Problemas con el término de error

V. PROBLEMAS CON EL TÉRMINO DE ERROR

Objetivos

1. Detectar los incumplimientos de las hipótesis básicas acerca del término de perturbación aleatoria.
2. Evaluar y juzgar qué ocurre cuando no se cumple alguna de las hipótesis básicas acerca del término de perturbación aleatoria.
3. Conocer cuáles son las causas de los incumplimientos.
4. Conocer qué se puede hacer para solucionar la violación de una de las hipótesis básicas sobre el término de perturbación aleatoria.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

Índice

V. PROBLEMAS CON EL TÉRMINO DE ERROR

Índice (Cont.)

V. PROBLEMAS CON EL TÉRMINO DE ERROR

1. Introducción

- $E[e_i] = 0$ (media nula de los errores)
- $Var(e_i) = \sigma^2 \ \forall i$ (homoscedasticidad)
- $Cov(e_i, e_j) = 0$ para $i \neq j$ (ausencia de correlación)
- $e_i \sim N$ (la perturbación aleatoria e sigue una distribución normal)

- Las **tres primeras hipótesis** se resumen diciendo que las **perturbaciones son *esféricas***
- En conjunto, todas las propiedades implican que las perturbaciones son:
  - **normales independientes (*i.i.d.*) con parámetros 0 y σ^2, $e_i \sim N(0,\sigma^2)$**
  - **en forma matricial, $e \sim N(0,\sigma^2 I_T)$).**

En este tema veremos:

- *qué ocurre cuando no se cumple alguna de las hipótesis citadas,*
- *cuáles son las causas de dicho incumplimiento y, finalmente,*
- *qué se puede hacer cuando se produce.*

V. PROBLEMAS CON EL TÉRMINO DE ERROR

2. No normalidad de los errores y estimación robusta

2.1. Distribución no normal de los errores

- Estimador MCO de β: lineal, insesgado y de mínima varianza entre los estimadores lineales e insesgados (teorema de Gauss-Markov)
  (sin utilizar la hipótesis de normalidad del vector de errores e).
- Para obtener su distribución ➔ normalidad de las perturbaciones

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

Permite:

a) demostrar su eficiencia y

b) realizar inferencias sobre el vector β y sus componentes β_i, i=1,…,k

<u>Cuando *e* no es normal</u>

- El estimador MCO de β *deja de ser eficiente* y no pueden, en principio, realizarse inferencias por desconocerse su distribución exacta.
- Los intervalos de confianza y los contrastes t y F dejan de ser válidos.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

2.1. Distribución no normal de los errores (cont.)

¿Satisface el modelo la hipótesis inicial de normalidad de los errores?

Test de Jarque-Bera
H_0: {e sigue una distribución normal}
H_1: {e sigue una distribución de la familia Pearson no normal}

$$JB = (n-k)\left(\frac{A^2}{6} + \frac{(K-3)^2}{24}\right) \xrightarrow[as]{H_0} \chi^2_2$$

A: coeficiente de asimetría
K: coeficiente de curtosis (curvatura)

- Rechazar H_0 ➔ resultados de contrastes t y F de la estimación MCO no son válidos
- Puede demostrarse que en condiciones muy generales se verifica

$$\frac{\hat{\beta}_i - \beta_i}{s(\hat{\beta}_i)} \xrightarrow[as]{} N(0,1)$$

a) **Con un tamaño de muestra suficientemente grande, puede seguir utilizándose el estadístico *t***, utilizando las tablas de la distribución Normal para contrastar hipótesis sobre los parámetros individuales.

b) **En lugar del test *F*, puede utilizarse el test de Wald**, válido asintóticamente, para contrastar hipótesis de restricciones lineales.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

2.2. Estimación robusta

Cuando la perturbación de un modelo de regresión **no se distribuye normalmente** (pero al menos tiene varianza finita)

- **el estimador MCO de los parámetros β** es aún el mejor estimador lineal insesgado **(MELI) y** es **consistente**, pero no necesariamente eficiente ni asintóticamente eficiente.
- **el estimador propuesto para σ^2** es **insesgado** bajo errores no normales, y además es **consistente**.

Por otro lado (**teorema central del límite**) los estimadores MCO de β son asintóticamente normales y los intervalos de confianza y los contrastes de hipótesis siguen siendo asintóticamente válidos aún si los errores no son normales.

Conclusión

Todos los resultados del modelo de regresión lineal general valen de forma aproximada para muestras de tamaño elevado sea cual sea la distribución del término de perturbación.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

4. Matriz de covarianzas no escalar

- ***Cov(e)*** **≠ $\sigma^2 I_n$** (Propiedades del estimador de MCO pueden no ser válidas). Tales situaciones ocurren cuando se da alguna (o ambas a la vez) de las condiciones siguientes:

1) $Var(e_i/X_i) = Var(e_i) = \sigma_i^2$, existiendo $i \neq j$ tal que $\sigma_i^2 \neq \sigma_j^2$.

Situación conocida como ***heteroscedasticidad*** en los errores.

- Frecuente en datos de corte transversal.
- Matriz *Cov*(*e*) de covarianzas:

$$Cov(e)=\Sigma=\begin{bmatrix}\sigma_1^2 & 0 & \dots & 0\\ 0 & \sigma_2^2 & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \sigma_n^2\end{bmatrix}=\sigma_1^2\begin{bmatrix}1 & 0 & \dots & 0\\ 0 & \sigma_2^2/\sigma_1^2 & \dots & 0\\ \dots & \dots & \dots & \dots\\ 0 & 0 & \dots & \sigma_n^2/\sigma_1^2\end{bmatrix}=\sigma^2\Omega$$

2) Covarianzas no nulas entre las distribuciones condicionadas de los errores;

$\exists\, t \neq t'$ con $Cov(e_t, e_{t'}) \neq 0$. Se dirá que existe ***autocorrelación*** en los errores.

- Frecuente con datos de series temporales.
- La matriz *Cov*(*e*) toma la forma general

$$Cov(e)=\Sigma=\sigma^2\begin{bmatrix}1 & \rho_1 & \dots & \rho_{T-1}\\ \rho_1 & 1 & \dots & \rho_{T-2}\\ \dots & \dots & \dots & \dots\\ \rho_{T-1} & \rho_{T-2} & \dots & 1\end{bmatrix}=\sigma^2\Omega$$

V. PROBLEMAS CON EL TÉRMINO DE ERROR

4.1. Propiedades del estimador MCO de β cuando $Cov(e) = \Sigma = \sigma^2\Omega$.

- El estimador de MCO **sigue siendo insesgado**
- La **matriz de covarianzas ya no es $\sigma^2(X'X)^{-1}$** su expresión viene dada por

$$Cov(\hat\beta)=E\left[(\beta-\hat\beta)(\beta-\hat\beta)'\right]=$$
$$=E\left[(X'X)^{-1}X'ee'X(X'X)^{-1}\right]=$$
$$=\sigma^2(X'X)^{-1}X'\Omega X(X'X)^{-1}$$

- Por tanto, si existe heteroscedasticidad y/o autocorrelación utilizar **$\sigma^2(X'X)^{-1}$** (o su estimación), como matriz de covarianzas del estimador MCO,, es incorrecto, pudiendo llevar su utilización a inferencias o predicciones erróneas.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

4.2. El estimador de mínimos cuadrados generalizados, MCG

- Sea $y = X\beta + e$, con $Cov(e) = \Sigma = \sigma^2\Omega$,

(siendo válidas el resto de las hipótesis del modelo de regresión lineal general)

- Para obtener un estimador adecuado en estas circunstancias ➔

se transforma el modelo de forma que se satisfagan las hipótesis MCO

- Como Ω es simétrica y definida-positiva, existe una matriz cuadrada no singular, V, tal que $\Omega = VV'$, y por tanto,

$$V^{-1}\Omega(V^{-1})' = I_T \quad \text{y} \quad \Omega^{-1} = (V^{-1})'V^{-1} = P'P, \text{ donde } P = V^{-1}.$$

- Transformando el modelo original (premultiplicando por la matriz P),

$$y^* = Py, \quad X^* = PX, \quad e^* = Pe,$$

se llega al nuevo modelo $y^* = X^*\beta + e^*$, cumpliéndose:

$$Cov(e^*) = Cov(Pe) = E(Pee'P') = PE(ee')P' = P\Sigma P' = \sigma^2(V^{-1})\Omega(V^{-1}) = \sigma^2 I$$

El modelo transformado satisface las hipótesis MCO

V. PROBLEMAS CON EL TÉRMINO DE ERROR

Estimadores para β y su matriz de covarianzas

- El modelo transformado satisface las hipótesis MCO, luego el estimador MCO del mismo, conocido como ***estimador de MCG o de Aitken,*** viene dado por

$$\hat{\beta}_{MCG} = (X^{*\prime}X^*)^{-1}X^{*\prime}y^* = (X'P'PX)^{-1}X'P'Py = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$$

Estimador lineal insesgado y de mínima varianza de β
(pueden aplicarse los resultados estándar al modelo transformado).

- Para estimar la matriz de covarianzas de hay que tener en cuenta que:

$$Cov(\hat{\beta}_{MCG}) = E\left[(\beta - \hat{\beta}_{MCG})(\beta - \hat{\beta}_{MCG})'\right] = \sigma^2(X^{*\prime}X^*)^{-1} = \sigma^2\left(X'\Omega^{-1}X\right)^{-1}$$

<u>Estimador de σ^2</u>

- Estimación MCG de la varianza residual: $\hat{\sigma}^2_{MCG} = \dfrac{\hat{e}^{*\prime}\hat{e}^*}{T-K} = \dfrac{\hat{e}'P'P\hat{e}}{T-K} = \dfrac{\hat{e}'\Omega^{-1}\hat{e}}{T-K}$

Estimador insesgado de σ^2

- Si las perturbaciones del modelo original son normales, se tiene que

$\hat{\beta}_{MCG} \approx N(\beta, \sigma^2(X'\Omega^{-1}X)^{-1})$ ➔Estimador lineal, insesgado y eficiente de β.

Se puede demostrar que la estimación de MV coincide con el estimador de MCG

Puede utilizarse el **test de razón de verosimilitudes** (para contrastar hipótesis sobre β)

V. PROBLEMAS CON EL TÉRMINO DE ERROR

Coeficiente de determinación

- **La interpretación del R^2** obtenido como medida de ajuste del modelo transformado **no es válida**, ya que no es aplicable a las variables originales del modelo.
- Por otra parte, como *el modelo transformado puede no tener término independiente*, ya *no está acotado necesariamente entre 0 y* 1.

- Una medida del ajuste podría basarse en los **residuos del modelo original** una vez que se ha obtenido el estimador MCG; es decir, se puede aplicar la fórmula original del R^2 con β substituido por $\hat{\beta}_{MCG}$
- Esta medida no está acotada en el intervalo unidad y, además, no puede ser usada de forma fiable para comparar modelos.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

4.3. Mínimos cuadrados generalizados factibles, MCGF.

- MCG requiere que la matriz Ω sea conocida ➔ se hace necesaria su estimación

Ω (matriz de covarianzas) ➔ es definida-positiva y simétrica

contiene $T(T+1)/2$ parámetros a estimar y, se dispone de T observaciones,

⬇

Estimar los parámetros estructurales β y la matriz Ω *no es factible.*

⬇

Se han de establecer algunas **hipótesis adicionales,**

o restricciones sobre los parámetros de Ω, que hagan posible la estimación.

- ***MCGF*** (Mínimos cuadrados generalizados factibles) parte de la hipótesis:
  - la matriz Ω es, en general, función de un pequeño número de parámetros;
  - es decir $\boldsymbol{\Omega = \Omega(\theta)}$, donde $\boldsymbol{\theta}$ es un vector formado por un nº reducido de parámetros.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

4.3. Mínimos cuadrados generalizados factibles, MCGF (cont.).

- *MCGF* parte de la hipótesis de que la matriz Ω es, en general, función de un pequeño número de parámetros; es decir $\boldsymbol{\Omega} = \boldsymbol{\Omega}(\boldsymbol{\theta})$, donde $\boldsymbol{\theta}$ es un vector formado por un número reducido de parámetros.

Por ejemplo, se tiene que

$$\Omega = \begin{bmatrix} 1 & \rho & \dots & \rho^{n-1} \\ \rho & 1 & \dots & \rho^{n-2} \\ \dots & \dots & \dots & \dots \\ \rho^{n-1} & \rho^{n-2} & \dots & 1 \end{bmatrix} = \Omega(\rho)$$

caso de autocorrelación de primer orden (aptdo. 6).

- Si $\hat{\theta}$ es un estimador consistente de θ, parece lógico usar $\hat{\Omega} = \Omega(\hat{\theta})$ en lugar de Ω en la expresión del estimador de MCG y de su matriz de covarianzas
- Se puede demostrar que *el estimador al que da lugar dicha substitución*, conocido como de ***MCGF***, *mantiene asintóticamente las buenas propiedades del estimador MCG* original, aunque dichas propiedades no están garantizadas si se trabaja con muestras de reducido tamaño.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

5. Heteroscedasticidad

- $y = X\beta + e$; Existe heteroscedasticidad si $Var(e_i) = \sigma_i^2$ (resto de hipótesis MCO válidas)

  $Cov(e) = \Sigma = \sigma^2\Omega = diag(\sigma_1^2,\dots, \sigma_n^2)$

- Habitual en el caso de **datos de corte transversal** (menos común en series temporales)
- **Puede generarse** (sobre todo con datos de series temporales) por una **especificación errónea** o por la **presencia de cambio estructural**. (por ej., la omisión de una variable explicativa relevante hace que su efecto sea asumido por el error y, por tanto, hace que éste varíe a lo largo de la muestra.

5.1. Consecuencias de la heteroscedasticidad

Conclusiones respecto a la validez del estimador MCO $\hat{\beta}$ en presencia de heterosced.

- Mantiene las propiedades de **insesgadez, consistencia y normalidad asintótica**.
- Las **estimaciones de las varianzas** son **sesgadas**, siendo probable que los **intervalos de confianza** se hagan **anormalmente grandes** y que los **coeficientes tienden a ser no significativos** por la mayor varianza y errores estándar estimados.
- Los **estimadores MCO** (corrigiendo adecuadamente la matriz de covarianzas) **no son eficientes** en relación a los de MCG.
- Resultados **inválidos** al usar **los estadísticos *t* y *F*** de la estimación MCO (si no se corrige adecuadamente la matriz de varianzas-covarianzas).

5.1. Consecuencias de la heteroscedasticidad (cont.)

Se hace necesaria la **utilización del estimador MCG** (aún cuando este método no agota la vía de soluciones al problema)

Necesidad de reducir el número de parámetros a estimar, $(\sigma_1,\ldots,\sigma_n)$ y β,

↓

establecer hipótesis adicionales sobre el posible comportamiento heteroscedástico de los errores.

Un comportamiento habitual de las varianzas es que su crecimiento está asociado a:

- alguna de las variables explicativas o regresores del modelo, o
- variables no incluidas en el mismo.

5.2. Detección de la heteroscedasticidad

Examen gráfico de los residuos MCO

Gráfico de los residuos MCO, $\hat{e}_i$, de su valor absoluto o de sus cuadrados, **respecto de cada una de las variables explicativas del modelo**.

- La identificación de modelos sistemáticos en dichas gráficas puede verse como evidencia de la violación de la hipótesis de homoscedasticidad
- Ejemplo: la evolución creciente de una gráfica puede deberse a un crecimiento de la variabilidad asociado a la correspondiente variable.

Diferentes *contrastes* bajo la hipótesis nula de homoscedasticidad

H_0: $\{\sigma_1^2 = \ldots = \sigma_n^2\}$, **frente a una alternativa específica** que depende:

- del procedimiento de estimación considerado y
- del patrón de comportamiento asumido para las varianzas de los errores.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

5.2. Detección de la heteroscedasticidad (cont.)

Contraste general de heteroscedasticidad de White (1980)

- Es un *contraste general* (es decir, existe heteroscedasticidad o no),

$$\boxed{H_0: \{\sigma_i^2 = \sigma^2, \forall i\} \quad vs. \quad H_1: \{\exists i \neq j / \sigma_i^2 = \sigma_j^2\}}$$

- Es *básicamente un test de mala especificación.*

  En esencia consiste en comparar las matrices de covarianzas del estimador MCO tanto bajo la hipótesis nula como bajo la alternativa:

$\hat{\sigma}^2(X'X)^{-1}$ es un *estimador inconsistente si hay heteroscedasticidad.*

$Cov(\hat{\beta}) = \sigma^2(X'X)^{-1}X'\Omega X(X'X)^{-1}$ *puede ser estimada consistentemente por*

$$C\hat{o}v(\hat{\beta}) = n(X'X)^{-1}S_0(X'X)^{-1}$$

Para ello:

1. Regresión de $\hat{e}_i^2$ frente a las variables originales, sus cuadrados y sus productos cruzados
2. $\boxed{nR^2 \xrightarrow{H_0} \chi_p^2}$

   - hipótesis nula de varianzas iguales,
   - p es el número de regresores de esta regresión auxiliar (sin incluir la constante).

V. PROBLEMAS CON EL TÉRMINO DE ERROR

Contraste general de heteroscedasticidad de White (1980) (cont.)

Ejemplo

1. Modelo $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + e_i$, siendo $\hat{e}_i$ los residuos de su estimación MCO

La regresión auxiliar que debe utilizarse viene dada por

$$\hat{e}_i^2 = \alpha_1 + \alpha_2 x_i + \alpha_3 z_i + \alpha_4 x_i^2 + \alpha_5 z_i^2 + \alpha_6 x_i z_i + u_i$$

2. Debe compararse el valor del estadístico nR^2 de esta regresión con el valor crítico de una variable χ_5^2

Problemas que presenta

- Si se rechaza la hipótesis de homoscedasticidad (nR^2 es significativo) no existe información sobre la naturaleza de la heteroscedasticidad.
- El contraste puede revelar heteroscedasticidad, pero también puede estar detectando alguna clase de error de especificación en su lugar (omisión de alguna variable relevante, cambio estructural, etc.).

V. PROBLEMAS CON EL TÉRMINO DE ERROR

Contraste de Goldfeld-Quandt (1965)

- Contraste muy habitual.
- Relación del tipo $\sigma_i^2 = f(z_i)$

z : puede ser una de las variables explicativas del modelo o una variable que no se encuentre en la regresión, y
f : función monótona creciente o decreciente, generalmente desconocida.

Se supone que **la relación entre la varianza y la variable *z* es positiva**.

1. Se ordenan las observaciones en orden creciente de z.
2. Se divide la muestra en dos grupos con n_1 y n_2 observaciones en cada grupo.
3. Se estima por MCO el modelo para cada uno de los dos grupos de observaciones, y se calculan la SCR de ambos modelos, dadas por $SCR1 = \hat{e}_1'\hat{e}_1$ y $SCR2 = \hat{e}_2'\hat{e}_2$.
4. Se calcula el cociente entre ambas sumas residuales: $$F = \frac{SCR2/(n_2-k)}{SCR1/(n_1-k)} = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2}$$

- **La hipótesis a contrastar es: $H_0:\{\sigma_1^2 = \sigma_2^2\}$ *vs.* $H_1:\{\sigma_1^2 < \sigma_2^2\}$.**
- Si hay heteroscedasticidad es de esperar $\sigma_2^2 > \sigma_1^2$, y, bajo la hipótesis nula, se cumple que $F \underset{H_0}{\longrightarrow} F_{n_2-k,n_1-k}$

- Si se rechaza la H_0 ($F > Fc(\alpha)$) se acepta el modelo heteroscedástico
- Si no se rechaza H_0 no se puede asegurar que no existe heteroscedasticidad, sino sólo que no existe heteroscedasticidad asociada a la variable *z*.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

Variante del contraste de Goldfeld-Quandt

- **Excluir *p* observaciones centrales**, tomándose entonces dos grupos de $m = (n-p)/2$ observaciones. Si se toma un **valor de *p* elevado** :
  - se pierden grados de libertad con lo cual las estimaciones son menos precisas y se pierde potencia en el contraste.
  - al estar las observaciones más separadas es más fácil detectar la heteroscedasticidad.
- **Algunos autores consideran óptimo tomar $p = n/3$**, (*n* número de observaciones). Puede adoptarse el principio de no eliminar más de la tercera parte de las observaciones, dependiendo *p* del número total de observaciones.

- Si se ha supuesto que $\sigma_i^2=\sigma^2 z_i^2$, la matriz de covarianzas de los errores vendrá dada por $Cov(e) = \Sigma = \sigma^2\Omega = \sigma^2 diag(z_1^2,\ldots,z_n^2)$, y la solución de MCGF consistiría en transformar el modelo original dividiendo las observaciones originales por z_i.
- En este caso, el modelo resultante sería $y^* = X^*b + e^*$, donde $y_i^* = y_i/z_i$, $x_{1i}^*=1/z_i,\ldots,x_{ki}^*=x_{ki}/z_i$, $e_i^* = e_i/z_i$, el cual satisface la hipótesis de MCO [el modelo transformado no tiene término independiente].
- Este método es conocido como **MC ponderados**, siendo los inversos de los valores z_i las correspondientes ponderaciones.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

Contraste de Breusch-Pagan (1979) de multiplicadores de Lagrange

- H_0:$\{\sigma_i^2 = f(\alpha_0) = \sigma^2\}$ frente a H_1:$\{\sigma_i^2 = f(\alpha_0 + \alpha_1 z_{1i} + ... + \alpha_p z_{pi})\}$,
  - *f* desconocida y
  - *z* un vector de variables exógenas o independientes que causan la heteroscedasticidad en **forma aditiva**.

1. Se estima el modelo original por MCO y se calculan los errores estimados $\hat{e}_i$
2. Regresión auxiliar $\hat{e}_i^2 = \alpha_0 + \alpha_1 z_{1i} + ... + \alpha_p z_{pi} + u_i$
3. Se calcula el valor del estadístico $\lambda = nR^2$.

$$\boxed{\lambda \underset{\text{asint}}{\overset{H_0}{\to}} \chi_p^2}$$

p es el número de variables en la regresión auxiliar (sin incluir el término constante).

- Contraste de B-P es muy sensible a la hipótesis de normalidad de los errores

V. PROBLEMAS CON EL TÉRMINO DE ERROR

5.3. Corrección de la heteroscedasticidad

5.3.1. Solución general de White (MCO corregidos).

- *Tomar el estimador MCO para β y ajustar su matriz de covarianzas.*

$$\Sigma = \frac{1}{n}\sigma^2 X'\Omega X = \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 x_i x'_i$$ se estima consistentemente por $$S_0 = \frac{1}{n-k}\sum_{i=1}^{n}\hat{e}_i^2 x_i x'_i$$

$$Cov(\hat{\beta}) = \sigma^2 (X'X)^{-1} X'\Omega X (X'X)^{-1}$$ puede estimarse por $$C\hat{o}v(\hat{\beta}) = n(X'X)^{-1} S_0 (X'X)^{-1}$$

Utilizando el estimador de White

- los **contrastes habituales de restricciones lineales** y, en particular, el de significación conjunta de las variables, **no pueden realizarse con el test *F*** (dado que éste se basa en la hipótesis de homoscedasticidad).
- Por tratarse de un estimador con propiedades asintóticas debe utilizarse el test de Wald, basado en el estadístico

$$W = (R\hat{\beta})'\{RC\hat{o}v(\hat{\beta})R'\}^{-1}(R\hat{\beta}) \underset{\text{asint}}{\overset{H_0:\{R\beta=0\}}{\to}} \chi_q^2$$

q: número de restricciones.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

5.3. Corrección de la heteroscedasticidad (Cont.)

5.3.2. Mínimos cuadrados ponderados (MCP).

Si $y = \mathbf{X}b + e$, y $\exists$ heteroscedasticidad ($Var(e_i)=\sigma_i^2= \sigma^2 w_i$), (resto de hipótesis válidas)

$$Cov(e) = \Sigma = \sigma^2\Omega = diag(\sigma_1^2,\dots,\sigma_n^2) = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sigma_n^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} w_1 & 0 & \dots & 0 \\ 0 & w_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & w_n \end{bmatrix}$$

- El estimador de *MCG* se obtiene dividiendo cada una de las observaciones del modelo original por el correspondiente w_i.

Matriz de transformación del modelo → $P = diag(1/\sqrt{w_1},\dots,1/\sqrt{w_n})$

- Estimador obtenido → ***mínimos cuadrados ponderados (MCP)***

(asigna a cada observación con menor varianza una mayor ponderación).

Ejemplo → función de consumo $C_i = \beta_1+\beta_2 Y_i+e_i$, y $Var(e_i) = \sigma_i^2 = \sigma^2\ Y^2_i$ *(proporcional al cuadrado de la renta),*

$$\frac{C_i}{Y_i} = \beta_1\frac{1}{Y_i} + \beta_2 + \frac{e_i}{Y_i}$$

← Modelo transformado

→ Estimaciones MCO: estimaciones MCP del modelo original, (ponderaciones dadas por $p_i = 1/Y_i$)

V. PROBLEMAS CON EL TÉRMINO DE ERROR

6. Autocorrelación. 6.1. Naturaleza y consecuencias

- Frecuente con **datos de series temporales**
- Se origina generalmente por una **mala especificación de la dinámica de interacción** entre los errores. También puede deberse a situaciones como:
  - Naturaleza inercial del comportamiento de los agentes económicos.
  - Especificación errónea del modelo de la variable dependiente, y_t, y los regresores x_t. (omisión de variables relevantes, sobre todo).
- Semejanzas de comportamiento (proximidad espacial,...)
- 'Manipulación' de datos (agregación, desestacionalización, eliminación de la tendencia), etc.

<u>**Consecuencias**</u>

Estimadores MCO **insesgados**, pero **los errores estándar MCO estarán sesgados** (cualquier inferencia que se haga sobre la base de dichas estimaciones será inválida)

$y_t = \beta_1 + \beta_2 x_{2t} + \dots + \beta_k\ x_{kt} + e_t$, y supongamos que los errores e_t están serialmente correlacionados, es decir, $\boldsymbol{Cov(e_t, e_{t'}) \neq 0}$ para $t \neq t'$.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

6. Autocorrelación. 6.1. Naturaleza y consecuencias

- **Autocovarianzas** ➔ $\gamma_s = E[e_t e_{t+s}]$ para $s = 0, \pm 1, \pm 2, \ldots$ (*cuando* $s = 0$, $\gamma_0 = \sigma^2$).

- **Coeficiente de autocorrelación** ➔ (suponiendo homoscedasticidad)

$$\rho_s = \frac{Cov(e_t e_{t+s})}{\sqrt{Var(e_t)Var(e_{t+s})}} = \frac{\gamma_s}{\gamma_0}$$

$$Cov(e) = \Sigma = \sigma^2\Omega = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{T-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{T-2} \\ \cdots & \cdots & \cdots & \cdots \\ \gamma_{T-1} & \gamma_{T-2} & \cdots & \gamma_0 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{T-1} \\ \rho_1 & 1 & \cdots & \rho_{T-2} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{T-1} & \rho_{T-2} & \cdots & 1 \end{bmatrix}$$

Si no se conoce *Cov(e)* ➔ estimar *T*+*k* parámetros (*k* de la regresión y *T* de la matriz de covarianzas) **con sólo *T* observaciones**

Establecer alguna hipótesis adicional sobre la naturaleza de la autocorrelación que se base en un pequeño número de parámetros.

- **Procesos que habitualmente describen el comportamiento de errores autocorrelacionados ➔ ARMA(*p*,*q*)**, (autorregresivos y de media móvil)

  (incluyen los modelos autorregresivos de orden *p*, **AR(*p*)**, y los de medias móviles **MA(*q*)**, como casos particulares).

V. PROBLEMAS CON EL TÉRMINO DE ERROR

6.1. Naturaleza y consecuencias (Cont.)

Este Tema se centra básicamente en el análisis de procesos AR(1): $e_t = \rho e_{t-1} + \varepsilon_t$

(muchos de los resultados son fácilmente extrapolables al caso general de modelos AR(*p*))

6.2. Detección de la autocorrelación

Examen gráfico de los residuos MCO

Representación gráfica frente al tiempo.

- **Indicios de autocorrelación positiva de orden uno:** residuos MCO agrupados en rachas de un signo de un signo determinado (positivos o negativos), seguidos de rachas de signo contrario.
- **Indicios de autocorrelación negativa de orden uno:** los residuos MCO van alternando su signo con respecto a la observación anterior.

Representación de $\hat{e}_t$ frente a $\hat{e}_{t-1}$.

- **Indicio de correlación positiva** en el tiempo de los errores ➔ agrupación de los puntos del gráfico alrededor de una línea creciente sería
- **Síntoma de autocorrelación negativa** ➔ agrupación alrededor de una recta con pendiente negativa.

V. PROBLEMAS CON EL TÉRMINO DE ERROR

6.2. Detección de la autocorrelación. Contraste de Durbin-Watson (1950, 1951, 1971)

- **Sólo** contrasta procesos autorregresivos de primer orden **-AR(1)-**

→ $e_t = \rho\, e_{t-1} + e_t$ $\quad H_0: \{\rho = 0\}$ *vs* $H_1: \{\rho \neq 0\}$

$$DW = d = \frac{\sum_{t=2}^{n} (\hat{e}_t - \hat{e}_{t-1})^2}{\sum_{t=1}^{n} \hat{e}_t^2} \approx 2(1-\hat{\rho}) \quad \rightarrow \quad \hat{\rho} = \sum_t \hat{e}_t \hat{e}_{t-1} \Big/ \sum \hat{e}_t^2$$

Estimación de ρ basada en la estimación MCO de la ecuación $\hat{e}_t = \rho\,\hat{e}_{t-1} + u_t$

- El estadístico ***d*** tiene recorrido entre 0 y 4.
- Si ***d<2 (d>2)*** →indicativo de ***autocorrelación positiva (negativa)*** de los errores e_t.
- Si ***d≈2*** → ***ausencia de correlación***.

d depende de la matriz de datos **X** (y también de ***n*** y de ***k***) → no existen tablas de valores críticos para el test exacto *d*.

Durbin y Watson → test aproximado, tabulando los puntos críticos del contraste para diferentes valores de *n* y *k*.

Dado, *a*, se ha de comparar *d* con d_L y d_U [(para *T* observaciones y *k'* variables explicativas (no se incluye el término independiente)]

V. PROBLEMAS CON EL TÉRMINO DE ERROR

6.2. Detección de la autocorrelación. Contraste de Durbin-Watson (Cont.)

- **Autocorrelación positiva**: $H_0: \{\rho = 0\}$ *vs.* $H_1: \{\rho > 0\}$
  - Si $\boldsymbol{d < d_L}$ se rechaza H_0
  - Si $\boldsymbol{d > d_U}$ no se rechaza H_0
  - Si $\boldsymbol{d_L < d < d_U}$ no se puede extraer ninguna conclusión (***zona de indecisión***)
- **Autocorrelación negativa:** $H_0: \{\rho = 0\}$ *vs.* $H_1: \{\rho < 0\}$
  - Si $\boldsymbol{d > 4-d_L}$ se rechaza H_0
  - Si $\boldsymbol{d < 4-d_U}$ no se rechaza H_0
  - Si $\boldsymbol{4-d_U < d < 4-d_L}$ no se puede decidir

$\boldsymbol{d_L}$ y $\boldsymbol{d_U}$ están calculadas bajo los supuestos:

1) **Existe término independiente** en la regresión
2) **No** se incluye entre las variables explicativas a la variable **dependiente retardada**

En ambas situaciones no es aplicable el contraste D-W

V. PROBLEMAS CON EL TÉRMINO DE ERROR

Detección de la autocorrelación. Test *h* de Durbin (1970)

- Si la **variable dependiente retardada** aparece entre los regresores

→ *d* sesgado hacia 2

Estadístico *h* de Durbin ← Utilizar otros contrastes

$$h = \left(1 - \frac{d}{2}\right)\sqrt{\frac{T}{1 - T\,\hat{Var}(\hat{\alpha})}} \xrightarrow{H_0:\{\rho=0\}} N(0,1)$$

Varianza estimada del coeficiente de y_{t-1} en $y_t = \alpha\, y_{t-1} + x_t'\beta + e_t$

Si $\hat{Var}(\hat{\alpha}) > 1/T$ ⇒ El test ***h*** no puede calcularse

alternativa → Test de Breusch-Godfrey

V. PROBLEMAS CON EL TÉRMINO DE ERROR

Detección de la autocorrelación. Contraste de Breusch-Godfrey (1978)

- Válido para detectar autocorrelación generada por procesos AR(1) y AR(p) ($p > 1$)
- Adecuado cuando el test de Durbin-Watson no lleva a ninguna conclusión o resulta inaplicable

Pasos a seguir:

1. Se aplica MCO y se obtienen los residuos, $\hat{e}_t$, de $y_t = x_t'\beta + e_t$
2. Se estima: $\hat{e}_t = \alpha_0 + \alpha_1\,\hat{e}_{t-1} + \ldots + \alpha_p\,\hat{e}_{t-p} + \delta_1\,x_{1t} + \ldots + \delta_k\,x_{kt} + u_t$ (rellenando los valores omitidos para los residuos retardados con ceros)
3. Bajo H_0: $\{\alpha_1 = \alpha_2 = \ldots = \alpha_p = 0\}$, $BG = TR^2 \sim \chi_p^2$ (asintóticamente)

V. PROBLEMAS CON EL TÉRMINO DE ERROR

Contrastes basados en la función de correlación. [Box-Pierce (1970) y Ljung-Box(1979)]

- Se basan en el supuesto de existencia de un modelo AR(p) en los errores, y contrastan la significación conjunta de las p primeras correlaciones.
- En economía no son habituales autocorrelaciones significativas de orden superior a 2 [en ocasiones se encuentran correlaciones superiores significativas debido a una mala especificación del modelo de regresión propuesto].

1. Se aplica MCO y se obtienen los residuos, $\hat{e}_t$, de $y_t = x_t'\beta + e_t$
2. Se estiman las m primeras autocorrelaciones (para $s = 1,\ldots,m$) mediante $r_s = \left(\sum_{t=s+1}^{T} \hat{e}_t\hat{e}_{t-s}\right) / \left(\sum_{t=1}^{T} \hat{e}_t^2\right)$
3. Se calculan los estadísticos Q y/o Q' que vienen dados, bajo la hipótesis nula $H_0 : \{\rho_i = 0;\ i = 1,\ldots,m\}$ por la expresiones

$$Q = T\,\Sigma_k\, r_k^2 \sim \chi_m^2 \text{ (Box-Pierce)}$$

$$Q' = (T+2)T\,\Sigma_j\, r_j^2/(T-j) \sim \chi_m^2 \text{ (Ljung-Box)}$$

4) Se rechazará H_0 a un nivel de significación α *si* Q (o Q') > $\chi_m^2(\alpha)$

V. PROBLEMAS CON EL TÉRMINO DE ERROR

6.3.2. Mínimos Cuadrados Generalizados Factibles (MCGF)

- Garantiza teóricamente la obtención de estimadores lineales ***insesgados*** y ***de mínima varianza***
- Desconocimiento de los parámetros del proceso requiere ***procedimientos en dos etapas*** (**Mínimos Cuadrados Generalizados Factibles**):
  - 1ª etapa se obtienen los parámetros del modelo de autocorrelación
  - 2ª etapa se estima el modelo de regresión.
- También se pueden utilizar ***procedimientos iterativos*** (se repite el proceso a partir del modelo estimado en la segunda etapa).

Si ∃ autocorrelación del tipo AR(1) ➔ MCG consiste en tomar **cuasi-diferencias** (transformar el modelo original mediante $z_t^* = z_t - \rho z_{t-1}$)
z representa a alguna de las variables del modelo (dependiente o regresores)

ρ es desconocido ➔ utilizar una estimación consistente (MCGF).

- En $y^* = X^*\beta + e^*$ (modelo transformado), el primer coeficiente es $\boldsymbol{\beta_1(1-\rho)}$.
- **Tomando cuasi-diferencias *se pierde la primera observación* ➔** puede subsanarse mediante la corrección de Prais-Winsten (multiplicar la ecuación para la primera observación en el modelo de regresión original por el factor $(1-\rho^2)^{-1/2}$)

V. PROBLEMAS CON EL TÉRMINO DE ERROR

6.3.2. Mínimos Cuadrados Generalizados Factibles (MCGF) (Cont.)

Alternativas para llevar a cabo la estimación de modelos con autocorrelación

Método de Cochrane-Orcutt (1949)

- **En dos etapas**
  - Se estima el modelo original por MCO.
  - Se utilizan $\hat{e}_t$, para obtener una estimación de ρ:
    bien sea mediante: $\hat{e}_t = \rho\, \hat{e}_{t\text{-}1} + \varepsilon_t$, ó
    a partir del estadístico de Durbin-Watson: $\hat{\rho} = 1 - d/2$
  - Se obtienen las cuasi-diferencias estimadas → $\boxed{y_t^* = y_t - \hat{\rho}\, y_{t-1} \quad x_t^* = x_t - \hat{\rho}\, x_{t-1}}$
  - Se estima el modelo con las variables transformadas, obteniéndose $\beta_{(1)}$.
- **En varias etapas**
  - Se utiliza $\beta_{(1)}$ para generar una nueva serie de residuos, se estima de nuevo ρ y se generan las cuasi-diferencias. $\hat{e}_t^{(1)} = y_t - x_t\beta_{(1)}$; $\hat{e}_t^{(1)} = \rho_{(1)}\hat{e}_{t-1}^{(1)} + \varepsilon_t$

    $$y_t^{*(1)} = y_t - \hat{\rho}_{(1)}\, y_{t-1} \quad , x_t^{*(1)} = x_t - \hat{\rho}_{(1)}\, x_{t-1}$$
  - Se estima el modelo, y se itera hasta alcanzar un grado de convergencia prefijado (por ejemplo, que la diferencia entre la estimación (i) y la (i+1) sea inferior al 1%).

V. PROBLEMAS CON EL TÉRMINO DE ERROR

7. Heteroscedasticidad autorregresiva condicional (opcional)

- Existen casos (sobre todo en el ámbito de la econometría financiera), de **modelos de series temporales en los que la varianza no es estable en el tiempo**.
- Engle (1982) y Cragg (1982), analizando modelos de inflación, encontraron prueba de una forma de heteroscedasticidad en la que la varianza del error de predicción dependía del tamaño de la perturbación precedente, y denominaron a este tipo de modelos procesos de **heteroscedasticidad Autorregregresiva Condicional**, o **modelos ARCH**.

- Dado $y_t = x_t'\beta + e_t$, la versión más simple de un modelo ARCH es

  $$e_t = u_t[\alpha_0 + \alpha_1 e_{t-1}^2]^{1/2}$$ donde los errores $u_t \sim N(0,1)$ e independientes.
- En este modelo ➔ $Var[e_t|e_{t\text{-}1}] = E[e_t^2|e_{t\text{-}1}] = E[u_t^2][\alpha_0 + \alpha_1 e_{t\text{-}1}^2] = \alpha_0 + \alpha_1 e_{t\text{-}1}^2$ es decir, condicionado sobre $e_{t\text{-}1}$, e_t es heteroscedástico.
- Sin embargo, $Var[e_t] = \alpha_0/(1\text{-}\alpha_1)$ ➔ el modelo cumple todas las hipótesis clásicas ➔ MCO es el MELI de β.

¿Cuál es el efecto de la heteroscedasticidad condicional?
Existe un **estimador *no lineal* más eficiente que MCO**: el estimador de **MV**

V. PROBLEMAS CON EL TÉRMINO DE ERROR

7. Heteroscedasticidad autorregresiva condicional (opcional) (Cont.)

Extensiones

- Modelos ARCH(q), en los que $Var[e_t / e_{t-1}, e_{t-2}, \ldots, e_{t-q}] = \sigma_t^2$ donde

$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \ldots + \alpha_q e_{t-q}^2$$

- Modelos ARCH generalizados (GARCH(p,q)).

Ejemplo: GARCH(1,1) $\quad \sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \delta \sigma_{t-1}^2$

- **Contraste presencia de ARCH(1)** (extensión a ARCH(q) es inmediata)

a) estimar $y_t = x_t' \beta + e_t$ *por MCO y obtener los residuos* $\hat{e}_t$

b) estimar $\hat{e}_t^2 = \alpha_0 + \alpha_1 \hat{e}_{t-1}^2 + v_t$

c) $\boxed{TR^2 \xrightarrow{H_0: \{\alpha_1 = 0\}} \chi_1^2}$

d) Si el estadístico resulta significativo ➔ se puede seguir utilizándo MCO, pero MV es más eficiente (y por tanto, más fiable: estimaciones más precisas, intervalos de predicción más estrechos, etc.)

V. PROBLEMAS CON EL TÉRMINO DE ERROR

8. Modelos con autocorrelación espacial (opcional)

- Correlación en los errores en el ámbito de datos de corte transversal
- Existen muchos procesos económicos que pueden dar lugar a aspectos espaciales en los datos. Ejemplos:
  a) variables de tipo inobservable (como el clima, la calidad y riqueza del suelo o la disponibilidad de ciertas materias primas o de bienes substitutivos) pueden estar relacionadas 'espacialmente' y, por tanto, producir correlación espacial en los errores;
  b) patrones de mimetismo en el comportamiento económico de las unidades económicas pueden originar correlación espacial en las ecuaciones describiendo dicho comportamiento.
- En los casos en los que la presencia de un proceso espacial es la responsable de los efectos espaciales, un modelo que tenga en cuenta de forma explícita tal correlación será, en general, más eficiente que un modelo con efectos fijos.
- **Los patrones de comportamiento espacial pueden describirse en términos de correlación espacial**: correlación positiva o negativa entre los valores de una variable sobre los individuos o las regiones al completo de una superficie (tal superficie debe estar dividida en zonas geográficas que completen la misma y que no se solapen).