Universidad del País Vasco
Euskal Herriko Unibertsitatea

Tema 7

El Modelo de Regresión Lineal General

Heterocedasticidad y Autocorrelación

Pilar González y Susan Orbe

Dpto. Economía Aplicada III (Econometría y Estadística)

Objetivos de aprendizaje

- Conocer los conceptos de heterocedasticidad y autocorrelación.
- Identificar las consecuencias de la presencia de heterocedasiticad y/o autocorrelación en las propiedades del estimador MCO.
- Identificar las consecuencias de la presencia de heterocedasiticad y/o autocorrelación sobre la inferencia basada en el estimador MCO.
- Detectar la posible existencia de existencia de heterocedasiticad y/o autocorrelación.
- Realizar inferencia robusta a la heterocedasiticad y/o autocorrelación basada en el estimador MCO.

Contenido

1. Especificación del Modelo de Regresión Lineal General.
2. Heterocedasticidad.
    - Concepto.
    - Consecuencias.
    - Detección de la heterocedasticidad.
3. Autocorrelación.
    - Concepto
    - Consecuencias.
    - Detección de la autocorrelación.
4. Inferencia con el estimador MCO.
5. Actividades propuestas: A7.
6. Ejercicios propuestas: E7.1, E7.2 y E7.3.

Contenido

Especificación del MRLG.

Supuestos básicos del MRLG.

S.1 El modelo en la población se puede escribir como:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_k X_{ki} + u_i \qquad i = 1, .., N.$$

S.2 Ausencia de colinealidad perfecta.

S.3 Media condicionada cero: $E(u_i|X_2, X_3, \ldots, X_k) = 0 \quad \forall i = 1, 2, \ldots, N.$

S.4 Homocedasticidad (varianza constante):

$\text{Var}(u_i|X_2, X_3, \ldots, X_k) = E(u_i^2|X_2, X_3, \ldots, X_k) = \sigma^2 \quad \forall i = 1, 2, \ldots, N.$

S.5 Ausencia de autocorrelación: $\text{Cov}(u_i, u_j|X_2, X_3, \ldots, X_k) = 0 \quad \forall i \neq j.$

S.6 Normalidad: las perturbaciones u_i son independientes de X y están idénticamente distribuidas.

Especificación del MRLG.

En forma matricial.

S.1 El modelo en la población se puede escribir como: $Y = X\beta + u$.

S.2 Ausencia de colinealidad perfecta.

S.3 Media condicionada cero: $E(u|X) = 0$.

S.4 + S.5 Homocedasticidad (varianza constante) y Ausencia de autocorrelación:

$$V(u|X) = E_X(uu') = \begin{bmatrix} \sigma^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \sigma^2 \end{bmatrix} = \sigma^2 \, I_N .$$

S.6 Las perturbaciones siguen una distribución normal.

Especificación del MRLG.

Bajo estos supuestos básicos, condicionado a X:

A. El estimador $\hat{\beta}_{MCO}$, es

- Lineal.
- Insesgado.
- Eficiente en el sentido de Gauss-Markov.

B. El estimador de la varianza de la perturbación, $\hat{\sigma}^2 = \frac{SCR}{(N-k)}$, es insesgado.

C. $\hat{\beta} \sim N(\beta, \sigma^2(X'X)^{-1})$

D. Estadísticos para contraste de hipótesis:

$$t = \frac{\hat{\beta}_j - \beta_j^0}{\hat{\sigma}_{\hat{\beta}_j}} \overset{H_0}{\sim} t(N-k) \qquad F = \frac{(SCR_R - SCR_{NR})/q}{SCR_{NR}/(T-k)} \overset{H_0}{\sim} \mathcal{F}(q, N-k)$$

Contenido

Concepto.

Homocedasticidad : la varianza de la perturbación u_i condicionada a X_i es constante para $i = 1, 2, \ldots, N$.

Heterocedasticidad : la varianza de la perturbación u_i es diferente en distintos segmentos de la población, dependiendo de los valores de uno o más de los regresores del modelo.

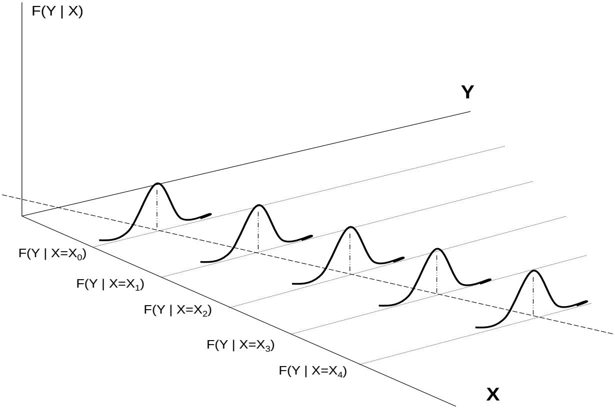
$$Var(u_i|X) = \sigma_i^2 \quad i = 1, 2, \ldots, N$$

Se espera la presencia de heterocedasticidad sobre todo cuando se trabaja con datos de sección cruzada.

Supongamos un modelo de regresión del consumo en función de la renta. Es de esperar que la varianza de los factores no observables que afectan al consumo (u) aumente conforme aumenta la renta.

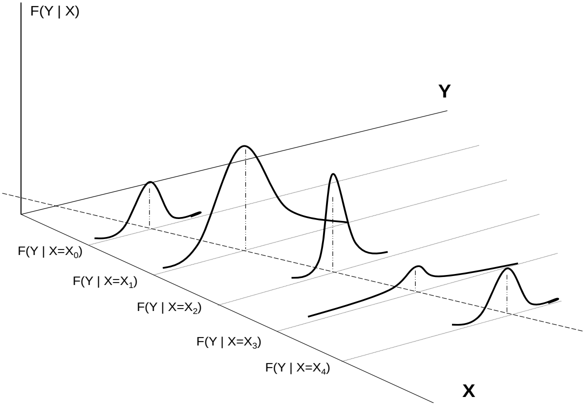
Heterocedasticidad.

Perturbaciones homocedásticas



Heterocedasticidad.

Perturbaciones heterocedásticas.



Concepto.

Heterocedasticidad

$$Var(u_i|X) = \sigma_i^2 \quad i = 1, 2, \ldots, N$$

La matriz de covarianzas de la perturbación es:

$$V_X(u) = E_X(uu') = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \ldots & 0 \\ 0 & \sigma_2^2 & 0 & \ldots & 0 \\ 0 & 0 & \sigma_3^2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & \sigma_N^2 \end{bmatrix} = \Sigma$$

Heterocedasticidad.

Propiedades del estimador MCO condicionado a X.

- Lineal: $\hat{\beta} = (X'X)^{-1}X'Y = \beta + (X'X)^{-1}X'u$
- Insesgado:

$$E_X(\hat{\beta}) = E_X(\beta + (X'X)^{-1}X'u) = \beta + (X'X)^{-1}X'E_X(u) = \beta$$

- NO es eficiente: : como no se satisface el supuesto básico S.4, no se cumple el Teorema de Gauss-Markov y el estimador MCO $\hat{\beta}$ NO tiene la varianza más pequeña dentro de los estimadores lineales e insesgados.

  Se puede demostrar que es posible obtener un estimador mejor que el MCO (en el sentido de tener una varianza menor) cuando se conoce la forma de la heterocedasticidad.

  $\rightarrow$ Estimador de Mínimos Cuadrados Generalizados (MCG): utiliza para estimar los coeficientes β la información de que se dispone sobre la matriz de covarianzas de u. El estudio de este estimador queda fuera del alcance de este libro.

Inferencia con el estimador MCO.

Si se cumplen los supuestos básicos del MRLG: $\hat{\beta} \sim N(\beta, \sigma^2(X'X)^{-1})$.

Si la matriz de covarianzas de la perturbación es $V_X(u) = \Sigma$ porque hay heterocedasticidad,

¿¿¿ cúal es la distribución del estimador MCO ???

$$\boxed{\hat{\beta} \sim N[\beta, (X'X)^{-1}X'\Sigma X(X'X)^{-1}]}$$

Demostración:

$$\begin{aligned} V_X(\hat{\beta}) &= E_X[\hat{\beta} - E(\hat{\beta})][\hat{\beta} - E(\hat{\beta})]' = E_X[\hat{\beta} - \beta)][\hat{\beta} - \beta)]' = \\ &= E_X[(X'X)^{-1}X'u][(X'X)^{-1}X'u]' = E_X[(X'X)^{-1}X'uu'X(X'X)^{-1}] = \\ &= (X'X)^{-1}X'E_X(uu')X(X'X)^{-1} = (X'X)^{-1}X'\Sigma X(X'X)^{-1} \end{aligned}$$

Inferencia con el estimador MCO.

La inferencia basada en los estadísticos de contrastes habituales t y F NO es adecuada y puede conducir a conclusiones incorrectas.

El problema es que el estimador habitual de la varianza del estimador MCO, $V_X(\hat{\beta}) = \hat{\sigma}^2(X'X)^{-1}$ no es adecuado porque está sesgado.

¿¿¿ Cúal es la distribución de los estadísticos t y F ???

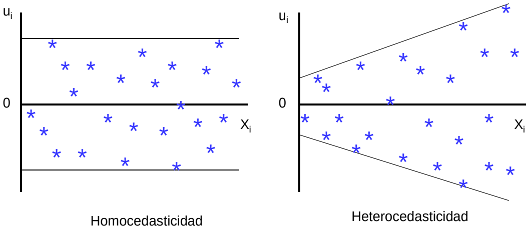
$$t = \frac{\hat{\beta}_j - \beta_j^0}{\hat{\sigma}_{\hat{\beta}_j}} \overset{H_0}{\sim} ??? \qquad F = \frac{(SCR_R - SCR_{NR})/q}{SCR_{NR}/(N-k)} \overset{H_0}{\sim} ???$$

Los estadísticos t y F no siguen las distribuciones habituales.

Detección de la heterocedasticidad.

A. Análisis gráfico.

El siguiente gráfico visualiza el comportamiento de la perturbación bajo los dos supuestos de homocedasticidad y heterocedasticidad.



Procedimiento:

- Estimar el modelo por MCO y obtener los residuos.
- Representar gráficamente los residuos MCO contra aquellos regresores que sospechemos pueden ser causantes de heterocedasticidad.

Detección de la heterocedasticidad.

B. Contrastes de heterocedasticidad.

Modelo de regresión lineal general:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \ldots + \beta_k X_{ki} + u_i \quad i = 1, 2, \ldots, N.$$

De los contrastes existentes para detectar la heterocedasticidad, explicaremos:

- Contraste de Goldfeld y Quandt.
- Contraste de Breusch y Pagan.
- Contraste de White.

El planteamiento de estos contrastes es siempre el siguiente:

$$H_0 : \text{Homocedasticidad:} \quad V(u_i) = \sigma_i^2 = \sigma^2 \quad \forall i$$
$$H_a : \text{Heterocedasticidad:} \quad V(u_i) = \sigma_i^2$$

Aplicación en el **Ejemplo 7.1.**

Contraste de Goldfeld-Quandt.

Este sencillo contraste se aplica cuando se cree que una única variable, usualmente uno de los regresores, es el causante de la heterocedasticidad, es decir, supongamos que la heterocedasticidad toma la forma: $\sigma_i^2 = E(u_i)^2 = h(Z)$ siendo creciente (decreciente) en Z.

Procedimiento:

1. Identificar la variable Z (usualmente uno de los regresores) con la que σ_i^2 está relacionadad.
2. Reordenar las observaciones según Z de forma creciente (decreciente).
3. Divide la muestra en tres partes: las N_1 primeras observaciones, las N_2 últimas observaciones y elimina las observaciones centrales. Se suelen eliminar alrededor de un tercio del total las observaciones.
4. Estimar separadamente por MCO dos regresiones para las N_1 primeras y las N_2 últimas observaciones.
5. Estadístico de contraste: $$GQ = \frac{SCR_2/(N_2 - k_2)}{SCR_1/(N_1 - k_1)} \overset{H_0}{\sim} \mathcal{F}((N_2 - k_2), (N_1 - k_1))$$

   donde SCR_1 y SCR_2 son, respectivamente, la suma de cuadrados de residuos de la primera y segunda regresión.
6. Se rechaza la hipótesis nula de homocedasticidad a un nivel de significación de $\alpha\,\%$ si:

$$GQ > \mathcal{F}_\alpha((N_2 - k_2), (N_1 - k_1))$$

Contraste de Breusch-Pagan.

Supone que la heterocedasticidad toma la siguiente forma: $\sigma_i^2 = E(u_i)^2 = h(z_i' \alpha)$,

- $z_i' = [1 \; z_{1i}, \ldots, z_{pi}]$ es un vector de variables conocidas,
- $\alpha = [\alpha_0 \; \alpha_1, \ldots, \alpha_p]$ es un vector de coeficientes desconocidos y
- $h(\cdot)$ es alguna función que debe tomar sólo valores positivos.

La hipótesis nula de homocedasticidad es: $H_0 : \alpha_1 = \alpha_3 = \ldots = \alpha_p = 0$

Procedimiento:

1. Estima el modelo por MCO y obtén los residuos $\hat{u}$ así como la estimación de la varianza de la perturbación $\tilde{\sigma}^2 = \sum \hat{u}^2/N$.
2. Regresa $\hat{u}_i^2/\tilde{\sigma}^2$ sobre z_i' por MCO y calcula la suma de cuadrados explicada (SCE).
3. Estadístico de contraste: $LM = \frac{1}{2} SCE \overset{H_0,a}{\sim} \chi^2(p)$
4. La hipótesis nula de homocedasticidad se rechaza a un nivel de significación de α % si:

$$LM = SCE/2 > \chi^2_\alpha(p)$$

Contraste de White.

Este contraste es muy flexible porque no es necesario hacer ningún supuesto sobre la estructura específica de la heterocedasticidad. En este sentido se dice que es un contraste robusto.

Procedimiento:

1. Estima el modelo por MCO y obtén los residuos $\hat{u}$.
2. Regresa $\hat{u}_i^2$, sobre el conjunto original de regresores, sus cuadrados y sus productos cruzados. Obtén el coeficiente de determinación de esta regresión auxiliar, R^2.
3. Estadístico de contraste:

$$LM = NR^2 \overset{H_0, a}{\sim} \chi^2(p)$$

   donde R^2 es el coeficiente de determinación de la regresión auxiliar, N es el tamaño muestral, p es el número de coeficientes estimados en la regresión auxiliar menos 1.

4. Se rechaza la hipótesis nula a un nivel de significación α % si: $NR^2 > \chi^2_\alpha(p)$.

Contenido

Concepto.

Autocorrelación : presencia de relación lineal entre las perturbaciones para distintas observaciones:

$$cov_X(u_t u_s) = E_X(u_t u_s) \neq 0 \quad \forall t, s \quad t \neq s$$

Se espera la presencia de autocorrelación cuando se trabaja con datos de series temporales.

La matriz de covarianzas de la perturbación es:

$$V_X(u) = E_X(uu') = \begin{bmatrix} \sigma^2 & \sigma_{12} & \sigma_{13} & \dots & \sigma_{1T} \\ \sigma_{21} & \sigma^2 & \sigma_{23} & \dots & \sigma_{2T} \\ \sigma_{31} & \sigma_{32} & \sigma^2 & \dots & \sigma_{3T} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{T1} & \sigma_{T2} & \sigma_{T3} & \dots & \sigma^2 \end{bmatrix} = \Sigma$$

Autocorrelación.

Consecuencias de la autocorrelación.

A. Propiedades de los estimadores MCO condicionados a X:

El estimador MCO es lineal, insesgado, pero NO es eficiente, en el sentido de varianza mínima, porque no se cumple el teorema de Gauss-Markov.

Existe otro estimador mejor que el MCO, es decir, con menor varianza: el estimador de Mínimos Cuadrados Generalizados.

B. Inferencia con el estimador MCO.

Cuando el supuesto S.5 no se cumple y hay autocorrelación la verdadera matriz de covarianzas del estimador MCO es:

$$V_X(\hat{\beta}) = (X'X)^{-1} X' \Sigma X (X'X)^{-1}$$

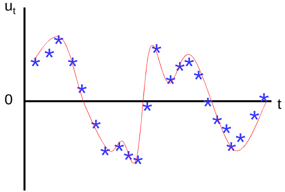
En consecuencia no se conoce la distribución de los estadísticos de contraste habituales t y F y la inferencia basada en estos estadísticos no es adecuada y puede conducir a conclusiones incorrectas.

Detección de la autocorrelación.

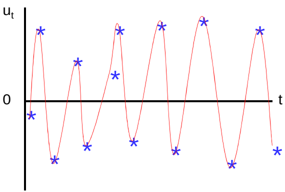
A. Análisis gráfico

El siguiente gráfico visualiza la evolución temporal de las perturbaciones cuando hay:

- Autocorrelación positiva: largas rachas de perturbaciones positivas y negativas.
- Autocorrelación negativa: las perturbaciones alternan signo positivo y negativo.



Autocorrelación positiva



Autocorrelación negativa

Procedimiento: estimar el modelo por MCO, obtener los residuos $\hat{u}$ y hacer un gráfico de series temporales de los mismos, para analizar si presentan algún tipo de regularidad en su comportamiento lo que sería síntoma de autocorrelación.

Detección de la autorrelación.

B. Contrastes de autocorrelación.

MRLG: $Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \ldots + \beta_k X_{kt} + u_t \quad t = 1, 2, \ldots, T.$

El planteamiento de los contrastes de autocorrelación existentes es:

H_0 : Ausencia de autocorrelación $\quad (cov(u_t u_s) = E(u_t u_s) = 0, \quad t \neq s)$

H_a : Autocorrelación $\quad (cov(u_t u_s) = E(u_t u_s) \neq 0, \quad t \neq s)$

Ahora bien, en cada uno de los contrastes es necesario especificar la estructura específica de la autocorrelación.
De los contrastes existentes para detectar la autocorrelación, vamos a presentar:

1. Contraste de Durbin-Watson.
2. Contraste de Breusch-Godfrey.

Aplicación en el **Ejemplo 7.2.**

Autocorrelación.

Contraste de Durbin-Watson.

En este contraste la estructura específica que se supone para la autocorrelación es un proceso autorregresivo de orden 1:

AR(1): $\quad u_t = \rho\, u_{t-1} + v_t \quad v_t \sim NID(0, \sigma_v^2) \quad |\rho| < 1$

La existencia de autocorrelación depende, por lo tanto, de ρ.

$$\begin{cases} \rho = 0 : & \text{No hay autocorrelación} \\ \rho > 0 : & \text{Autocorrelación positiva} \\ \rho < 0 : & \text{Autocorrelación negativa} \end{cases}$$

Contraste de Durbin-Watson. Autocorrelación positiva.

Procedimiento:

1. $$\begin{cases} H_0 : \rho = 0 & \text{(No hay autocorrelación)} \\ H_a : u_t \; = \; \rho\, u_{t-1} + v_t \quad \rho > 0 & \text{(Autocorrelación positiva)} \end{cases}$$

2. Estadístico de contraste $$DW = \frac{\sum_{t=2}^{T} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{T} \hat{u}_t^2} \overset{H_0}{\sim} ???$$

   donde $\hat{u}_t$ son los residuos MCO del modelo de regresión original.

   Se puede demostrar que: $DW \approx 2\,(1 - \hat{\rho})$ donde: $$\hat{\rho} = \frac{\sum_{t=2}^{T} \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^{T} \hat{u}_{t-1}^2}$$

Contraste de Durbin-Watson. Autocorrelación positiva.

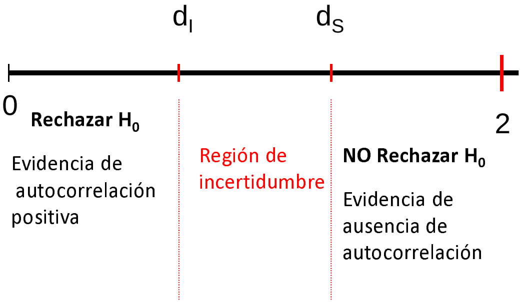
Por lo tanto: $$\begin{cases} \hat{\rho} = 0 \ \rightarrow \ DW \approx 2 \\ \hat{\rho} \in (0,1] \ \rightarrow \ DW \in [0,2) \end{cases}$$

3. Regla de decisión:

Durbin y Watson no derivaron la distribución exacta del estadístico DW bajo la hipótesis nula porque esta distribución depende de la muestra. Pero sí estimaron los valores críticos dependiendo del tamaño muestral y del número de regresores incluidos en el modelo original para un nivel de significación dado. Sólo estimaron un límite superior (d_S) y un límite inferior (d_I).

- Si $DW < d_I$, se rechaza la hipótesis nula: hay evidencia en la muestra a favor de la hipótesis alternativa de autocorrelación positiva.
- Si $DW > d_S$, no se rechaza la hipótesis nula: no hay evidencia en la muestra a favor de la hipótesis alternativa de autocorrelación positiva.
- Si $d_I < DW < d_S$, estamos en la región de incertidumbre.

Contraste de Durbin-Watson.



Contraste de Durbin-Watson. Autocorrelación negativa.

Procedimiento:

1. $$\begin{cases} H_0 : \rho = 0 & \text{(No hay autocorrelación)} \\ H_a : u_t = \rho\, u_{t-1} + v_t \quad \rho < 0 & \text{(Autocorrelación negativa)} \end{cases}$$

2. Estadístico de contraste $$DW = \frac{\sum_{t=2}^{T} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{T} \hat{u}_t^2} \overset{H_0}{\sim} ???$$

   Se puede demostrar que: $DW \approx 2\,(1 - \hat{\rho})$

   Por lo tanto: $$\begin{cases} \hat{\rho} = 0 \;\rightarrow\; DW \approx 2 \\ \hat{\rho} \in (0, 1] \;\rightarrow\; DW \in (2, 4] \end{cases}$$

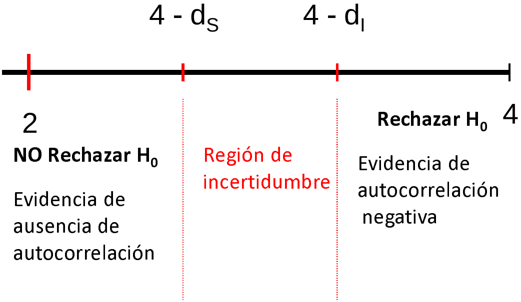
Contraste de Durbin-Watson. Autocorrelación negativa.

Durbin-Watson estimaron los valores críticos dependiendo del tamaño muestral y del número de regresores incluidos en el modelo original para un nivel de significación dado. Pero sólo estimaron un límite superior (d_S) y un límite inferior (d_I).

- Si $DW > 4 - d_I$, se rechaza la hipótesis nula: hay evidencia en la muestra a favor de la hipótesis alternativa de autocorrelación negativa.
- Si $2 < DW < 4 - d_S$, no se rechaza la hipótesis nula: no hay evidencia en la muestra a favor de la hipótesis alternativa de autocorrelación negativa.
- Si $4 - d_S < DW < 4 - d_I$, estamos en la región de incertidumbre.

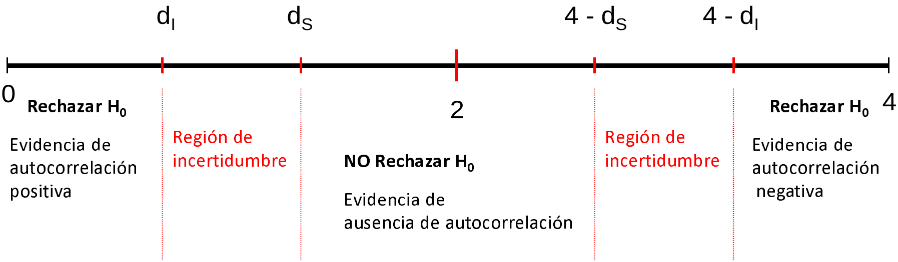
Autocorrelación.

Contraste de Durbin-Watson. Autocorrelación negativa.



Autocorrelación.

Contraste de Durbin-Watson.



Autocorrelación.

Contraste de Breusch-Godfrey.

Este contraste esta diseñado para contrastar la presencia de autocorrelación de orden más alto. Supongamos que la perturbación sigue un proceso autorregresivo de orden q:

$$\begin{cases} H_0 : \rho_1 = \rho_2 = \ldots = \rho_q = 0 \quad \text{(No hay autocorrelación)} \\ \\ H_a : \text{Autocorrelación de orden q} \end{cases}$$

Procedimiento:

1. Estima por MCO el modelo original y obtén los residuos MCO, $\hat{u}$.
2. Estima la regresión auxiliar de $\hat{u}_t$ sobre: $X_{2t}, X_{3t}, \ldots, X_{kt}, \hat{u}_{t-1}, \hat{u}_{t-2}, \ldots, \hat{u}_{t-q}$
3. Contrastar la significación conjunta de $\hat{u}_{t-1}, \hat{u}_{t-2}, \ldots, \hat{u}_{t-q}$ mediante el estadístico de contraste del multiplicador de Lagrange:

$$LM = T\,R^2 \overset{H_0,a}{\sim} \chi^2(q)$$

donde R^2 es el coeficiente de determinación de la regresión auxiliar. Se rechaza la hipótesis nula a un nivel de significación α % si: $LM > \chi^2_\alpha(q)$.

Contenido

Inferencia con el estimador MCO.

¿Existe algún procedimiento para poder hacer inferencia utilizando el estimador MCO?

$$\boxed{\hat{\beta} \sim N[\beta, (X'X)^{-1}X'\Sigma X(X'X)^{-1}]}$$

Idea : Estimar la verdadera matriz de covarianzas del estimador MCO.

$$V_X(\hat{\beta}) = (X'X)^{-1}X'\Sigma X(X'X)^{-1}$$

Esto no es fácil porque habitualmente no se sabe mucho sobre la matriz Σ .

Algunos económetras $\left\{ \begin{array}{l} \text{White (heterocedasticidad)} \\ \text{Newey-West (autocorrelación)} \end{array} \right\}$ han derivado expresiones para estimar la matriz $V_X(\hat{\beta})$ con buenas propiedades para muestras grandes.

Inferencia con el estimador MCO.

Denotemos por $\widehat{V}_X^R(\hat{\beta})$ al estimador de la verdadera matriz de covarianzas del estimador MCO de β. Se le denomina también: estimador robusto a la heterocedasticidad y/o a la autocorrelación.
Se puede demostrar que:

$$t = \frac{\hat{\beta}_j - \beta_j^0}{\hat{\sigma}^R_{\hat{\beta}_j}} \quad \overset{H_0,a}{\sim} \quad N(0,1)$$

donde $\hat{\sigma}^R_{\hat{\beta}_j}$ es la desviación típica robusta a la heterocedasticidad y/o a la autocorrelación. Se rechaza la hipótesis nula a un nivel de significación α % si: $|t| > N_{\alpha/2}(0,1)$.
Para contrastar restricciones generales el estadístico habitual debe multiplicarse por el número de restricciones pra lograr así un estadístico $q \times F$ que se distribuirá como una $\chi^2(q)$ en muestras grandes. Se rechaza la hipótesis nula a un nivel de significación α % si: $q \times F \quad > \quad \chi^2_\alpha(q)$.

Contenido

Contenido