ocurre cuando m=1? El conjunto de variables se reduce a una sola, $\{y_t\}$ y de acuerdo con la definición, en un proceso estacionario este conjunto ha de tener la misma distribución independiente del valor de índice t. Es decir que, la esperanza y la varianza de las variables y_t deben ser independientes del tiempo, ya que son iguales para todas las variables y_t. Cuando m=2 la definición implica que la distribución conjunta de pares de variables {yt,yt-1} también debe ser invariante en el tiempo, y en particular, los valores de las funciones de autocovarianza y autocorrelación para k=1 deben ser independientes del tiempo. Un proceso para el que los valores de la función de autocorrelación son diferentes para distintos sub-períodos muestrales será, por tanto, no estacionario. Novales A 230-231.

Síntomas

Errores estándar estandarizados se alejan de los errores estándar de l regresión.
Durbin Watson se aleja del valor 2.
Tiende a no haber normalidad en los residuales.
El modelo tiende de a perder estabilidad.

Pruebas para detectar

Método gráfico.
Prueba de rachas.
Prueba de Durbin Watson.
Correlograma Q (FACP Función de Autocorrelación Parcial).
Breusch-Godfrey.
Prueba de Berenblutt-Webb.

Consecuencias

Las varianzas y los errores estándar son altos.
Los estimadores dejan de ser eficientes.
Si se estima en presencia de autocorrelación los pronósticos no son confiables.

Corrección

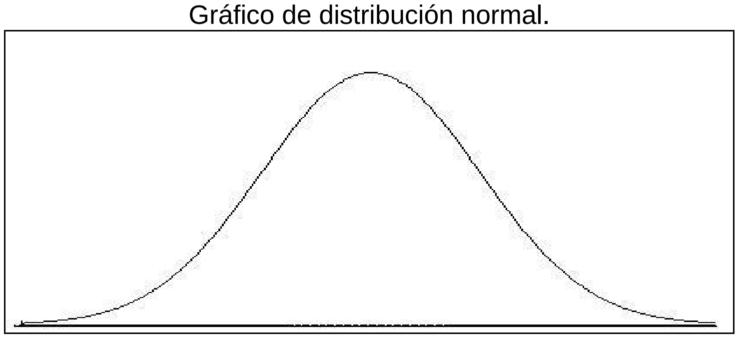
Si la autocorrelación es pura utilizar el método de mínimos cuadrados generalizados.
Para muestras grandes se utiliza el método de Newey-West.
Cuando ρ es conocido. Transformación Prais-Winsten.
Cuando ρ es desconocido. Método de la primera diferencia.
Método iteractivo para estimar ρ (Cochrane-Orcutt).
Promedios móviles.
Método autorregresivo de Koyck.
Método de rezagos distribuidos de Almon.

4. Normalidad

Al hablar de normalidad es necesario recordar que es una distribución continua, que tiene forma de campana y está determinada por su media y su desviación estándar, además es simétrica respecto a su media y para cada combinación de la media y la desviación estándar especifica una distribución normal única.

Si $E(\mu_i)=0$ y $E(u_i=u_j)$ el estimador es insesgado
A partir de esto sabemos que μ_i $N(0,\sigma^2)$ por lo que hay normalidad en los residuales
Gráficamente tenemos la distribución normal:

Figura 12
Gráfico de distribución normal.



Se sabe que los momentos de una distribución normal estandarizada son:
Primer Momento (media): $\mathrm{E}(X) = 0$

Segundo Momento (varianza): $\mathrm{E}(X - \bar{X})^2 = \sigma^2$

Tercer Momento(sesgo): $\mathrm{E}(X - \bar{X})^3 = 0$

Cuarto Momento (kurtosis): $\mathrm{E}(X - \bar{X})^4 = 3$

Este supuesto no es esencial si el objetivo es solamente la estimación, los estimadores mínimo cuadráticos siguen siendo los mejores sin importar si las μ_i, están normalmente distribuidas o no. Es posible establecer que los estimadores MCO de los coeficientes de la regresión siguen la distribución normal, que (n-k) $\frac{\hat{\sigma}^2}{\sigma^2}$ y siguen la distribución χ^2 y que podrían utilizarse las pruebas t y F para verificar diversas hipótesis estadísticas, sin importar el tamaño de la muestra.
Si no están normalmente distribuidas las μ_i ¿qué sucede? El profesor Henri Theil, en su libro Introduction to Econometrics, expone “Si las perturbaciones μ_i son independientes e idénticamente distribuidas con media cero y varianza constante σ^2 y si las variables explicativas son constantes en muestras repetidas, los estimadores MCO de coeficientes son asintóticamente[17] normales con medias iguales a las β correspondientes”

Pero que ocurre cuando las muestras son pequeñas o finitas, el supuesto de normalidad se hace muy importante para prueba de hipótesis y predicción. Los efectos del incumplimiento de la normalidad y de los temas relacionados es analizado bajo el tema de estimación robusta en la teoría.

Los estimadores robustos, conocidos también como estimadores no paramétricos o estimadores libres de distribución, son estimadores libres de la asunción de una forma de distribución de la población de la cual se extrae la muestra.

Los estimadores clásicos o paramétricos, tienen asociada un tipo de distribución de la población. Así, por ejemplo, a la media aritmética se le asocia la distribución normal o mesocurtica; a la mediana se le asocia la distribución leptocurtica; y a la media de los extremos se le asocia la distribución uniforme o rectangular.

4.1. Cómo detectar normalidad

Histograma de residuos. Es un gráfico que se utiliza para saber la forma de la distribución de probabilidad de una variable aleatoria.
Gráfica de probabilidad normal. Utiliza el papel de probabilidad normal (especialmente diseñado para gráficas). Sobre el eje X’s se grafican los residuos y sobre el eje vertical se grafica el valor esperado de la variablesi estuviera normalmente distribuido.
Prueba de Anderson-Darling. La hipótesis nula supone que la variable a estimar está normalmente distribuida.
Prueba Jarque-Bera. Es una prueba asintótica, basada en los residuos MCO. Calcula primero la asimetría y la kurtosis y utiliza el siguiente estadístico de prueba.

$$JB = n\left[\frac{S^2}{6} + \frac{(K-3)^2}{24}\right] \tag{30}$$

donde:
n= tamaño de muestra
S= coeficiente de asimetría
K= Coeficiente de kurtosis.
Para una variable normalmente distribuida con S=0 y K=3, en este caso se espera que el valor del estadístico sea cero[18].
De acuerdo con la hipótesis nula en que los residuos están normalmente distribuidos, el estadístico sigue una distribucióbn χ^2 con 2 grados de libertad (5.99 valor crítico).
Prueba Kolmogorov-Smirnov.

[17] Es decir muestra grande.
[18] Para abundar más sobre este tema se puede consultar a Sánchez, B. La estadística aplicada al análisis económico, FE-UNAM 2007, en el capítulo II.

4.2. Medidas remediales

Incrementar el número de datos en las variables, ya que con esto y siguiendo el teorema de límite central justifica el supuesto de normalidad.

4.3. Soluciones

Cuando se trata de una muestra finita, se debe realizar una prueba explícita para el supuesto de normalidad. Considerando bondad de ajuste χ^2 y prueba Jarque-Bera.

Se estima la ecuación PIB en función de la importaciones y de la formación bruta de capital, que están disponibles para su consulta en el anexo número 5, se obtiene la siguiente regresión.

Cuadro 35
Regresión del PIB vs. Importaciones y Formación Bruta de Capital

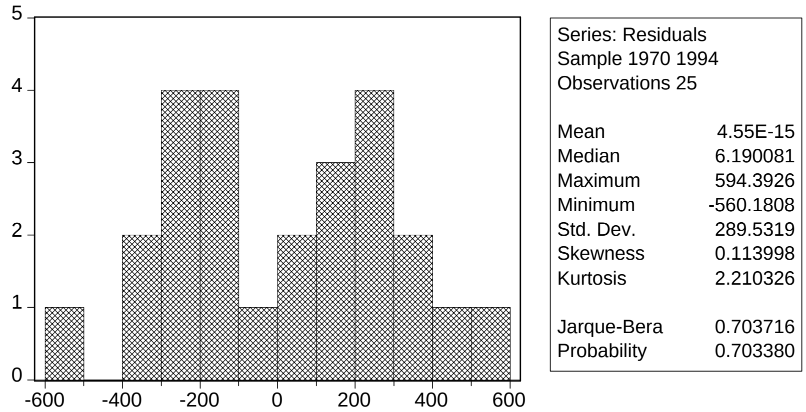
LS // Dependent Variable is PIB
Sample: 1970 1994
Included observations: 25

Variable	Coefficient	Std. Error	T-Statistic	Prob.
C	-646.0891	351.2111	-1.839603	0.0794
IMPORTACION	-8.203239	1.131877	-7.247468	0.0000
FBC	15.64716	1.517894	10.30846	0.0000

R-squared	0.928946	Mean dependent var	4272.402
Adjusted R-squared	0.922486	S.D. dependent var	1086.179
S.E. of regression	302.4062	Akaike info criterion	14.37359
Sum squared resid	2011889	Schwartz criterion	14.51985
Log likelihood	-176.6698	F-statistic	143.8113
Durbin-Watson stat	1.375246	Prob(F-statistic)	0.000000

Se prueba el estadístico Jarque Bera, bajo la siguiente hipótesis:
H_0: Existe normalidad
H_a. No hay normalidad
Se observa que el valor es de 0.7037 contrastado con el χ^2 crítico (5.99 con 2 gl) se decide aceptar la H_0 y se concluye que el modelo presenta normalidad. Ahora si se observa la probabilidad asociada al estadístico se tiene un valor de 0.7033 contrastado con la probabilidad de 0.05 se corrobora la aceptación de la hipótesis nula.

Gráfica 11
Histograma de normalidad



Aplicando la prueba de Kolmogorov a cada una de las series se observa que las dos primeras variables no tienen un comportamiento normal no así la formación bruta de capital, esto puede ocurrir más sin embargo al realizar el análisis de regresión y aplicar la prueba Jarque Bera los resultados pueden ser contrario, o sea que los residuales tengan un comportamiento normal.

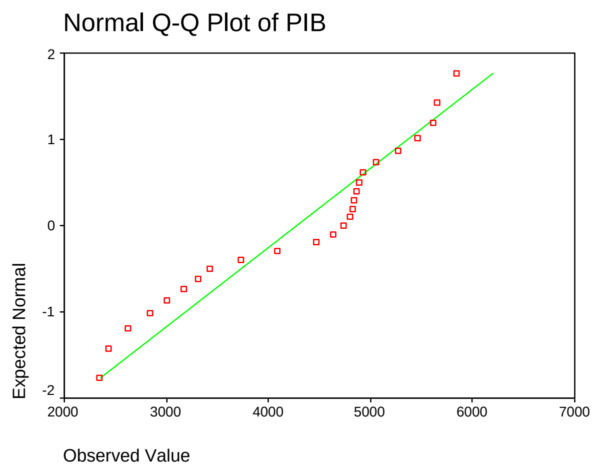
Es importante y vital señalar que la anterior prueba fue aplicada a los residuales y deben cumplirse los supuestos del modelo de regresión lineal clásico, para estas pruebas (kolmogorov y Shapiro) es aplicada a la serie de datos lo que llamamos datos crudos u originales; esto es aplicado a los cuadros 36, 37 y 38.

Cuadro 36
Pruebas de normalidad

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Estadístico	gl	Probabilidad	Estadístico	gl	Probabilidad
PIB	.189	25	.022	.916	25	.046

En las gráficas 12,13 y 14 de cada variable se muestra que el comportamiento de cada serie se distribuye de manera homogénea con respecto a su línea de regresión, es decir, que los valores se distribuyen de manera normal, no hay ninguna observación que perturbe o “jale” los datos hacia un comportamiento atípico.

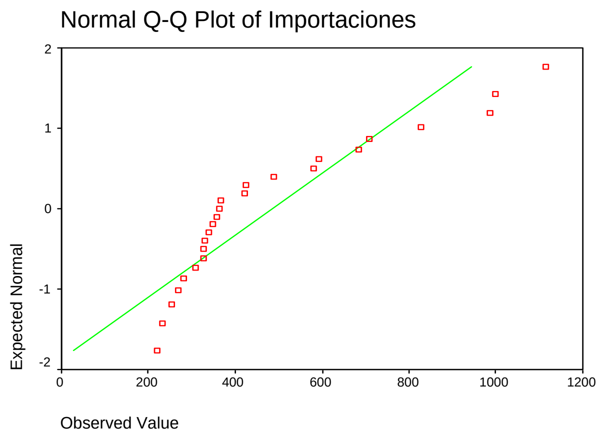
Gráfica 12
Cuartil-Cuartil del PIB



Cuadro 37
Pruebas de normalidad para importaciones

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Estadístico	gl	Probabilidad	Estadístico	gl	Probabilidad
Importaciones	.236	25	.001	.830	25	.010

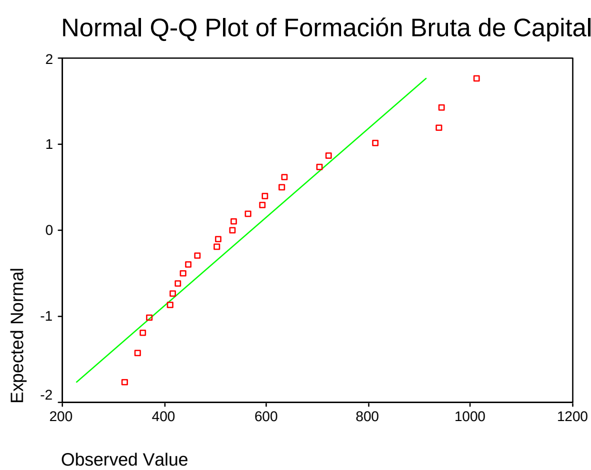
Gráfica 13
Cuartil-Cuartil de Importaciones



Cuadro 38
Pruebas de normalidad para formación bruta de capital

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Estadístico	gl	Probabilidad	Statistic	gl	Probabilidad
Form. Bruta de K.	.126	25	.200	.907	25	.030

Gráfica 14
Cuartil-Cuartil de Formación bruta de capital



Cuadro 39
Propiedades deseadas de los estimadores

Hipótesis incumplidas	Propiedades deseables de los estimadores					Contraste de significación	Predicción
	Insesgados	Eficientes	Consistentes	Suficientes	Lineal		
Multicolinealidad	-	-	-	-	-	Afecta al término independiente (media constante)	Predicciones no eficientes
Heterocedasticidad	-	Ineficiencia según comportamiento heterocedasticidad	-	-	-	Varianzas de los parámetros generalmente infravaloradas y constantes teóricos no validos	Predicciones no eficientes
Autocorrelación	-	Ineficiencia según comportamiento de autocorrelación	-	-	-	Varianzas de los parámetros generalmente infravaloradas y constantes teóricos no validos	Predicciones no eficientes
Normalidad	-	-	-	-	-	Invalidados o reducidos en su significación	Predicciones no eficientes

Ejercicio en el que se aplican las violaciones a los supuestos.
Cuadro 40
Dependent Variable: X
Method: Least Squares

Sample: 1980:1 2005:4
Included observations: 104

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-8721.057	3780.913	-2.306601	0.0232
PIB	-0.000114	2.30E-05	-4.972278	0.0000
TC	1204.688	157.8829	7.630262	0.0000
M	0.550443	0.051089	10.77418	0.0000
PP	331.1358	24.17063	13.69992	0.0000
IED	0.113380	0.093703	1.209993	0.2292
PIBEU	0.000776	0.000660	1.176636	0.2422

R-squared	0.994894	Mean dependent var	24581.79
Adjusted R-squared	0.994579	S.D. dependent var	17672.59
S.E. of regression	1301.241	Akaike info criterion	17.24496
Sum squared resid	1.64E+08	Schwarz criterion	17.42295
Log likelihood	-889.7379	F-statistic	3150.268
Durbin-Watson stat	0.761598	Prob(F-statistic)	0.000000

Es claro que presenta multicolinealidad el modelo (IED y PIBEU) estadísticamente no significativos R^2 muy alto y la prueba global es estadísticamente significativa.
El error estándar de la regresión y la suma de errores cuadrados nos indica que las varianzas son grandes por lo tanto es posible pensar que hay heteroscedasticidad.
El estadístico DW es muy pequeño (se aleja del valor dos) por lo que se sospecha la presencia de autocorrelación.
Por el número de observaciones es posible afirmar que el modelo presenta normalidad y que los datos se distribuyen de una forma normal.

Ahora al estimar el siguiente modelo:
Cuadro 41
Dependent Variable: X
Method: Least Squares
Date: 06/04/06 Time: 14:29
Sample: 1980:1 2005:4
Included observations: 104

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4254.883	389.0385	-10.93692	0.0000
PIB	-0.000132	1.60E-05	-8.216363	0.0000
TC	1317.703	129.9919	10.13681	0.0000
M	0.606772	0.029894	20.29721	0.0000
PP	321.9466	23.60114	13.64115	0.0000

R-squared	0.994751	Mean dependent var	24581.79
Adjusted R-squared	0.994539	S.D. dependent var	17672.59
S.E. of regression	1306.028	Akaike info criterion	17.23425
Sum squared resid	1.69E+08	Schwarz criterion	17.36139
Log likelihood	-891.1810	F-statistic	4690.153
Durbin-Watson stat	0.698365	Prob(F-statistic)	0.000000

Se corrige la multicolinealidad, (se eliminaron 2 variables) todas las variables independientes son estadísticamente significativas.

Al aplicar la prueba de White se observa que la heteroscedaticidad desaparece del modelo
Cuadro 42(a)
White Heteroskedasticity Test:

F-statistic	1.862682	Probability	0.074987

Obs*R-squared	14.10128	Probability	0.079163

Aplicando la prueba White para datos cruzado con una mayor holgura se muestra que no hay presencia de heteroscedasticidad.

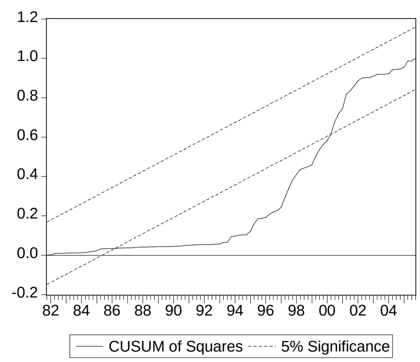
Cuadro 42(b)

White Heteroskedasticity Test:

F-statistic	1.443737	Probability	0.150070
Obs*R-squared	19.24765	Probability	0.155701

Al haber eliminado la heteroscedasticidad del modelo no es garantía que el mismo sea estable, esto se observa con la gráfica de cusum cuadrado y es claro que cuando la línea sale de las bandas ocurre lo que conocemos en economía como cambio estructural y esto le quita estabilidad al modelo ver gráfica 15.

Gráfica 15



La estimación que aparece en el cuadro es utilizando el método de Koyck que dice que la variable dependiente puede pasar como variable independiente, pero rezaga, entonces estamos hablando de una econometría dinámica en la todo se mueve y nada permanece estático, el estadístico Durbin Watson ya no es funcional para este modelo, por eso se utiliza la prueba Breusch Godfrey Serial.

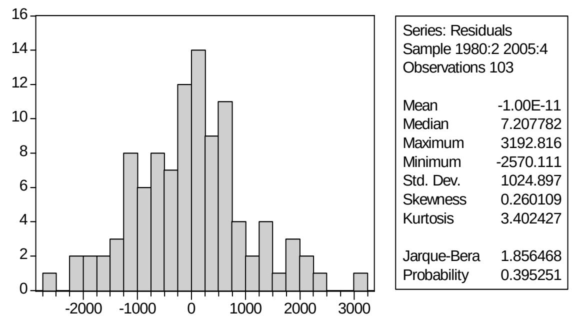
Cuadro 43

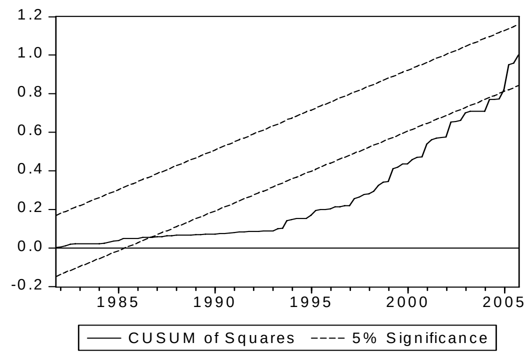
Dependent Variable: X
Method: Least Squares
Sample(adjusted): 1980:2 2005:4
Included observations: 103 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2897.899	364.2209	-7.956432	0.0000
PIB	-9.09E-05	1.49E-05	-6.109028	0.0000
TC	807.4465	125.1478	6.451944	0.0000
M	0.365815	0.040957	8.931694	0.0000
PP	215.7790	23.99648	8.992111	0.0000
X(-1)	0.401548	0.054865	7.318807	0.0000
R-squared	0.996632	Mean dependent var		24764.03
Adjusted R-squared	0.996459	S.D. dependent var		17660.53
S.E. of regression	1050.980	Akaike info criterion		16.80932
Sum squared resid	1.07E+08	Schwarz criterion		16.96280
Log likelihood	-859.6800	F-statistic		5740.954
Durbin-Watson stat	1.826997	Prob(F-statistic)		0.000000

Es importante señalar que en el examen de signos todas las variables independientes se comportan de acuerdo a la teoría económica, es decir, los signos son los esperados a excepción del PIB, que tenemos un signo negativo, esto ocurre debido a que los datos pueden estar no deflactados, o la bese es diferente, pero en nuestro caso los datos están “limpios” no hay problema de unidades ni de fuente, entonces la pregunta surge ¿qué debemos hacer para solucionar este problema? , bueno vamos a estimar otro modelo y veamos si se corrige el signo.

Gráfica 16 y Gráfica 16(b)





En la gráfica 16 muestra que los residuales se comportan de manera normal, a un nivel de confianza del 95%, la gráfica 16(b) llamada cusum muestra que existe un cambio estructural en la variable que estamos estimando, esto puede ser debido a una devaluación, a que dejamos de ser competitivos en el mercado mundial y las exportaciones ya no fueron atractivas, este cuadro se presta más para un análisis económico.

Cuadro 44

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.630202	**Probability**	**0.429239**
Obs*R-squared	0.671744	Probability	0.412444

El cuadro 44 solo esta probando la no existencia de autocorrelación la cual es probada a un nivel de confianza del 95% y estamos aceptando la hipótesis nula y se observa que no presenta autocorrelación,

Cuadro 45

ARCH Test:

F-statistic	6.734839	**Probability**	**0.010876**
Obs*R-squared	6.436076	Probability	0.011183

Cuadro 45(a)

White Heteroskedasticity Datos No cruzadosTest:

F-statistic	3.017984	**Probability**	**0.002424**
Obs*R-squared	25.44220	Probability	0.004567

Cuadro 45(b)

White Heteroskedasticity Datos CruzadosTest:

F-statistic	3.487474	**Probability**	**0.000033**
Obs*R-squared	47.34248	Probability	0.000526

Los cuadros 45, 45(a) y 45(b) prueban la homoscedasticidad en el modelo, en los tres es clara la existencia de heteroscedasticidad a un nivel de confianza del 95%. Se concluye que el modelo presenta heteroscedasticidasd.
La nueva estimación se presenta en el cuadro 46, en la que se incluye en método llamado autorregresivo (que en el siguiente capítulo se desarrolla la teoría llamada Box-Jenkins) se busca que con esta nueva estiación el modelo mejore en autocorrelación y eliminemos la heteroscedasticidad.

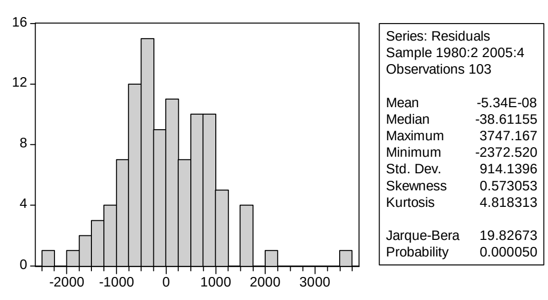
Cuadro 46
Dependent Variable: X
Method: Least Squares
Sample(adjusted): 1980:2 2005:4
Included observations: 103 after adjusting endpoints
Convergence achieved after 14 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2499.154	915.2850	-2.730465	0.0075
PIB	-8.84E-05	4.12E-05	-2.147433	0.0343
TC	1043.738	177.2348	5.889012	0.0000
M	0.683999	0.036337	18.82360	0.0000
PP	194.4679	32.40945	6.000346	0.0000
AR(1)	0.792016	0.067429	11.74595	0.0000

R-squared	0.997321	Mean dependent var	24764.03
Adjusted R-squared	0.997183	S.D. dependent var	17660.53
S.E. of regression	937.4039	Akaike info criterion	16.58059
Sum squared resid	85236433	Schwarz criterion	16.73407
Log likelihood	-847.9005	F-statistic	7221.361
Durbin-Watson stat	1.984777	Prob(F-statistic)	0.000000
Inverted AR Roots	.79		

Regresando al signo con el PIB se ve que persiste el problema, aún cuando muestra una significancia estadística, se realizará otra estimación más y veremos si así es como se puede corregir este signo.

Gráfica 17



La gráfica prueba si existe normalidad en los residuales, esta vez con la nueva estimación dejo de ser normal, entonces el modelo presenta la no aceptación de la hipótesis nula y se conluye que tiene problemas de normalidad.

Cuadro 47
Breusch-Godfrey Serial Correlation LM Test:

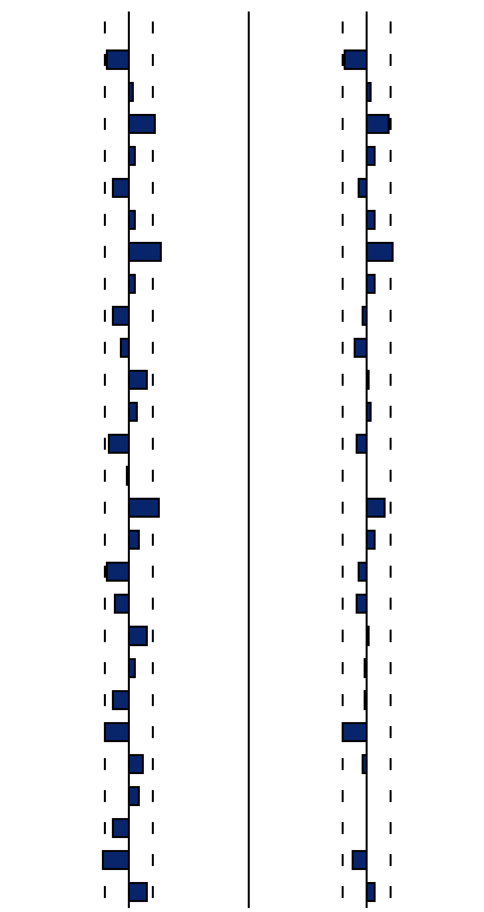
F-statistic	0.002925	**Probability**	**0.956981**
Obs*R-squared	0.003138	Probability	0.955326

El cuadro 47 prueba la no existencia de autocorrelación y de acuerdo a la probabilidad asociada al estadístico F (0.956981) se concluye que el modelo no presenta autocorrelación.

Gráfica 18

Correlogram of Residuals

Sample: 1980:2 2005:4
Included observations: 103
Q-statistic probabilities adjusted for 1 ARMA term(s)



	AC	PAC	Q-Stat	Prob
1	-0.004	-0.004	0.0019	
2	-0.193	-0.193	3.9871	0.046
3	0.035	0.034	4.1174	0.128
4	0.226	0.196	9.7001	0.021
5	0.049	0.068	9.9640	0.041
6	-0.140	-0.072	12.140	0.033
7	0.056	0.061	12.487	0.052
8	0.283	0.223	21.581	0.003
9	0.044	0.061	21.806	0.005
10	-0.138	-0.038	24.022	0.004
11	-0.069	-0.099	24.585	0.006
12	0.151	0.021	27.304	0.004
13	0.074	0.036	27.967	0.006
14	-0.165	-0.089	31.286	0.003
15	-0.014	-0.006	31.311	0.005
16	0.260	0.156	39.718	0.001
17	0.089	0.076	40.714	0.001
18	-0.188	-0.063	45.192	0.000
19	-0.125	-0.098	47.194	0.000
20	0.155	0.010	50.329	0.000
21	0.055	-0.019	50.728	0.000
22	-0.134	-0.021	53.114	0.000
23	-0.213	-0.219	59.228	0.000
24	0.127	-0.035	61.423	0.000
25	0.090	0.000	62.542	0.000
26	-0.145	0.001	65.475	0.000
27	-0.225	-0.118	72.708	0.000
28	0.149	0.080	75.892	0.000

Cuadro 48
ARCH Test:

F-statistic	0.015087	Probability	0.902489
Obs*R-squared	0.015387	Probability	0.901282

Cuadro 48 (a)
White Heteroskedasticity Daton No Cruzados Test:

F-statistic	1.978839	Probability	0.057466
Obs*R-squared	14.84615	Probability	0.062207

Cuadro 48 (b)

White Heteroskedasticity Datos Cruzados Test:			
F-statistic	1.354140	Probability	0.193299
Obs*R-squared	18.25642	Probability	0.195350

Los cuadros 48, 48(a) y 48(b) prueban la existencia de homoscedasticidad y de acuerdo a la probabilidad asociada al estadístico F en las tres pruebas es claro que no tiene problemas de heteroscedasticidad, es decir mejor de acuerdo al modelo del cuadro 43.

En la gráfica 18 se complementa la estimación del cuadro 46 y se estará probando la existencia de no autocorrelación es claro que algunas barras salen de las bandas de confianza y además la probabilidad asociada es menor a 0.05, por lo que la gráfica indica que hay autocorrelación, pero puede ocurrir que la prueba Breusch-Godfrey Serial indique o contrario como es este caso y la duda es que decidimos, la respuesta es que la prueba Breusch-Godfrey es más contundente por que utiliza las variables independientes además de los residuales y el correlograma solo muestra el comportamiento de los residuos.

En el cuadro 49 se estima el modelo agregando un modelo de medias móviles (MA) de orden 2, (de largo plazo) la intención aparte de mantener la no autocorrelación tanto en el correlograma como en la prueba de Breusch-Godfrey, también de mantener la homoscedasticidad es corregir el problema del signo que se viene arrastrando desde la estimación del cuadro 43.

Cuadro 49

Dependent Variable: X
Method: Least Squares
Sample(adjusted): 1980:2 2005:4
Included observations: 103 after adjusting endpoints
Convergence achieved after 21 iterations
Backcast: 1979:4 1980:1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2220.592	1273.570	-1.743597	0.0844
PIB	-7.14E-05	4.36E-05	-1.636933	0.1049
TC	908.5775	177.3947	5.121786	0.0000
M	0.725029	0.032844	22.07509	0.0000
PP	162.7022	32.61061	4.989241	0.0000
AR(1)	0.912242	0.055014	16.58191	0.0000
MA(2)	-0.310588	0.122582	-2.533711	0.0129

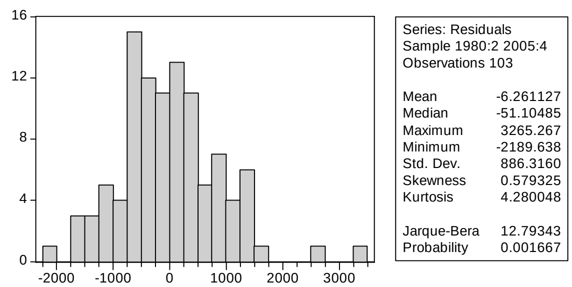
R-squared	0.997481	Mean dependent var	24764.03
Adjusted R-squared	0.997324	S.D. dependent var	17660.53
S.E. of regression	913.6166	Akaike info criterion	16.53824
Sum squared resid	80130754	Schwarz criterion	16.71730
Log likelihood	-844.7194	F-statistic	6336.264
Durbin-Watson stat	2.044954	Prob(F-statistic)	0.000000

Inverted AR Roots	.91	
Inverted MA Roots	.56	-.56

Hay que señalar que todas las variables independientes son estadísticamente significativa a excepción del PIB y el problema del signo sigue presente, entonces la pregunta es que hacer con ese signo que va contra lo que nos dice la teoría. Lo que hay que revisar primero es que si la teoría dice que si la especificación del modelo es la correcta ya se tiene un gran avance, después ver si hay algunas otras variables que puedan complementar la estimación, si ya no es posible agregar algo al modelo entonces el fundamento teórico es el que le da el sustento al modelo aunque estadísticamente no se cumplan los supuestos, esto es, que el problema de estadística es un problema de datos, de muestra y la teoría económica es la que va a dar respuesta a las interrogantes planteadas en el modelo. Desde 1992 a la fecha que estoy trabajando con la econometría y he revisado gran cantidad de libros de estadística, econometría, series de tiempo y de teoría económica, que ninguno de ellos habla sobre este problema del signo, este ejemplo tiene la intención de mostrar que en la vida profesional se presentan estos problemas y a veces no se sabe que respuesta dar.

Otra alternativa es cambiar las variables independientes para dar respuesta a la variable dependiente, esto claro que lleva un poco más de tiempo, más las dificultades intrínsecas que lleva la nueva estimación, el estimar de nueva cuenta el modelo con las nuevas variables puede ser otra opción, por ejemplo sabemos que para Keynes el consumo está en función del ingreso y eso nadie lo discute o lo refuta, pero que ocurre si el signo que esperamos no es el correcto, entonces que debemos hacer pues sustentarlo teóricamente y ya quedo. Pero si en la vida profesional ocurre esto, se debe dar respuesta entonces se usan otras variables que puedan explicar al consumo, como por ejemplo la inversión, el tipo de cambio, el mismo ingreso, el nivel de empleo, en fin buscar nuevas variables para que se de un resultado que nos están pidiendo, aunque puede también ocurrir que esto vaya en contra de lo que dice la teoría.

Gráfica 19



La gráfica 19 muestra que los residuales no se comportan de manera normal, esto se puede solucionar de acuerdo al teorema de límite centra que dice que a medida que se aumenta la muestra el comportamiento de la misma tiende a ser normal.

Cuadro 50

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.259150	**Probability**	**0.611885**
Obs*R-squared	0.275033	Probability	0.599975

Se observa que el modelo no presenta autocorrelación.

Cuadro 51

ARCH Test:

F-statistic	0.490409	Probability	0.485371
Obs*R-squared	0.497776	Probability	0.480479

Cuadro 51(a)

White Heteroskedasticity Datos No Cruzados Test:

F-statistic	2.425111	Probability	**0.019868**
Obs*R-squared	17.62148	Probability	0.024250

Cuadro 51(b)

White Heteroskedasticity Datos Cruzados Test:

F-statistic	1.695675	Probability	0.070733
Obs*R-squared	21.88272	Probability	0.081056

Los cuadro 51 y 51(b) muestras la existencia de homoscedasticidad, no así el cuadro 51(a) para datos no cruzados, este problema se presenta, puede ser que pase dos pruebas de heteroscedasticidad y una no entonces ¿qué hay que hacer? La prueba White para datos no cruzados la podemos evaluar a un nivel de confianza del 90% y entonces si estaría pasando esa violación al supuesto.

Gráfica 19

Correlogram of Residuals

Sample: 1980:2 2005:4
Included observations: 103
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.043	-0.043	0.1928	
		2	-0.028	-0.030	0.2747	
		3	0.006	0.003	0.2783	0.598
		4	0.153	0.153	2.8464	0.241
		5	0.035	0.050	2.9818	0.394
		6	-0.148	-0.140	5.4273	0.246
		7	0.027	0.013	5.5092	0.357
		8	0.197	0.179	9.9097	0.129
		9	0.012	0.020	9.9254	0.193
		10	-0.158	-0.128	12.826	0.118
		11	-0.071	-0.090	13.410	0.145
		12	0.076	0.003	14.103	0.168
		13	0.062	0.070	14.559	0.204
		14	-0.161	-0.077	17.695	0.125
		15	0.006	0.010	17.699	0.169
		16	0.209	0.157	23.119	0.058
		17	0.076	0.076	23.853	0.068
		18	-0.145	-0.077	26.525	0.047
		19	-0.088	-0.077	27.523	0.051
		20	0.116	0.020	29.262	0.045
		21	-0.018	-0.062	29.304	0.061
		22	-0.076	0.010	30.085	0.068
		23	-0.205	-0.193	35.764	0.023
		24	0.136	0.009	38.292	0.017
		25	0.023	0.015	38.365	0.023
		26	-0.070	0.050	39.050	0.027
		27	-0.209	-0.139	45.279	0.008
		28	0.133	0.070	47.822	0.006

En está gráfica se observa que las barras ahora ya no salen de las bandas de confianza, lo que nos está indicando que la presencia de autocorrelación es nula.

3. CAPITULO

3.1. Pronóstico y series de tiempo[19]

Cabe subrayar la importancia que el hombre siempre le ha dado al conocimiento del futuro y su gran preocupación por establecer leyes y patrones de comportamiento, a partir de la observación de regularidades empíricas en los fenómenos que estudia. De hecho, aun cuando en algunas áreas del conocimiento se puede llegar a anticipar el resultado de alguna acción, en ocasiones prácticamente sin error (como ocurre en las ciencias físicas o naturales), en las ciencias sociales el hombre ha notado su incapacidad para lograrlo. Aunque no por ello se declara incompetente, sino simplemente falto de conocimientos, ya sea de la teoría correspondiente al fenómeno social o de las técnicas apropiadas para analizar dicho fenómeno.

Conviene advertir la existencia de términos semejantes al de pronóstico para referirse al conocimiento del futuro; tal es el caso de la *profecía,* que recurre a la inspiración divina para anticipar el futuro sin posibilidad de equivocación. También se usa el término de *previsión,* con una connotación en la cual, independientemente de cómo se llegue a obtener el conocimiento del futuro, se excluye la posibilidad de error. De igual manera se usan los términos *proyección* y *extrapolación* para indicar el uso mecánico de instrumentos que anticipen la ocurrencia de algún evento futuro (tal es el caso de las ecuaciones puramente matemáticas) en donde tampoco intervienen elementos de error aleatorio. Por su lado, el término *predicción* está más íntimamente ligado con el de pronóstico en tanto que su uso generalmente incorpora, de manera implícita el error atribuible a la aleatoriedad en los fenómenos y, por ende el uso de técnicas estadísticas para tener en cuenta dicho error.

Pronosticar, según el Diccionario de la Real Academia Española, significa "Conocer por algunos indicios, lo futuro". En la antigüedad este término era tomado en su acepción más estrecha, de tal manera que el pronóstico sólo podía ser realizado por unos cuantos seres privilegiados, a quienes les era dado conocer el futuro con certeza. En la actualidad, el término pronóstico sigue estando asociado con el conocimiento del futuro, mas ya no de una manera completamente cierta, sino aceptando explícitamente márgenes de error. Esto se debe, en particular, a que las personas que pronostican hoy en día, no tienen porqué. Ser necesariamente videntes del futuro y más bien son analistas, que buscan indicios del futuro en la historia o en la teoría correspondiente al fenómeno que se estudia.

La necesidad del pronóstico en nuestros días, es clara en cuanto se reconoce que los tomadores de decisiones, ya sean funcionarios públicos, ejecutivos de empresa o jefes de familia, tienen que decidir sus acciones con base en lo que prevean para el futuro. Es decir, pronosticar no es un fin en sí mismo, sino únicamente un instrumento que apoya a la planeación, puesto que permite reducir la incertidumbre asociada con el futuro. Así pues, con el pronóstico se trata de anticipar eventos que podrían ocurrir en el futuro e incorporar su posible impacto en las acciones que se tomen en el presente. Para tener en cuenta la compleja situación actual y los distintos factores que influyen sobre un determinado fenómeno que se desea pronosticar, se han desarrollado diversas técnicas que tienen usos específicos, de tal manera que para usarlas se debe saber en qué casos y bajo qué condiciones son aplicables, y no proceder a utilizarlas indiscriminadamente.

Al realizar un pronóstico, se requiere conocer primero las condiciones presentes, ya que de nada sirve tener conocimiento de los cambios que se darán en el futuro, sin saber dónde se está en el momento actual. Por ello es importante allegarse datos relativos al estado actual del fenómeno en estudio y analizados, a fin de extraer de ellos la información acerca del futuro. Si dichos datos son de carácter numérico, la manera óptima de analizarlos es mediante el uso de procedimientos estadísticos. Dentro de estos procedimientos, sobresalen aquellos que están basados en modelos,

[19] Serie temporal o cronológica, es una sucesión de valores observados de una variable referidos a momentos o a periodos de tiempo diferentes.

ya que es en los modelos donde se resumen los conocimientos teóricos y los supuestos acerca del fenómeno de interés. Esto último permite discriminar entre los diversos procedimientos disponibles, para decidir cuál es el más apropiado, dependiendo del contexto dentro del cual se desea obtener el pronóstico.

Puede suceder que los datos de que se disponga sean insuficientes para construir un modelo o bien sean de tipo no-numérico. En esas condiciones, las técnicas de pronóstico que se pueden utilizar, no se fundamentan en modelos estadísticos y son más bien de carácter cualitativo. Entre otras cosas, las distinciones respecto al tipo y al volumen de los datos disponibles, conducen a efectuar una división dentro de las técnicas de pronóstico. Por un lado se tienen las técnicas cualitativas, que son en esencia subjetivas, aun cuando pretendan sistematizar mediante algún camino lógico y formal, los datos y las conjeturas del pronosticador. Por el otro están las técnicas cuantitativas, que utilizan herramientas matemáticas y pretenden ser objetivas en tanto que surgen del análisis de los datos observados, de tal forma que dos pronosticadores que utilicen el mismo conjunto de datos y la misma técnica cuantitativa de pronóstico, deben llegar esencialmente al mismo pronóstico.

El conjunto básico de datos que se supondrá disponible, es el que se conoce como una serie de tiempo, es decir, una sucesión de observaciones sobre una cierta variable, efectuadas a intervalos equidistantes en el tiempo y ordenadas cronológicamente, la cual se denotará genéricamente por {Zt}, para hacer explícita el número de observaciones de la serie, también se podrá escribir como $\{Z_1,Z_2,...,Z_N\}$. Al hacer referencia a la serie de esta manera, de antemano deberá especificarse la variable Z involucrada, la separación temporal entre observaciones consecutivas (o sea, la periodicidad del registro de los datos), el número, de datos (N) y el origen a partir del cual se cuentan las observaciones (es decir, el momento al que corresponde la observación Z_1).

Una serie de tiempo puede considerarse generalmente constituida por varios componentes, que aunque no sean observables directamente por separado, sí se puede suponer su presencia en la serie, ya sea por el conocimiento teórico que de ella se tenga a por la inspección preliminar que se haya realizado de la misma. Independientemente de que una serie pueda estar constituida por otros elementos, el componente que siempre estará presente es el que, desde un punto de vista estadístico, se considera variación aleatoria y que justifica la aplicación de técnicas estadísticas para el análisis de la serie. Dicho de otro modo, si la serie de tiempo estuviese formada tan solo por componentes deterministas, no existiría aleatoriedad y el problema del pronóstico se reduciría a la mera extrapolación de tales componentes, sin tener en cuenta la herramienta estadística propiamente dicha, sino a lo más, las fórmulas matemáticas involucradas en la representación de los componentes existentes en la serie.

3.2. Análisis gráfico de las series

Casi siempre es bueno comenzar los proyectos de pronóstico con un análisis gráfico, ya que en muchos aspectos el ojo humano es una herramienta mucho más poderosa para analizar y modelar datos que la más sofisticada de las herramientas de modelado. Pero debe quedar claro que esto no quiere decir que se sacará el trabajo sólo con el análisis gráfico (las gráficas tienen muchas limitaciones propias) sin embargo son lo mejor para comenzar.

El conjunto de datos que se muestra en el cuadro 36 permite apreciar nítidamente el poder de las gráficas.

Como [Diebold F. 1999] muestra que cada conjunto de datos esta formado por once observaciones de dos variables. Si tan sólo se echa una mirada a estos se comprende poco. Naturalmente que nos preguntaremos para eso contamos con modernas técnicas estadísticas. Por ejemplo la regresión lineal, por lo tanto se determina la regresión de **Y** a partir de **X** para cada uno de los conjuntos de datos.

Cuadro 51
Cuatro conjuntos de datos

Y_1	X_1	Y_2	X_2	Y_3	X_3	Y_4	X_4
8.04	10	9.14	10	7.46	10	6.58	8
6.95	8	8.14	8	6.77	8	5.76	8
7.58	13	8.74	13	12.74	13	7.71	8
8.81	9	8.77	9	7.11	9	8.84	8
8.33	11	9.26	11	7.81	11	8.47	8
9.96	14	8.1	14	8.84	14	7.04	8
7.24	6	6.13	6	6.08	6	5.25	8
4.26	4	3.1	4	5.39	4	12.5	19
10.84	12	9.13	12	8.15	12	5.56	8
4.82	7	7.26	7	6.42	7	7.91	8
5.68	5	4.74	5	5.73	5	6.89	8

Cuadro 51(b)
Resultado de las regresiones de los cuatro conjuntos de datos

LS//Dependent Variable is Y1

Variable	Coefficient	Std. Error	T-Statistic
C	3.000091	1.124747	2.667348
X1	0.500091	0.117906	4.241455

R-squared 0.666542 S.E. of regression 1.236603

LS // Dependent Variable is Y2

Variable	Coefficient	Std. Error	T-Statistic
C	3.000909	1.125302	2.666758
X2	0.500000	0.117964	4.238590

R-squared 0.666242 S.E. of regression 1.237214

LS // Dependent Variable is Y3

Variable	Coefficient	Std. Error	T-Statistic
C	3.002455	1.124481	2.670080
X3	0.499727	0.117878	4.239372

R-squared 0.666324 S.E. of regression 1.236311

Continua cuadro 37

LS // Dependent Variable is Y4

Variable	Coefficient	Std. Error	T-Statistic
C	3.001727	1.123921	2.670763
X4	0.499909	0.117819	4.243028
R-squared	0.666707	S.E. of regression	1.235695

Es interesante que aunque los cuatro conjuntos de datos contienen valores numéricos distintos de datos, el resultado de la regresión lineal es idéntico para todos.

En primer lugar, la línea de regresión ajustada es igual en cada caso **y = 3+1/2x**.

En segundo lugar, los coeficientes y los errores estándar, también son iguales en cada conjunto de datos, en consecuencia los estadísticos **t**.

En tercer lugar, el coeficiente de determinación R^2 .

En cuarto lugar, la suma de los residuos elevados al cuadrado y por consiguiente, el error estándar de la regresión es igual.

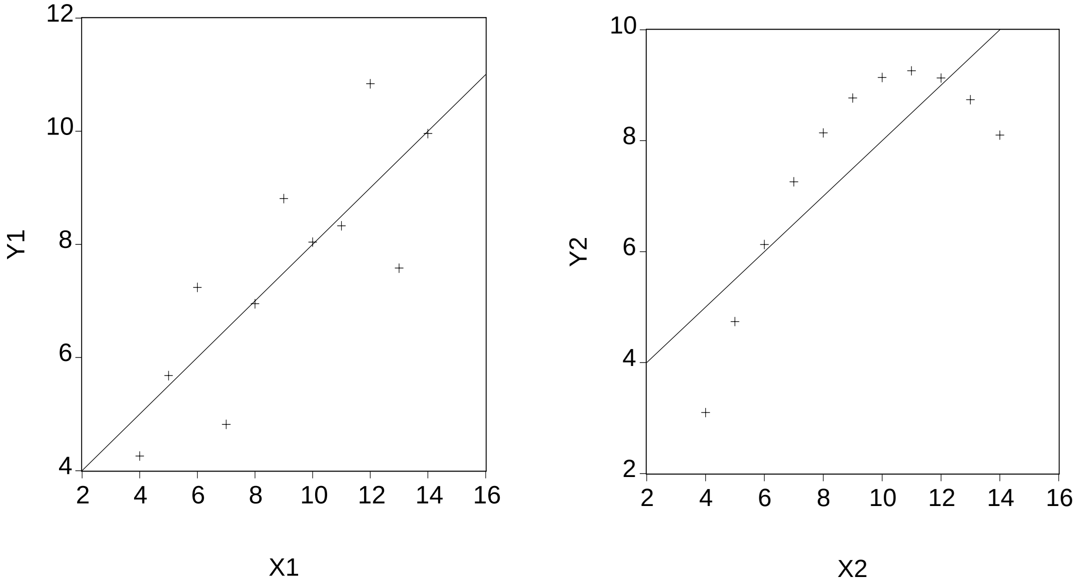
Hasta aquí parece que todo marcha bien, se podría decir que la relación entre **Y** y **X** simplemente es igual en cada conjunto de datos aun cuando los datos específicos difieren debido a influencias aleatorias. Pero un análisis gráfico de los datos indica que no es correcta esta aseveración. En la gráfica 10 se muestran las líneas de **Y** en función de **X** (diagramas de dispersión)

En el primer conjunto de datos parece estar todo bien, las variables tienen correlación positiva, y parecen apegarse muy bien a una relación lineal.

En el segundo conjunto de datos la gráfica indica que si hay una relación entre las variables pero se aprecia que no es una relación lineal. Por consiguiente no se aconseja el uso del modelo de regresión lineal.

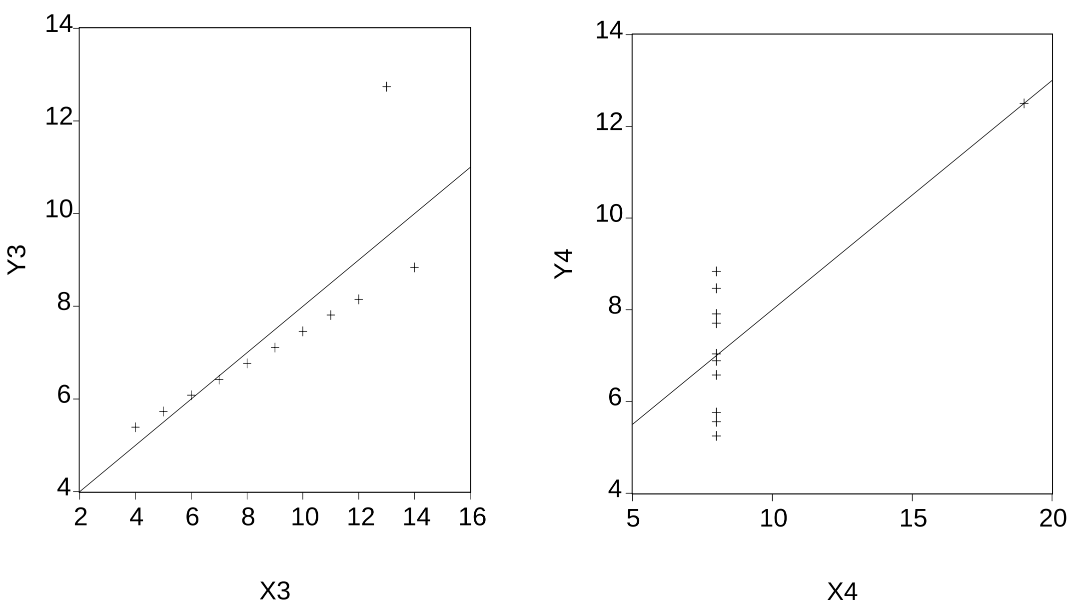
En el conjunto de datos tres, la gráfica indica que las variables parecen apegarse a una relación lineal, hay un par que no apega bien a ella. Lo más probable es que no nos percatamos de este dato cuando examinamos los datos en forma tabular, a pesar de que cuando se emplean las gráficas visualmente es obvio.

Gráfica 20
Diagrama de dispersión conjunto 1 y 2



En el conjunto cuatro las variables aparecen en una serie vertical, a excepción de un punto y éste ejerce una gran influencia sobre la línea de regresión. Nuevamente la gráfica hace aparente la naturaleza atípica de esta situación.

Gráfica 20 (continua)
Diagrama de dispersión conjunto 3 y 4



Con este ejemplo se observa la gran importancia que tiene el empezar un análisis con el uso de gráficas, resumiendo lo que se explica acerca del poder de las misma es: qué permiten resumir y revelar pautas en los datos así como a identificar anomalías en los mismos además facilitan y promueven las comparaciones, finalmente nos presentan e interpretan cantidades inmensas de datos en un espacio pequeño.

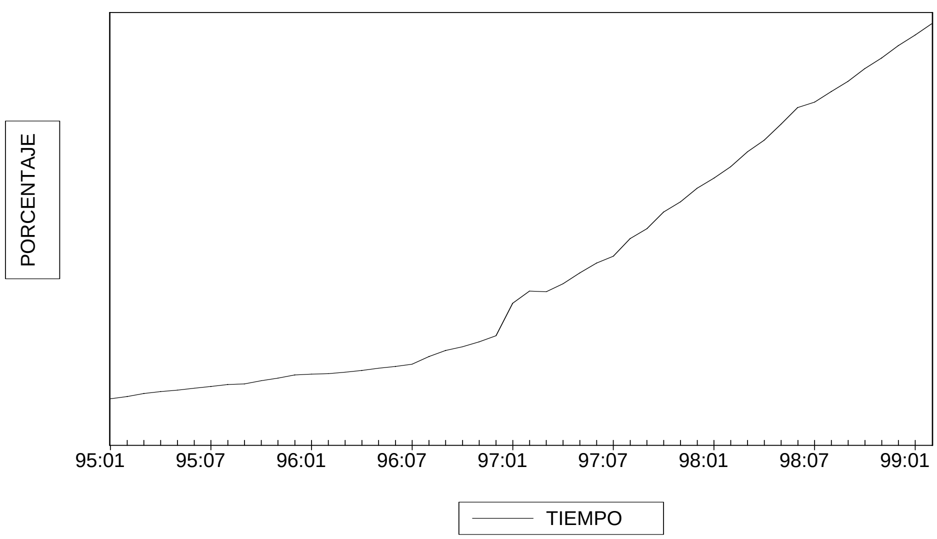
Dentro del enfoque para una sola variable, el análisis clásico de series de tiempo comprende cada serie temporal integrada por cuatro componentes:

3.3. Características y/o componentes de una serie de tiempo

Tendencia. Es la evolución lenta y a largo plazo de las variables que deseamos modelar en las cuales es posible tomar mediciones para un análisis de series de tiempo cada hora, día, semana, mes o año o cualquier intervalo. Por lo general los datos de series de tiempo muestran variaciones aleatorias, es probable que muestren a un así cambios o movimientos hacia valores relativamente más altos o más bajos para un período prolongado. El cambio gradual de la serie de tiempo, que por lo general se debe a factores a largo plazo como cambios en la población, cambios en tecnología y cambios en las preferencias de los consumidores es lo que se denomina la tendencia de una serie de tiempo.

Las series económicas presentan a menudo evoluciones regulares a largo plazo, que se denominan tendencias. La existencia de una tendencia a simple vista es obvia. Como se puede ver la tendencia de la gráfica 21 es ascendente.

Gráfica 21
Porcentaje de la población femenina en la fuerza laboral

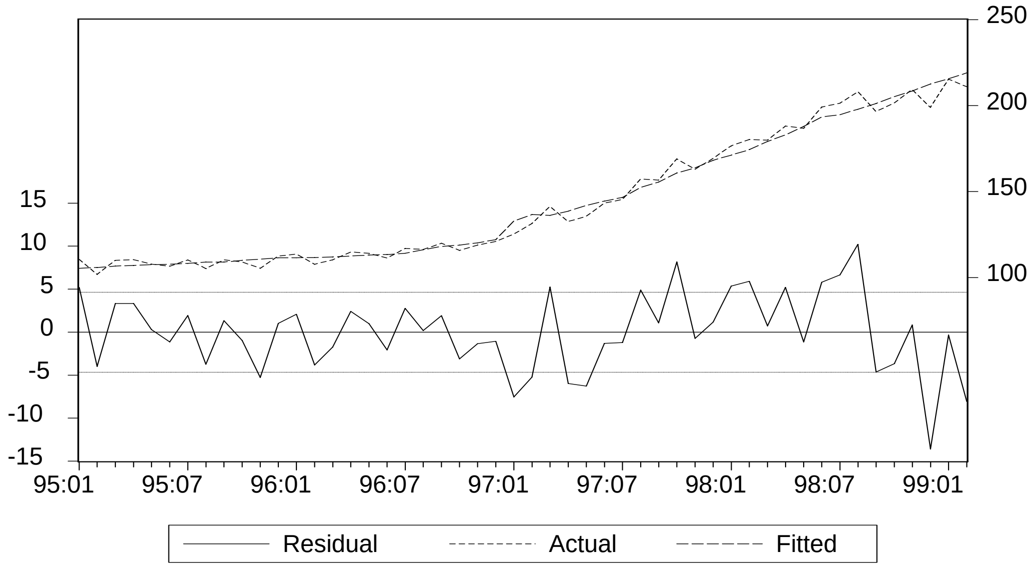


En ocasiones la tendencia parece ser no lineal o curva, por ejemplo cuando una variable aumenta con rapidez cada vez mayor o menor, en este caso no será necesario que las tendencias sean lineales, sino sólo uniformes. En la gráfica 12 se observa como el volumen aumenta con rapidez cada vez mayor, la tendencia es lineal. Esto indica que aumenta o disminuye como si fuera una línea recta. Es decir una función sencilla del tiempo.

$$.T_t = \beta_0 + \beta_1 \text{ Tiempo}_t$$

La variable tiempo se forma artificialmente y se llama indicador de tiempo. El tiempo es igual a 1 en el primer periodo de la muestra, 2 en el segundo y así sucesivamente.

Gráfica 22
Tendencia lineal



Ciclicidad. Es probable que una serie de tiempo puede mostrar un patrón gradual de cambio o tendencia para intervalos de tiempo, no se puede esperar que todos los valores futuros de las series sean exactamente correspondientes a la recta de tendencia. Es frecuente que las series de tiempo muestran secuencias alternantes de puntos por encima o por debajo de la recta de tendencia. Cualquier patrón regular de secuencias de puntos que duren más de un año es atribuible al componente cíclico de la serie de tiempo. Algunas series de tiempo muestran conducta cíclica con corridas normales de observaciones por debajo (descensos o valles) y por encima (ascensos o cima) de la línea de tendencia. La consideración general es que este componente de la serie representa movimientos cíclicos de varios años en la economía. Por ejemplo períodos de inflación moderada seguidos por períodos de inflación rápida, las ventas de juguetes por el día de reyes, la venta de muñecos de peluche en el mes de febrero o la venta de flores en el mes de mayo estas series pueden conducir a alternar por arriba o por debajo de una línea de tendencia.

Estacionalidad. Muchas series de tiempo muestran un patrón regular de variabilidad en períodos menores de un año. No debe sorprender que se denomine al componente de la serie de tiempo que representa la variabilidad de los datos debido a variaciones estacionales el componente estacional. Por lo general se piensa que el movimiento estacional ocurre en períodos menores de un año, se puede también utilizar para representar cualquier patrón que se repite con regularidad y que tiene menos de un año de duración. Por ejemplo los datos que indican el volumen de tráfico muestran la conducta estacional diaria, teniendo niveles pico en las horas de entrada y salida de la escuela y del trabajo, un flujo moderado el resto del día. Para poder apreciar mejor el elemento aleatorio de las series, conviene aislar primero aquellos componentes más cercanamente deterministas. La idea es llegar a obtener una serie reducida esencialmente a fluctuaciones aleatorias, no-sistemáticas, alrededor de un nivel constante, para poder estudiar la estructura de variabilidad que guardan estas fluctuaciones entre sí. Con este propósito en mente, conviene definir la noción de equilibrio estadístico. De hecho, se dice que la serie de tiempo $\{Z_1,Z_2,...,Z_N\}$ tiene equilibrio en su nivel (porque éste es constante) o bien que es una serie estacionaria de primer orden, si satisface que su primer momento (la media) no depende del tiempo, o sea, si

$$E(Z_t) = \mu \text{ para } t = 1, 2,...,N$$

De igual manera una serie estacionaria de segundo orden es aquella que no sólo tiene su nivel constante sino que su segundo momento también es independiente del tiempo. Al referirse al segundo momento de una serie de tiempo se debe considerar a la varianza y a las *autocovarianzas* de la serie que están dadas por la expresión

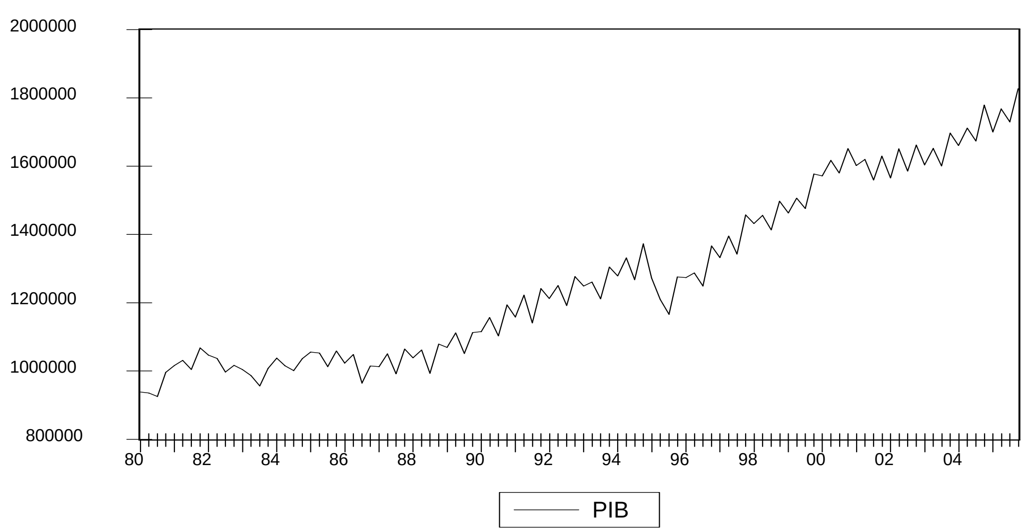
$$\gamma_k = Cov(Z_t, Z_{t+k}) = E[(Z_t - m)(Z_{t+k} - m)] \text{ para k=0,1,...} \tag{31}$$

La γ_o proporciona la varianza de la serie, es decir, $\gamma_k = Var(Z_t)$ y que γ_k depende sólo de la separación, k, que hay entre Z_t y Z_{t+k} mas no de *t*. Así pues para que una serie se pueda considerar estacionaria de segundo orden se requiere que tanto su nivel como su varianza, permanezcan constantes en el tiempo y que las autocovarianzas tampoco cambien con el tiempo. En la práctica se prefiere trabajar con la versión estandarizada de las autocovarianzas. a las cuales se conoce como autocorrelaciones y que dan origen a la Función de Autocorrelación (FAC) definida mediante

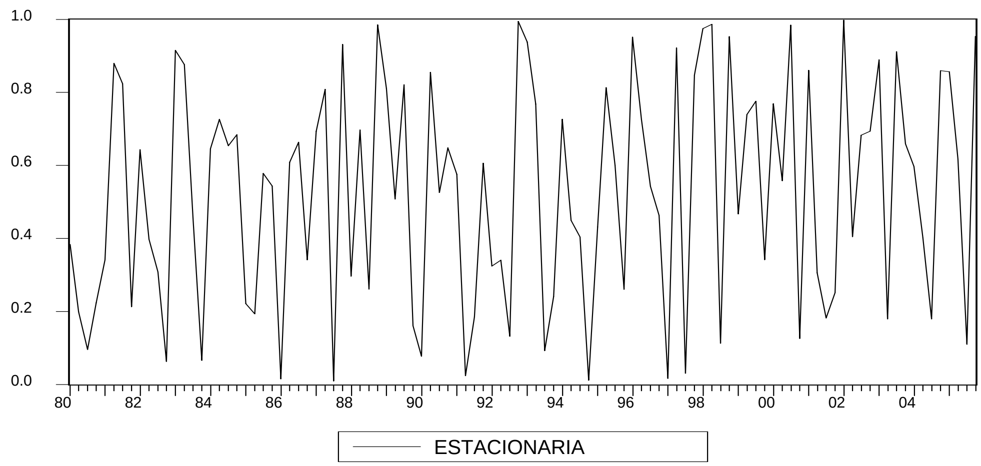
$$\rho_k = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)Var((Z_{t+k})}} = \frac{\gamma_k}{\gamma_0} \text{ para k=0,1....} \tag{32}$$

Si además de que la serie sea estacionaria de segundo orden se puede suponer razonablemente que las observaciones individuales de la serie siguen una distribución normal en cuyo caso la distribución conjunta de toda la serie $\{Z_1, Z_2, ..., Z_N\}$ es una normal multivariada entonces dicha distribución conjunta queda completamente caracterizada por sus dos primeros momentos, que se resumen a través de la media, la varianza y la FAC. En tal caso la serie se denomina estrictamente estacionaria, ya que la distribución conjunta de cualquier subconjunto de observaciones de la serie es la misma, independientemente del momento en que se observe. En las gráficas 23 a 25 se muestran algunos comportamientos típicos de series, tanto estacionarias como no-estacionarias. Es de hacer notar que, con frecuencia, se dice simplemente que una serie es estacionaria, sin especificar si la estacionariedad es de primer orden, de segundo orden o estricta; en tales casos, lo usual es que la serie sea estacionaria de segundo orden. Además, como sinónimo de estacionaria de segundo orden, se usa débilmente estacionaria, mientras que estacionariedad fuerte es sinónimo de estacionariedad estricta,

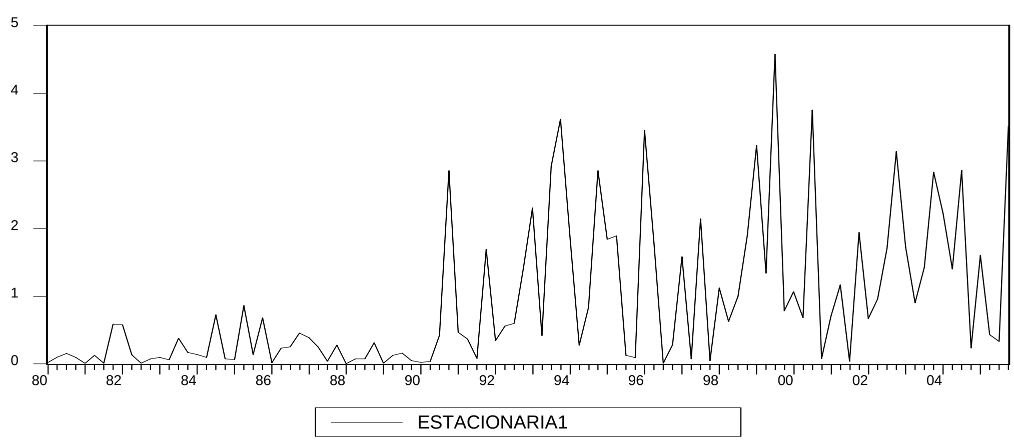
Gráfica 23
No Estacionaria



Gráfica 24
Estacionaria de 2° orden



Gráfica 25
Estacionaria de 1° orden solamente



Una clase particularmente importante de series de tiempo, la forman aquellas cuya media y varianza son constantes y además sus observaciones son independientes entre sí. A ese tipo de series, se les considera generadas por un mecanismo denominado proceso de ruido blanco, y la serie es conocida como ruido blanco. De hecho, un ruido blanco[20] es un conjunto de variables aleatorias independientes e idénticamente distribuidas, ordenadas cronológicamente. Si la distribución asociada con las observaciones es una normal, entonces el que sus autocorrelaciones (para $k \geq 1$) sean iguales a cero, permite concluir que hay independencia entre las observaciones. Asimismo, para caracterizar por completo a un ruido blanco cuya distribución es normal, es suficiente con especificar su media y su varianza, puesto que todas las autocorrelaciones serán cero, para $k \geq 1$.

Si se desea profundizar sobre el tema de la estacionalidad se puede ver el trabajo del profesor [Guerrero V. I-1998].

[20] Se definió en el capítulo 2.

Irregularidad Se puede definir como el factor residual o que abarca el resto y da cuenta de las desviaciones de los valores reales de la serie de tiempo con respecto a lo que se esperaría dados los efectos de los componentes de tendencia, ciclo y estacional. Da cuenta de la variabilidad aleatoria en la serie de tiempo. El componente irregular es ocasionado por factores a corto plazo, no previstos y no recurrentes. Este componente representa la variabilidad aleatoria de la serie resulta impredecible.

Existen tres modelos de descomposición de la serie Y que suelen usarse alternativamente en las aplicaciones, según se comporte la serie analizada. Estos tres modelos son el aditivo $Y_t = T_t + C_t + E_t + I_t$ el multiplicativo $Y_t = T_t * C_t * E_t * I_t$, y el mixto $Y_t = T_t(1 + C_t)(1 + E_t) + I_t$,

“Aun cuando se dispone de fuerte evidencia empírica de la presencia de tales componentes en muchas series temporales económicas, debe tenerse bien presente que cualquier clase de descomposición que se haga en la práctica constituye la aplicación de un modelo y que los modelos pueden ser útiles para representar algunos comportamientos, pero pueden fallar en otros casos.

El análisis clásico de series temporales consiste en describir las pautas de regularidad que sigue cada uno de los componentes de la serie, se hace uso, a su vez, de modelos matemáticos. Estos modelos ignoran las causas de las variaciones de cada componente; tan sólo reproducen su evaluación temporal de acuerdo con una pauta de regularidad que se sigue cada uno de los componentes de la serie, a fin de reproducir en última instancia el comportamiento de ésta. En las predicciones se supone que las regularidades observadas en el pasado se mantendrán en el futuro”[Otero J.1993].

3.4. Criterios para validar un pronóstico

Una vez construído el modelo, la utilidad que aquí se enfatiza del mismo es la del pronóstico, por ello es conveniente hacer uso de criterios que validen el empleo del modelo, para este fin en particular. De hecho, un esquema general para la obtención de pronósticos debe considerar, entre otras cosas, la validación del modelo mediante la contrastación de los pronósticos que de él surjan, con las observaciones de la serie, tanto en la parte histórica (muestra) como en períodos postmuestrales. También se debe tener en cuenta el mejoramiento de los pronósticos, buscando incorporar en ellos información adicional, posiblemente externa al modelo. El esquema que se sugiere seguir para obtener pronósticos, tiene como base tres niveles de análisis el primero, denominado análisis preliminar, es en el que se conocen y preparan los datos para su análisis definitivo; el segundo es propiamente el de la construcción del modelo, y el tercero es el de monitoreo y mejoramiento continuo de los pronósticos, en particular mediante el uso de información externa al modelo.

Las etapas para la elaboración del pronóstico se deben realizar de manera secuencial y lógica para que los resultados que se obtengan sean lo más cercanos a la “realidad” para este análisis el profesor Guerrero V. 1991, sugiere los siguientes pasos para la elaboración de un buen pronóstico.

3.4.1. Análisis preliminar

“Las etapas que se realizan en este nivel de análisis, son esencialmente las siguientes dos:
1) *Conocimiento de la información.* Aquí se debe tener en cuenta el fenómeno en estudio y las necesidades que debe cubrir el modelo, para decidir cuál es la serie de tiempo a pronosticar. Asimismo hay que considerar la disponibilidad de los datos y la posibilidad de que existan variables explicativas, para lo cual se debe uno documentar respecto a la teoría correspondiente y a estudios previos sobre el tema.

2) *Preparación de los datos para su análisis definitivo.* Con esto se entiende la aplicación de procedimientos estadísticos, fundamentalmente para: i*)* presentar los datos apropiadamente a través de cuadros-resumen y gráficas, *ii)* descubrir patrones de comportamiento relevantes, *iii)*

decidir la escala de expresión y/o alguna transformación de los datos que simplifique su estructura, y *iv)* determinar una clase general de modelos, que se considere adecuada de antemano para generar los pronósticos requeridos.

3.4.2. Construcción del modelo

Para realizar esta labor se distinguen cuatro etapas básicas:

1) *Selección de un modelo.* Dentro de la clase general de modelos que se haya decidido utilizar se debe especificar uno de manera explícita, de forma tal que se aprecien tanto sus parámetros como los supuestos que lo sustentan. Esto constituye en esencia, la identificación del modelo.

2) *Asignación de valores a los parámetros.* Mediante algún criterio de tipo estadístico, se debe decidir el valor numérico de los parámetros, para que el modelo represente adecuadamente a la serie observada. En esta etapa se realiza la estimación del modelo.

3) *Determinación de la validez del modelo,* La adecuación del modelo se puede juzgar respecto al cumplimiento de los supuestos, para esto se realiza la verificación del modelo mediante: *i)* análisis de residuos y *ii)* validación de criterios de bondad de ajuste. Si el modelo no es válido, se debe buscar la causa y seleccionar un modelo alternativo, en caso contrario se procede con la siguiente etapa.
4) *Obtención de los pronósticos.* Se supone que el modelo es construido precisamente con este fin, así que esta etapa corresponde al uso del modelo, con lo cual se termina el análisis a este nivel.

3.4.3. Monitoreo y mejoramiento

En este nivel de análisis, se sugiere realizar las siguientes etapas:

1) *Contrastación con la realidad* Para que los pronósticos puedan ser considerados válidos, se requiere que reflejen el comportamiento real de la variable en estudio. La adecuación de los pronósticos debe medirse de acuerdo con algún criterio previamente definido y mediante un monitoreo continuo de los resultados. Si los pronósticos son razonablemente adecuados, se continúa con la etapa de documentación del análisis, pero si esto no es así, se debe procurar mejorar los pronósticos, principalmente mediante el uso de información externa, como se indica en la etapa siguiente. De una manera más drástica, si el modelo no brinda pronósticos razonables, se debe proceder a construir un modelo alternativo, que ponga remedio a las deficiencias encontradas.

2) *Uso de información externa al modelo.* La idea de emplear información adicional a la originalmente utilizada en la construcción del modelo, surge del deseo de mejorar la calidad de los pronósticos. Dicha información comúnmente es del tipo de conjeturas acerca de los valores futuros de la variable o bien son pronósticos alternativos de la misma serie, obtenidos con algún otro modelo u otra técnica.

3) *Documentación del análisis.* Esto es útil para tener un registro del trabajo realizado y de los resultados obtenidos en las diferentes etapas realizadas. Se debe poner énfasis especialmente en las posibles limitantes y debilidades del modelo (y de los pronósticos mismos). Además, un seguimiento futuro de los pronósticos podrá realizarse sólo si se cuenta con la documentación apropiada.

4) *Uso de los pronósticos.* El objetivo último de la obtención de pronósticos, no son los pronósticos en sí mismos, sino el empleo de ellos para algún fin en particular (usualmente para la planeación). Así pues, los pronósticos deben ser empleados para satisfacer la necesidad que se tenía en mente al inicio del estudio, con lo cual se concluye el análisis".

3.5. Modelos de pronóstico[21]

La mayoría de la gente se sorprende al saber que gran parte de las decisiones que toma tiene que ver con los pronósticos. Éstas están dirigidas a sucesos futuros y requieren predicciones que rodean a ese ambiente de incertidumbre.

Esto se vive diariamente en las decisiones personales, por ejemplo cuándo y cómo invertir los ahorros; y para las decisiones críticas y complejas que afectan a una organización, como puede ser la construcción de una nueva línea del metro. "Al adoptar la decisión de construir un nuevo aeropuerto se necesitan, entre otros muchos elementos, los pronósticos relativos a la demanda futura, las innovaciones tecnológicas, costos, precios, planes de los competidores, la mano de obra y la legislación"[Balbuena, R. 1997].

La gran parte de los pronósticos que se necesitan para la toma de decisiones se maneja por lo general en forma intuitiva, a menudo sin separar de manera explícita la tarea de elaborar los pronósticos y la toma de decisiones.

3.5.1. ¿Hasta donde es posible pronosticar?

Durante mucho tiempo se ha escuchado un gran número de críticas en relación con la poca habilidad y creatividad de los pronosticadores para preveer los sucesos y cambios por venir. Además, los amantes de los pronósticos se han desilusionado debido a los grandes errores de predicción que han propiciado que la planificación y toma de decisiones pierdan el rumbo (el error de 1994, las crisis financieras). "Es importante hacer notar que a medida que han aumentado las quejas acerca de los pronósticos, también se han ampliado el número de peticiones y el interés por obtener predicciones adicionales. Reflexionando un poco, esto no es nada sorprendente, ya que cuando existe poca incertidumbre en el medio y las cosas se dan como se esperaban, hay mucha menos necesidad de pronósticos formales, mientras que en ambientes inestables con alto grado de incertidumbre, la necesidad de tales predicciones es mayor"[22].

Muchas de las críticas hechas a los pronósticos han estado bien fundamentadas. Los avances inesperados, los sucesos pronosticados que nunca tuvieron lugar, los grandes errores de los pronósticos y los errores en la oportunidad e intensidad de los cambios pronosticados son sólo algunos de los problemas más comúnmente mencionados. Pero los usuarios de los pronósticos deben aceptar su parte de la culpa porque, al igual que los que buscan un consejo de parte de los adivinos o de los brujos de Catemaco, frecuentemente sus expectativas se apartan de la realidad, en particular en relación con los pronósticos sin errores.

Los enfoques predictivos sistemáticos pueden proporcionar ventajas sustanciales cuando se usan apropiadamente, pero es ilusorio creer que tales enfoques han descubierto poderes universales. Por ejemplo con la llegada del año 2000 se rumuró que el fin del mundo llegaría, pero ya estamos en el 2004 y no pasa nada, hay que recordar que el calendario azteca dice que el cuarto sol se acabará en el año 2008 y es entonces cuando el fin del mundo llegará, pero nadie se atreve a afirmar tal aseveración, solo se especula o se cometa, entonces tenemos que esperar esa fecha y ver que ocurre.

Para comprender las ventajas y limitaciones de los pronósticos, es importante reconocer que todos los tipos, formas y técnicas de predicción son por naturaleza extrapolativos. (o sea, el predecir dentro de los datos existentes). Cuando se dispone de información histórica cuantitativa, los métodos de predicción utilizados se llaman cuantitativos. De no ser así, generalmente se les conoce con el nombre de métodos cualitativos-tecnológicos o discrecionales-subjetivos.

Los usuarios de las técnicas cuantitativas de pronosticado no disponen de una forma sencilla y contable para pronosticar lo que sucederá cuando cambien patrones o relaciones establecidos. Debido a que los métodos cuantitativos basan sus pronósticos en extrapolaciones de patrones e

[21] Si se desea consultar la historia de los pronósticos consultar Hanke, J. *Pronósticos en los negocios*.
[22] El financiero, 13 de enero del 2000, páginas 1 y 17.

interrelaciones pasados, tienen buenos resultados solamente cuando el futuro es similar al pasado o cuando sucede que los cambios (por casualidad) se eliminan.

Cuando las técnicas cuantitativas no funcionan, la alternativa para pronosticar el impacto del cambio es el juicio humano con un grado adecuado de apoyo y organización (por experiencia). No obstante, puesto que los métodos cualitativos también basan sus pronósticos en la observación de tendencias existentes, los cambios en esas tendencias y la magnitud del cambio futuro, están sujetas a cierto número de deficiencias. Pero la ventaja de los enfoques de predicción humana fundamentados es que pueden identificar el cambio sistemático con más rapidez e interpretar mejor el efecto de dicho cambio sobre el futuro.

En contraposición a tal situación se encuentra el hecho de que todos nosotros tenemos intereses definidos que a menudo dominan el buen juicio; nuestro deseo por un resultado o suceso específico matiza nuestra visión de lo que puede ser el resultado más probable.
A pesar de estas deficiencias, algunos pronosticadores se han ganado una reputación de ser sumamente confiables. Valdría la pena identificar algunas de las formas por las que los pronosticadores "experimentados" han logrado mantener lo que parecen ser calificaciones extremadamente buenas (pero este no es el objetivo de la investigación, si se desea abundar más sobre este tema, el profesor Nervole Marc, presenta una serie de opciones que hacen que los pronosticadores sean más eficientes al prever situaciones). Sin embargo, para ubicar la cuestión en una perspectiva adecuada, es importante dejar claro desde el principio que los autores no tienen noticia de ninguna evidencia que indique que cierto método de pronóstico en particular haya sido consistente y significativamente más preciso que algún otro.

Cuadro 52
Ventajas y desventajas de los métodos cuantitativos frente a los cualitativos para la predicción

Consideración	Métodos cuantitativos	Métodos cualitativos
Elección del método o modelo que se usa	El método o modelo empleado no se puede seleccionar con bases meramente estadísticas	La elección del método o modelo tiene un impacto importante en las predicciones pero requiere del juicio humano para ser seleccionado
Habilidad para predecir los cambios de los patrones o relaciones establecidos	Los cambio futuros no pueden pronósticarse	Los cambios futuros se pueden predecir, pero también pueden ignorarse, o la gente puede reaccionar exageradaamente a ellos.
Utilización de información o datos disponibles	No se está empleando toda la información en los datos	La gente puede usar información o conocimiento interior, pero también es selctiva, sesgada e inconsistente al usar tal información
Asignación o ajuste de las predicciones una vez que se han identificado los cambios	La modificación de los pronósticos como resultado de los cambios depende del método específico que se este usando	Una vez que los cambios han sido confirmados y aceptados, las personas pueden evaluar sus efectos y modificar las predicciones
Introducción de objetividad en los pronósticos	La objetividad se logra con base en algunos criterios de selección (minimación del error cuadrático medio), lo cual debe decidirse discrecionalmente	Los pronósticos están en gran parte influidos por consideraciones personales y políticas, así como por óptimismo o pesimismo indebidos
Determinación de la incertidumbre futura	La mayoría de los métodos generalmente subestiman la incertidumbre futura	La incertidumbre futura generalmente es subestimada por una gran cantidad, en su mayoría del lado optimista
Necesidad de predicciones respectivas	Los pronósticos son consistentes si un método se usa una o 100,000 veces	La gente fácilmente se aburre en situaciones de pronósticos repetitivos, lo cual introduce inconsistencia en las predicciones
Costo del método empleado	No es caro usar los métodos hoy, cuando el cálculo computarizado no es costoso	Las personas y reuniones son caras, lo mismo que las predicciones discrecionales.

3.6. Función e importancia de los pronósticos en economía

Si se analizan los logros de una variedad de pronosticadores (CAPEM, BANAMEX, J. MORGAN, BANXICO, CEMPE, REU), queda claro que algunos de ellos han tenido más éxito (han sido más exactos) que otros, al menos cuando se trata de periodos cortos. No obstante, desde un punto de vista predictivo, los hechos son igualmente claros con respecto a que no existe manera de saber *a priori* quién o cuál método será más preciso. La exactitud de los pronósticos no se correlaciona con la fama de los pronosticadores, sus antecedentes, sus experiencias o sus éxitos pasados. Aunque nuestro deseo podría ser diferente, el reconocer este hecho básico es fundamental para comprender las ventajas y limitaciones de diversos enfoques para pronosticar, y así desarrollar expectativas que puedan ser de gran ayuda en los procesos de planificación y toma de decisiones.

A pesar de la afirmación anterior, “es posible considerar el conocimiento y la experiencia propios hasta cierto punto en desacuerdo con la idea de no tener la habilidad para identificar quién hará el mejor trabajo de pronosticado en el futuro. Existen algunas características que ayudan a seleccionar los pronosticadores y a los métodos de predicción que tengan la más alta probabilidad de proporcionar la información deseada en una situación dada. Sin embargo, no es posible elegir del grupo la que funcionará consistentemente mejor, al menos no con base en las corroboraciones empíricas actualmente disponibles; debido a que los que venden métodos para pronosticar y pronósticos quisieran que los clientes potenciales pensaran de otra manera, es conveniente comprender algunas de las técnicas de selección que pueden hacer que el pronosticador profesional parezca especialmente efectivo en un ambiente dado.”[Mann P., 1995].

Cuadro 53 Clasificación de los métodos de pronósticos, campos de desarrollo y áreas de aplicación

Método de predicción			Campo en el que se desarrollo el método							Areas principales de aplicaciones empresariales																			
																Ventas													
																Agregadas			Desagregadas										
			Estadística	Investigación de operaciones	Economía	Psicología	Planificaciónde largo alcance	Práctica	Mercadotecnia	Planificación de la producción	Programación de la producción	Inventarios	Requerimiento de material	Programación del personal	Planificación de personal	Corto plazo	Mediano plazo	Largo plazo	Corto plazo	Mediano plazo	Largo plazo	Fijación de precios	Publicidad y promoción	Presupuestación anual	Productos nuevos	Proyectos de investigación y desarrollo	Presupuestación de capital	Análisis competitivo	Estrategia
Cuantitativos	Series de tiempo	Ingenuo						X				X							X			X							
		Suavizamiento		X				X		X	X	X	X	X	X				X										
		Descomposición			X					X	X			X	X		X			X				X					
		Promedio Móvil autorregresivo	X							X	X			X	X	X													
	Explicativos	Autorregresivos vectoriales	X		X											X	X					X	X	X					
		Regresión	X		X											X	X					X	X	X					
		Econometría	X		X											X	X					X	X	X					
	Enfoques de monitoreo		X	X				X			X	X	X	X					X										
Discrecionales	Predicción produc nuevos		X						X							X	X	X				X	X	X	X	X		X	X
	Individuales	Juicio individual				X		X		X				X		X	X	X				X	X	X	X	X	X	X	X
	Grupal	Reglas de decisión				X					X	X	X	X															
		Estimación de la fuerza de ventas						X	X										X	X				X					
Agregado		Jurados de opinión gerencial						X	X								X					X	X	X	X	X	X	X	X
		Actuación				X																						X	
	Agregados	Sondeos anticipados						X	X								X							X					
		Investigación de mercado							X										X	X		X	X		X	X			
		Programas piloto y programas de premercado							X										X	X	X	X	X		X				
Tecnológicos	Extrapolativos	Curvas de crecimiento	X				X														X					X	X	X	X
		Comparaciones independientes de tiempo					X														X					X	X	X	X
		Analógias históricas y de otro tipo					X														X					X	X	X	X
	Expertos	Delphi				X														X					X	X	X	X	
		Futurista				X														X					X	X	X	X	
		Matrices de impacto cruzado				X														X					X	X	X	X	

Se analizan las técnicas de pronóstico que son apropiadas para las series de tiempo, El objetivo del método de pronóstico es alisar o suavizar el componente irregular de la serie mediante un proceso promediado.

3.7. Técnica de promedios móviles simples

Se basa en el hecho de que al tenerse disponibles nuevas observaciones de la serie de tiempo, se reemplaza la observación más antigua y se calcula un nuevo promedio; como resultado, el promedio cambia, o se mueve, al disponerse de nuevas observaciones.

Este método utiliza el promedio de los valores de los **n** datos más recientes en la serie de tiempo como pronóstico para el período siguiente. En forma matemática el cálculo de los promedios móviles se realiza de la siguiente manera:

$$S_{t+1} = \frac{X_t + X_{t-1} + X_{t-2} + \ldots + X_{t-N+1}}{n} \qquad (33)$$

Limitaciones de la técnica

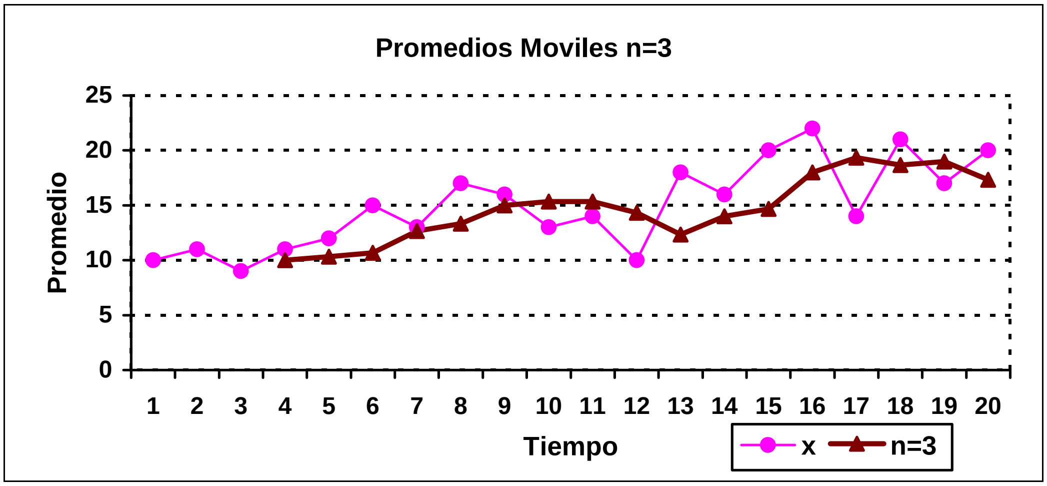
a) Requiere de tantas observaciones como valores de “n” posea la muestra.

b) No toma en cuenta cambios repentinos o se adopta fácilmente a las características de una serie.

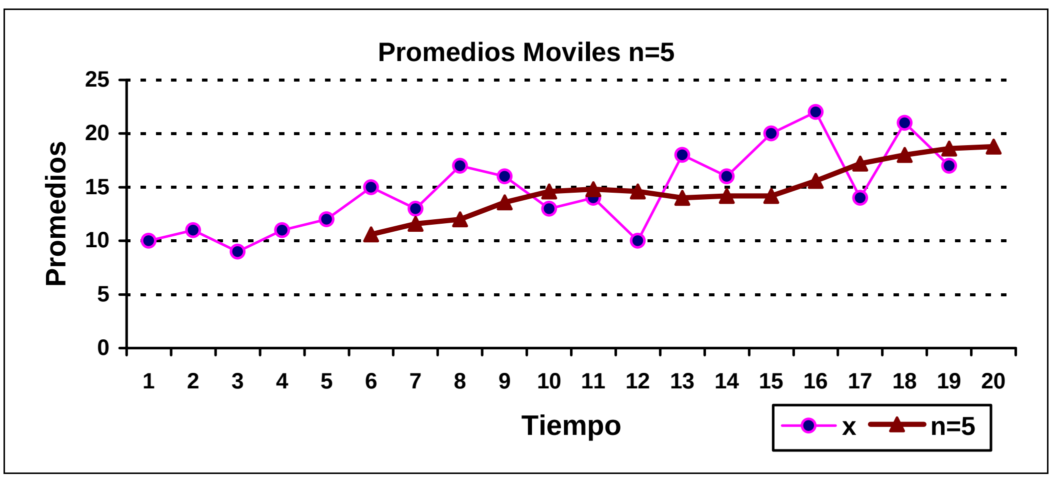
Cuadro 54
Aplicación de la Técnica de Promedios Móviles Simples

t	X	S_t N=3	S_t N=5
1	10		
2	11		
3	9		
4	11	10.00	
5	12	10.33	
6	15	10.67	10.60
7	13	12.67	11.60
8	17	13.33	12.00
9	16	15.00	13.60
10	13	15.33	14.60
11	14	15.33	14.80
12	10	14.33	14.60
13	18	12.33	14.00
14	16	14.00	14.20
15	20	14.67	14.20
16	22	18.00	15.60
17	14	19.33	17.20
18	21	18.67	18.00
19	17	19.00	18.60
20		17.33	18.80

Gráfica 26
Promedios Móviles n=3



Gráfica 27
Promedios Móviles n=5



Cuadro 55
Aplicación dela Técnica de Promedios Móviles Trimestral

PROMEDIO MÓVIL TRIMESTRAL						PROMEDIO MÓVIL PENTAMENSUAL			
t	**X**	**PRONOSTICO**	**ERROR**	**Error Absoluto**	**Error al Cuadrado**	**PRONOSTICO**	**ERROR**	Error Absoluto	Error al Cuadrado
1	10								
2	11								
3	9								
4	11	10	1	1	1				
5	12	10	2	2	3				
6	15	11	4	4	19	11	4	4	16
7	13	13	0	0	0	12	1	1	1
8	17	13	4	4	13	12	5	5	25
9	16	15	1	1	1	14	2	2	4
10	13	15	-2	2	5	15	-2	2	4
11	14	15	-1	1	2	15	-1	1	1
12	10	14	-4	4	19	15	-5	5	25
13	18	12	6	6	32	14	4	4	16
14	16	14	2	2	4	14	2	2	4
15	20	15	5	5	28	14	6	6	36
16	22	18	4	4	16	16	6	6	36
17	14	19	-5	5	28	17	-3	3	9
18	21	19	2	2	5	18	3	3	9
19	17	19	-2	2	4	19	-2	2	4
20		17				19			
				45	173			46	190
				2.81[a]	10.81[b]			3.29[a]	13.57[b]
[a] Desviación Estandar del Error de Pronóstico									
[b] Valor Medio del Error del									

3.8. Técnica de promedios móviles dobles

El método consiste en calcular un conjunto de promedios móviles y después se calcula un segundo conjunto como promedio móvil del primero.

El promedio móvil doble se ubica por debajo del primer conjunto en proporción similar a la que se ubica el primer conjunto por debajo de los valores reales. Estas diferencias entre ambos conjuntos se emplea para pronosticar los valores reales.
Se puede definir el concepto de promedio móvil doble en los siguientes términos:

$$S'_{t+1} = \frac{S_t + S_{t-1} + S_{t-2} + \ldots + S_{t-N+1}}{n} \qquad (34)$$

El ajuste para este método se obtiene introduciendo dos parámetros "**a**" y "**b**" que se definen como:

$$a = 2S_{t+1} - S'_{t+1} \qquad b = \frac{2(S_{t+1} - S'_{t+1})}{n-1} \qquad (35)$$

Para realizar el pronóstico de ***m*** periodos futuros, el pronóstico final será: $S'_{t+m} = a + b \bullet m$
Limitaciones del método
Son las que se mencionaron en la técnica de promedios móviles simples, pero esta requiere de una cantidad mayor de información.
Se recomienda para muestras grandes n>30 .

La técnica le otorga el mismo peso o importancia a todas las observaciones históricas de la serie. Para fines didácticos solo se utilizan 20 datos.

Cuadro 56
Aplicación dela Técnica de Promedios Móviles Dobles

T	X	St+1 Pronosticos	S't+1	a	b	St-m	ERROR	Error Absoluto	Error al Cuadrado
	0								
1	10								
2	11								
3	9	7							
4	11	10							
5	12	10	9	12	1	13	10	10	107
6	15	11	10	11	0	11	11	11	114
7	13	13	11	14	1	16	13	13	160
8	17	13	12	14	1	16	13	13	178
9	16	15	14	16	1	18	15	15	225
10	13	15	15	16	1	17	15	15	235
11	14	15	15	15	0	16	15	15	235
12	10	14	15	14	-1	13	14	14	205
13	18	12	14	11	-2	9	12	12	152
14	16	14	14	14	0	15	14	14	196
15	20	15	14	16	1	17	15	15	215
16	22	18	16	20	2	23	18	18	324
17	14	19	17	21	2	23	19	19	374
18	21	19	19	19	0	19	19	19	348
19	17	19	19	19	0	19	19	19	361
20		17	18	16	-1	15			
								223	3429.8889
								14.87	228.66

3.9. Métodos de suavización exponencial simple

La suavización exponencial es una técnica de pronóstico en la que se utiliza un promedio ponderado de una serie de valores anteriores o pasados (haciendo esto de manera decreciente) para pronosticar el valor de la serie de tiempo en el período siguiente. Las observaciones se ponderan, asignando mayor peso a las más recientes.
La técnica se define a través de un pronóstico inicial a partir del cual podemos generalizar:
Primer pronóstico o valor inicial
$S_2 = X_1$
Término genérico del pronóstico
$$S_{t+1} = S_t + \alpha(X_t - S_t) \quad (36)$$
Donde “α” y “t” son:
$0 \leq \alpha \leq 1$ y $t \geq 2$
y el término $(X_t - S_t)$ corresponde al error de pronóstico anterior y “α” es el factor de ponderación(arbitrario) que estamos utilizando.
$\alpha \leq 1$ el pronóstico le otorga mayor importancia a los valores recientes de la serie.
$\alpha \geq 0$ el pronóstico le otorga mayor importancia a los valores más antiguos de la serie.
Limitaciones del método.

Es difícil aplicar esta técnica debido a que no se conoce el valor de "α" que haga mínimo el error y la varianza

Cuadro 57
Suavización Exponencial Simple

t	X	0.1	0.5	0.9
		f	f	f
1	10			
2	11	10	10	10
3	9	10.1	10.5	10.9
4	11	10	9.8	9.2
5	12	10.1	10.4	10.8
6	15	10.3	11.2	11.9
7	13	10.8	13.1	14.7
8	17	11	13	13.2
9	16	11.6	15	16.6
10	13	12	15.5	16.1
11	14	12.1	14.3	13.3
12	10	12.3	14.1	13.9
13	18	12.1	12.1	10.4
14	16	12.7	15	17.2
15	20	13	15.5	16.1
16	22	13.7	17.8	19.6
17	14	14.5	19.9	21.8
18	21	14.5	16.9	14.8
19	17	15.1	19	20.4

Cuadro 58
Continua

	0.1			0.5			0.9	
error	error abs	error cuad	error	error abs	error cuad	error	error abs	error cuad
-1	1	1	-1	1	1	-1	1	1
1	1	1	2	2	2	2	2	4
-1	1	1	-1	1	2	-2	2	3
-2	2	4	-2	2	3	-1	1	1
-5	5	22	-4	4	15	-3	3	10
-2	2	5	0	0	0	2	2	3
-6	6	36	-4	4	16	-4	4	15
-4	4	20	-1	1	1	1	1	0
-1	1	1	3	3	6	3	3	9
-2	2	4	0	0	0	-1	1	0
2	2	5	4	4	17	4	4	15
-6	6	35	-6	6	35	-8	8	58
-3	3	11	-1	1	1	1	1	2
-7	7	49	-4	4	20	-4	4	15
-8	8	69	-4	4	18	-2	2	6
1	1	0	6	6	35	8	8	60
-7	7	43	-4	4	16	-6	6	39
-2	2	3	2	2	4	3	3	11
Valor medio del		310.1	Valor medio del		191	Valor medio del		253
error de pronóstico		17.23	error de pronóstico		10.62	error de pronóstico		14.04

3.10. Método de suavización exponencial doble o método de brown

Es utilizada para pronósticar series de tiempo que tienen una tendencia lineal. Aquí eliminamos el problema de almacenar **2n** observaciones históricas y la uniformidad de asignación de pesos a las observaciones de la serie.
Se repiten los tres pasos del Método de Suavización Exponencial
Primer pronóstico o valor inicial

$$S_2 = X_1$$

Término genérico del pronóstico

$$S_{t+1} = S_t + \alpha(X_t - S_t) \quad (37)$$

Donde "α" y "t" son:

$0 \leq \alpha \leq 1$ y $t \geq 2$

Se realiza otra suavización exponencial

$$S'_{t+1} = S'_{t+\alpha}(S_{t+1} - S'_t) \quad (38)$$

Agregamos los parámetros "**a**" y "**b**" para obtener el ajuste

$$a = 2S_t - S'_t \quad (39)$$

$$b = \frac{\alpha}{1-\alpha}(S_t - S'_t) \quad (40)$$

De manera que el término genérico de pronóstico es: $S_{t+m} = a + bm$
donde:
"α" es el factor de ponderación

“m” es el número de períodos futuros a pronosticar.
Limitaciones del método
Las mismas que en el Método de Suavización Exponencial Simple

3.11. Método box-jenkins[23]

Es un proceso iteractivo de construcción de modelos que se puede utilizar mediante las siguientes fases o etapas:

1. Identificación del modelo
2. Estimación de los parámetros involucrados y
3. Verificación de los supuestos en que se fundamenta el modelo.

Estas etapas se repiten tantas veces como sea necesario hasta que la verificación indique una satisfacción aproximada de los supuestos; entonces decimos que tenemos construido un modelo para la serie de tiempo que es útil ya sea para pronósticar, simular o simplemente para explicar el comportamiento de la variable en estudio.

Cabe señalar que en la etapa de identificación del modelo es importante notar que se supone que la serie es estacionaria, esto implica que la serie fluctúa aleatoriamente alrededor de un nivel medio, el cual permanece constante en el tiempo así como la variación con respecto a este valor medio.

La metodología de Box-Jenkins se basa en el diagrama esquemático mostrado en la figura 15. Primero se postula una clase general de modelos de predicción; después, se proponen tres etapas:

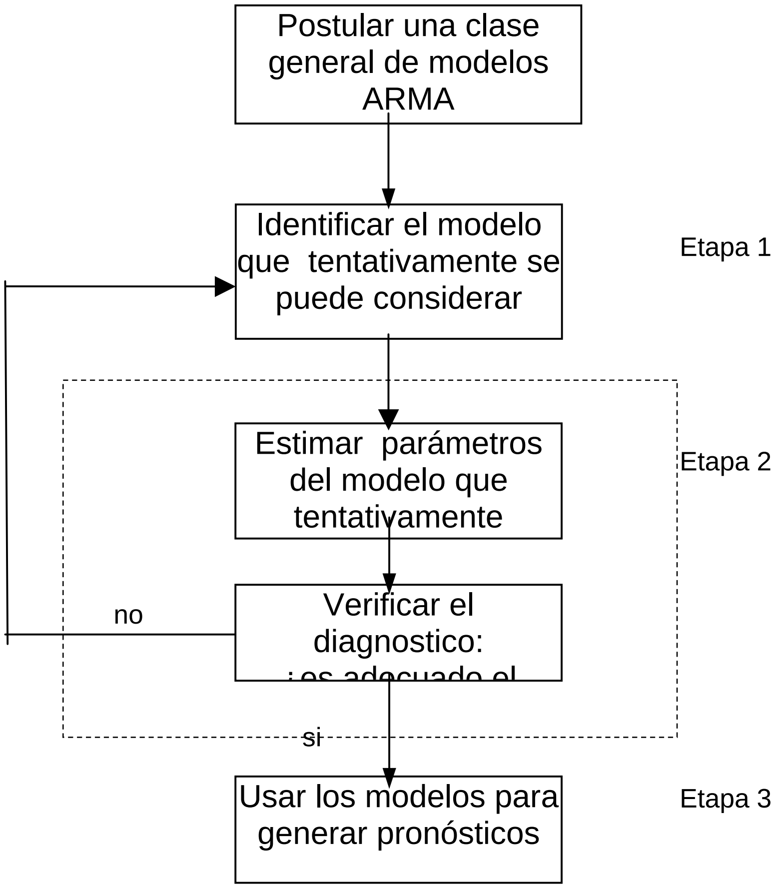
En la etapa 1 se identifica un modelo especifico que puede considerarse tentativamente como el modelo de predicción más apropiado a la situación.

La etapa 2 consiste en ajustar dicho modelo a los datos históricos disponibles y en realizar una verificación para determinar si es adecuado. Si no, el enfoque se regresa a la etapa 1 y se identifica un método alternativo de los disponibles en la clase general.

Cuando ya se ha aceptado un modelo adecuado, la etapa 3 el desarrollo de un pronóstico para algún periodo futuro se lleva a cabo. A continuación se describen estás tres etapas.

[23] Vid, Box, G.E. y G.M. Time Series Analysis, Second Edition, Holden-Day, 1976

Figura 15
Método de predicción Box-Jenkins



Etapa 1: Identificar tentativamente un modelo apropiado
Desde el punto de vista teórico, la ecuación

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \ldots - \theta_q e_{t-q} \qquad (41)$$

puede ajustarse a cualquier patrón de datos. Sin embargo, los valores de p (0,1,2, ...) y q (0,1,2,...) se deben especificar antes que se aplique el método. Por ejemplo, si p=1 y q=0, el modelo ARMA es

$$Y_t = \phi_1 Y_{t-1} + e_t \qquad (42)$$

La ecuación (42) se llama modelo AR de primer orden (también llamado de Markov) y se escribe como AR(1) o ARMA(1,0).

SI p=2 y q=0, el modelo correspondiente es un AR(2) o un modelo ARMA (2,0) de la forma

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t \qquad (43)$$

La ecuación (43) es un proceso AR de segundo orden.

Cuando p = 0 y q=1, se trata de un modelo MA de primer orden y se escribe como MA(1) o ARMA(0,1)

$$Y_t = e_t - \theta_1 e_{t-1} \qquad (44)$$

Cuando p=0 y q=2, el modelo es un MA (2) o ARMA (0,2) de la forma

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} \qquad (45)$$

Por último, p y q pueden ser diferentes de 0, en cuyo caso existe un modelo ARMA mixto. Por ejemplo, cuando p=1 y q=1, el modelo es un modelo ARMA (1,1) y es de la forma

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t - \theta_1 e_{t-1} \qquad (46)$$

Es interesante darse cuenta que en la ecuación (46) e_t, está influida por $\mathbf{Y_{t-1}}$ al mismo tiempo por $\theta\, e_{t-1}$, lo que hace la ecuación (46) no lineal y por lo tanto suficientemente general para escribir una variedad de patrones de datos.

Obviamente **p** y **q** puede adoptar muchos otros valores, aun cuando raramente sin mayores que 2. Sin embargo, el problema que permanece es cómo elegir un orden apropiado para p y q de suerte que se pueden ajustar un modelo a los datos sin tener que recurrir a todas posibles combinaciones un proceso que toma tiempo. Esto se logra con el examen de los coeficientes de autocorrelación y una medida relacionada, las autocorrelaciones parciales.

No es exagerado decir que la parte más difícil de la metodología de Box – Jenkins es identificar un modelo apropiado. Así pues, se describe el proceso implicado sin intentar profundizar en la teoría del por qué se hace de esa manera. Se sugiere al lector interesado consultar las obras escritas sobre el tema, (Granger, C., Kling, J., Newbold, P., y Parzen, E.) que se encuentran en la bibliografia de este trabajo.

La tarea identificar un modelo se puede subdividir en tres pasos:
1. Logro de estacionariedad.
2. Elección de **p** y **q** para datos no estacionales
3. Selección de **p** y **q** así como los parámetros estacionales para los datos estacionales.

Consecuencia de estacionariedad mediante diferenciación

La teoría básica detrás de los modelos ARMA se aplica solamente a datos estacionarios. Así, antes de intentar identificas p y q en la ecuación (41), los datos deben ser sin tendencia. Si la serie de datos contiene una tendencia, se puede remover mediante tomas de diferencias sucesivas de los datos o por algún otro procedimiento para remover tendencias. En el cuadro 59 se muestra como se hace la diferenciación y cómo los datos de una tendencia lineal se convierten en horizontales[24] después de diferenciar.

La nueva serie en el cuadro 59 se incorpora estas diferencias es constante porque los datos originales no eran aleatorios. Si, además de lo no aleatoria, están presentes otros patrones, la diferenciación remueve la tendencia sin afectar los patrones restantes.

Si los datos tienen un patrón estacional, además del tipo de diferenciación como se observa en el cuadro 59 (llamada corta o de periodo a periodo), también se necesitara la diferencia estacional o

[24] En terminología estadística estacionarios.

larga (digamos, de enero de un año a enero del siguiente) par asegurar la estacionariedad de los datos.

Cuadro 59
Datos con tendencia líneal y sus diferencias

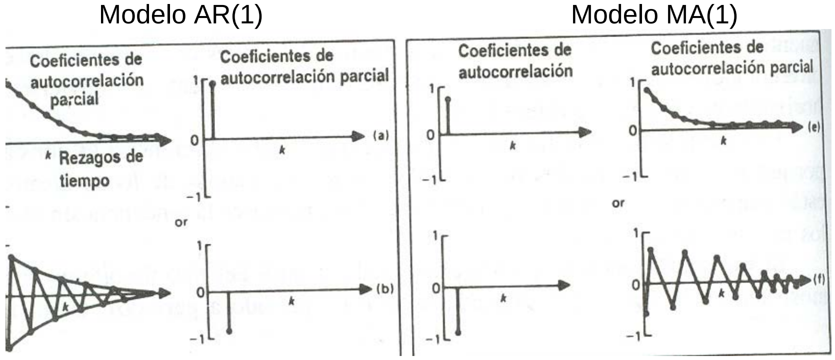
Series de datos	Primeras diferencias					Nueva serie
2	4	-	2	=	2	2
4	6	-	4	=	2	2
6	8	-	6	=	2	2
8	10	-	8	=	2	2
10	12	-	10	=	2	2
12						

3.11.1. Identificación de p y q

Una vez que los datos son estacionarios, se identifican p y q por la observación de las autocorrelaciones y las autocorrelaciones parciales de los datos diferenciados. La figura 16 muestra varios perfiles de autocorrelaciones y autocorrelaciones parciales así como los modelos ARMA apropiados que corresponden a tales autocorrelaciones. Como regla general, cuando las autocorrelaciones se acercan exponencialmente a 0, el modelo es AR y su orden se determina por el número de autocorrelaciones parciales que son significativamente diferentes de 0. Si las autocorrelaciones parciales se acercan exponencialmente a 0, el modelo es MA y su orden se determina por el número de autocorrelaciones estadísticamente significativa. Cuando tanto las autocorrelaciones como las autocorrelaciones parciales se acercan exponencialmente a 0, el modelo es ARMA mixto.

Figura 16

Modelo AR(1) Modelo MA(1)



Modelo AR(2) Modelo MA(2)(

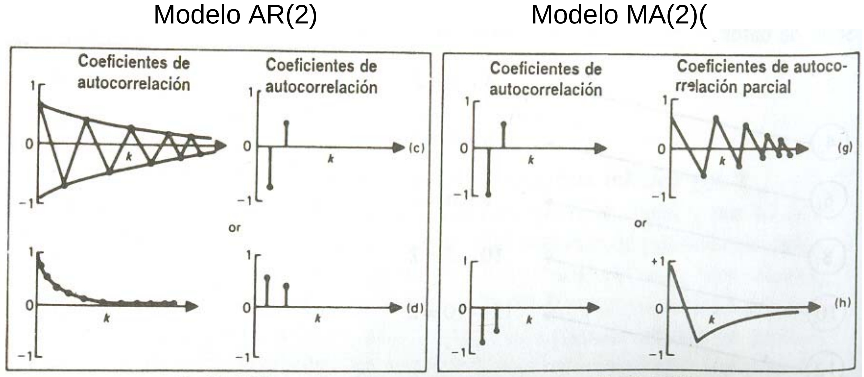
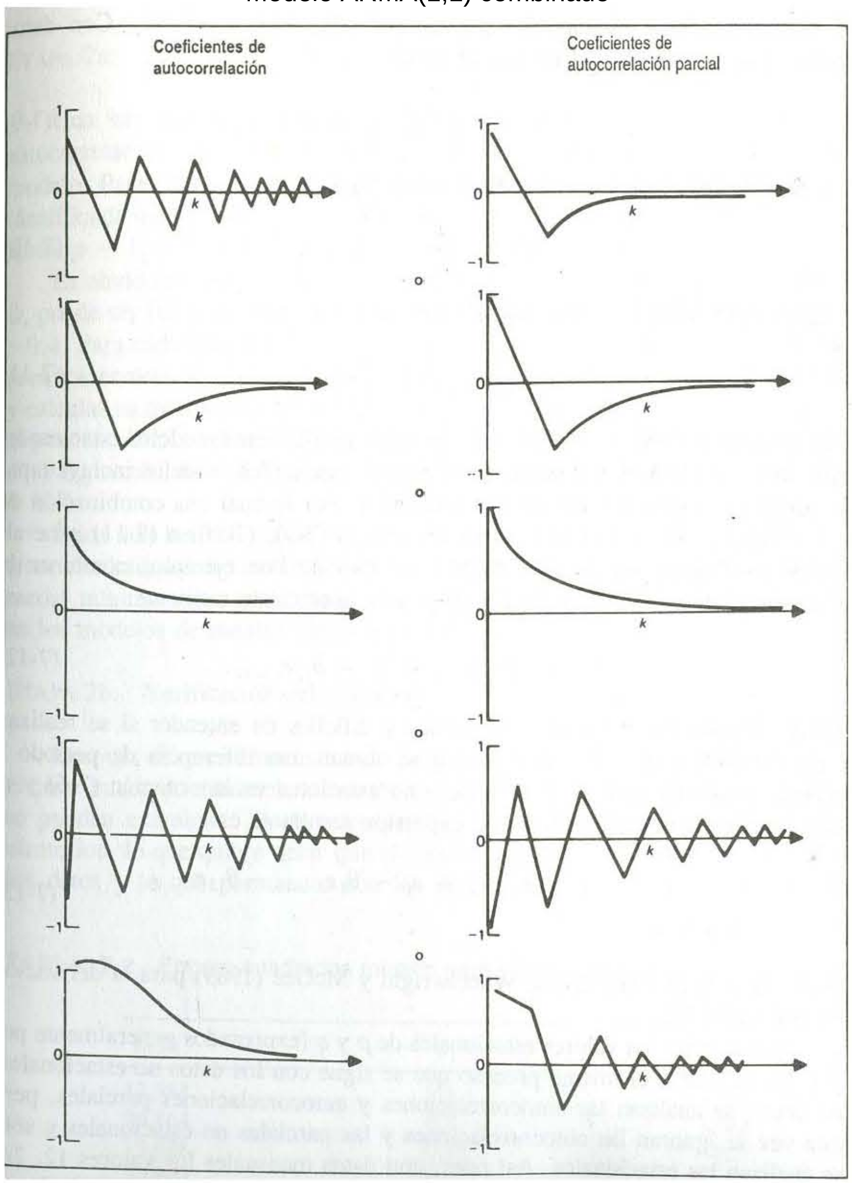


Figura 17
Modelo ARMA(1,1) combinado



Datos estacionales

Cuando los datos estacionales, la ecuación (41) podría no ser suficiente y debe completarse con parámetros estacionales. Como las ecuaciones (42) a (46), los modelos estacionales pueden ser AR, MA o ARMA. Con datos mensuales, un modelo AR estacional sería:

$$Y_t = \phi_{12} Y_{t-12} + e_t \qquad (47)$$

Un modelo MA estacional tendría la forma siguiente:

$$Y_t = e_t - \theta_{12} e_{t-12} \qquad (48)$$

Y un modelo ARMA presentaría la siguiente ecuación:

$$Y_t = \phi_{12} Y_{t-12} + e_t - \theta_{12} e_{t-12} \qquad (49)$$

Las ecuaciones (47), (48) y (49) representa solamente modelos estacionales AR, MA o ARMA. En realidad, la mayor parte de los modelos incluye tanto parámetros no estacionales como estacionales. Por lo cual la combinación de las ecuaciones (42) a (46) con las ecuaciones (47), (48) o (49) generalmente conforman un modelo ARMA estacional. Por ejemplo, combinar la ecuación (42) con la ecuación (48) genera la ecuación siguiente:

$$Y_t = \phi_1 Y_{t-1} + e_t - \theta_{12} e_{t-12} \qquad (50)$$

Tales expresiones se hacen complicadas y difíciles de entender si se realizan otras combinaciones. Por ejemplo, si se toman una diferencia de periodo a periodo y una estacional, y el modelo no estacional es la ecuación (44) y el estacional es la ecuación (48), la expresión resultante es

$$Y_t = Y_{t-1} + Y_{t-12} + Y_{t-13} + e_t - \theta_1 e_{t-1} - \theta_{12} e_{t-13} \qquad (51)$$

Para estimar los valores estacionales de p y q (expresados generalmente por P y Q), se utiliza el mismo el proceso que se sigue con los datos no estacionales. Es decir, se analizan las autocorrelaciones y autocorrelaciones parciales, pero esta vez se ignoran las autocorrelaicones y las parciales estacionales y sólo se analizan las estacionales. Así pues, con datos mensuales los valores 12, 24, 36, 48 y así sucesivamente se examinan para patrones similares a los de la figura 15. Sin embargo, desafortunadamente no existen suficientes autocorrelaciones y parciales disponibles para facilitar un proceso de identificación preciso como el que es posible en el caso de los datos no estacionales. Dada esta situación, se requiere de buen juicio y de ensayo y error. Afortunadamente, P y Q generalmente son 0 ó 1, lo que hace que el trabajo de selección sea relativamente más fácil.

Etapa 2ª Estimación de parámetros

Una vez que se ha elegido el modelo tentativo mediante el análisis de autocorrelación y autocorrelaciones parciales, se estiman los parámetros del modelo. Para ejemplificar el proceso, supóngase que la ecuación (50) ha sido identificada tentativamente como el modelo más adecuado para los datos. Por lo tanto, p=1, q=0, P=0, Q=1 y no hay diferenciación.

Es obvio que ϕ_1 y θ_{12} puede tomar muchos valores posibles. Por ejemplo, ϕ_1 puede ser 0.5 y θ_{12} = -0.3; ϕ_1 =0.2 y θ_{12}= 0.4; o ϕ_1 =-0.7 y θ_{12}=-0.4. Para cada conjunto de valores de ϕ_1 y θ_{12} habrá errores correspondientes para los errores e_t. Cada uno de los valores de e_t se pueden elevar al cuadrado y calcular su total.

Etapa 2b: Verificación del diagnóstico

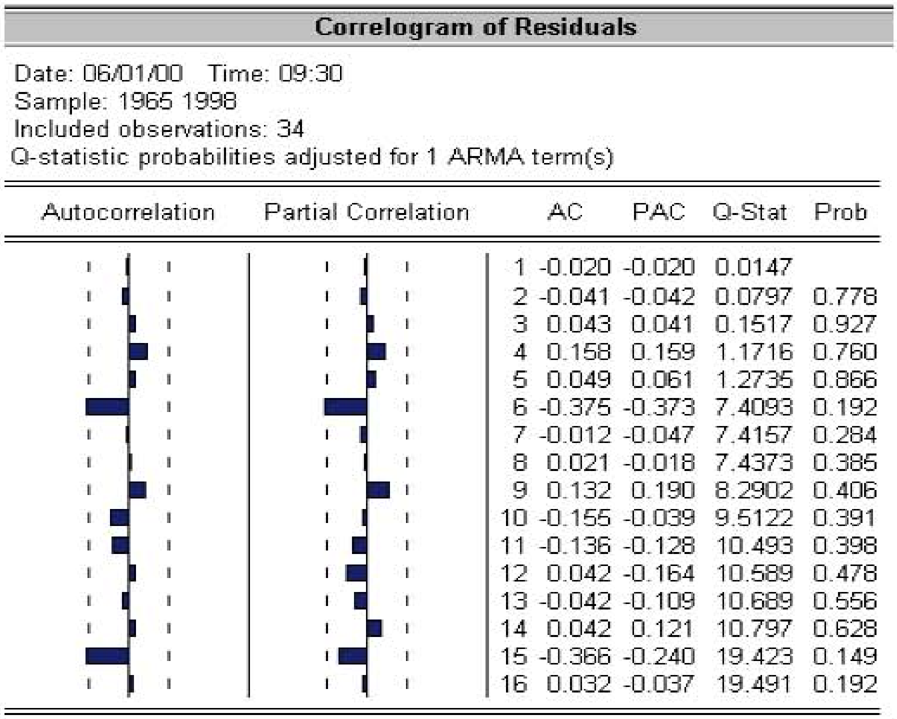
Después e que se han estimado los parámetros óptimos (lo cual proporciona el error cuadrado mínimo), los errores e_t [véase ecuación (50), por ejemplo] se puede examinar. Se dan dos posibles resultados:

1. Que los errores son aleatorios, lo que quiere decir que el modelo ajustado ha eliminado el patrón de los datos y lo que permanece son los errores aleatorios, o
2. Que el modelo tentativamente identificado no ha removido todo el patrón, como queda indicado por el hecho que el error e_t, no sea aleatorio.

Afortunadamente, se puede calcular fácilmente las autocorrelaciones y los residuos. Puesto que las autocorrelaciones puede decir cómo se relaciona unos a otros los valores de los residuos, si son aleatorios, entonces ninguna autocorrelación debe ser significativamente diferente de cero. Esto es válido para la figura 18, la que muestra que el modelo identificado es adecuado. La figura 19, por otra parte, tiene varias autocorrelaciones fuera del intervalo de confianza del 95%, lo cual implica que el modelo ajustado no ha eliminado todo lo del patrón a partir de los datos. De aquí que otro modelo sería mejor si se removiera algo del patrón remanente. Por lo tanto, el modelo que proporcionó los resultados de la figura 16 no es apropiado y se debe identificar uno nuevo.

Figura 18

Autocorrelación en los residuos de un modelo adecuado.



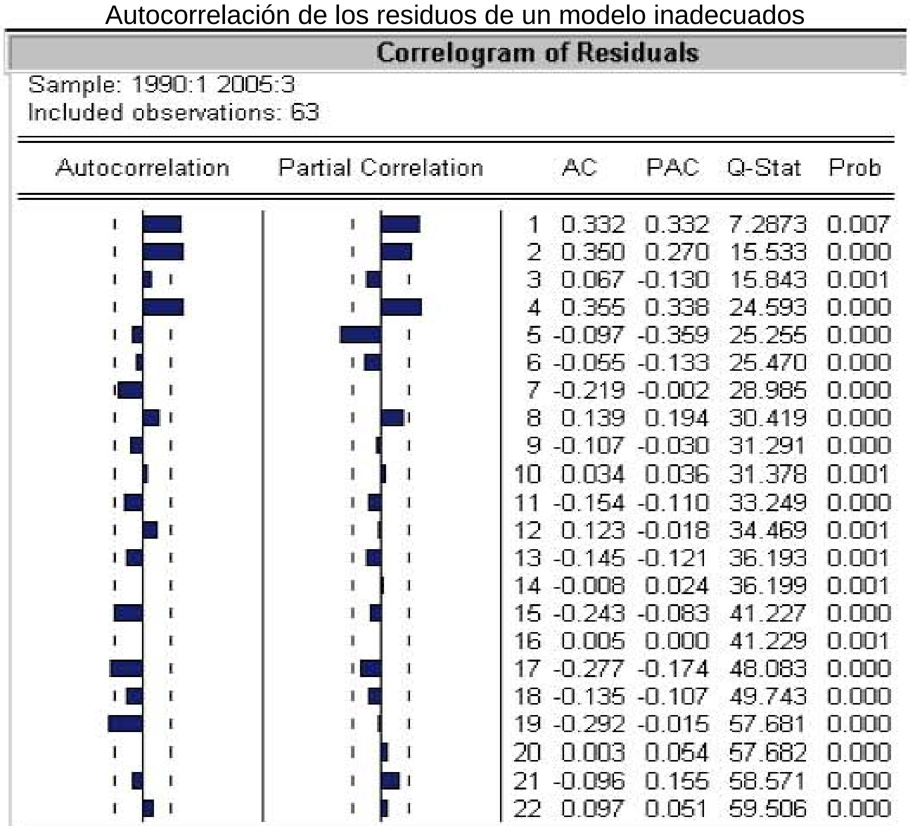
Correlogram of Residuals

Date: 06/01/00 Time: 09:30
Sample: 1965 1998
Included observations: 34
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.020	-0.020	0.0147	
		2	-0.041	-0.042	0.0797	0.778
		3	0.043	0.041	0.1517	0.927
		4	0.158	0.159	1.1716	0.760
		5	0.049	0.061	1.2735	0.866
		6	-0.375	-0.373	7.4093	0.192
		7	-0.012	-0.047	7.4157	0.284
		8	0.021	-0.018	7.4373	0.385
		9	0.132	0.190	8.2902	0.406
		10	-0.155	-0.039	9.5122	0.391
		11	-0.136	-0.128	10.493	0.398
		12	0.042	-0.164	10.589	0.478
		13	-0.042	-0.109	10.689	0.556
		14	0.042	0.121	10.797	0.628
		15	-0.366	-0.240	19.423	0.149
		16	0.032	-0.037	19.491	0.192

Figura 19

Autocorrelación de los residuos de un modelo inadecuados



Correlogram of Residuals

Sample: 1990:1 2005:3
Included observations: 63

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.332	0.332	7.2873	0.007
		2	0.350	0.270	15.533	0.000
		3	0.067	-0.130	15.843	0.001
		4	0.355	0.338	24.593	0.000
		5	-0.097	-0.359	25.255	0.000
		6	-0.055	-0.133	25.470	0.000
		7	-0.219	-0.002	28.985	0.000
		8	0.139	0.194	30.419	0.000
		9	-0.107	-0.030	31.291	0.000
		10	0.034	0.036	31.378	0.001
		11	-0.154	-0.110	33.249	0.000
		12	0.123	-0.018	34.469	0.001
		13	-0.145	-0.121	36.193	0.001
		14	-0.008	0.024	36.199	0.001
		15	-0.243	-0.083	41.227	0.000
		16	0.005	0.000	41.229	0.001
		17	-0.277	-0.174	48.083	0.000
		18	-0.135	-0.107	49.743	0.000
		19	-0.292	-0.015	57.681	0.000
		20	0.003	0.054	57.682	0.000
		21	-0.096	0.155	58.571	0.000
		22	0.097	0.051	59.506	0.000

ETAPA 3: Preparación de pronósticos

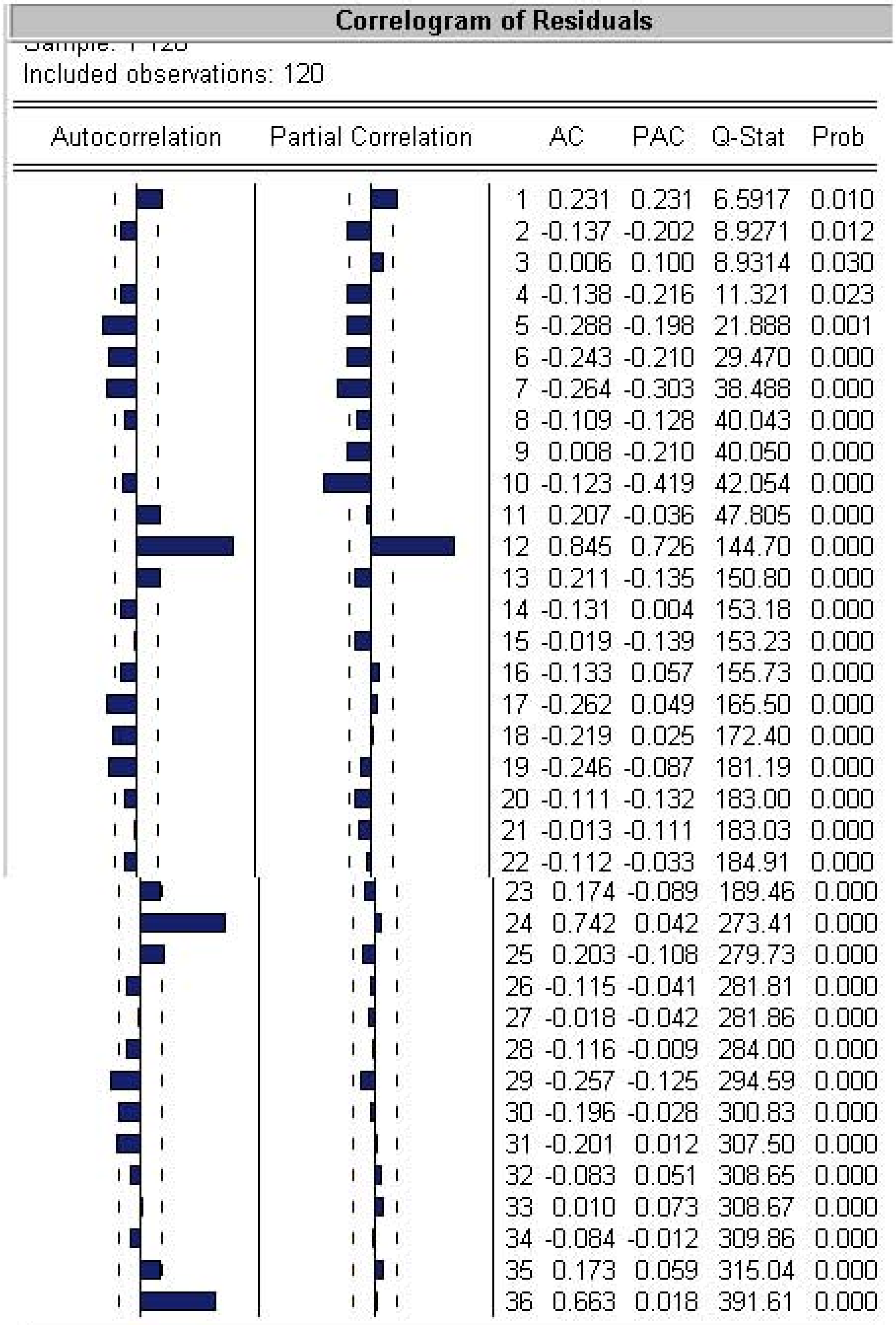
Una vez que el modelo ha sido identificado, se estiman los parámetros y los residuos muestran ser aleatorios, la predicción con tal modelo es una cuestión directa mecánica. El programa de computadora necesarios para realizar el los cálculos de la identificación y estimación puede proporcionar tantos pronósticos como el gerente (administrador) desee, junto con intervalos de confianza del 95% o 99% para los mismos

3.12. Aplicación de la métodología de box-jenkins

Para mostrar el uso del método de Box-Jenkins, se aplicará a la estimación de las ventas al menudeo de papel para imprenta y de escritura ver anexo 7, tomado de Makridakis Spyros y Wheelwright Steven. (1983). Forecasting Methods and Applications, 2ª Edición Nueva York. Tabla 9.2 página 104.

Se requiere un modelo tentativo para empezar el modelo de Box-Jenkins. Se puede determinar una forma inicial para este modelo si se examina los coeficientes de autocorrelación y autocorrelación parcial de los datos. En la figura 20 se puede observar que la autocorrelación más grande se da cada 12 meses, lo que sugiere un patrón estacional de 12 meses en los datos. Este es el tipo de información revelada por las autocorrelaciones que debe utilizarse para seleccionar el mejor modelo. También revela la existencia de tendencia en los datos, lo cual se observa por el hecho de casi todas las autocorrelaciones son mayores que cero. (No existe tendencia o estacionariedad si las autocorrelaciones fluctúan alrededor de cero después del segundo o tercer rezago de tiempo). Esto da a entender que se necesita cierta diferenciación para que los datos se hagan estacionarios. La diferencia puede ser periodo a periodo, estacional, es decir cada 12 periodos, o una combinación de las dos.

Figura 20
VENTAS C T



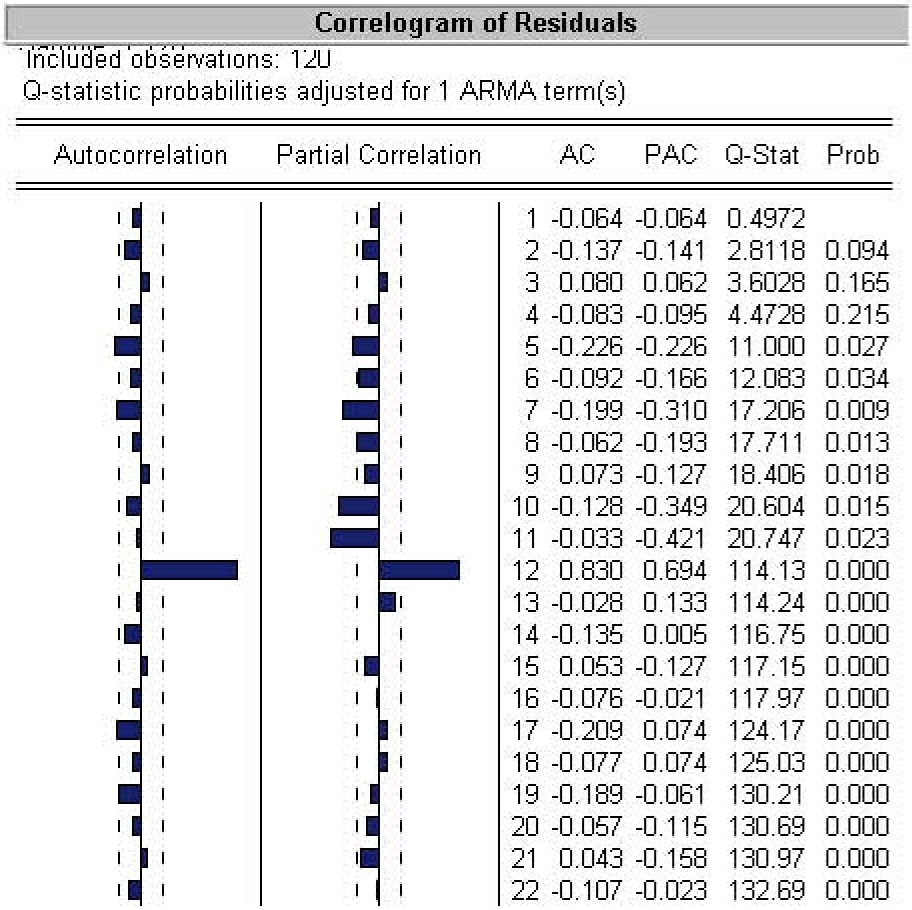
Correlogram of Residuals

Sample: 1 120
Included observations: 120

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.231	0.231	6.5917	0.010
		2	-0.137	-0.202	8.9271	0.012
		3	0.006	0.100	8.9314	0.030
		4	-0.138	-0.216	11.321	0.023
		5	-0.288	-0.198	21.888	0.001
		6	-0.243	-0.210	29.470	0.000
		7	-0.264	-0.303	38.488	0.000
		8	-0.109	-0.128	40.043	0.000
		9	0.008	-0.210	40.050	0.000
		10	-0.123	-0.419	42.054	0.000
		11	0.207	-0.036	47.805	0.000
		12	0.845	0.726	144.70	0.000
		13	0.211	-0.135	150.80	0.000
		14	-0.131	0.004	153.18	0.000
		15	-0.019	-0.139	153.23	0.000
		16	-0.133	0.057	155.73	0.000
		17	-0.262	0.049	165.50	0.000
		18	-0.219	0.025	172.40	0.000
		19	-0.246	-0.087	181.19	0.000
		20	-0.111	-0.132	183.00	0.000
		21	-0.013	-0.111	183.03	0.000
		22	-0.112	-0.033	184.91	0.000
		23	0.174	-0.089	189.46	0.000
		24	0.742	0.042	273.41	0.000
		25	0.203	-0.108	279.73	0.000
		26	-0.115	-0.041	281.81	0.000
		27	-0.018	-0.042	281.86	0.000
		28	-0.116	-0.009	284.00	0.000
		29	-0.257	-0.125	294.59	0.000
		30	-0.196	-0.028	300.83	0.000
		31	-0.201	0.012	307.50	0.000
		32	-0.083	0.051	308.65	0.000
		33	0.010	0.073	308.67	0.000
		34	-0.084	-0.012	309.86	0.000
		35	0.173	0.059	315.04	0.000
		36	0.663	0.018	391.61	0.000

La figura 21 muestra lo que sucederá, por ejemplo, si se hubiera tomado sólo la diferencia estacional (larga, MA). Los residuos resultantes no son estacionarios i (todos son mayores que ceros), porque una diferencia larga por sí misma no remueve la tendencia. Esto también pudo haberse descubierto al analizar las autocorrelaciones de los datos después de haber tomado una diferencia corta. En este caso las autocorrelaciones fluctúan alrededor de cero, lo cual indica que éste es el nivel apropiado de diferenciación.

Figura 21
VENTAS C T MA(1)

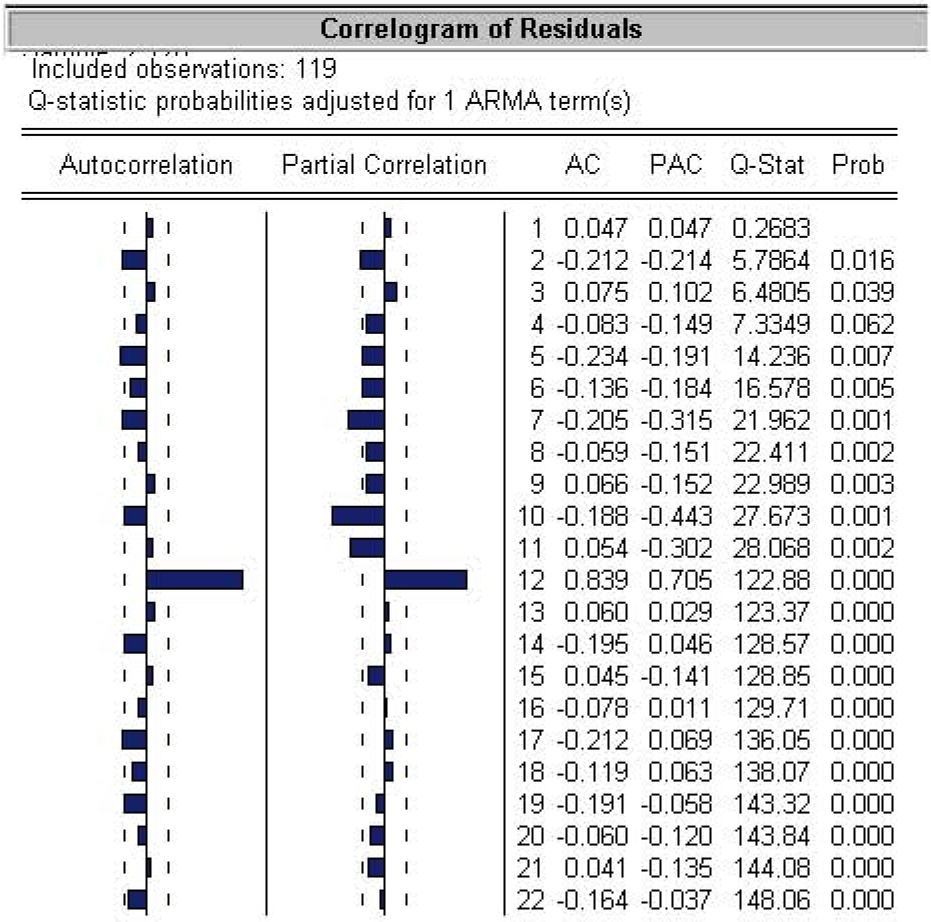

Correlogram of Residuals

Included observations: 120
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.064	-0.064	0.4972	
		2	-0.137	-0.141	2.8118	0.094
		3	0.080	0.062	3.6028	0.165
		4	-0.083	-0.095	4.4728	0.215
		5	-0.226	-0.226	11.000	0.027
		6	-0.092	-0.166	12.083	0.034
		7	-0.199	-0.310	17.206	0.009
		8	-0.062	-0.193	17.711	0.013
		9	0.073	-0.127	18.406	0.018
		10	-0.128	-0.349	20.604	0.015
		11	-0.033	-0.421	20.747	0.023
		12	0.830	0.694	114.13	0.000
		13	-0.028	0.133	114.24	0.000
		14	-0.135	0.005	116.75	0.000
		15	0.053	-0.127	117.15	0.000
		16	-0.076	-0.021	117.97	0.000
		17	-0.209	0.074	124.17	0.000
		18	-0.077	0.074	125.03	0.000
		19	-0.189	-0.061	130.21	0.000
		20	-0.057	-0.115	130.69	0.000
		21	0.043	-0.158	130.97	0.000
		22	-0.107	-0.023	132.69	0.000

La figura 22 presenta las autocorrelaciones de los datos después de tomar una diferencia corta (es decir, periodo a periodo) y una diferencia larga (o sea estacional). Se puede ver que las autocorrelaciones parciales para ese nivel de diferenciación (ver la figura 23) a continuación se identifica un modelo adecuado mediante el análisis tanto de las autocorrelaciones como de las parciales (figuras 22 y 23) y su comparación con las teóricas mostradas en la figura 17.

Figura 22
VENTAS C T AR(1)



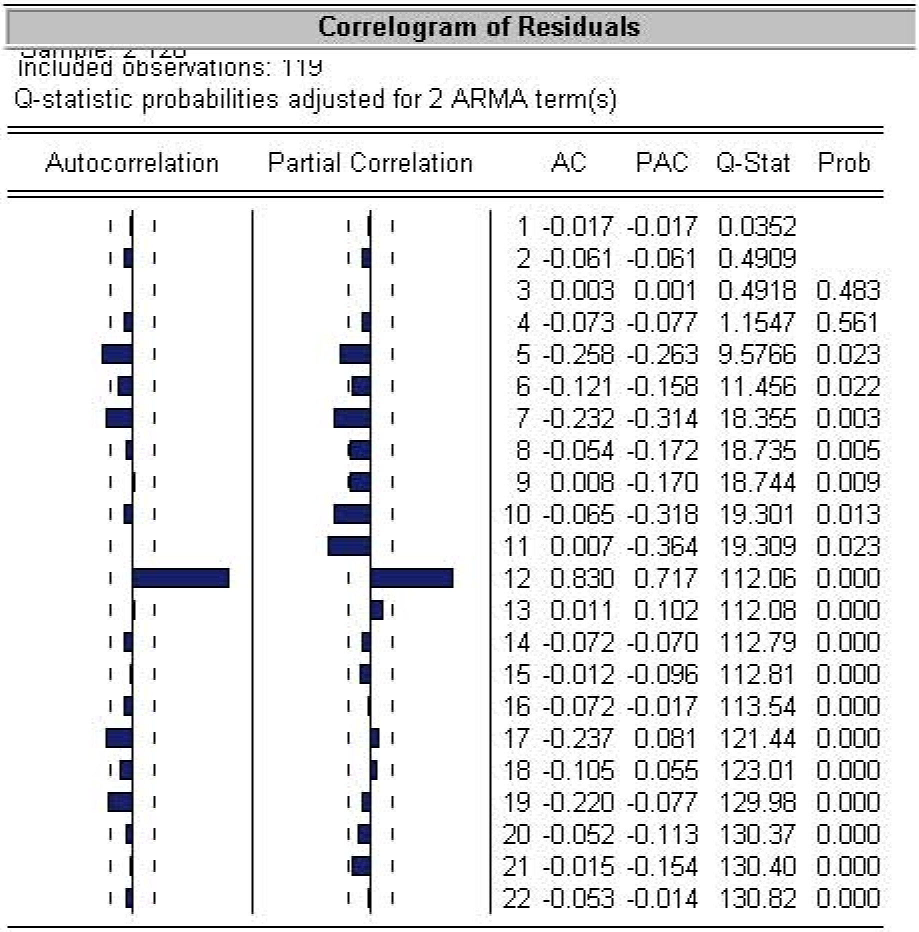
Correlogram of Residuals

Included observations: 119
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.047	0.047	0.2683	
		2	-0.212	-0.214	5.7864	0.016
		3	0.075	0.102	6.4805	0.039
		4	-0.083	-0.149	7.3349	0.062
		5	-0.234	-0.191	14.236	0.007
		6	-0.136	-0.184	16.578	0.005
		7	-0.205	-0.315	21.962	0.001
		8	-0.059	-0.151	22.411	0.002
		9	0.066	-0.152	22.989	0.003
		10	-0.188	-0.443	27.673	0.001
		11	0.054	-0.302	28.068	0.002
		12	0.839	0.705	122.88	0.000
		13	0.060	0.029	123.37	0.000
		14	-0.195	0.046	128.57	0.000
		15	0.045	-0.141	128.85	0.000
		16	-0.078	0.011	129.71	0.000
		17	-0.212	0.069	136.05	0.000
		18	-0.119	0.063	138.07	0.000
		19	-0.191	-0.058	143.32	0.000
		20	-0.060	-0.120	143.84	0.000
		21	0.041	-0.135	144.08	0.000
		22	-0.164	-0.037	148.06	0.000

Al analizar las figuras 22 y 23, lo primero que hay que buscar es cual de los dos patrones se acerca especialmente a cero. Este ejemplo es la figura 23, los parciales, la que lo hace, lo que significa que el modelo es apropiado es uno del promedio móvil. Una vez que esto se ha identificado, se observa la figura 22 (sólo autocorrelaciones no estacionales) y se determina que solamente existe una autocorrelación que es significativa diferente de cero estadísticamente hablando (es decir, fuera de las líneas punteadas paralelas al intervalo del 95%.

Figura 23
VENTAS C T AR(1) MA(1)


Correlogram of Residuals

Sample: 2 128
Included observations: 119
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.017	-0.017	0.0352	
		2	-0.061	-0.061	0.4909	
		3	0.003	0.001	0.4918	0.483
		4	-0.073	-0.077	1.1547	0.561
		5	-0.258	-0.263	9.5766	0.023
		6	-0.121	-0.158	11.456	0.022
		7	-0.232	-0.314	18.355	0.003
		8	-0.054	-0.172	18.735	0.005
		9	0.008	-0.170	18.744	0.009
		10	-0.065	-0.318	19.301	0.013
		11	0.007	-0.364	19.309	0.023
		12	0.830	0.717	112.06	0.000
		13	0.011	0.102	112.08	0.000
		14	-0.072	-0.070	112.79	0.000
		15	-0.012	-0.096	112.81	0.000
		16	-0.072	-0.017	113.54	0.000
		17	-0.237	0.081	121.44	0.000
		18	-0.105	0.055	123.01	0.000
		19	-0.220	-0.077	129.98	0.000
		20	-0.052	-0.113	130.37	0.000
		21	-0.015	-0.154	130.40	0.000
		22	-0.053	-0.014	130.82	0.000

Con sólo una autocorrelación significativa, se concluye que el promedio movil es 1, o MA(1) que es el modelo representado por la ecuación (44). Puesto que los datos son estacionales, se debe identificar también un parámetro estacional. Como se mostró anteriormente, no se encuentra disponible mucha información para ser utilizada en esto. Por lo que el mejor enfoque es manejar tanto el modelo MA como el AR. Generalmente, es mejor empezar con el modelo MA estacional, porque después de tomar una diferencia larga, este es el modelo apropiado. Por lo tanto, se supone un modelo MA estacional, como el mostrado por la ecuación (48). Así pues, el modelo global incluye una diferencia larga y una corta; es un MA(1) en la porción no estacional y un MA(1) en la estacional. Esto es precisamente el modelo representado en al ecuación (51)

Los parámetros ϕ_1 y θ_{12} del modelo representado por la ecuación (51) se estima empleando un algoritmo de optimización no lineal. Calcular los errores residuales e_t (es decir entre los valores reales y los pronosticados) permite verificar si son aleatorios, lo que significa que es un modelo apropiado, o no aleatorios, lo que sugiere que el moldeo no es adecuado.

La figura 24 presenta las autocorrelaciones de los residuales e_t, para los datos de papel de imprenta y de escritura, todos los residuales se hallan dentro de las líneas punteadas, de manera que la hipótesis de que todos provienen de una población cuya media es 0 y cuyos valores son aleatorios esa apoyada y puede aceptarse. Así el modelo seleccionado es apropiado y se puede aceptarse. Así el modelo seleccionado es el apropiado y se puede utilizar para pronosticar.

Figura 24
VENTAS C T AR(12) MA(12)

Correlogram of Residuals

Included observations: 108
Q-statistic probabilities adjusted for 2 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.026	0.026	0.0755	
		2	0.210	0.209	5.0162	
		3	0.174	0.172	8.4432	0.004
		4	0.089	0.048	9.3507	0.009
		5	0.069	-0.001	9.8949	0.019
		6	0.127	0.077	11.767	0.019
		7	0.068	0.039	12.306	0.031
		8	0.058	0.007	12.711	0.048
		9	0.095	0.045	13.804	0.055
		10	0.084	0.053	14.665	0.066
		11	-0.007	-0.054	14.670	0.100
		12	-0.068	-0.142	15.240	0.124
		13	-0.076	-0.119	15.963	0.143
		14	-0.058	-0.038	16.388	0.174
		15	-0.193	-0.165	21.157	0.070
		16	-0.158	-0.160	24.390	0.041
		17	-0.021	0.056	24.449	0.058
		18	-0.114	0.019	26.154	0.052
		19	-0.153	-0.106	29.278	0.032
		20	-0.091	-0.059	30.393	0.034
		21	-0.120	0.010	32.361	0.028
		22	-0.043	0.111	32.620	0.037

La metodología de Box-Jenkins para modelos ARMA es general. Para todo fin práctico, puede ajustar un modelo a todos los tipos de patrones de datos. Aunque es difícil de utilizar y requiere un compromiso individual para determinar el mejor modelo ARMA, el método es teóricamente elegante y proporciona un medio excelente para ejemplificar la aplicación de la teoría estadística a una situación practica de predicción.

CONCLUSIONES

Como primera conclusión para entender y conocer la econometría es necesario tener conocimientos de estadística tanto descriptiva como inferencial, ya que los métodos econométricos se apoyan en la estadística y si es mayor nuestro dominio de la estadística mejor será el avance y comprensión que tengan con la econometría.

Lo que he vivido en el aula de clase junto a muchos compañeros alumnos, es el miedo a esta materia y al uso de la computadora, de acuerdo al libro profesor no entiendo, el índice de reprobación de materias que llevan operaciones matemáticas es alto debido a que se tiene miedo de abordar estos temas que parecen complicados pero no lo son, claro siempre y cuando el alumno ponga de su parte y estudie sus lecciones.

Para encontrar un modelo que tenga los menor problemas posibles es necesario hacer una limpieza de datos, esto es deben trabajarse (de ser posible) las mismas unidades, contar (de preferencia) con una base de datos amplia que permita manipular los periodos de tiempo y sobre todo lo más importantes es especificar correctamente el modelo.

La violación a los supuestos, son métodos que permiten modificar el modelo desde el punto de vista estadístico, más no así, la realidad que representa o a la teoría, por eso si se viola algún supuesto no del todo el modelo ya no sirve o no debe ser utilizado, es importante saber que problema presenta y así poder aislar ese efecto ahora si no es posible se debe tomar con sus reservas, hacer el análisis, por ejemplo si va a ser utilizado para pronosticar y digamos presenta heteroscedasticidad es muy seguro que los intervalos de confianza sean grandes así como sus varianzas, porque los estimadores dejan de ser eficientes, entonces si queremos pronosticar el PIB para el siguiente año la probabilidad de que el dato que digamos sea erróneo es alto ya que no tenemos confianza en ese estimador.

Ahora de acuerdo a la experiencia es importante enfatizar que cuando se nos solicite un pronóstico a menos que se tenga una gran certeza decir el valor que encontró o pronóstico, pero sino, hay que dar un intervalo o sea una oscilación por ejemplo para este año se estima que el PIB este en 3.6 y si no se cumple ese pronóstico entonces nos podrán decir que nuestras estimaciones carecen de certeza, pero si decimos que el PIB va a estar entre 3.4 y 3.6 tenemos un margen de maniobra para que la estimación sea la correcta, es importante tomar encuenta el aspecto político que vamos a vivir, el clima social y la economía norteamericana para ver como nos va a influir. El pronóstico toma en cuenta todo lo que lo rodea.

Para realizar un pronóstico de cualquier variables, lo más importante es conocer esa variable, (conocerla hasta sus entrañas) debido a que no es posible conocer lo que ocurrirá el día de mañana, pero si puedo aproximarme lo más cercano posible a un resultado futuro si conozco el medio que rodea a esa variable, esto es, las técnicas de pronóstico son las mismas en cualquier parte del mundo; la pregunta sería porque no tenemos los mismo resultados que Engle, Granger, Morgan o las consultoras financieras de los bancos, la respuesta es que tal vez no contams con la suficiente información para que la estimación sea completa, o bien que la variable a estudiar no la conocemos a profundidad. Cito un ejemplo en clase de evaluación de proyectos con el profesor Rubén Valbuena se deseba elaborar el proyecto de un rastro GIF; se contaba con el terreno, se estimaron los costos de instalación de materia y equipo y de las personas que laborarían. Al cuestionar sobre cuanto cuesta el kilo de cerdo por metro para estimar la producción, resulta que el responsable de elaborar el proyecto no conocía a profundidad las razas de los cerdos ni el tiempo en que deben ser alimentados para crecer y para engordar y así cuando sean adultos poder ser utilizados para cría o bien para consumo humano, pero un dato importante que nos facilito un compañero que labora en Bancomer dice que manejan una estadística de metro por kilo de acuerdo a la raza y al lugar (clima y terreno) para estimar la demanda y la capacidad de producción, esto lo hacen porque es necesario para otorgar un crédito, entonces así es posible realizar un pronóstico y la pregunta es si ambos compañeros presenta una estimación de la demanda y producción para el próximo semestre, la pregunta es ¿cual de los dos pronósticos será el mejor?.

ANEXO ESTADÍSTICO

Anexo 1

Observaciones	Demanda de billetes	Precio del boleto	Gastos
1	4,475.592	12,000	1,082
2	7,766.554	12,500	1,125
3	4,474.026	11,500	1,036
4	4,901.003	12,000	1,082
5	4,966.902	12,600	1,129
6	4,684.715	13,000	1,169
7	4,934.257	13,000	1,172
8	5,239.256	13,400	1,208
9	5,099.800	14,000	1,259
10	5,125.852	13,600	1,219
11	5,586.101	14,200	1,274
12	5,483.044	14,500	1,302
13	5,488.272	14,400	1,296
14	5,548.987	15,000	1,346
15	6,159.040	15,000	1,352
16	5,944.903	15,600	1,406
17	6,154.661	15,900	1,434
18	6,006.833	15,600	1,400
19	6,250.133	16,000	1,441
20	6,038.611	16,500	1,485
21	5,944.998	16,500	1,483
22	6,016.141	16,600	1,495
23	6,148.402	17,000	1,533
24	6,341.689	17,000	1,530
25	6,202.944	17,500	1,574
26	6,280.347	17,600	1,581
27	6,415.922	17,700	1,594
28	6,293.523	17,600	1,582
29	6,311.806	17,800	1,603
30	6,756.047	18,400	1,654
31	6,823.044	18,600	1,670
32	6,908.101	18,800	1,695
33	6,938.690	19,000	1,711
34	6,937.003	19,000	1,708
35	6,897.003	19,500	1,754
36	7,231.595	19,700	1,768
37	7,172.691	20,000	1,800
38	7,186.839	20,100	1,808
39	7,246.436	20,000	1,796
40	7,392.408	20,400	1,835
41	7,249.284	20,300	1,826
42	7,556.200	20,600	1,853
43	7,942.424	20,600	1,852
44	7,600.025	20,800	1,872
45	7,663.419	21,000	1,888
46	7,746.627	21,000	1,889
47	7,901.372	21,500	1,934
48	7,794.934	22,000	1,980
49	7,781.061	22,000	1,981
50	8,204.235	22,500	2,019

Pérez A. Teodosio y Amorós G. Pablo. Ejercicios de Econometría Empresarial. Universidad Complutense de Madrid. 1992.

Anexo 2

Ingreso	Consumo	Gasto	Gex
302.565	180.236	51.189	274.346
107.689	51.850	27.642	100.310
348.540	165.985	75.735	330.602
14.619	7.991	3.813	13.253
17.812	8.693	5.356	16.000
117.007	59.046	30.838	115.951
134.899	75.042	30.844	126.969
74.856	43.100	17.484	70.744
761.142	389.019	193.713	708.598
241.131	118.762	60.946	225.651
14.817	7.806	4.242	13.773
186.509	92.236	43.276	176.606
337.445	194.993	82.411	323.954
64.181	36.941	12.675	59.007
170.364	102.965	38.456	168.279
1051.196	568.539	254.043	1037.175
50.517	28.624	10.255	46.462

Pena Trapero J Bernardo. (1995). Cien Ejercicios de Econometria. Mc Graw Hill.

Anexo 3

Observaciones	Consumo	Ingreso	Indice de precios
1986.01	110.5822	121.303	100.000
1986.02	101.8749	122.364	100.852
1986.03	109.9688	123.765	101.509
1986.04	110.3758	124.674	102.269
1986.05	107.7555	125.357	102.541
1986.06	106.5921	126.197	103.505
1986.07	110.2095	127.006	103.646
1986.08	105.2085	127.957	103.681
1986.09	110.3003	128.217	104.204
1986.10	108.9894	129.720	104.437
1986.11	105.3561	130.856	104.950
1986.12	112.5339	132.380	105.581
1987.01	113.5705	132.755	106.428
1987.02	107.7264	132.910	106.608
1987.03	110.2678	133.602	106.775
1987.04	114.8640	134.427	107.207
1987.05	114.0675	135.492	107.609
1987.06	111.2953	136.300	108.500
1987.07	116.8521	137.362	108.658
1987.08	116.3594	140.855	109.950
1987.09	119.9550	143.673	110.461
1987.10	115.7285	145.419	111.848
1987.11	118.8287	147.741	112.887
1987.12	120.9660	150.533	113.355
1988.01	125.2613	165.570	113.637
1988.02	131.3489	171.154	114.418
1988.03	141.3871	170.896	115.134
1988.04	132.5623	174.599	115.995
1988.05	135.6305	179.649	116.825
1988.06	143.2009	184.145	118.782
1988.07	145.3711	187.348	119.501
1988.08	157.2821	195.510	119.762
1988.09	156.6270	200.083	120.165
1988.10	168.9906	207.788	121.089
1988.11	162.9529	212.585	122.875
1988.12	169.3464	218.896	123.112
1989.01	176.5547	223.551	124.102
1989.02	180.2491	228.760	126.007
1989.03	179.8900	235.798	126.746
1989.04	188.0911	241.125	127.143
1989.05	186.7479	248.507	128.162
1989.06	199.1130	256.207	128.548
1989.07	201.3457	258.758	129.954
1989.08	208.1099	263.586	130.774
1989.09	196.5700	268.352	131.224
1989.10	201.5204	274.242	131.985
1989.11	209.2899	279.196	132.946
1989.12	198.9554	284.896	133.059
1990.01	215.2799	289.792	134.457
1990.02	210.9201	295.056	135.706

Pérez A. Teodosio y Amorós G. Pablo. Ejercicios de Econometría Empresarial. Universidad Complutense de Madrid. 1992.

Anexo 4

obs	CP	PIBPM
1985	4,498,034	28,200,885
1986	4,740,221	32,323,992
1987	5,159,905	36,143,972
1988	5,368,137	40,158,739
1989	5,813,462	45,044,128
1990	6,197,776	50,145,195
1991	6,543,696	54,927,320
1992	6,808,095	59,104,986
1993	6,971,511	60,952,584
1994	6,948,140	64,811,535
1995	7,074,014	69,780,058
1996	7,141,101	73,743,261
1997	7,239,097	77,896,586

Martín R. Gullermina y Labeaga A. José. (1997). Introducción a la Econometría. Prentice Hall.

Anexo 5

AÑOS	PIB	IMPORT	FBC
1970	2,340.75	230.681	323.258
1971	2,428.82	220.785	348.066
1972	2,628.68	254.870	359.018
1973	2,835.33	307.869	369.950
1974	2,999.12	362.738	415.714
1975	3,171.40	348.563	426.317
1976	3,311.50	327.437	445.983
1977	3,423.78	267.655	411.993
1978	3,730.45	326.578	437.703
1979	4,092.23	422.760	534.420
1980	4,470.08	579.961	630.373
1981	4,862.22	682.726	703.001
1982	4,831.69	424.297	596.833
1983	4,628.94	280.906	464.968
1984	4,796.05	330.945	501.599
1985	4,920.43	367.298	562.951
1986	4,735.72	339.442	504.297
1987	4,823.60	356.910	536.792
1988	4,883.68	487.947	591.774
1989	5,047.21	591.666	636.057
1990	5,271.54	708.518	720.467
1991	5,462.73	827.372	814.356
1992	5,615.96	999.992	943.161
1993	5,649.67	987.484	937.597
1994	5,848.47	1,115.315	1,011.976

www.inegi.gob.mx datos anuales para el período 1970-1994

Anexo 6

Periodos	PIB	TC	X bys	M bys	IED	PP dlls por	PIB EU
1980 01	40,966,614.54	0.02	5,810.90	7,146.50	400.8	29.8	5,221,300
1980 02	40,849,834.63	0.02	6,059.90	8,614.30	509.8	30.7	5,115,900
1980 03	40,053,909.96	0.02	6,263.50	9,442.10	680.6	31.8	5,107,400
1980 04	42,729,067.30	0.02	6,726.30	10,091.90	498.6	31.9	5,202,100
1981 01	42,668,174.37	0.02	8,156.30	10,741.10	530.4	36.2	5,307,500
1981 02	42,259,859.18	0.02	8,132.10	12,009.60	970.4	33.1	5,266,100
1981 03	39,843,785.44	0.03	6,988.60	12,067.30	541.9	31.8	5,329,800
1981 04	40,733,621.45	0.03	7,834.40	12,534.20	1,033.10	31.6	5,263,400
1982 01	23,099,712.16	0.05	7,018.30	11,111.80	584.4	30.2	5,177,100
1982 02	21,687,966.19	0.05	7,708.70	10,232.30	620.8	28.2	5,204,900
1982 03	14,239,045.06	0.07	8,038.10	8,413.20	607.9	28.3	5,185,200
1982 04	10,557,071.58	0.10	8,207.00	7,104.80	87.2	28.5	5,189,800
1983 01	9,298,983.67	0.11	7,546.80	5,908.50	824.3	27.5	5,253,800
1983 02	8,220,332.85	0.12	8,115.70	6,706.70	682.6	25.6	5,372,300
1983 03	7,245,502.55	0.13	8,270.10	7,235.80	454.9	26.0	5,478,400
1983 04	7,014,264.09	0.14	8,996.00	7,218.10	229.9	26.7	5,590,500
1984 01	6,665,564.63	0.16	9,768.90	7,595.10	665.4	26.7	5,699,800
1984 02	6,065,484.25	0.17	9,502.10	8,258.00	451.5	26.9	5,797,900
1984 03	5,582,879.73	0.18	9,288.20	8,996.90	259.9	26.9	5,854,300
1984 04	5,393,418.37	0.19	9,271.40	8,797.10	164.2	26.9	5,902,400
1985 01	5,061,517.79	0.21	9,245.20	9,070.70	601.6	26.4	5,956,900
1985 02	4,626,170.14	0.23	8,379.10	8,901.30	835.4	26.1	6,007,800
1985 03	3,336,279.12	0.30	9,110.30	8,796.30	433.6	24.6	6,101,700
1985 04	2,874,674.89	0.37	9,124.30	8,291.20	112.9	24.4	6,148,600
1986 01	2,174,808.75	0.47	7,517.30	7,917.10	281.6	14.9	6,207,400
1986 02	1,833,877.68	0.57	7,342.10	8,194.00	702.7	10.1	6,232,000
1986 03	1,291,503.84	0.75	7,025.00	7,658.80	400.1	10.6	6,291,700
1986 04	1,108,266.28	0.92	8,043.50	7,531.50	1,016.30	12.5	6,323,400
1987 01	903,251.40	1.12	8,720.40	7,341.80	352.5	15.1	6,365,000
1987 02	778,688.27	1.35	9,548.50	8,025.90	632.5	17.0	6,435,000
1987 03	633,690.36	1.57	9,402.50	8,871.70	497	17.0	6,493,400
1987 04	481,661.54	2.21	9,697.00	8,890.10	1,152.60	14.9	6,606,800
1988 01	455,346.11	2.28	10,339.20	9,580.00	470.3	12.8	6,639,100
1988 02	465,317.03	2.28	10,749.40	10,713.00	857.8	13.3	6,723,500
1988 03	435,455.50	2.28	10,343.10	11,796.60	594.5	11.8	6,759,400
1988 04	472,870.58	2.28	10,664.10	12,381.80	957.5	10.9	6,848,600
1989 01	451,153.58	2.37	11,696.90	12,727.30	618.3	14.5	6,918,100
1989 02	451,871.96	2.46	12,347.00	13,617.90	842.4	16.3	6,963,500
1989 03	411,958.85	2.55	11,708.20	13,769.70	883.3	15.3	7,013,100
1989 04	421,017.90	2.64	12,351.40	13,809.70	831.6	16.5	7,030,900
1990 01	408,038.64	2.73	13,121.80	15,312.70	561	15.6	7,112,100
1990 02	410,448.44	2.82	12,468.70	14,127.50	615.6	11.9	7,130,300
1990 03	381,529.60	2.89	14,203.80	15,982.10	578.9	22.2	7,130,800
1990 04	405,179.80	2.95	16,276.50	18,099.50	877.8	25.4	7,076,900

Periodos	PIB	TC	X bys	M bys	IED	PP dlls por	PIB EU
1991 01	388,307.75	2.98	13,390.60	15,533.60	1,649.40	14.4	7,040,800
1991 02	404,798.76	3.02	14,807.40	18,659.30	20.00	14.3	7,086,500
1991 03	373,100.90	3.06	14,606.20	18,763.80	702.1	15.2	7,120,700
1991 04	404,134.31	3.07	15,283.00	19,777.20	1,263.40	14.4	7,154,100
1992 01	393,009.72	3.08	14,460.20	19,611.50	1,042.70	12.5	7,228,200
1992 02	400,299.87	3.12	15,467.30	21,365.10	1,180.20	15.5	7,297,900
1992 03	382,303.39	3.12	15,484.30	22,339.20	1,275.00	16.3	7,369,500
1992 04	409,586.21	3.12	16,257.20	22,791.50	894.9	15.1	7,450,700
1993 01	403,126.72	3.10	15,628.30	21,289.40	1,163.70	14.1	7,459,700
1993 02	403,803.66	3.12	16,951.90	22,596.90	954.4	14.0	7,497,500
1993 03	388,600.85	3.12	16,683.30	23,349.40	550.1	12.9	7,536,000
1993 04	419,886.94	3.11	18,488.60	23,915.50	1,720.60	11.8	7,637,400
1994 01	380,331.58	3.36	18,061.80	24,843.30	3,152.00	11.6	7,715,100
1994 02	392,545.27	3.39	19,406.30	26,882.10	3,283.40	14.3	7,815,700
1994 03	372,322.65	3.40	19,460.70	27,369.10	2,813.90	14.6	7,859,500
1994 04	257,679.31	5.33	21,442.90	28,939.20	1,723.20	15.0	7,951,600
1995 01	186,614.09	6.82	23,017.40	24,372.10	1,982.80	15.8	7,973,700
1995 02	191,633.28	6.31	24,056.10	23,699.80	2,913.60	16.8	7,988,000
1995 03	181,568.69	6.42	24,413.10	24,863.60	2,254.70	14.8	8,053,100
1995 04	166,903.17	7.64	25,542.60	25,670.60	2,375.20	15.5	8,112,000
1996 01	168,666.52	7.55	26,626.20	26,789.60	2,027.70	17.0	8,169,200
1996 02	169,154.53	7.61	28,416.70	28,161.50	1,779.90	18.0	8,303,100
1996 03	165,662.58	7.54	29,104.40	29,968.90	2,004.40	19.1	8,372,700
1996 04	174,029.98	7.85	31,168.80	32,903.70	3,373.50	21.6	8,470,600
1997 01	168,750.64	7.89	30,415.40	30,635.20	2,109.20	18.5	8,536,100
1997 02	175,333.01	7.96	32,658.90	33,842.80	2,594.50	15.9	8,665,800
1997 03	171,619.58	7.82	33,334.00	35,924.20	5,595.00	16.0	8,773,700
1997 04	180,282.60	8.08	34,909.80	38,580.90	2,530.80	15.6	8,838,400
1998 01	168,127.96	8.52	33,738.10	36,983.20	2,538.60	10.8	8,936,200
1998 02	161,004.58	9.04	35,441.10	38,868.20	3,468.90	10.5	8,995,300
1998 03	139,803.49	10.11	34,207.00	38,920.90	3,006.80	10.3	9,098,900
1998 04	151,738.71	9.87	36,682.60	41,368.80	3,006.80	9.2	9,237,100
1999 01	153,717.00	9.52	35,406.30	39,084.30	3,237.90	9.2	9,315,500
1999 02	158,767.59	9.49	39,268.40	42,124.90	3,298.90	13.8	9,392,600
1999 03	157,669.47	9.36	40,820.90	43,990.20	2,907.60	18.5	9,502,200
1999 04	165,774.89	9.51	43,444.30	47,741.20	3,708.80	21.3	9,671,100
2000 01	170,180.75	9.23	44,417.20	49,166.40	4,223.34	24.4	9,695,600
2000 02	162,456.29	9.95	47,553.50	50,947.30	4,458.34	24.4	9,847,900
2000 03	167,872.92	9.41	49,760.90	53,501.60	2,784.72	26.3	9,836,600
2000 04	172,531.22	9.57	51,549.40	57,832.60	5,443.14	23.5	9,887,700
2001 01	167,923.24	9.54	46,821.10	51,647.70	3,141.94	19.6	9,875,600
2001 02	178,752.25	9.06	47,764.20	51,381.30	5,136.57	19.8	9,905,900
2001 03	163,650.95	9.53	45,626.30	49,121.10	14,968.31	19.9	9,871,100
2001 04	178,189.89	9.14	45,390.00	51,609.50	4,473.95	14.9	9,910,000

Periodos	PIB	TC	X bys	M bys	IED	PP dlls por	PIB EU
2002 01	173,313.45	9.03	43,198.70	46,586.00	2,809.50	17.0	9,977,300
2002 02	165,052.22	10.00	48,485.60	51,264.70	4,546.52	22.4	10,031,600
2002 03	155,926.28	10.17	48,168.90	51,315.00	3,078.72	23.8	10,090,700
2002 04	161,116.67	10.31	48,003.30	52,773.20	4,890.48	22.7	10,095,800
2003 01	148,956.54	10.77	46,400.10	48,462.40	3,056.48	26.5	10,138,600
2003 02	157,617.84	10.48	48,032.60	49,662.60	3,958.32	23.0	10,230,400
2003 03	146,454.07	10.93	49,340.80	51,494.10	2,395.65	24.6	10,410,900
2003 04	150,983.98	11.24	51,388.70	54,691.70	2,253.17	25.1	10,502,600
2004 01	148,567.67	11.2	51,720.2	53,152.0	8,646.16	26.8	10,612,500
2004 02	148,441.37	11.5	56,590.1	56,792.1	3,294.33	30.2	10,704,100
2004 03	146,925.64	11.4	57,325.7	58,219.2	2,554.89	33.5	10,808,900
2004 04	159,543.12	11.1	58,539.7	63,283.0	3,414.36	33.3	10,897,100
2005 01	150,550.42	11.3	56,377.8	59,491.5	4,356.59	34.6	10,999,300
2005 02	164,200.45	10.8	63,903.2	63,719.5	3,847.86	41.5	11,089,200
2005 03	162,119.85	10.7	64,601.5	65,023.3	4,689.42	49.3	11,202,300
2005 04	170,761.53	10.7	69,866.2	72,586.5	3,000.00	45.4	11,233,500

www.reu.com.mx, datos trimestrales para el período 1980.1 a 2005.4.

Anexo 7

Ventas al menudeo de papel para imprenta y de escritura					
Periodo	Observación	Periodo	Observación	Periodo	Observación
1	562.674	41	701.108	81	742.000
2	599.000	42	790.079	82	847.152
3	668.516	43	594.621	83	731.675
4	597.798	44	230.716	84	898.527
5	579.889	45	617.189	85	778.139
6	668.233	46	691.389	86	856.075
7	499.232	47	701.067	87	938.833
8	215.187	48	705.777	88	813.023
9	555.813	49	747.636	89	783.417
10	586.935	50	773.392	90	828.110
11	546.136	51	813.788	91	657.311
12	571.111	52	766.713	92	310.032
13	634.712	53	728.875	93	780.000
14	639.283	54	749.197	94	860.000
15	712.182	55	680.954	95	780.000
16	621.557	56	241.424	96	807.993
17	621.000	57	680.234	97	895.217
18	675.989	58	708.326	98	856.075
19	501.322	59	694.238	99	893.268
20	220.286	60	772.071	100	875.000
21	560.727	61	795.337	101	835.088
22	602.530	62	788.421	102	934.595
23	626.379	63	889.968	103	832.500
24	605.508	64	797.393	104	300.000
25	646.783	65	751.000	105	791.443
26	658.442	66	821.255	106	900.000
27	712.906	67	691.605	107	781.729
28	687.714	68	290.655	108	880.000
29	723.916	69	727.147	109	875.024
30	707.183	70	868.355	110	992.968
31	629.000	71	812.390	111	976.804
32	237.530	72	799.556	112	968.697
33	613.296	73	843.038	113	871.675
34	730.444	74	847.000	114	1006.852
35	734.925	75	941.952	115	832.037
36	651.812	76	804.309	116	345.587
37	676.155	77	840.307	117	849.528
38	748.183	78	871.528	118	913.871
39	810.681	79	656.33	119	868.746
40	729.363	80	370.508	120	993.733

Makridakis Spyros y Wheelwright Steven. (1983). Forecasting Methods and Applications, 2ª Edición Nueva York. Tabla 9-3 página 434.

BIBLIOGRAFIA BÁSICA

1. Alcaide Angel y Alvarez Nelson. (1992). Econometría, Modelos Deterministas y Estocásticos. Editorial Centro de Estudios Ramón Areces S.A.. Madrid España.
2. Amemiya Takeshi. Advanced Econometrics. Harvard.
3. Barceinas Fernando y Mendoza Miguel A. (1998). Tópicos en Economía Matemática y Econometría. UNAM-UAM Azcapotzalco.
4. Barrera G. Genoveva. (1997). Apuntes de clase series de tiempo. Facultad de Economía. UNAM.
5. Calero V. Arístides. (1998). Estadística II. IPN México - Ministerio de Educación superior Cuba.
6. Castro Cesar; Loría Eduardo; Mendoza Miguel A. (1997). Eudoxio: Modelo Macroeconométrico de la Economía Mexicana. Facultad de Economía UNAM.
7. Charemza Wojciech y Deadman Derek. (1997). New Directions in Econometric Practice General to Specifie Modelling, Cointegration and Vector Autoregression. 2ª Edición Edward Elgar.
8. Fernández Ana y González Pilar. (1995). Ejercicios de Econometria. Mc Graw Hill.
9. Granger C.W. y Micon G.E. Nonstationary Time Series Analysis and Cointegration. Oxford University Press. Edited by Colinn P. Hargreaves.
10. Green William. (1999). Análisis Econométrico. 3ª Edición. Prentice Hall.
11. Griliches Zvi. (1997). Handbook of Econometrics. Edit Amsterdam: North-Holland Volumen II Parte 5 Capítulos 17 - 22.
12. Guerrero Victor M. Obtención de Pronósticos Óptimos Sujetos a Restricciones con Modelos ARIMA, Documento de Trabajo Número 1-1988.
13. Guerrero Victor M. (1991). Análisis Estadístico de Series de Tiempo Económicas. UAM-Iztapalapa. México.
14. Gujarati Domadar. (1995). Econometría Básica. Mc Graw Hill 3ª Edición.
15. Hanke Jonh y Reitsch Arthur. (1996). Pronósticos en los Negocios. Prentice Hall Hispanoamérica S.A. 5ª edición.
16. Intriligator Michael. (1990). Modelos Econométricos, Técnicas y Aplicaciones. Fondo de Cultura Económica.
17. Johnston J. (1979). Métodos de Econometría. Editorial Vincens Vives. España.
18. Judge George. (1980). The Theory and Practice of Econometrics. 2ª Edición.
19. Kelejian Harry y Oates Wallace. (1995). Introducción a la Econometría, Principios y Aplicaciones.. Editorial Harla. 3ª Edición.
20. Kmenta. (1979). Elementos de Econometría. Editorial Vincens Vives. España.
21. Kohler Heinz. (1998). Estadística para Negocios y Economía. Parte IV y V. Compañía Editorial Continental, S.A. de C.V. México.
22. Ludlow W. Jorge. (1997). Modelos, Pronósticos y Volatilidad de las Series de Tiempo Generadas en la Bolsa Mexicana de Valores. UAM-Azcapotzalco.
23. Luis P. Octavio. (1992). La Multicolinealidad en Econometría, Diagnóstico y Corrección del Problema. IPN-SITESA.
24. Maddala G. (1996). Introducción a la Econometría. Prentice Hall Hispanoamérica S.A. 2ª Edición.
25. Maddala. (1985). Econometría. Mc Graw Hill.
26. Makridakis Spyros y Wheelwright Steven. (1983). Forecasting Methods and Applications, 2ª Edición Nueva York.
27. Martín R. Gullermina y Labeaga A. José. (1997). Introducción a la Econometría. Prentice Hall.
28. Montgomery Douglas y Johnson Lynwood. (1990). Forcasting and Time Series Analysis. 2ª Edición. Mc Graw Hill.
29. Montgomery Douglas y Runger George. (1996). Probabilidad y Estadística Aplicadas a la Ingeniería. Mc Graw Hill..
30. Nervole Marc. Análisis de Series Temporales Económicas. Fondo de Cultura Económica.1991.
31. Novales C. Alfonso. Econometría. Mc Graw Hill. 1988.
32. Novales C. Alfonso. Estadística y Econometría. Mc Graw Hill. 1997.
33. Nuñez Z. Rafael. Una Introducción a la Econometría: Enfoques Clásicos y Contemporáneos. Microediciones. 1996.
34. Otero José. Econometría, Series Temporales y Predicción. Edición A.C.1993.

35. Padilla D. José. Estadística Inferencial y Econometría. IPN. 1991.
36. Pérez A. Teodosio y Amorós G. Pablo. Ejercicios de Econometría Empresarial. Universidad Complutense de Madrid. 1992.
37. Pindyck Robert y Rubinfeld Daniel. Econometric Models and Economic Forecasts. Mc Graw Hill IRWIN.4ª Edición. 1998.
38. Pulido Antonio. Modelos Econométricos. Ediciones Pirámide, S.A. de C.V. 1998.
39. Salas Javier. Econometría Aplicada a los Países en Desarrollo el Caso Mexicano. SEP- Fondo de Cultura Económica.1990.
40. Salvatore Dominick. Econometría. Mc Graw Hill. 1991.
41. Sánchez B. Genaro. Introducción a la Econometría. Facultad de Economía UNAM. 1999.
42. Spanos Aris. Probability Theory and Statistical Inference Econometric Modeling with Observational Data. Cambridge. 1999.
43. Spanos Aris. Statistical Foundations of Econometric Modelling, Cambridge. 1986.
44. Uriel Ezequiel y Contreras Dulce. Econometría El Modelo Líneal. Editorial A.C. Libros Científicos y Técnicos Madrid.

BIBLIOGRAFIA COMPLEMENTARIA

1. Aceituno G y Máttar J. Modelos Macroeconómicos en México: Un Análisis Comparativo. Editorial Trillas. México 1986.
2. Ackoff Russell. El Arte de Resolver Problemas. Editores Limusa-Noriega. 1999.
3. Anderson David y Sweeney Dennis. Introducción a los Modelos Cuantitativos para Administración.. Grupo Editorial Iberoamérica. Capítulo 16 - 19. 1993.
4. Balbuena Alvarez R. Formulación y Evaluación de Proyectos. 1997. Material de clase.
5. Barbacho Alfonso. Fundamentos y Posibilidades de la Econometría. Editorial Ariel. España 1976.
6. Barry Nalebuff y Avinash Dixit. Pensar Estratégicamente una Arma Decisiva en los Negocios, la Política y la Vida Diaria. Editor Antoni Bosch. 1992.
7. Canavos George. Probabilidad y Estadística, Aplicaciones y Métodos. Mc Graw Hill. 1988.
8. Christ Carl. Modelos y Métodos Econométricos. Editorial Limusa. 1974.
9. Dagun Camilo. Introducción a la Econometría. Editorial Siglo XXI. 1971.
10. Hanke Jonh y Reitsch Arthur. Estadística para Negocios. Mc Graw Hill IRWIN. 2ª Edición. 1997.
11. Levin Richard. Estadística para Administradores. Prentice Hall. Segunda Edición 1988. Capítulos 14 y 15.
12. Pasinetti Luigi. Lecciones de Teoría de la Producción. Fondo de Cultura Económica.1987.
13. Sánchez B. Genaro. La Estadística Aplicada al Análisis Económico. Facultad de Economía UNAM. 1998.
14. Shao Stepen. Estadística para Administradores de Empresas. Editorial Herrero Hermanos SUCS S.A.. Parte 5 Capítulos 17 - 20. 1988.
15. Sun-Tzu. El Arte de la Guerra. 3ª Edición. Grupo Editorial Tomo, S.A. de C.V. 1998.
16. Wonnacott Tomas. Introductory Statistics for Business and Economics. 4ª edición. Capítulos 15 y 24.

TESIS

1. Barreto V. Gabriel. La Metodología ARIMA para la Formulación de los Modelos de Series de Tiempo. Facultad de Economía UNAM.1992.
2. Fernández R. Fernando. El Problema de la Predicción en Series Temporales: Aplicaciones del Caos Determinista. Universidad de Las Palmas de Gran Canaria. 1995.

SOFTWARE

1. Eviews User Guide Versión Dos. Quantitative Micro Software Irvine, California. 1995. Teléfono (714) 856-33-68 fax (714) 856-20-44.
2. Hall Robert y Johnston Jack. Micro TSP User's Manual Versión 7.0. Quantitative Micro Software Irvine, California. 1990. Teléfono (714) 856-33-68 fax (714) 856-20-44.
3. Vargas S. Gustavo. Aplicaciones Econométricas con TSP. Facultad de Economía UNAM. 1996.

REVISTAS

1. Cassoni Adriana. Pruebas de Diagnóstico en el Modelo Econométrico. Documentos de Trabajo, CIDE 1991.
2. Feliz Raúl y Welch John. Cointegración y Pruebas de un Modelo Clásico de Inflación en Argentina, Bolivia, Brasil, México y Perú. Documento de trabajo 2 CIDE.1992.
3. Vargas Laura y Feliz Raúl. Una Prueba Econométrica del Enfoque Moderno de la Cuenta Corriente en México. Documento de trabajo Número 34. CIDE. 1998.
4. "Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models" Journal of the American Statistical Association, vol. 65, (1970), pp. 1509-1526.

DIRECCIONES ELECTRONICAS

http://www.eviews.com Página WEB de Eviews
http://www.mathworks.com Página WEB de Matlab
http://www.estima.com Página WEB de RATS
http://www.ststs.com Página WEB de programa estadístico con enlace a SPSS versión 7.0 y 8.0.
http://weatherhead.cwru.edu/forecasting International Institute of Forecasters, Case Western Reserve University.
http://weatherhead.cwru.edu/ forecasting/ijf.html International Journal of Forecasting
http://journals.wiley.com/wilcat-bin/ops/ID0012094/0277-6693/prod Journal of Forecasting
http://www.lancs.ac.uk/users/mansch/manageme/research/forecast.htom Centre for Forecasting, Lancaster University, GB.