Econometria

Matemàtiques per als instruments financers
UAB

Curs 2016/2017

Lluís Ferreras

Índex

1. Introducció a l’anàlisi economètrica dels mercats financers i del Risc de Crèdit

1.1. Què és l’econometria?

1.1.1. Presentació

Etimològicament, la paraula econometria significa medició econòmica. L’objectiu que persegueix l’econometria és trobar i quantificar les relacions econòmiques fent ús de les eines estadístiques.

L’econometria va néixer a principis del segle XX amb la introducció dels mètodes estadístics i matemàtics en la medició dels cicles econòmics i la teoria de la demanda.

De totes les definicions d‘econometria que es poden trobar a la literatura es poden destacar les següents:

‘L’ econometria és el camp de l’ economia relacionat amb l’aplicació de l’estadística matemàtica i les eines d’ inferència estadística, a les medicions empíriques de relacions postulades per l’economia teòrica.’

Greene (1998). Econometric Analysis

‘Anàlisi quantitativa de dades econòmiques basant-se en l’ús de teoria econòmica i d’observació.’

P.A. Samuelson, T.C. Koopmans i J.R. Stone (1954),
“Report of the Evaluative Committee for Econometrica”,
Econometrica, vol.22(2): 141-146

‘L’art i la ciència d’utilitzar mètodes estadístics per mesurar relacions entre variables econòmiques.’

Chow(1983) Econometrics

Els trets més rellevants que caracteritzen l’econometria són els següents:

- L’econometria és la branca de la ciència econòmica que s’ocupa de l’anàlisi quantitativa dels fenòmens econòmics.

- L’econometria està relacionada amb altres disciplines com ara la teoria econòmica, l’estadística i les matemàtiques

- L’econometria es basa en un enfocament probabilístic de la realitat.

1.1.2. Variables, relacions i paràmetres

En tot model es distingeixen dos tipus de variables:

- La variable endògena, també anomenada variable dependent o variable a explicar
- Les variables explicatives, també anomenades variables independents o variables exògenes

La variable endògena és aquella que estem interessats a conèixer i a explicar-ne el comportament. Les variables explicatives són les que, d’acord amb els postulats de la teoria econòmica, permeten explicar el comportament de la variable endògena.

Depenent del nombre de variables explicatives que es considerin, una o més d’una, el model s’anomenarà model simple o model múltiple.

Entre la variable endògena i les variables explicatives hi ha una relació de causalitat unidireccional.

Els comportaments de les variables explicatives causen el de la variable endògena. Això permet formular un model, però la relació pot ser de molts tipus: lineal, quadràtica, logarítmica, … En el moment d'especificar el model cal determinar la forma funcional que adopta la relació entre la variable endògena i les explicatives.

Els paràmetres, que estan associats a cada variable explicativa, quantifiquen la relació entre la variable endògena i cadascuna de les variables explicatives. Són, allò que es desconeix i que cal estimar.

1.2. La metodologia de l'econometria

1.2.1. Presentació

Un estudi economètric es pot dividir en les següents fases:

1. Definició d'una teoria o hipòtesi
2. Recopilació de les dades
3. Especificació del model matemàtic de la teoria econòmica
4. Especificació del model economètric
5. Estimació dels paràmetres
6. Contrastació de l'especificació del model
7. Contrastació de les hipòtesis derivades del model
8. Aplicació del model per fer previsions

1.2.2. Definició d'una teoria o hipòtesi

El punt de partida consisteix en saber què és el que afirma la teoria econòmica sobre el tema que es vol estudiar.

Per exemple, ens podem preguntar si les condicions econòmiques afecten a les decisions de la gent per entrar en el mercat laboral. Com a mesura de les condicions econòmiques es pot utilitzar la taxa d'atur (Atur) i ens preguntem com afecta a la taxa d'activitat de l'economia (TA).

La teoria econòmica ens diu que hi ha 2 hipòtesis:

1. Hipòtesi del treballador desanimat: Quan augmenta la taxa d'atur molts treballadors deixen de buscar feina i surten del mercat laboral.

2. Hipòtesi del treballador afegit: Quan l'atur augmenta, molts treballadors que estaven fora del mercat laboral (ex: mares amb fills) decideixen incorporar-se al mercat laboral.

Que la taxa d'activitat (TA) augmenti o disminueixi dependrà de quin d'aquests efectes tingui més força.

1.2.3. Recopilació de les dades

En el nostre exemple necessitem recopilar informació quantitativa sobre la taxa d'atur i la taxa d'activitat. Bàsicament hi ha tres tipus de dades:

1. Dades de tall transversal: Fan referència a una variable mesurada per diferents unitats econòmiques per un període de temps determinat.

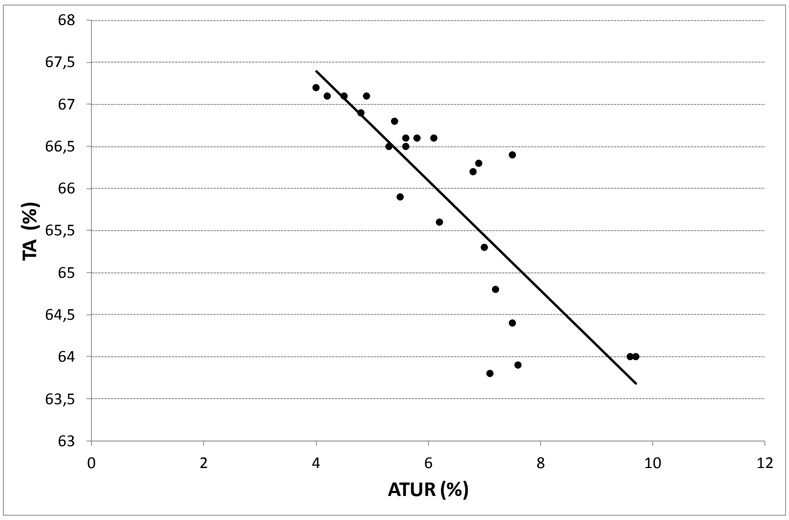
2. Dades de sèries temporals: Fan referència a una variable mesurada per una unitat econòmica en diferents períodes de temps.

3. Combinació dels 2 tipus anteriors. També reben el nom de panells (panel data)

Respecte a la fase de recopilació de dades és important tenir en compte que l'èxit de l'estudi economètric dependrà de la qualitat i quantitat de les dades que siguem capaços d'obtenir.

1.2.4. Especificació del model matemàtic de la teoria econòmica

Per veure com es comporta la taxa d'activitat (TA) respecte al nivell d'atur (ATUR) el primer que hem de fer és dibuixar un gràfic de dispersió d'aquestes dues variables:



El dibuix mostra que hi ha una relació inversa entre l'atur i la taxa d'activitat, implicant que té més força l'efecte del treballador desanimat. Com a primera aproximació es pot escriure el següent model matemàtic lineal:

$$\mathrm{TA} = \beta_1 + \beta_2 \cdot \mathrm{ATUR}$$

Si l'efecte del treballador afegit domina, el valor de β_2 serà positiu. Si, en canvi és l'efecte del treballador desanimat el que domina en la relació, el valor serà negatiu.

1.2.5. Especificació del model economètric

El model economètric afegeix el component aleatori al model econòmic anterior, per mitjà del terme d'error u.

$$\mathrm{TA} = \beta_1 + \beta_2 \cdot \mathrm{ATUR} + \mathrm{u}$$

El terme d'error u s'introdueix per a recollir:

1. L'efecte de totes les altres variables que expliquen el comportament de la variable endògena (taxa d'activitat) però que no han quedat explicitades com a regressors. Moltes d'aquestes variables poden ser petits factors dels quals no es disposa de dades, i se suposa que el seu efecte conjunt sobre la variable endògena és nul.
2. El mateix comportament aleatori que hi ha en la conducta humana en particular, i en les relacions econòmiques i socials en general.
3. Els errors de mesura en les variables incloses en el model i els errors en l'equació.

1.2.6. Estimació dels paràmetres

Els paràmetres β_1 i β_2 de l'equació de dalt s'estimen mitjançant mètodes com el de mínims quadrats ordinaris (MQO) o el de màxima versemblança (MV).

Utilitzant les mateixes dades de la taxa d'atur (ATUR) i de la taxa d'activitat (TA) amb les que s'ha dibuixat el gràfic de dispersió de dalt i utilitzant el mètode MQO s'obté la següent equació estimada:

$$\widehat{\mathrm{TA}} = 69.996 - 0.651 \cdot \mathrm{ATUR}$$

Per tant, si la taxa d'atur augmenta en una unitat (un punt percentual), ceteris paribus, la taxa d'ocupació s'espera que disminueixi, de mitjana, 0,65 punts percentuals.

Per altra banda, si la taxa d'atur fos del 0%, la taxa d'activitat seria, en mitjana, del 70%.

1.2.7. Contrastació de l'especificació del model

Ens podem preguntar per la validesa del model estimat. És cert que una persona tindrà en compte les condicions del mercat laboral abans d'entrar en aquest mercat?

Per exemple, hi ha altres factors que poden afectar a la decisió de participar en la població activa, per exemple el salari per hora. Per tenir en compte l'efecte d'aquesta variable (IH), es pot analitzar el següent model:

$$\mathrm{TA} = \beta_0 + \beta_1 \cdot \mathrm{ATUR} + \beta_2 \cdot \mathrm{IH} + \mathrm{u}$$

Això és un exemple de model de regressió lineal múltiple, en front de l'anterior que s'anomena de regressió lineal simple perquè té una única variable explicativa.

L'estimació d'aquest model per MQO (després d'afegir les dades d'ingressos per hora) és la següent:

$$\widehat{\mathrm{TA}} = 80.901 - 0.673 \cdot \mathrm{ATUR} - 1.404 \cdot \mathrm{IH}$$

Aquest model inclou l'anterior, perquè conté una variable més i per tant es pot utilitzar aquest perquè és més complet. Es podrien anar incloent més variables, encara que no sempre és el millor afegir masses, ja que el model economètric podria arribar a tenir poca utilitat pràctica.

1.2.8. Contrastació de les hipòtesis derivades del model

Una vegada seleccionat el model podem estar interessats a fer un contrast d'hipòtesis. Per exemple, la hipòtesi del treballador desanimat postula una relació negativa entre la participació al mercat laboral i la taxa d'atur. Els nostres resultats semblen ajustar-se a aquesta hipòtesi ja que el coeficient estimat de la taxa d'atur és negatiu.

A vegades, la contrastació d'hipòtesis pot ser molt complexa. Per exemple, un estudi anterior del coeficient de la taxa d'atur va concloure que era aproximadament -1. Si s'utilitza un model o un altre es pot arribar a diferents conclusions. La resposta a una hipòtesi en particular pot dependre del model que hem decidit utilitzar.

1.2.9. Aplicació del model per fer previsions

Una vegada estimat el model, es voldrà utilitzar per fer prediccions o previsions. Per exemple, si es disposa d'una estimació de quina serà la taxa d'atur i el salari per hora en 2017, substituint aquests valors a l'equació de dalt es pot predir quina serà la taxa d'activitat per l'any 2017.

Quan disposem de les dades reals de taxa d'activitat per l'any 2017 es podrà comparar el valor previst amb el valor real. La discrepància entre aquests dos valors representarà l'error de predicció. Intentarem que aquest error de predicció sigui el més petit possible.

1.3. L'econometria als mercats financers

1.3.1. Algunes àrees rellevants dins l'econometria dels mercats financers

Algunes de les aplicacions de l'econometria en els mercats financers són les següents:

-Predictibilitat del rendiment dels actius: Fins a quin punt són predictibles els preus dels actius financers ? L'evidència empírica sembla suggerir que els rendiments dels actius financers són predictibles en un cert grau.

-Eficiència en el mercat (market efficiency) : Si un mercat és eficient els preus sempre reflexen tota la informació disponible. L'eficiència dels mercats implica que és impossible predir el rendiment dels actius. Reflexen els preus dels actius completament tota la informació disponible ?

-Microestructura de mercats (Market Microstructure): Es refereix a l'estudi de l'estructura institucional en la que es determina el rendiment dels actius financers. Per exemple, estudia com afecten els costos de transacció als rendiments dels actius.

-Anàlisi d'estudi d'esdeveniments (Event study analysis): És l'estudi de l'impacte d'un esdeveniment en un determinat actiu.

-Caracterització del risc dels actius (Econometria del risc) : Estudi de les mesures del Risc via volatilitat.

-Models de rendiment dels actius: Com es pot explicar la conducta dels preus o rendiments d'un actiu?

1.3.2. Algunes característiques de les dades financeres

Les dades financeres presenten diferències respecte les dades macroeconòmiques:

Dades Macroeconòmiques	Dades Financeres
-Manca de dades: Les sèries són curtes	-Moltes dades: Se'n registren molt sovint
-Errors de mesura	-Errors de mesura petits
-Tenen poc soroll	-Tenen molt soroll: És més difícil extreure el comportament d'una sèrie
-Ex: PIB, Inflació, IPC, Tipus de canvi, Taxa d'atur	-Ex: Cotitzacions d'actius

1.3.3. Algunes característiques de les sèries temporals financeres

Algunes característiques de les sèries temporals financeres són les següents :

-La volatilitat de les dades és alta

-La volatilitat de les dades varia en el temps : Aquest fenòmen també es coneix amb el nom de 'Volatility Clustering'. En una sèrie de temps de preus d'actius s'observa que la variància dels rendiments és alta per períodes llargs i després baixa també per períodes llargs. Això fa que els models que suposen que els rendiments estan idènticament distribuïts al llarg del temps no siguin convincents.

-Volatilitats asimètriques: La volatilitat és més alta en períodes de baixades que en períodes de pujades en els mercats.

-Risk Spillovers: El moviment dels actius afecta a altres actius.

-Absència de normalitat: El rendiment dels actius no té una distribució normal. Les cues de les distribucions són més gruixudes.

-La memòria de les sèries és bastant gran: Les sèries financeres presenten autocorrelació alta.

1.3.4. Exemple d'aplicació de tècniques economètriques als mercats financers : Event study analysis

Estem interessats en saber si un determinat esdeveniment afecta a un títol. Hi ha un article conegut que explica aquest procés:

'Event Studies in Economics and Finance'
A.Craig MacKinlay
Journal of Economic Literature (March 1997)

Suposem que tenim una sèrie dels rendiments del títol. Es pot separar el temps en 3 finestres:

1: Finestra de l'estimació
2: Finestra de l'esdeveniment
3: Finestra post-esdeveniment

Es defineixen els rendiments anormals com:

$$AR_j = R_j - R_j^{Normals}$$

R_j són els rendiments observats, $R_j^{Normals}$ són els rendiments que haguéssim obtingut si no hagués passat l'esdeveniment. Si es vol saber si l'esdeveniment no ha afectat als rendiments es fa la hipòtesis nul.la:

$$H_0: \quad AR_j = 0$$

Els rendiments anormals es calculen a la finestra de l'esdeveniment. A aquest període es diu τ
Per tant, el model és

$$AR_{j\tau} = R_{j\tau} - R_{j\tau}^{Normals}$$

Per calcular $R_{j\tau}^{Normals}$ necessitem un model que ens expliqui els rendiments

model: $R_j = \alpha_j + \beta_j R_m + \varepsilon_j$

R_m : Rendiment de la cartera de mercat
α_j : Part sistemàtica i específica del rendiment del títol
$\beta_j R_m$: Part sistemàtica i no específica del rendiment del títol
ε_j : Part no sistemàtica i específica del rendiment del títol

S'agafa com a rendiment normal del títol, el component sistemàtic del títol, és a dir;

$$R_j^{Normal} = E(R_j)$$

Per tant s'han d'estimar α_j i β_j utilitzant les dades de la finestra d'estimació. A partir d'aquí s'han de calcular els rendiments normals fent una predicció puntual.
L'últim pas és calcular $AR_{j\tau}$

$$R_{j\tau}^{Normal} = \hat{\alpha}_j + \hat{\beta}_j \cdot R_{m\tau}$$

$$A\hat{R}_{j\tau} = R_{j\tau} - \hat{\alpha}_j - \hat{\beta}_{j\tau} \cdot R_{m\tau}$$

A partir d'aquí s'ha de contrastar la hipòtesi nul.la.

Si la finestra d'estimació és prou gran, sota la hipòtesi nul.la H0

$$A\hat{R}_{j\tau} \sim N\left(0, \sigma^2_{\varepsilon_j}\right)$$

Si agreguem per tots els dies de la finestra del event

$$C\hat{A}R_j(\tau_1, \tau_2) = \sum_{\tau=\tau_1}^{\tau_2} A\hat{R}_{j\tau} \qquad \sigma^2_j(\tau_1, \tau_2) = (\tau_2 - \tau_1 + 1) \cdot \sigma^2_{\varepsilon_j}$$

$$C\hat{A}R_j(\tau_1, \tau_2) \sim N\left(0, \sigma^2_j(\tau_1, \tau_2)\right)$$

1.4. Aplicacions de l'econometria al Risc de Crèdit

1.4.1. Introducció al Risc de Crèdit

El risc de crèdit és la probabilitat de que un acreedor no pugui fer front a les seves obligacions. Tot contracte té un risc d'impagament, el qual ha de cobrir l'entitat. Generalment, aquest és el risc més important al que ha de fer front una entitat financera.

La pèrdua d'un contracte es defineix de la següent forma:

$$P_i = 1_{D_i} \cdot EAD_i \cdot LGD_i$$

La pèrdua del contracte serà:

- Pèrdua = 0 si el contracte no entra en default
- Pèrdua = EAD x LGD si el contracte entra en default

on

- EAD (Exposure at default): Exposició en el moment d'entrada en mora.
- LGD (Loss given default): Severitat. És la part no recuperada del capital prestat.

El risc de crèdit té dos components principals:

- Pèrdua esperada: És la pèrdua mitjana de la cartera creditícia a la que s'exposa la entitat. Pot considerar-se un cost addicional del negoci, i conceptualment es cobreix mitjançant dotacions regulars al fons de provisions corresponents, impactant en els resultats de la entitat.
- Pèrdua inesperada: És la possible pèrdua derivada d'efectes adversos. Conceptualment aquestes pèrdues es cobreixen amb el capital, recursos propis de l'entitat destinats a absorvir les pèrdues i protegir-la de la fallida.

La Pèrdua esperada d'una exposició és la mitjana (esperança matemàtica) de la distribució de pèrdues.

$$PE = E[P_i] = E\left[1_{D_i} \cdot EAD_i \cdot LGD_i\right] = PD_i \cdot EAD_i \cdot LGD_i$$

on

- PD (Probability of default) : Probabilitat d'entrada en mora

1.4.2. Eines de medició del risc de crèdit

Les eines de Rating i Scoring permeten ordenar les operacions/clients en funció de la seva qualitat creditícia. Existeixen diferents tipus d'eines :

- Scoring sociodemogràfic : Eines de qualificació d'operacions que es basen en informació objectiva del client i de l'operació (edat, estat civil, import,…). S'utilitzen per la concessió d'operacions. Es calculen en el moment de l'admissió.
- Scoring de comportament : Eines de qualificació de clients que es basen en dades històriques de comportament financer (saldos, descoberts). Tenen una elevada capacitat predictiva i s'utilitzen pel seguiment del risc i per la concessió de noves operacions. Es solen calcular mensualment.
- Rating : Eines de qualificació d'empreses a partir de dades financeres (estats financers), sector, dades qualitatives de l'empresa …

Aquests models es desenvolupen utilitzant eines estadístiques (anàlisi discriminant, regressió logística, …).

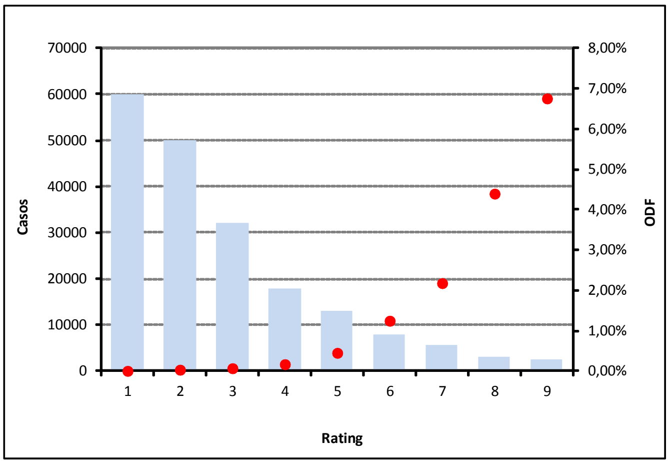
Les eines de qualificació assignen una puntuació a cada acreditat. Això permet estimar una probabilitat d'entrada en mora, en funció de la puntuació assignada pel model, ajustant una corba a les freqüències de default històriques observades.

1.4.3. Estimació de PD

Les eines de qualificació assignen una puntuació a cada acreditat/contracte, el que permet estimar una probabilitat d'entrada en mora a partir d'aquesta puntuació.

Si es disposa de dades observades del comportament d'una cartera, es pot estimar la probabilitat d'entrada en mora de la cartera, en funció de la puntuació assignada pel model, ajustant una corba a les freqüències de default observades.

Rating	Casos	Morosos	Buenos	ODF
1	60000	4	59996	0,01%
2	50000	18	49982	0,04%
3	32000	23	31977	0,07%
4	18000	31	17969	0,17%
5	13000	59	12941	0,45%
6	8000	100	7900	1,25%
7	5500	120	5380	2,18%
8	3000	132	2868	4,40%
9	2500	169	2331	6,76%



1.4.4. Estimació de EAD

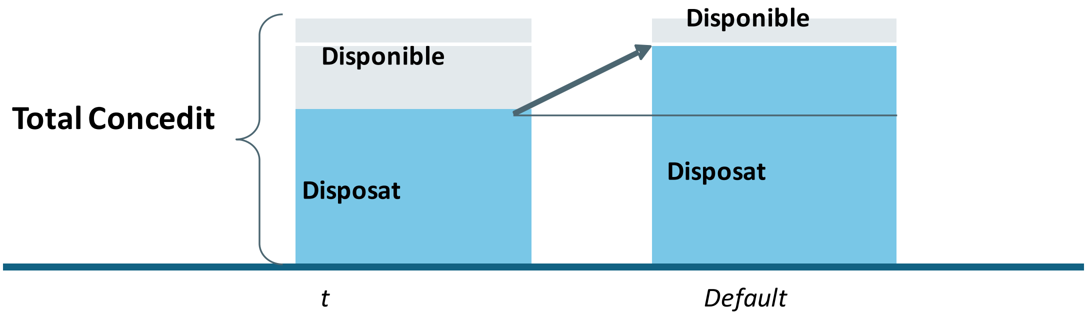
La EAD és una estimació de l'import adeutat pel client en el moment d'entrar en default.

$$EAD = D_m = D_n + CCF \cdot (C_n - D_n)$$

on

- D_m és el disposat en mora
- D_n és el disposat abans d'entrar en mora
- C_n és el concedit abans d'entrar en mora

- CCF:Credit conversion factor. Indica l'augment de disposat al entrar en mora



1.4.5. Estimació de LGD

La severitat o LGD és una estimació del percentatge d'exposició que l'entitat perdrà si el client entra en mora.

$$LGD = 1 - \%Recuperació = 1 - \frac{Recuperació - Costos}{EAD}$$

Generalment els models de LGD es desenvolupen a partir de la tipologia del titular de l'operació i les garanties aportades.

En els últims anys s'han desenvolupat models quantitatius especialment rellevants per les carteres hipotecàries, introduint risk drivers com el *Loan to Value*.

1.4.6. Exemple d'aplicació de tècniques economètriques en la mesura del Risc de Crèdit : Desenvolupament d'un model sociodemogràfic.

Els passos a seguir en la construcció d'un model són els següents :

1. Fase inicial
   a. Definir i concretar l'objectiu i abast del model, així com la metodologia a seguir.
   b. Petició de dades
2. Obtenció de les mostres de desenvolupament
   a. Anàlisi de dades
   b. Detecció de valors extrems (outliers) i blancs
   c. Tractament d'aquests valors
   d. Selecció de la mostra de construcció i de validació
3. Anàlisi univariant
   a. Definició dels factors a incorporar al model
   b. Estudi descriptiu de les variables
   c. Anàlisi de la capacitat predictiva dels factors
   d. Pre-selecció de les variables que participaran en el model
4. Anàlisi multivariant / Disseny de l'algoritme de càlcul
   a. Anàlisi de la correlació dels factors
   b. Selecció de variables definitives
   c. Selecció de la metodologia de construcció
   d. Disseny de l'algoritme de càlcul
5. Calibració (estimació de la probabilitat de default)
6. Validació : Verificar el correcte funcionament del model amb la mostra de validació (test out-of-sample) i amb una mostra d'un altre moment del temps (test out-of-time)
   a. Validació del model
   b. Validació de la calibració

2. Extensions al model de regressió clàssic

2.1. Estimació màxim versemblant

2.1.1. Presentació

El principi que hi ha darrera de l'estimació màxim versemblant és el de triar els paràmetres que maximitzen la probabilitat d'obtenir la mostra observada. La diferència amb el mètode d'estimació MQO és que es basa en les hipòtesis que s'estableixen per a la distribució de les variables aleatòries que apareixen en el model.

La funció de densitat conjunta

$$f(y_1, y_2, ..., y_n / \theta) = f(y / \theta)$$

donats uns paràmetres ens dóna la probabilitat de la mostra que hem observat.

La funció de versemblança L donada una mostra en concret valora els possibles valors dels paràmetres. Matemàticament és igual que la funció de densitat però la interpretació és diferent.

$$L(\theta / y) = f(y / \theta)$$

L'estimador màxim versemblant es defineix com

$$\hat{\theta}_{MV} = \arg \underset{\theta}{Max}\, L(\theta, y)$$

2.1.2. Derivació

Distribució normal conjunta: Si z és un vector de variables aleatòries, distribuides com una normal multivariant amb E(z)=μ i var(z)=Σ, llavors

$$f(z / \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(z-\mu)' \Sigma^{-1} (z-\mu) \right)$$

Sabem que

$$\left. \begin{array}{r} Y = X\beta + u \\ u \sim N(0, \sigma^2 I_n) \end{array} \right\} \quad \Rightarrow \quad Y \approx N(X\beta, \sigma^2 I_n)$$

Per tant, aplicant el resultat anterior la funció de densitat de Y és

$$f(Y / \beta, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta) \right)$$

La funció de versemblança serà

$$L(\beta, \sigma^2 / Y) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta) \right)$$

Per trobar el màxim és maximitza el logaritme neperià de la funció de versemblança, igualant a cero les derivades primeres respecte els paràmetres

$$\ln L\left(\beta,\sigma^2 / Y\right) = -\frac{n}{2}\ln 2\pi\sigma^2 - \frac{1}{2\sigma^2}\left(Y - X\beta\right)'\left(Y - X\beta\right)$$
$$\frac{\partial \ln L}{\partial \beta} = -\frac{1}{2\tilde{\sigma}^2_{MV}}\left(-2X'Y + 2X'X\tilde{\beta}_{MV}\right) = 0$$
$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{\tilde{\sigma}^2_{MV}} + \frac{1}{2\tilde{\sigma}^4_{MV}}\left(Y - X\tilde{\beta}_{MV}\right)'\left(Y - X\tilde{\beta}_{MV}\right) = 0$$

La solució de les condicions de primer ordre anteriors és

$$\tilde{\beta}_{MV} = \left(X'X\right)^{-1}X'Y$$
$$\tilde{\sigma}^2_{MV} = \frac{\left(Y - X\tilde{\beta}_{MV}\right)'\left(Y - X\tilde{\beta}_{MV}\right)}{n}$$

Per tant s'observa que quan hi ha pertorbacions normals l'estimador MQO i el màxim versemblant coincideixen en el cas de les β. En el cas de σ^2 es té

$$\tilde{\sigma}^2_{MV} = \frac{n-k}{n}\hat{\sigma}^2$$

Per tant l'estimador màxim versemblant de σ^2 té biaix.

2.1.3. Teorema de Cramer-Rao

Teorema: Si tenim un vector de variables aleatòries $z = \left(z_1, z_2, ..., z_n\right)$ amb una funció de densitat conjunta $f\left(z / \theta\right)$ i tenim un estimador sense biaix del vector de paràmetres $\theta, \tilde{\theta}$ llavors

$\mathrm{var}\left(\tilde{\theta}\right) \geq I\left(\theta\right)^{-1}$ (desigualtat de Cramer-Rao)

on $I\left(\theta\right) = -E\left[\frac{\partial^2 \ln L\left(\theta\right)}{\partial^2 \theta}\right]$ (matriu d'informació de Fisher)

$I\left(\theta\right)^{-1}$ es coneix com límit inferior de Cramer-Rao

Si es calcula la matriu d'informació de Fisher pel model de regressió lineal clàssic s'obté

$$I(\beta,\sigma^2)^{-1} = -E\begin{bmatrix} \frac{\partial^2 \ln L}{\partial\beta\partial\beta'} & \frac{\partial^2 \ln L}{\partial\beta\partial\sigma^2} \\ \frac{\partial^2 \ln L}{\partial\sigma^2\partial\beta'} & \frac{\partial^2 \ln L}{\partial\sigma^2\partial\sigma^2} \end{bmatrix}^{-1} =$$

$$= -E\begin{bmatrix} -\frac{1}{\sigma^2}X'X & -\frac{1}{2\sigma^4}(2X'Y - 2X'X\beta) \\ -\frac{1}{2\sigma^4}(2X'Y - 2X'X\beta) & \frac{n}{2\sigma^4} - \frac{(Y - X\beta)'(Y - X\beta)}{\sigma^6} \end{bmatrix}^{-1} =$$

$$\begin{bmatrix} \frac{1}{\sigma^2}X'X & 0 \\ 0 & \frac{n}{2\sigma^4} \end{bmatrix}^{-1} = \begin{bmatrix} \sigma^2(X'X)^{-1} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{bmatrix}$$

Com sabem que per MQO $\text{var}(\hat{\beta}) = \sigma^2(X'X)^{-1}$ aquest resultat ens permet dir que si u té distribució normal, aleshores $\hat{\beta}$ és eficient. Si u no té distribució normal, l'únic que podem dir és que $\hat{\beta}$ és el millor estimador lineal sense biaix (BLUE).

2.2. Propietats asimptòtiques

2.2.1. Presentació

Les propietats asimptòtiques fan referència a com canvia la funció de densitat quan li donem més informació. Les principals propietats asimptòtiques dels estimadors són les següents:

(i) Consistència (feble):

Un estimador $\hat{\theta}_n$ de θ és consistent si $\hat{\theta}_n \xrightarrow{p} \theta$

També s'escriu $p\lim \hat{\theta}_n = \theta$

Una condició suficient per tenir consistència és:

(1) $\lim_{n\to\infty} E(\hat{\theta}_n) = \theta$ (i.e sense biaix asimptòtic)

(2) $\lim_{n\to\infty} \text{var}(\hat{\theta}_n) = 0$

(ii) Normalitat asimptòtica (CAN) (Consistent and asymptotically normal)

Un estimador $\hat{\theta}_n$ de θ és consistent i asimptòticament normal (CAN) si:

(1) $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0,Q)$ on Q és una matriu finita definida positiva

De (1) en podem derivar una definició operacional de la variància asimptòtica de $\hat{\theta}_n$:

$a\,\text{var}(\hat{\theta}_n) \equiv Q$

Així la distribució asimptòtica operacional de $\hat{\theta}_n$ és: $\hat{\theta}_n \xrightarrow{d} N(\theta,Q)$

(iii) Eficiència asimptòtica:

Un estimador $\hat{\theta}_n$ de θ és asimptòticament eficient si $a\operatorname{var}(\tilde{\theta}_n) - a\operatorname{var}(\hat{\theta}_n)$ és una matriu definida no negativa, on $\tilde{\theta}_n$ és qualsevol altre estimador CAN.

2.2.2. Propietats asimptòtiques dels estimadors MQO i MV

L'estimador MQO té les següents propietats asimptòtiques:

(1) Consistència

(2) Normalitat asimptòtica (CAN)

L'estimador MV sota certes condicions de regularitat té les següents propietats asimptòtiques:

(1) Consistència

(2) Normalitat asimptòtica (CAN)

(3) Eficiència asimptòtica : La versió asimptòtica del Teorema de Cramer-Rao diu que si $\tilde{\theta}_n$ és qualsevol estimador CAN, llavors $a\operatorname{var}(\tilde{\theta}_n) \geq I(\theta_\infty)^{-1}$. A la matriu $I(\theta_\infty)$ se l'anomena matriu d'informació asimptòtca.

2.3. Estimació amb restriccions

2.3.1. Presentació

En aquest tema aprendrem a estimar els paràmetres del model imposant una informació que tenim a priori dels paràmetres.

Hi ha diferents tipus de restriccions:

-Exactes vs Aleatòries: Una restricció exacta és, per exemple, que el paràmetre està entre 0 i 1. Una restricció aleatòria seria, per exemple, que el paràmetre β_2 està entorn de 1.

-Lineals vs no lineals

-Igualtat vs desigualtat

Nosaltres farem restriccions exactes, lineals i d'igualtat.

Si suposem que l'informació que tenim és bona l'estimació amb restriccions millora l'eficiència ja que la variància és més petita i es pot estimar amb més precisió.

2.3.2. Algunes propietats importants sobre distribucions

1) Sigui $z = (z_1, z_2, ..., z_q)$ un vector de q variables aleatòries i A una matriu de coeficients. Llavors:

Si $z \sim N(\mu, \Sigma) \quad \Rightarrow \quad Az \sim N(A\mu, A\Sigma A')$

2) Sigui $z = (z_1, z_2, ..., z_q)$ un vector de q variables aleatòries i A una matriu de coeficients idempotent de rang r. Llavors:

Si $z \sim N(0, I) \quad \Rightarrow \quad z'Az \sim \chi^2(r)$

3) Sigui $z = (z_1, z_2, ..., z_q)$ un vector de q variables aleatòries. Llavors:

Si $z \sim N(0, \Sigma) \quad \Rightarrow \quad z'\Sigma^{-1}z \sim \chi^2(q)$

4) Siguin z_1 i z_2 dues variables aleatòries independents on:

$$z_1 \sim N(0,1) \quad i \quad z_2 \sim \chi^2(q)$$

Llavors: $\dfrac{z_1}{\sqrt{z_2/q}} \sim t(q)$

5) Siguin z_1 i z_2 dues variables aleatòries independents on:

$$z_1 \sim \chi^2(q_1) \quad i \quad z_2 \sim \chi^2(q_2)$$

Llavors: $\dfrac{z_1/q_1}{z_2/q_2} \sim F(q_1, q_2)$

2.3.3. Estimació amb restriccions lineals, exactes, d'igualtat

Tenim un model $Y = X\beta + u$. Les restriccions seran de la forma $R\beta = r$

Exemple:

$$\left.\begin{aligned} \beta_1 + \beta_2 &= 1 \\ \beta_3 &= 2 \end{aligned}\right\} \quad \Leftrightarrow \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

L'estimació MQO (minimització de la suma de quadrats) amb restriccions s'anomena MQR (mínims quadrats restringit).

La derivació de l'estimador MQR és la següent:

$$\left.\begin{aligned} Min\ SQR = (Y - X\beta)'(Y - X\beta) \\ s.a: R\beta = r \end{aligned}\right\} \quad \Rightarrow \quad L = (Y - X\beta)'(Y - X\beta) - 2\lambda'(r - R\beta)$$

Es deriva el Lagrangià L respecte tots els paràmetres,

$$D_\beta L(\beta, \lambda) = -2X'Y + 2X'X\hat{\beta}_R + 2R'\hat{\lambda} = 0$$
$$D_\lambda L(\beta, \lambda) = R\hat{\beta}_R - r = 0$$

Per tant,

$$\begin{bmatrix} X'X & R' \\ R & 0 \end{bmatrix} \begin{bmatrix} \hat{\beta}_R \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} X'Y \\ r \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} \hat{\beta}_R \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} X'X & R' \\ R & 0 \end{bmatrix}^{-1} \begin{bmatrix} X'Y \\ r \end{bmatrix}$$

Resultat: Inversa d'una matriu particionada

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \Rightarrow A^{-1}=\begin{bmatrix} \left(A_{11}-A_{12}A_{22}^{-1}A_{21}\right)^{-1} & -A_{11}^{-1}A_{12}\left(A_{22}-A_{21}A_{11}^{-1}A_{12}\right)^{-1} \\ -\left(A_{22}-A_{21}A_{11}^{-1}A_{12}\right)^{-1}A_{21}A_{11}^{-1} & \left(A_{22}-A_{21}A_{11}^{-1}A_{12}\right)^{-1} \end{bmatrix}$$

$$Si\ F=\left(A_{22}-A_{21}A_{11}^{-1}A_{12}\right)^{-1} \quad \Rightarrow A^{-1}=\begin{bmatrix} A_{11}^{-1}\left(I+A_{12}FA_{21}A_{11}^{-1}\right) & -A_{11}^{-1}A_{12}F \\ -FA_{21}A_{11}^{-1} & F \end{bmatrix}$$

utilitzant l'expressió de la inversa d'una matriu particionada,

$$\begin{aligned} \hat{\beta}_R &= \left\{(X'X)^{-1}+(X'X)^{-1}R'\left[-R(X'X)^{-1}R'\right]^{-1}R(X'X)^{-1}\right\}X'Y+ \\ &\quad +\left\{(X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}\right\}r= \\ &=(X'X)^{-1}X'Y+(X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}\left\{r-R(X'X)^{-1}X'Y\right\} \end{aligned}$$

$$\hat{\beta}_R=\hat{\beta}+(X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}\left\{r-R\hat{\beta}\right\}$$

A la pràctica, per calcular MQR a vegades és més fàcil integrar la restricció al model que utilitzar l'expressió anterior com es pot veure en el següent exemple:

Donat el model $\left.\begin{array}{c} y_i=\beta_1+\beta_2 x_{2i}+\beta_3 x_{3i}+u_i \\ \beta_2+\beta_3=1 \end{array}\right\}$

es construeix el següent model restringit posant la restricció a dins:

$$y_i=\beta_{R1}+\beta_{R2}x_{2i}+\left(1-\beta_{R2}\right)x_{3i}+u_{iR}$$

La pertorbació es comporta igual que en el model inicial

$$y_i-x_{3i}=\beta_{R1}+\beta_{R2}\left(x_{2i}-x_{3i}\right)+u_{iR}$$

La variable dependent ara és diferent. S'estima el model de dalt per MQO i s'obtenen $\hat{\beta}_{R1}$ i $\hat{\beta}_{R2}$. Després es troba $\hat{\beta}_{R3}=1-\hat{\beta}_{R2}$

2.3.4. Propietats de l'estimador MQR

$$\begin{aligned} E\left(\hat{\beta}_R\right) &= E\left(\hat{\beta}\right)+(X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}\left\{r-RE\left(\hat{\beta}\right)\right\}= \\ &=\beta+(X'X)^{-1}R'\left[R(X'X)^{-1}R'\right]^{-1}\left\{r-R\beta\right\} \end{aligned}$$

-Si la informació és certa, i.e $r=R\beta$, l'estimador MQR no té biaix. En cas contrari l'estimador MQR presenta biaix.

També és compleix $\operatorname{var}\left(\hat{\beta}_R\right)<\operatorname{var}\left(\hat{\beta}\right)$

La variància sempre es redueix però l'eficiència és per estimadors sense biaix. Si les restriccions són certes no hi ha biaix i per tant MQR és més eficient.

2.4. *Contrast d'hipòtesis de restriccions lineals exactes*

2.4.1. Presentació

Al tema anterior haviem vist que la t de student la podiem utilitzar per fer contrast d'una sola combinació lineal. Exemple:

$$H_0 : \beta_2 + \beta_3 = 0 \quad \rightarrow \quad t = \frac{\hat{\beta}_2 + \hat{\beta}_3}{\hat{\sigma}_{\hat{\beta}_2 + \hat{\beta}_3}} \sim t_{n-k}$$

Si ara afegim la restricció $\beta_4 = 0$ ja no podem utilitzar la t de Student.

2.4.2. Derivació de l'estadístic de contrast

Es pot derivar un estadístic de contrast per qualsevol q (número de restriccions) Per

$$H_0 : R\beta = r$$

Sabem que $\hat{\beta} \sim N\left(\beta, \sigma^2 (X'X)^{-1}\right)$ (k variables aleatòries)

Llavors $R\hat{\beta} \sim N\left(R\beta, \sigma^2 R(X'X)^{-1} R'\right)$ (q variables aleatòries)

Sota H_0
$$R\hat{\beta} \sim N\left(r, \sigma^2 R(X'X)^{-1} R'\right) \Rightarrow R\hat{\beta} - r \sim N\left(0, \sigma^2 R(X'X)^{-1} R'\right) \Rightarrow$$
$$\left(R\hat{\beta} - r\right)\left[\sigma^2 R(X'X)^{-1} R'\right]^{-1}\left(R\hat{\beta} - r\right) \sim \chi^2(q)$$

Donat que $\dfrac{\hat{u}'\hat{u}}{\sigma^2} \sim \chi^2(n-k)$

Llavors,

$$F \equiv \frac{\dfrac{\left(R\hat{\beta} - r\right)\left[\sigma^2 R(X'X)^{-1} R'\right]^{-1}\left(R\hat{\beta} - r\right)}{q}}{\dfrac{\hat{u}'\hat{u}}{\sigma^2} \Big/ n-k} =$$

$$= \frac{\dfrac{\left(R\hat{\beta} - r\right)\left[R(X'X)^{-1} R'\right]^{-1}\left(R\hat{\beta} - r\right)}{q}}{\dfrac{\hat{u}'\hat{u}}{n-k}} \sim F(q, n-k)$$

Es pot derivar una expressió equivalent d'aquest estadístic de contrast,

Estimació sense restriccions MQO $\hat{Y} = X\hat{\beta} \quad \rightarrow \quad \hat{u} \equiv Y - \hat{Y}$

Estimació incorporant restriccions MQR $\hat{Y}_R = X\hat{\beta}_R \quad \rightarrow \quad \hat{u}_R \equiv Y - \hat{Y}_R$

Es redefineix $\hat{u}_R \equiv Y - \hat{Y}_R = Y - X\hat{\beta}_R = Y - X\hat{\beta} + X\hat{\beta} - X\hat{\beta}_R = \hat{u} - X\left(\hat{\beta}_R - \hat{\beta}\right)$

Llavors $\hat{u}'_R \hat{u}_R = \hat{u}'\hat{u} + \left(\hat{\beta}_R - \hat{\beta}\right)(X'X)\left(\hat{\beta}_R - \hat{\beta}\right)$ (ja que $X'\hat{u} = 0$) (1)

Sabem que $\hat{\beta}_R - \hat{\beta} = (X'X)^{-1} R'\left[R(X'X)^{-1} R'\right]^{-1}\left(r - R\hat{\beta}\right)$ (2)

Substituint (2) a (1): $\hat{u}'_R \hat{u}_R = \hat{u}'\hat{u} + \left(r - R\hat{\beta}\right)\left[R(X'X)^{-1}R'\right]^{-1}\left(r - R\hat{\beta}\right)$

Aleshores l'estadístic de contrast F es pot re-escriure com:

$$F = \frac{\left(R\hat{\beta} - r\right)\left[R(X'X)^{-1}R'\right]^{-1}\left(R\hat{\beta} - r\right) / q}{\hat{u}'\hat{u} / n-k} = \frac{\left(\hat{u}'_R \hat{u}_R - \hat{u}'\hat{u}\right) / q}{\left(\hat{u}'\hat{u}\right) / n-k} = \frac{\left(SQR_R - SQR\right) / q}{SQR / n-k} = \frac{\left(R^2 - R_R^2\right) / q}{\left(1 - R^2\right) / n-k}$$

2.4.3. Cas particular: Test de significació global

En aquest test es tracta de veure si totes les variables explicatives conjuntament expliquen el comportament de Y.

$$y_i = \beta_1 + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + u_i$$

La hipòtesis nul.la és

$$H_0 : \beta_2 = \beta_3 = \ldots = \beta_k = 0$$

per tant la matriu de restriccions R té k-1 files. L'estadístic de contrast és

$$F = \frac{\dfrac{SQR_R - SQR}{k-1}}{\dfrac{SQR}{n-k}}$$

El model restringit és $y_i = \beta_{1R} + u_{iR}$

Com $\hat{\beta}_{1R} = \bar{y}$ les y's ajustades seran $\hat{y}_{iR} = \bar{y}$

Per tant $SQR_R = \sum_{i=1}^{N} \hat{u}_{iR}^2 = \sum_{i=1}^{N} \left(y_i - \bar{y}\right)^2 = SQT$

Per tant l'estadístic F es pot escriure com

$$F = \frac{\dfrac{SQT - SQR}{k-1}}{\dfrac{SQR}{n-k}} = \frac{\dfrac{R^2}{k-1}}{\dfrac{1-R^2}{n-k}} \sim F(k-1, n-k)$$

2.5. Anàlisi de la permanència estructural. Contrast de Chow

2.5.1. Presentació

Una de les hipòtesis bàsiques del model de regressió estàndard és la hipòtesi de permanència estructural dels paràmetres del model. El significat de l'imcompliment d'aquesta hipòtesi depèn del tipus de dades amb que tractem:

- Si treballem amb dades de sèrie temporal, que hi hagi un canvi d'estructura dins el període mostral vol dir que s'han produït un o més canvis en l'estructura interna del procés generador de les dades.

- Si treballem amb dades de tall transversal, un canvi estructural vol dir que la mostra conté dos o més grups d'individus que presenten comportaments diferents entre ells.

Si no es té en compte el canvi estructural, l'ajust no és bo i els estimadors que s'obtenen són aproximadament una mitjana ponderada dels estimadors que s'obtindrien si s'haguessin ajustat dues regressions diferents, una per cada una de les submostres en què es pot dividir la mostra total.

Les prediccions no són fiables, i els contrastos de significació són erronis, ja que si els estimadors són ineficients, les seves variàncies estan inflades. Augmenta la probabilitat de no rebutjar la hipòtesi nul.la de significació individual d'un paràmetre pel període mostral total, per la qual cosa es pot considerar com estadísticament no significatiu, encara que en realitat ho pugui ser per cada una de les submostres per separat.

Si es conclou que hi ha canvi estructural, per a estimar correctament el model s'haurà d'especificar una regressió diferent per a cadascuna de les submostres detectades.

2.5.2. Contrast de Chow de permanència estructural

La hipòtesi nul.la és

H_0: permanència estructural

El test consisteix en comparar les SQR de la regressió per a tota la grandària mostral amb les SQR de les regressions per a cadascuna de les submostres fixades.
Suposant que únicament hi ha dues submostres, el contrast de Chow es fa de la següent manera:

1) Es descompon el període mostral en les dues submostres possibles detectades.

2) S'especifiquen les 3 regressions següents:
   a. Una per a tota la grandària mostral: $Y_t = X\beta_T + u_t \quad t = 1,2,...,T$
   b. Una segona per a la primera submostra: $Y_i = X\beta_1 + u_i \quad i = 1,2,...,T_1$
   c. Una tercera per a la segona submostra: $Y_j = X\beta_2 + u_j \quad j = 1,2,...,T_2$

3) S'estimen per MQO cadascuna de les 3 regressions anteriors i es calculen les SQR associades. Així s'obtenen les SQR associades a tot el període mostral (SQR_T), les associades a la primera submostra (SQR_1) i a la segona submostra (SQR_2)

Una vegada s'ha fet això es construeix l'estadístic de prova, que es defineix com

$$F = \frac{\dfrac{SQR_T - (SQR_1 + SQR_2)}{k}}{\dfrac{SQR_1 + SQR_2}{T-2k}} \sim F(k, T-2k)$$

Si el valor de l'estadístic de prova, F, és superior al valor crític en taules de la F de Snedecor, es rebutja la hipòtesi nul.la, és a dir, es tenen indicis que en la població hi ha canvi estructural.
Si hi haguessin 3 submostres, l'estadístic de prova està determinat per l'expressió següent:

$$F = \frac{\dfrac{SQR_T - (SQR_1 + SQR_2 + SQR_3)}{2k}}{\dfrac{SQR_1 + SQR_2 + SQR_3}{T-3k}} \sim F(2k, T-3k)$$

2.5.3. Limitacions del Contrast de Chow

A continuació es senyalen 4 limitacions o inconvenients del Test de Chow per a detectar el possible incompliment de la hipòtesi de permanència estructural:

1) Es pot detectar un canvi estructural espuri, que té lloc quan es detecta un canvi estructural com a conseqüència d'una especificació errònia del model en la seva part determinista.
2) És sensible a l'existència d'heterocedasticitat en el terme de pertorbació. Per tant, s'ha de corregir abans d'aplicar el test.
3) Perd potència (probabilitat d'acceptar la hipòtesi alternativa si és certa) a mesura que el punt on es produeix el canvi estructural s'apropa a un dels extrems de la mostra.
4) De vegades, no es disposa d'informació suficient per a conèixer el punt exacte en què es produeix el possible canvi estructural. Aquest inconvenient, es dóna principalment quan es treballa amb dades de sèrie temporal. En aquest cas és recomanable utilitzar altres mètodes com el **Contrast de Hansen**.

2.6. Formes funcionals dels models de regressió

2.6.1. Presentació

Fins ara hem analitzat models lineals tant en els paràmetres com en les variables. Però per aplicar la teoria de la regressió per MQO només fa falta la linealitat en els paràmetres.

En aquest apartat es tracten alguns models utilitzats en l'anàlisi economètrica que bé són lineals en els paràmetres o es poden linealitzar realitzant les transformacions adients. Aquests models són els següents:

1. Models lineals en logaritmes o d'elasticitat constant
2. Models semilogarítmics
3. Models recíprocs

2.6.2. Models lineals en logaritmes o d'elasticitat constant

Són models de la forma:

$$y_i = \beta_0 \cdot x_i^{\beta_1} \cdot e^{u_i} \qquad \forall i = 1, 2, \dots, n$$

Si es prenen logaritmes es transformen en

$$\ln y_i = \ln\beta_0 + \beta_1 \cdot \ln x_i + u_i = \alpha_0 + \beta_1 \cdot \ln x_i + u_i \qquad \forall i = 1, 2, \dots, n$$

Es defineixen les noves variables

$$y_i^* = \ln y_i \qquad x_i^* = \ln x_i \qquad \forall i = 1, 2, \dots, n$$

El model transformat serà

$$y_i^* = \alpha_0 + \beta_1 \cdot x_i^* + u_i \qquad \forall i = 1, 2, \dots, n$$

Aquest model pot ser estimat per MQO. La característica més important d'aquests models és que la pendent estimada $\widehat{\beta_1}$ és una estimació de l'elasticitat de la variable dependent y respecte a la variable explicativa.

$$\hat{\beta}_1 = \frac{\partial \hat{y}^*}{\partial x^*} = \frac{\partial\left(\widehat{\ln y}\right)}{\partial(\ln x)} = \frac{\frac{\partial \hat{y}}{\hat{y}}}{\frac{\partial x}{x}}$$

Aquests models són útils quan es suposa que hi ha un coeficient d'elasticitat constant entre les variables.

2.6.3. Models semilogarítmics

Per models semilogarítmics ens referim a construccions en les que la transformació logarítmica afecta només a una de les variables, la dependent o la explicativa.

Els models de la forma

$$y_i = \beta_1 \cdot e^{(\beta_2 \cdot x_i + u_i)} \qquad \forall i = 1, 2, \dots, n$$

que alternativament es poden escriure com

$$\ln y_i = \alpha_1 + \beta_2 \cdot x_i + u_i \qquad \forall i = 1, 2, \dots, n$$

es diuen **models log-lin** i transformant com abans només la variable dependent es poden escriure com

$$y_i^* = \alpha_1 + \beta_2 \cdot x_i + u_i \qquad \forall i = 1, 2, \dots, n$$

Per aquests models la pendent estimada $\widehat{\beta_1}$ és una estimació del canvi relatiu en la variable dependent y respecte un canvi absolut en la variable explicativa x

$$\hat{\beta}_1 = \frac{\partial \hat{y}^*}{\partial x} = \frac{\partial\left(\widehat{\ln y}\right)}{\partial x} = \frac{\frac{\partial \hat{y}}{\hat{y}}}{\partial x} = \frac{\text{Canvi relatiu de } \hat{y}}{\text{Canvi absolut de x}}$$

Aquests models també reben el nom de models de creixement i són útils en els casos que la variable explicativa x sigui una variable de tendencia temporal.

Els models de la forma

$$y_i = \alpha_1 + \beta_2 \cdot \ln x_i + u_i \qquad \forall i = 1, 2, \dots, n$$

reben el nom de **models lin-log** que una vegada linealitzats, s’expressen com

$$y_i = \alpha_1 + \beta_2 \cdot x_i^* + u_i \qquad \forall i = 1, 2, \dots, n$$

Per aquests models la pendent estimada $\widehat{\beta_1}$ és una estimació del canvi absolut en el valor esperat de la variable dependent y respecte una variació relativa en la variable explicativa x.

$$\hat{\beta}_1 = \frac{\partial \hat{y}}{\partial x^*} = \frac{\partial \hat{y}}{\partial (\ln x)} = \frac{\partial \hat{y}}{\frac{\partial x}{x}} = \frac{\text{Canvi absolut de } \hat{y}}{\text{Canvi relatiu en x}}$$

Aquesta classe de models s’utilitza en els casos en els que s’espera que un canvi relatiu en la variable explicativa x produeixi un canvi absolut en la variable dependent y.

2.6.4. Models recíprocs

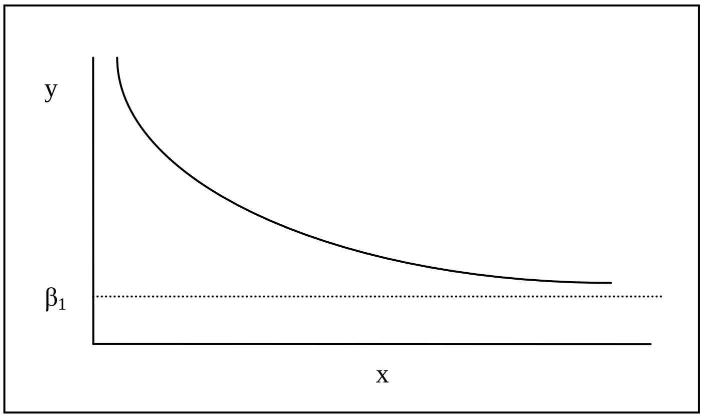
Els models recíprocs són de la forma

$$y_i = \beta_1 + \beta_2 \cdot \frac{1}{x_i} + u_i \qquad \forall i = 1, 2, \dots, n$$

Aquests models es caracteritzen pel fet que a mesura que la variable explicativa x augmenta indefinidament, la variable dependent y s’aproxima al valor asimptòtic β_1

La pendent d’aquests models és

$$\text{pendent} = \frac{\partial y}{\partial x} = -\beta_2 \cdot \frac{1}{x_i^2}$$



2.7. *Variables fictícies*

2.7.1. Presentació

En l'especificació del model de regressió lineal hem considerat fins ara que els regressors considerats són de caràcter quantitatiu.

Freqüentment a l'anàlisi de regressió, la variable dependent està afectada no només per variables quantitatives sinó també per variables qualitatives com poden ser el sexe, l'estat civil o la localització geogràfica.

El tractament d'aquestes situacions requereix la construcció d'unes variables artificials que permetin quantificar les variacions qualitatives dels diferents factors considerats. En termes economètrics aquestes variables es denominen variables fictícies o binàries (dummies) i només prenen els valors 0 i 1.

Els models de regressió que només inclouen com a regressors variables fictícies es diuen **ANOVA** (analysis of variance):

$$y_i = \beta_1 + \beta_2 \cdot D_i + u_i$$

on y representa la despesa anual en menjar i D pren el valor 0 si l'observació i correspon a una dona i el valor 1 si correspon a un home.

Els models ANOVA són freqüents en camps com la sociologia, psicologia o els estudis de mercat, però en economia la majoria de models de regressió inclouen variables explicatives quantitatives i qualitatives i es denominen **ANCOVA** (analysis of covariance):

$$y_i = \beta_1 + \beta_2 \cdot D_i + \beta_3 \cdot x_i + u_i$$

on y representa la despesa anual en menjar, x la renta disponible i D pren el valor 0 si l'observació i correspon a una dona i el valor 1 si correspon a un home.

2.7.2. Variables qualitatives politòmiques

Si la variable explicativa que es vol utilitzar té més de 2 categories codificades amb valors que van des de 1 fins a n, per exemple, en el cas de l'estat civil, 1 per solters, 2 per casats i 3 per la resta (vidus, separats, ...), no podem utilitzar directament aquesta codificació ja que imposaríem un ordre i una proporcionalitat que no desitgem.

En aquest cas es defineixen n-1 variables fictícies que prenen valors 0 o 1. A l'exemple de l'estat civil es definirien les 2 variables fictícies següents:

$$D_2 = \begin{cases} 1 & \text{casats} \\ 0 & \text{altres} \end{cases} \qquad D_3 = \begin{cases} 1 & \text{vidus, separats, ...} \\ 0 & \text{altres} \end{cases}$$

Els individus solters tindran el valor 0 a les 2 variables fictícies. La categoria que no s'utilitza per definir les variables fictícies s'anomena categoria de referència

2.7.3. Ús de variables fictícies en contrastos sobre discriminació

Per exemple, si sospitem que el salari d'un treballador depén de la seva antiguitat podem escriure el model següent:

$$w_i = \alpha + \beta e_i + u_i$$

Si volem introduir en el model la dependència del salari en si és home o dona, ho podem fer com en l'apartat anterior introduint una variable fictícia D que pren el valor 0 si l'observació i correspon al salari d'una dona i el valor 1 si la observació i correspon al salari d'un home. Aleshores, el model es defineix de la següent forma:

$$w_i = \alpha + \beta e_i + \gamma D_i + u_i$$

En aquest model, per veure si hi ha discriminació salarial per sexe es pot fer el contrast

$$H_0 : \gamma = 0$$

Aquest contrast serveix per veure si hi ha discriminació al entrar en el mercat laboral, però no en l'evolució professional. En efecte,

$E(w_i / e_i ; D_i = 1) = \alpha + \beta e_i + \gamma$ (salari promig que té un home amb e_i anys d'experiència)

$E(w_i / e_i ; D_i = 0) = \alpha + \beta e_i$ (promig que té una dona amb e_i anys d'experiència)

Si volem estudiar la discriminació en l'evolució professional, es pot proposar aquest model

$$w_i = \alpha + \beta e_i + \delta e_i D_i + u_i$$

Llavors,

$$E(w_i / e_i ; D_i = 1) = \alpha + \beta e_i + \delta e_i = \alpha + (\beta + \delta) e_i$$
$$E(w_i / e_i ; D_i = 0) = \alpha + \beta e_i$$

Si no volem condicionar cap tipus de discriminació, es pot utilitzar el següent model:

$$w_i = \alpha + \gamma D_i + \beta e_i + \delta e_i D_i + u_i$$

Aleshores,

$$E(w_i / e_i ; D_i = 1) = (\alpha + \gamma) + (\beta + \delta) e_i$$
$$E(w_i / e_i ; D_i = 0) = \alpha + \beta e_i$$

El contrast de no discriminació en aquest cas és $H_0 : \gamma = \delta = 0$. S'ha d'utilitzar l'estadístic F,

$$F = \frac{\frac{SQR_R - SQR}{2}}{\frac{SQR}{n-4}}$$

2.7.4. Ús de variables fictícies per tractar dades atípiques

Les dades atípiques poden distorsionar els resultats de l'estimació. Si es desitja conservar-les en la base de dades però eliminar-ne la influència sobre les estimacions, es poden emprar variables fictícies.

Si es disposa de N observacions i l'observació i_0-èsima és una dada atípica, es pot definir la següent variable fictícia:

$$D_i = \begin{cases} 1 & si \ \ i = i_0 \\ 0 & si \ \ i \neq i_0 \end{cases}$$

Si s'introdueix aquesta variable fictícia, s'elimina la influència d'aquesta observació. L'avantatge d'aquest procediment és que el residu de l'estimació serà zero per aquesta observació, ja que el paràmetre que acompanya la variable fictícia només té efecte per aquesta observació i s'ajusta sota el criteri MQO.

En conseqüència, el model ja ajustarà exactament aquesta observació i deixarà d'influir en l'estimació de la resta de paràmetres i els seus estadístics corresponents. Igualment, s'haurà aïllat el seu efecte en aquest únic paràmetre i es podrà estimar l'impacte d'aquesta observació o esdeveniment concret.

2.7.5. Ús de variables fictícies a l'anàlisi estacional

Quan es treballa amb dades en forma de sèries temporals moltes vegades s'ha d'eliminar el component estacional.

Si estudiem el comportament de la despesa en turisme de les unitats familiars en funció de la renta salarial i tenim observacions mensuals es pot construir el següent model utilitzant variables fictícies:

$$GT_i = \beta_1 + \beta_2 D_{i2} + \beta_3 D_{i3} + \beta_4 D_{i4} + \beta_5 Y_i + u_i$$

on les variables fictícies es defineixen de la següent manera:

$$D_{i2} = \begin{cases} 1 & si \ i \in \{segon \ trimestre\} \\ 0 & si \ i \notin \{segon \ trimestre\} \end{cases} \qquad D_{i3} = \begin{cases} 1 & si \ i \in \{tercer \ trimestre\} \\ 0 & si \ i \notin \{tercer \ trimestre\} \end{cases}$$

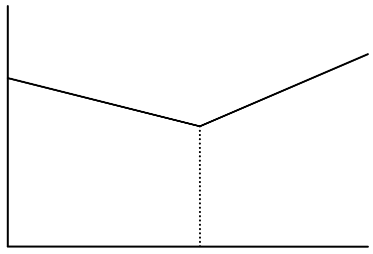
$$D_{i4} = \begin{cases} 1 & si \ i \in \{quart \ trimestre\} \\ 0 & si \ i \notin \{quart \ trimestre\} \end{cases}$$

No fa falta afegir una variable dummy pel primer trimestre per que ja tenim la constant que capta l'efecte d'aquestes observacions.

Els efectes estacionals diferencials sobre el primer trimestre venen recollits pels paràmetres $\beta_2, \beta_3, \beta_4$ i per tant, es poden fer contrastos d'hipòtesis sobre ells.

2.7.6. Ús de variables fictícies a la regressió lineal per trams

Suposem que volem estimar una recta on a partir d'un valor determinat de X, la pendent es fa més gran. Les variables dummy ens poden ajudar a fer-ho.



X_E

El valor de X on la pendent canvia d'inclinació es suposa conegut X_E. Aleshores es construeix el següent model:

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 (X_i - X_E) D_i + u_i$$

on la variable dummy es defineix $D_i = \begin{cases} 1 & si\ X_i > X_E \\ 0 & si\ X_i \leq X_E \end{cases}$

Aleshores pel primer tram tindrem,

$$E\left(Y_i / D_i = 0, X_i, X_E \right) = \beta_1 + \beta_2 X_i$$

i pel segon tram tindrem,

$$E\left(Y_i / D_i = 1, X_i, X_E \right) = (\beta_1 - \beta_3 X_E) + (\beta_2 + \beta_3) X_i$$

2.7.7. El model d'efectes fixos (panel data)

Quan tenim dades sobre N individus al llarg de T períodes, diem que tenim un panell de dades. Es vol especificar un model de regressió amb el total de N·T observacions i volem eliminar els efectes temporals i de discrepància entre les unitats.

Per fer-ho es poden definir N-1 i T-1 variables fictícies que identifiquin totes les unitats (excepte la de referència) i tots els períodes (excepte el de referència) i incloure-les additivament en el model de regressió. Aquest model s'anomena model d'efectes fixos.

Els paràmetres associats a les variables fictícies s'interpretaran com els efectes singulars deguts al període (o unitat), respecte al període (o unitat) de referència.

El principal problema del model d'efectes fixos és que la inclusió de les variables fictícies fa augmentar el nombre de paràmetres que cal estimar en el model i, en conseqüència, el nombre de graus de llibertat disminueix.

2.8. Variables endògenes qualitatives

2.8.1. Presentació

Quan volem explicar un event de natura qualitativa, es construeixen models d'elecció discreta. En aquests models la variable dependent Y pot prendre els valors 0 i 1.

Un exemple de model d'elecció discreta és el que vam veure al tema 1 en l'apartat d'aplicacions de l'econometria al risc de crèdit on s'intenta predir la probabilitat de que un acreditat entri en mora a partir de la puntuació del seu scoring o rating.

Els models d'aquesta classe més utilitzats són els següents:

1. Model de probabilitat lineal
2. Model probit
3. Model logit

2.8.2. Model de probabilitat lineal

El model de probabilitat lineal expressa la variable dicotòmica dependent Y com una funció lineal de les variables explicatives:

$$\mathrm{Y_i = \beta_1 + \beta_2 \cdot X_{2i} + \beta_3 \cdot X_{3i} + \cdots + \beta_k \cdot X_{ki} + u_i} \qquad \forall \mathrm{i} = 1, 2, \ldots, \mathrm{n}$$

Com Y només pot prendre els valors 0 i 1 i l'esperança de u és 0, la probabilitat d'èxit és

$$\mathrm{P_i = P(Y_i = 1) = E\left({}^{Y_i}/_{X_{2i}\ldots X_{ki}}\right)} = \beta_1 + \beta_2 \cdot X_{2i} + \cdots + \beta_k \cdot \mathrm{X_{ki}}$$

En aquest model, els coeficients de les pendents s'interpreten com el canvi en la probabilitat d'èxit donat un canvi unitari de la corresponent variable explicativa.

Aquest model es pot estimar per MQO, però presenta les següents limitacions:

1. Encara que Y només pot prendre els valors 0 i 1, no hi ha garanties que els valors estimats de Y es trobin a l'interval (0,1) ja que el model és lineal.
2. Com Y és una variable binària, u no té distribució normal sinó una distribució Bernoulli.
3. El terme d'error és heterocedàstic.
4. Com que Y només pot prendre els valors 0 i 1 el valor de R2 no té gaire sentit.
5. La probabilitat d'èxit canvia linealment amb el valor de la variable explicativa. Aquest és el principal problema, perquè aquest supòsit no és realista.

2.8.3. Models probit i logit

Els models probit i logit són models de regressió no lineal que han estat disenyats específicament per variables dependents binàries. Es tracta d'adoptar una formulació no lineal que obligui a que els valors estimats estiguin entre 0 i 1 i a més presentin un creixement no lineal.

El model **logit** utilitza la funció logística

$$\mathrm{f(z) = \frac{1}{1 + e^{-z}} = \frac{e^z}{1 + e^z}}$$

El model de regressió logit és de la forma

$$\mathrm{Y_i = f(Z_i) + u_i} \qquad \text{on } \mathrm{Z_i = \beta_1 + \beta_2 \cdot X_{2i} + \cdots + \beta_k \cdot X_{ki}} \qquad \forall \mathrm{i} = 1, 2, \ldots, \mathrm{n}$$

Igual que en el cas del model de probabilitat lineal, la probabilitat d'èxit es pot posar en funció de l'esperança de Y condicionada al valor de les variables explicatives

$$\mathrm{P_i = P(Y_i = 1) = E\left({}^{Y_i}/_{X_{2i}\ldots X_{ki}}\right) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 \cdot X_{2i} + \cdots + \beta_k \cdot X_{ki})}}}$$

El model de regressió **probit** es basa en la distribució de probabilitat acumulada d'una normal standard

$$f(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{s^2}{2}} ds$$

La probabilitat d'èxit en aquest cas pren la següent forma

$$P_i = P(Y_i = 1) = E\left({Y_i}/{X_{2i} \ldots X_{ki}} \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta_1 + \beta_2 \cdot X_{2i} + \cdots + \beta_k \cdot X_{ki}} e^{-\frac{s^2}{2}} ds$$

2.8.4. Estimació per MQO del model logit

El model logit es pot transformar de tal manera que sigui possible estimar-lo usant el mètode MQO.

Com s'ha vist abans, el model logit pren una forma d'aquest tipus:

$$Y_i = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_{2i} + \ldots + \beta_k X_{Ki})}} + u_i$$

En aquest model la probabilitat d'èxit és:

$$P_i = P(Y_i = 1) = E\left({Y_i}/{X_2 \ldots X_K} \right) = \frac{1}{1 + e^{-(\beta_1 + \beta_2 X_{2i} + \ldots + \beta_k X_{ki})}} \quad \forall i$$

Per aplicar el mètode de mínims quadrats es fa la següent transformació

$$L_i = \ln\left(\frac{P_i}{1 - P_i} \right) = \beta_1 + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + v_i \quad \forall i$$

2.9. *Errors d'especificació*

2.9.1. Presentació

Donat el model

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \ldots + \beta_k x_{ki} + u_i \quad i = 1, \ldots, n$$
$$u_i \sim iiN\left(0, \sigma^2\right)$$

Ens preguntem per els errors que hem pogut cometre en l'especificació de la part sistemàtica del model. Aquests es poden resumir en els següents:

a) Inclusió de variables irrellevants: Es comet aquest error quan s'afegeixen variables explicatives que són irrellevants en l'explicació de la variable endògena.

b) Omissió de variables rellevants: Es comet quan s'inclouen menys variables explicatives de les que s'haurien d'incloure.

c) Forma funcional errònia: Es dóna quan la funció especificada entre la variable endògena i les variables explicatives no és la correcta.

2.9.2. Inclusió de variables irrellevants

Com afecta al estimador MQO la inclusió de variables irrellevants?

Model real $\begin{array}{c} Y = X_1\beta_1 + u \\ u \sim \text{supòsits clàssics} \end{array}$ $\quad X_1 \quad n\,x\,k_1$

Model estimat $\begin{array}{c} Y = X_1\beta_1 + X_2\beta_2 + u \\ u \sim \text{supòsits clàssics} \end{array}$ $\quad X_2 \quad n\,x\,k_2$

En el model estimat s'han inclós k_2 variables irrellevants. L'estimador MQO serà

$$\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \left(\begin{bmatrix} X_1 & X_2 \end{bmatrix}'\begin{bmatrix} X_1 & X_2 \end{bmatrix}\right)^{-1}\begin{bmatrix} X_1 & X_2 \end{bmatrix}'Y = \begin{bmatrix} X'_1 X_1 & X'_1 X_2 \\ X'_2 X_1 & X'_2 X_2 \end{bmatrix}^{-1}\begin{bmatrix} X'_1 Y \\ X'_2 Y \end{bmatrix}$$

Per tant,

$$E\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = E\left(\begin{bmatrix} X'_1 X_1 & X'_1 X_2 \\ X'_2 X_1 & X'_2 X_2 \end{bmatrix}^{-1}\begin{bmatrix} X'_1 Y \\ X'_2 Y \end{bmatrix}\right) =$$

$$= E\left(\begin{bmatrix} X'_1 X_1 & X'_1 X_2 \\ X'_2 X_1 & X'_2 X_2 \end{bmatrix}^{-1}\begin{bmatrix} X'_1 (X_1\beta_1 + u) \\ X'_2 (X_1\beta_1 + u) \end{bmatrix}\right) =$$

$$= \begin{bmatrix} X'_1 X_1 & X'_1 X_2 \\ X'_2 X_1 & X'_2 X_2 \end{bmatrix}^{-1}\begin{bmatrix} X'_1 X_1 \\ X'_2 X_1 \end{bmatrix}\beta_1 = \begin{bmatrix} I_{K_1} \\ 0 \end{bmatrix}\beta_1 = \begin{bmatrix} \beta_1 \\ 0 \end{bmatrix}$$

Això ens diu que no tenim biaix

$$\operatorname{var}\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \sigma^2\begin{bmatrix} X'_1 X_1 & X'_1 X_2 \\ X'_2 X_1 & X'_2 X_2 \end{bmatrix}^{-1} = \sigma^2\begin{bmatrix} \left[X'_1 X_1 - X'_1 X_2 (X'_2 X_2)^{-1} X'_2 X_1\right]^{-1} & . \\ . & . \end{bmatrix}$$

Per tant, si comparem la variància del model estimat i del model real,

$$\begin{cases} \operatorname{var}(\hat{\beta}_1) \text{model estimat} = \sigma^2\left[X'_1 X_1 - X'_1 X_2 (X'_2 X_2)^{-1} X'_2 X_1\right]^{-1} \\ \operatorname{var}(\hat{\beta}_1) \text{model real} = \sigma^2\left[X'_1 X_1\right]^{-1} \end{cases}$$

La conclusió és que l'estimador del model estimat és menys eficient que si s'utilitza el model real (excepte si $X'_1 X_2 = 0$). El fet que la variància del model estimat sigui més gran que la del model real provoca que l'interval de confiança sigui més gran, l'error de Tipus II en el contrast sigui més alt i les t de Student siguin més petites.

Per altra banda, l'estimador MQO continúa sent consistent i sense biaix. També es pot demostrar que l'estimador MQO de la variància del terme de pertorbació que s'obtindrà no té biaix.

2.9.3. Exclusió de variables rellevants

En aquest cas el model real té més variables que el model que estimem,

Model real $\begin{array}{l} Y = X_1\beta_1 + X_2\beta_2 + u \quad X_1\ n\,x\,k_1, \quad X_2\ n\,x\,k_2, \quad k_1 + k_2 = k \\ u \sim \text{supòsits clàssics} \end{array}$

Model estimat $\begin{array}{l} Y = X_1\beta_1 + u^* \quad u^* = X_2\beta_2 + u \\ u \sim \text{supòsits clàssics} \end{array}$

Les implicacions per l'estimador MQO són les següents:

$$\hat{\beta}_1 = (X'_1 X_1)^{-1} X'_1 Y$$

$$E[\hat{\beta}_1] = E[(X'_1 X_1)^{-1} X'_1 Y] = E[(X'_1 X_1)^{-1} X'_1 (X_1\beta_1 + X_2\beta_2 + u)] =$$
$$= \beta_1 + (X'_1 X_1)^{-1} X'_1 X_2\beta_2$$

Per tant, tindrem biaix excepte si $X'_1 X_2 = 0$. En general no serà consistent excepte si $X'_1 X_2 = 0$

Es pot demostrar que els elements de la diagonal principal de la matriu de variàncies i covariàncies són més petits o com a màxim iguals, que els que s'haurien obtingut pels mateixos estimadors amb el model correctament especificat, però no cal mirar la eficiència perquè aquest estimador té biaix.

Per altra banda, l'estadístic de contrast t no serà vàlid perquè tenim biaix i no podem fer la construcció com en el cas sense biaix.

Com els estimadors són esbiaixats i el biaix no tendeix a zero quan augmenta la grandària, són estimadors inconsistents.

També es pot demostrar que l'estimador MQO de la variància és esbiaixat.

2.9.4. Forma funcional errònia

Aquest error es dóna quan la veritable relació és diferent de l'especificada (lineal, quadràtica, cúbica, exponencial, etc…).

Una especificació incorrecta es pot considerar, en alguns casos, com un error d'especificació assimilable a l'omissió de variables rellevants. Les conseqüències per tant són que els estimadors són esbiaixats i inconsistents.

En general, especificar una relació equivocada pot conduir a obtenir un terme de pertorbació amb heterocedasticitat i/o autocorrelació i que la distribució del terme de pertorbació no sigui normal.

Per detectar una forma funcional incorrecta es pot utilitzar el contrast reset.

$$H_0 : \text{Forma funcional correcta}$$

Aquest contrast es fa de la següent manera:

1) S'estima per MQO el model $y_i = \beta_1 + \beta_2 x_{2i} + ... + \beta_k x_{ki} + u_i$

2) S'estima per MQO el model auxiliar $y_i = \delta_1 + \delta_2 x_{2i} + ... + \delta_k x_{ki} + \gamma\hat{y}_i^2 + v_i$
   on s'ha afegit com a regressor la variable endògena ajustada al pas anterior elevada al quadrat

3) Es contrasta si el coeficient γ estimat és estadísticament diferent de zero, cas en què es rebutja la hipòtesi nul.la.

2.10. Col.linealitat

2.10.1. Presentació

La col.linealitat és un problema de les dades de les variables explicatives de la mostra. Es dóna quan les variables explicatives estan relacionades linealment.

Hi ha diferents tipus de col.linealitat:

-Col.linealitat perfecta: Linealitat exacta entre dos o més regressors.

Un exemple de col.linealitat perfecta és la “trampa de les variables fictícies”. Consisteix en introduir dues variables fictícies en comptes d’una.

$$D_{i1}\begin{cases}1 & Home\\ 0 & Dona\end{cases} \qquad D_{i2}\begin{cases}0 & Home\\ 1 & Dona\end{cases}$$

$$w_i = \beta_1 + \beta_2 x_i + \beta_3 D_{1i} + \beta_4 D_{2i} + u_i$$

Com $D_{1i} + D_{2i} = 1$, en aquest cas hi ha col.linealitat perfecta.
Amb col.linealitat perfecta no es pot trobar l’estimador MQO perquè el determinant de la matriu X és zero i per tant no es pot invertir.
La col.linealitat perfecta es detecta fàcilment.

-Col.linealitat alta:

La col.linealitat alta apareix sovint en models econòmics on les variables explicatives estan fortament correlacionades sense arribar a ser una combinació lineal.

La col.linealitat alta no afecta gaire a les propietats del estimador MQO ja que les propietats del estimador depenen de les pertorbacions.

La col.linealitat afecta sobretot a la variància dels estimadors, i per tant a la precisió i als intervals de confiança.

2.10.2. Problemes davant la presència de col.linealitat

-Col.linealitat perfecta :

En el cas de col.linealitat perfecta, no podem estimar els paràmetres de forma única. Per exemple, seguint amb el cas anterior,

$$\begin{aligned}
w_i &= \beta_1 + \beta_2 x_i + \beta_3 D_{1i} + \beta_4 D_{2i} + u_i\\
w_i &= \beta_1 + \beta_2 x_i + \beta_3 D_{1i} + \beta_4 (1 - D_{1i}) + u_i\\
w_i &= \beta_1 + \beta_4 + \beta_2 x_i + (\beta_3 - \beta_4) D_{1i} + u_i\\
w_i &= \gamma_1 + \beta_2 x_i + \gamma_2 D_{1i} + u_i
\end{aligned}$$

D’aquí es treuen només 3 paràmetres, per tant hi ha un paràmetre que queda indeterminat i la solució no és única.

-Col.linealitat alta:

En aquest cas els paràmetres del model no tindran la mateixa interpretació que tindrien si no existís col.linealitat.

Donat el model $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + ... + \beta_k x_{ki} + u_i$

Si no hi ha col.linealitat $\beta_2 = \dfrac{dy}{dx_2}$ es pot interpretar com la sensibilitat de y respecte x_2.

Si existeix col.linealitat, β_2 ja no es pot interpretar com la derivada total de y respecte x_2, ja que una variació en x_2 també afectarà a x_3. Per tant β_2 ara només serà la derivada parcial.

Ara suposem que x_j és la variable afectada per col.linealitat i veurem perquè $\operatorname{var}\left(\hat{\beta}_j\right)$ és gran.
Particionem la matriu X de la següent forma:

$$X = \begin{bmatrix} x_j & W \end{bmatrix}$$

Per tant,

$$(X'X)^{-1} = \begin{bmatrix} x'_j x_j & x'_j W \\ W' x_j & W'W \end{bmatrix}^{-1}$$

aplicant la fórmula de la inversa d'una matriu particionada,

$$\begin{aligned} \operatorname{var}\left(\hat{\beta}_j\right) &= \sigma^2 (X'X)^{-1}_{11} = \sigma^2 \left[x'_j x_j - x'_j W (W'W)^{-1} W' x_j\right]^{-1} = \\ &= \sigma^2 \left[x'_j \left(I - W(W'W)^{-1} W'\right) x_j\right]^{-1} = \sigma^2 \left[x'_j M_j x_j\right]^{-1} \end{aligned}$$

on $M_j \equiv I - W(W'W)^{-1} W'$ és una matriu idempotent
Recordem que en el model

$$Y = X\beta + u \rightarrow SQR = \hat{u}'\hat{u} = Y'MY \quad on\ M \equiv \left(I - X(X'X)^{-1} X'\right)$$

Així paral.lelament $x'_j M_j x_j = SQR_j$ (SQR de la regressió auxiliar $x_j = W\alpha + \nu$)

Llavors :

$$\operatorname{var}\left(\hat{\beta}_j\right) = \sigma^2 \left[x'_j M_j x_j\right]^{-1} = \sigma^2 \left[SQR_j\right]^{-1} = \sigma^2 \frac{1}{SQT_j} \frac{1}{\left(1 - R_j^2\right)}$$

on $SQT_j = \sum_i \left(x_{ji} - \bar{x}_j\right)^2, \qquad R_j^2 \equiv 1 - \dfrac{SQR_j}{SQT_j}$

Com R_j^2 és el coeficient de determinació de la regressió auxiliar i x_j presenta col.linealitat, R_j^2 serà proper a 1 i per tant $\operatorname{var}\left(\hat{\beta}_j\right)$ serà gran.
Això provoca poca precisió en l'estimació i que els estadístics t siguin petits. Per tant els erros de tipus II augmenten (es pot considerar que un paràmetre no és estadísticament significatiu quan relament sí que ho és).
A més, l'estabilitat de l'estimació és més petita, petits canvis en la mostra poden fer canviar moltíssim l'estimació dels paràmetres.

2.10.3. Com detectar la presència de col.linealitat

Factors que fan sospitar la presència de col.linealitat:

-Un coeficient de R^2 elevat amb pocs estadístics t-Student significatius.

-Petits canvis en les dades produeixen grans variacions en els estimadors dels paràmetres

-Variables que per la teoria haurien d'entrar en el model però no surten significatives (estadístics petits).

-Els coeficients poden tenir signes oposats als esperats o no tenir la magnitud esperada.

-Un element indicador de la presència de col.linealitat és el determinant de la matriu de coeficients de correlació simple.

$$(\text{multicol.linealitat perfecta})\, 0 \leq \begin{vmatrix} 1 & r_{23} & \dots & r_{2k} \\ r_{23} & 1 & \dots & r_{3k} \\ \dots & \dots & \dots & \dots \\ r_{2k} & \dots & \dots & 1 \end{vmatrix} \leq 1\, (\text{absència total de multicol.linealitat})$$

-També podem utilitzar l'expressió que hem trobat abans,

$$\operatorname{var}\left(\hat{\beta}_j\right) = \sigma^2 \frac{1}{SQT_j} \frac{1}{\left(1 - R_j^2\right)}$$

Una R_j^2 alta és indicador de col.linealitat.

2.10.4. Solucions a la col.linealitat

No hi ha regles generals. Es pot intentar fer el següent:

-Una solució al problema de col.linealitat és treure les variables que presenten col.linealitat. Però aleshores incorrem en errors d'especificació. Per tant com hem vist abans això farà que l'estimador tingui biaix.

-Una altra solució és intentar afegir més dades, encara que no sempre es poden trobar.

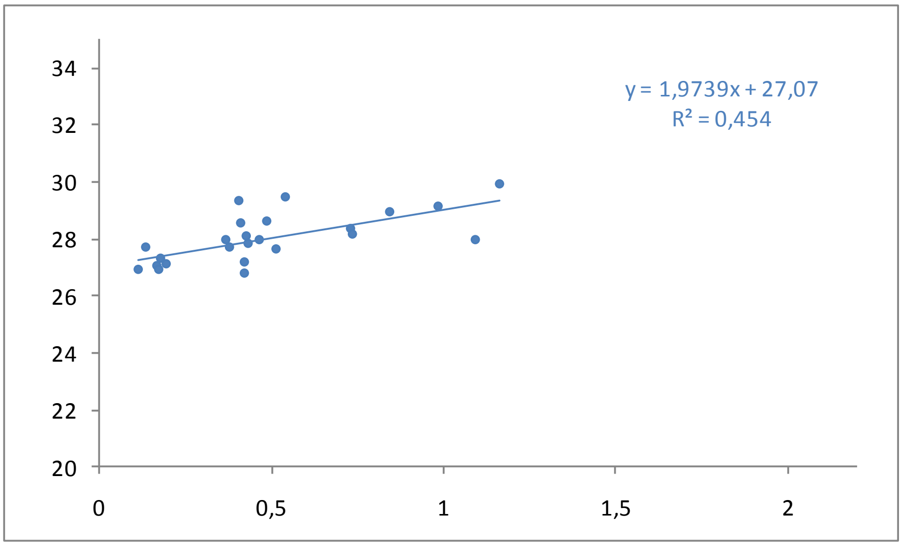
-La millor solució consisteix en compensar els efectes de la col.linealitat. Si la col.linealitat fa la variància gran, s'ha de fer alguna cosa per reduir la variància.
Una forma de reduir la variància és imposar que l'estimador compleixi restriccions. Per exemple, si sabem que un coeficient està en un rang determinat ho imposem. Si estem segurs de les restriccions que imposem, això és un procediment correcte. Si no estem segurs, reduirem la variància però afegirem biaix i per tant el test de significació global no es farà bé.

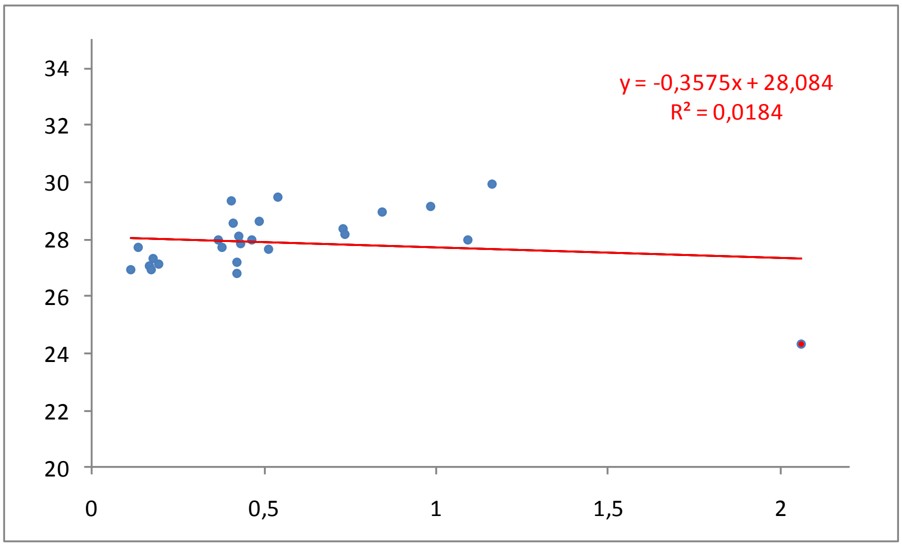
2.11. Presència de valors estranys (outliers)

2.11.1. Presentació

Els outliers són observacions individuals atípiques. Són rares en el sentit que són observacions que tenen un procés generador diferent del de la resta d'observacions.

L'anàlisi d'aquestes observacions és important ja que poden tenir un efecte molt gran en el resultat de l'estimació i en les variàncies dels estimadors com es mostra en els següents gràfics on l'única diferència és afegir un outlier sobre una mostra de 25 observacions.





2.11.2. Leverage d'una observació

Una observació presenta leverage (palanquejament) si està molt allunyada de la resta d'observacions pel que fa a les coordenades de les variables explicatives.

Hi ha un leverage associat a cada observació, h_{ii} , que és l'element i-èsim de la diagonal principal de la matriu H que es defineix de la següent manera:

$$H = X\left(X'X\right)^{-1}X'$$

El leverage h_{ii} també es pot escriure de la següent forma:

$$h_{ii} = \frac{1}{N}\left(1 + d_i\right) \quad \forall i = 1,\ldots,n$$

d_i és la distància de Mahalanobis, que ve donada per la següent expressió:

$$d_i = \left(X_i - \overline{X}\right)' S^{-1}\left(X_i - \overline{X}\right) \quad \forall i = 1,\ldots,n$$

on $\left(X_i - \bar{X}\right)' = \left(X_{1i} - \bar{X}_1, X_{2i} - \bar{X}_2, ..., X_{ki} - \bar{X}_k\right)$ i $S = \frac{X'X}{n}$

Es considera que una observació té leverage si $h_{ii} \geq 2\bar{h}$ $\quad \bar{h} = \frac{\sum_{i=1}^{N} h_{ii}}{n} = \frac{k}{n}$

El leverage és important ja que es pot demostrar que $Var\left(\hat{y}_i\right) = \sigma_u^2 h_{ii}$. Per tant, si una observació té associat un leverage alt, la variància de l'ajust obtingut per aquesta observació és gran i per tant el valor ajustat pel model és menys fiable.

2.11.3. Detecció dels outliers mitjançant els residus

Per detectar els outliers es poden emprar diferents mètodes:

1) Observació dels residus $\hat{u}_i$

   Si $\left|\hat{u}_i\right| \geq 2 \frac{\sum_{i=1}^{N} \left|\hat{u}_i\right|}{n}$ l'observació i-èsima es pot considerar un outlier

2) Anàlisi del residu estandarditzat

   El residu estandarditzat es defineix com $\frac{\hat{u}_i}{\sqrt{\frac{\hat{u}'\hat{u}}{n-k}}} = \frac{\hat{u}_i}{\hat{\sigma}_u}$

   Si $\frac{\hat{u}_i}{\hat{\sigma}_u} \geq 2$ l'observació i-èsima es pot considerar un outlier

3) Anàlisi del residu estudentitzat. A diferència del residu estandarditzat, pondera l'error de l'ajust MQO associat a l'observació i-èsima per la seva desviació estàndard, en lloc de fer-ho per la desviació estàndard del terme de pertorbació, és a dir,

$$r_i = \frac{\hat{u}_i}{\sqrt{\hat{\sigma}_u^2\left(1 - h_{ii}\right)}} = \frac{\hat{u}_i}{\sqrt{\frac{\hat{u}'\hat{u}}{n-k}\left(1 - h_{ii}\right)}}$$

   Si $r_i \geq t_{n-k,1-\alpha/2}$ l'observació i-èsima es pot considerar un outlier

2.11.4. Distància de Cook

La distància de Cook és una mesura que permet detectar l'estranyesa d'una observació valorant tant el leverage com la concordança del procés que l'ha generat amb el procés generador de la resta de la mostra.

Es defineix de la següent manera:

$$DC_i = \frac{\frac{\left(\hat{Y} - \hat{Y}_i\right)'\left(\hat{Y} - \hat{Y}_i\right)}{k}}{\frac{\hat{u}'\hat{u}}{n-k}}$$

on $\hat{Y}_i$ és l'estimació obtinguda treient de la mostra l'observació i

La distància de Cook es distribueix com una variable F de Snedecor amb k graus al numerador i n-k al denominador, per tant es pot fer el següent contrast:

Si $DC_i \geq F_{k,n-k}$ l'observació i-èsima té més influència en l'ajust del model que la resta.

També es pot demostrar la següent relació entre la distància de Cook DC_i, el residu estudentitzat r_i, el leverage h_{ii} i el nombre de regressors k:

$$DC_i = \frac{r_i^2}{K}\frac{h_{ii}}{1-h_{ii}}$$

2.12. Implicacions del incompliment dels supòsits clàssics

2.12.1. Presentació

Fem un repàs dels supòsits clàssics :

Referits al component determinista :

-El component determinista és lineal respecte els paràmetres : Si aquest supòsit falla és pot aplicar el model de mínims quadrats no lineal (MQNL)

-Regressors no estocàstics i linealment independents: Si els regressors són linealment dependents és el problema de la col.linealitat que ja hem vist. Si els regressors són estocàstics ho veurem a l'apartat següent.

Referits al component aleatori :

- $E(u)=0$ Si l'esperança de u no és 0, l'estimació del terme constant captaria aquest biaix, per tant sempre es pot fer aquest supòsit.
- $\mathrm{var}(u)=\sigma^2 I$ i.e. la variància és constant per totes les pertorbacions (homocedasticitat). El cas de variància no constant (heteroscedasticitat) l'estudiem al tema 4

-Supòsit de no autocorrelació: Aquest supòsit no és cert per sèries temporals econòmiques. L'estudiem al tema 4.

- $u \sim Normal$: Si això no es compleix ho veurem més endavant.

2.12.2. Regressors estocàstics

Si els regressors són estocàstics ens podem trobar amb relacions de dependència entre les X (regressors) i les u (pertorbacions).

Ens podem trobar amb tres casos:

-Els regressors i les pertorbacions són independents:

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

Per tant l'estimador no té biaix

$$V(\hat{\beta}) = E((X'X)^{-1}X'uu'X(X'X)^{-1}) = \sigma^2 E((X'X)^{-1})$$

No coneixem la distribució exacta dels estimadors MQO i per tant no podem utilitzar els estadístics t i F en mostres finites. Però es pot utilitzar la teoria asimptòtica ja que els estimadors no tenen biaix i són consistents.

-Els regressors i les pertorbacions estan correlacionades en el temps però no per un mateix valor de t:

En aquest cas l'estimador pot tenir biaix i tampoc coneixem la distribució exacta dels estimadors MQO, però es pot utilitzar la teoria asimptòtica perquè els estimadors són consistents.

-Els regressors i les pertorbacions estan correlacionats en un mateix moment del temps.

En aquest cas, a més dels problemes anteriors, tampoc podem utilitzar la teoria asimptòtica. En aquest cas s'han de buscar mètodes alternatius a MQO (variables instrumentals).

2.12.3. Pertorbacions no normals

Si les pertorbacions no són normals, aleshores:

1) L'estimador MQO no té biaix (com en el cas de pertorbacions normals)

2) El teorema de Gauss-Markov és vàlid: L'estimador MQO és el millor estimador lineal sense biaix

3) L'eficiència és perd. Si les pertorbacions no són normals, no podem dir que l'estimador MQO sigui eficient.

4) Contrastos: Ja no els podem utilitzar per mostres finites (l'estadístic no és Student). Per mostres grans es pot suposar normalitat asimptòtica i es poden utilitzar els tests.

Un test per saber si les pertorbacions són normals és el **Test de Jarque-Bera**. És un test asimptòtic que utilitza els coeficients d'asimetria i curtosis mostrals calculats a partir dels residus MQO:

$$S = \frac{\frac{\sum_{i=1}^{n} \hat{u}_i^3}{n}}{\left(\frac{\sum_{i=1}^{n} \hat{u}_i^2}{n}\right)^{3/2}} \qquad K = \frac{\frac{\sum_{i=1}^{n} \hat{u}_i^4}{n}}{\left(\frac{\sum_{i=1}^{n} \hat{u}_i^2}{n}\right)^{2}}$$

Aleshores,

$$JB = \frac{n}{6} \cdot \left[S^2 + \frac{(K-3)^2}{4}\right] \underset{a}{\sim} \chi_2^2$$

2.13. *Estadístics de contrast més generals: Test de Wald*

2.13.1. Presentació

L'estadístic de Wald serveix quan les pertorbacions no són normals i per si tenim restriccions no lineals.

La hipòtesis nul.la és de la següent forma:

$$H_0 : g(\beta) = 0$$

En aquest cas, g és una funció general, no necessàriament lineal.

Per tant, l'objectiu és trobar un estadístic de contrast amb distribució coneguda i que sigui calculable.

2.13.2. Derivació de l'estadístic de Wald

Derivarem l'estadístic de Wald pel cas que la restricció sigui lineal, de la forma,

$H_0 : R\beta = r$

Com les pertorbacions no són normals ja no podem suposar $\hat{\beta} \sim N\left(\beta, \sigma^2 (X'X)^{-1}\right)$
Fem servir la distribució asimptòtica:

$$\sqrt{n}\left(\hat{\beta} - \beta\right) \underset{a}{\sim} N\left(0, \sigma^2 Q_{XX}^{-1}\right) \quad on\ Q_{XX} = \lim_{n\to\infty} \frac{X'X}{n}$$

Per tant,

$$\sqrt{n}\left(R\hat{\beta} - R\beta\right) \underset{a}{\sim} N\left(0, \sigma^2 R Q_{XX}^{-1} R'\right)$$

Sota la hipòtesis nul.la H_0

$$\sqrt{n}\left(R\hat{\beta} - r\right) \underset{a}{\sim} N\left(0, \sigma^2 R Q_{XX}^{-1} R'\right)$$

Com $R\hat{\beta}$ és un vector aleatori de dimensió q, el resultat anterior implica

$$n\left(R\hat{\beta} - r\right)\left[\sigma^2 R Q_{XX}^{-1} R'\right]^{-1}\left(R\hat{\beta} - r\right) \underset{a}{\sim} \chi_q^2$$

Com que σ^2 i Q_{XX}^{-1} no són coneguts, els substituim per estimadors consistents de σ^2 i Q_{XX}^{-1}.
Per tant,

$$n\left(R\hat{\beta} - r\right)\left[\hat{\sigma}^2 R\left(\frac{X'X}{n}\right)^{-1} R'\right]^{-1}\left(R\hat{\beta} - r\right) \underset{a}{\sim} \chi_q^2$$

Per tant, l'estadístic de Wald quan les restriccions són lineals pren la següent forma,

$$w \equiv \left(R\hat{\beta} - r\right)\left[\hat{\sigma}^2 R(X'X)^{-1} R'\right]^{-1}\left(R\hat{\beta} - r\right) \underset{a}{\sim} \chi_q^2$$

Si tenim $H_0 : g(\beta) = 0$ l'estadístic de Wald pren la següent forma

$$w \equiv g(\hat{\beta})'\left\{H(\hat{\beta})\left[\sigma^2 (X'X)^{-1}\right]H(\hat{\beta})'\right\}^{-1} g(\hat{\beta}) \underset{a}{\sim} \chi_q^2 \quad on\ H(\beta) = \frac{\partial g(\beta)}{\partial \beta'}$$

Per utilitzar el contrast de Wald s'ha de tenir una mostra gran ja que estem utilitzant la distribució asimptòtica.

2.14. Bootstraping

2.14.1. Presentació

El Bootstraping és una tècnica que ens permet extreure informació sobre la distribució d'una variable aleatòria sense fer servir supòsits de normalitat ni de teoria assimptòtica.

El Bootstraping és semblant al mètode de Montecarlo en que també fa servir simulació. A partir d'un mecanisme generador de dades (dgp) es generen mostres. Per cada mostra es calcula un estimador MQO. Després amb tots aquests estimadors es dibuixa l'histograma i s'obté la funció de densitat empírica.

El procés de Bootstraping funciona de la següent forma:

1) Donat el model $y_i = \beta_1 + \beta_2 x_{2i} + u_i$, i una mostra, per MQO s'estimen $\hat{\beta}_1, \hat{\beta}_2$.

2) A partir de les estimacions i dels valors reals de la mostra, es calculen els residus $\hat{u}_i = y_i - \hat{y}_i = y_i - \left(\hat{\beta}_1 + \hat{\beta}_2 x_{2i}\right)$

3) Es genera la mostra s de residus prenent n observacions de forma aleatòria i amb reposició dels residus $\hat{u}_i$ calculats al pas 2

4) Per cada mostra s de residus, es calcula una mostra y^s de dimensió n de la següent forma: $y_i^s = \hat{\beta}_1 + \hat{\beta}_2 x_{2i} + \hat{u}_i^s$

5) Per cada mostra y^s s'estimen paràmetres $\hat{\beta}_1^s, \hat{\beta}_2^s$ per MQO.

6) Els pasos 3,4 i 5 es repeteixen milers de vegades per obtenir la distribució empírica dels estimadors $\hat{\beta}_1, \hat{\beta}_2$.

2.14.2. Bootstraping d'intervals de confiança

Donat el model $y_i = \beta_1 + \beta_2 x_{2i} + u_i$

Si es suposa normalitat de pertorbacions l'interval de confiança de β_2 és:

$$\left[\hat{\beta}_2 \pm t_{1-\alpha/2,n-2}\hat{\sigma}_{\hat{\beta}_2}\right] = I$$

Per tant un $(1-\alpha)\%$ de les vegades $\beta_2 \in I$

Utilitzant Bootstraping, es pot generar un interval de confiança sense fer cap supòsit preliminar sobre la distribució de les pertorbacions.

Seguint els pasos 1 a 6 de l'apartat anterior es generen milers de mostres $\hat{\beta}_2^s$. S'ordenen els valors obtinguts i es treuen el $(\alpha/2)\%$ de valors més petits i el $(\alpha/2)\%$ de valors més grans. Les observacions restants delimiten l'interval de Bootstraping.

Per exemple, si repetim el procés de Bootstraping 2000 vegades i el nivell de confiança és el 5% treuriem els 50 primers i els 50 últims.

3. Models amb pertorbacions no esfèriques

3.1. *El model de regressió generalitzat: presentació*

3.1.1. Descripció del model

El model de regressió general és:

$$\begin{cases} Y = X\beta + u \\ u \sim \left(0, \sigma^2 I\right) \quad \textit{pertorbacions esfèriques independents} \end{cases}$$

El model de regressió generalitzat és:

$$\begin{cases} Y = X\beta + u \\ u \sim \left(0, \Sigma\right) \quad \Sigma \textit{ matriu simètrica definida positiva} \end{cases}$$

Segons la forma de Σ tenim diferents casuístiques:

- Si $\Sigma = \sigma^2 I$: pertorbacions esfèriques (model de regressió general)
- Si $\Sigma \neq \sigma^2 I$: pertorbacions no esfèriques
- Si Σ diagonal amb elements a la diagonal diferents: heterocedasticitat
- Si Σ té elements fora de la diagonal diferents de zero: correlació

3.1.2. L'estimador MQO sota pertorbacions no esfèriques

$$\begin{cases} Y = X\beta + u \\ u \sim \left(0, \Sigma\right) \quad \Sigma \textit{ matriu simètrica definida positiva} \end{cases}$$

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

- $E\left(\hat{\beta}\right) = E\left(\beta + \left(X'X\right)^{-1} X'u\right) = \beta + E\left(\left(X'X\right)^{-1} X'u\right) = \beta$ (sense biaix)

- $\operatorname{var}\left(\hat{\beta}\right) = E\left(\hat{\beta} - \beta\right)\left(\hat{\beta} - \beta\right) = E\left(\left(X'X\right)^{-1} X'uu'X\left(X'X\right)^{-1}\right) =$
$= \left(X'X\right)^{-1} X'Euu'X\left(X'X\right)^{-1} = \left(X'X\right)^{-1} X'\Sigma X\left(X'X\right)^{-1} \neq \sigma^2\left(X'X\right)^{-1}$

Per tant les expressions dels estadístics t,F i χ^2 abans derivades no són vàlides

- $\hat{\beta}$ és consistent : No ha canviat res que modifiqui la demostració feta pel model general

- Si les pertorbacions estan distribuïdes normalment, llavors:

$$\hat{\beta} \sim N\left(\beta, \left(X'X\right)^{-1} X'\Sigma X\left(X'X\right)^{-1}\right)$$

- Encara que les pertorbacions no estiguin distribuïdes normalment, l'estimador MQO és asimptòticament normal

3.1.3. L'estimador Mínim Quadrat Generalitzat (MQG): derivació

Donat el model original

$$\begin{cases} Y = X\beta + u \\ u \sim (0, \Sigma) \quad \Sigma \textit{ matriu simètrica definida positiva} \end{cases}$$

es vol transformar per tal que compleixi els supòsits clàssics

$$\begin{cases} Y^* = X^*\beta + u^* \\ u^* \sim (0, \sigma^2 I) \end{cases}$$

L'estimador MQG serà l'estimador MQO del model transformat.

Per trobar el model transformat utilitzem la descomposició de Cholesky:
"Si A és una matriu simètrica, definida positiva, llavors A es pot descomposar com:

(i) A=U'U (on U és una matriu triangular superior)
(ii) A=LL' (on L és una matriu triangular inferior)"

Donat que Σ és una matriu simètrica, definida positiva, llavors Σ^{-1} també serà una matriu simètrica i definida positiva. Utilitzant la descomposició de Cholesky, sabem que sempre:

$$\exists\, P \,/\, P'P = \Sigma^{-1}$$

Aquest resultat implica

$$P\Sigma P' = I$$

Aquesta matriu P s'utilitza per trobar el model transformat

El model original $Y = X\beta + u$ es multiplica per P $\rightarrow PY = PX\beta + Pu$

$$Y^* = X^*\beta + u^*$$
$$E(u^*) = E(Pu) = PE(u) = 0$$
$$\operatorname{var}(u^*) = E(u^*u^{*}{}') = E(Puu'P') = PE(uu')P' = P\Sigma P' = I$$

Per tant, el model transformat és esfèric. L'estimador MQG és l'estimador MQO del model transformat:

$$\hat{\beta}_{MQG} = (X^*{}'X^*)^{-1} X^*{}'Y^* = ((PX)'PX)^{-1}(PX)'PY = (X'P'PX)^{-1}X'P'PY =$$
$$= (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$$

3.1.4. L'estimador MQG: Distribució i propietats

$$\hat{\beta}_{MQG} = (X^*{}'X^*)^{-1} X^*{}'Y^* = \beta + (X^*{}'X^*)^{-1} X^*{}'u^*$$

$$E(\hat{\beta}_{MQG}) = E\left(\beta + (X^*{}'X^*)^{-1} X^*{}'u^*\right) = \beta$$

$$\operatorname{var}(\hat{\beta}_{MQG}) = E\left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'\right] = E\left[(X^*{}'X^*)^{-1} X^*{}'u^*u^*{}'X^*(X^*{}'X^*)^{-1}\right] =$$
$$= (X^*{}'X^*)^{-1} X^*{}'E\left[u^*u^*{}'\right]X^*(X^*{}'X^*)^{-1} = (X^*{}'X^*)^{-1} =$$
$$= (X'P'PX)^{-1} = (X'\Sigma^{-1}X)^{-1}$$

-Tots els resultats anteriors sobre l'estimador MQO es poden aplicar aquí doncs l'estimador MQG s'ha derivat d'un model que compleix tots els supòsits clàssics, i així té totes les propietats desitjables.

En particular, l'estimador MQG és més eficient que l'estimador MQO i també és asimptòticament més eficient que MQO

- És fàcil veure que l'estimador MQG minimitza $(Y - X\beta)'\Sigma^{-1}(Y - X\beta)$

-Els contrastos definits per MQO són tots vàlids, sempre que es consideri X^* en lloc de X, i que quan es necessiti utilitzar σ^2 s'estableixi igual a 1.

3.1.5. L'estimador MQG a la pràctica

En general Σ no és coneguda. Per tant, $\hat{\beta}_{MQG}$ no és, en general, aplicable. Per tant es proposa,

Estimador MQG factible: $\hat{\beta}_{MQGF} = \left(X'\hat{\Sigma}^{-1}X\right)^{-1} X'\hat{\Sigma}^{-1}Y$

Estimar Σ implica estimar $n + (n^2 - n)/2$ paràmetres. Donat que el nombre de paràmetres a estimar augmenta més ràpidament que n, no podem estimar Σ consistentment, a no ser que s'introdueixin restriccions suficients (supòsits sobre com són els elements de Σ) per poder garantir una estimació consistent de Σ.

Si suposem que parametritzem Σ en funció de X i d'un vector de paràmetres θ:

$$\Sigma = \Sigma(X, \theta)$$

Si $\hat{\theta}$ és un estimador consistent de θ, aleshores $\hat{\Sigma} \equiv \Sigma(X, \hat{\theta})$ serà un estimador consistent de Σ (Teorema de Slutsky)

Per tant, l'estimador MQGF té les mateixes propietats asimptòtiques que l'estimador MQG :

1) Consistència
2) normalitat asimptòtica
3) eficiència asimptòtica (si pertorbacions distribuïdes normalment)

A la pràctica, els passos per trobar l'estimador MQGF són sempre els mateixos:

1) Trobar un estimador consistent de θ, i construir $\hat{\Sigma} \equiv \Sigma(X, \hat{\theta})$
2) Trobar la matriu de transformació $\hat{P}$ / $\hat{P}'\hat{P} = \hat{\Sigma}^{-1}$
3) Trobar el model transformat: $\hat{P}Y = \hat{P}X\beta + u^*$
4) Aplicar MQO al model transformat per trobar MQGF

3.2. Heteroscedasticitat

3.2.1. Presentació

Els models amb heterocedasticitat es presenten de la següent forma:

$$y_i = x'_i \beta + u_i \qquad i = 1, \ldots, n$$
$$u_i \sim indep\left(0, \sigma_i^2\right)$$

o alternativament en forma matricial

$$Y = X\beta + u \qquad u \sim (0, \Sigma) \qquad \text{on} \qquad \Sigma = \begin{bmatrix} \sigma_1^2 & 0 & . & 0 \\ 0 & \sigma_2^2 & . & 0 \\ . & . & . & . \\ 0 & . & . & \sigma_n^2 \end{bmatrix}$$

Els estimadors MQO aplicats sobre models amb heteroscedasticitat tenen les següents propietats:

1) Els estimadors MQO segueixen sent lineals
2) Segueixen no tenint biaix
3) Ja no són eficients, perquè ja no tenen variància mínima
4) Les fòrmules habituals per estimar les variàncies dels estimadors MQO donen variàncies esbiaixades
5) Les proves de contrastació d'hipòtesis i els intervals de confiança de les distribucions t i F ja no són fiables

3.2.2. Causes de l'heteroscedasticitat

Normalment l'heteroscedasticitat apareix en les dades de tall transversal on l'escala de la variable dependent varia al llarg de les observacions. Les principals causes són les següents:

1) Raons d'escala: Per exemple, si es vol explicar les despeses en innovació d'una empresa a partir del capital, és molt normal que les empreses més grans tinguin més variació a la despesa que les empreses petites. També es pot donar quan s'utilitzen dades macros de diferents països. Els països més grans presenten més variabilitat a les seves dades.

2) Observacions influents: Es surten de les pautes generals del model. S'han d'identificar per estudiar el seu impacte i identificar la seva causa. Si són conseqüència d'un error a l'entrada de dades es poden treure o corregir. Si són degudes a una situació extraordinària es poden tractar incloent variables fictícies. Si no sabem la causa de l'anomalia s'han de deixar a la mostra.

3) Problemes d'especificació: L'elecció d'una forma funcional incorrecta pel fenomen que s'està estudiant pot conduir a l'aparició d'heteroscedasticitat.

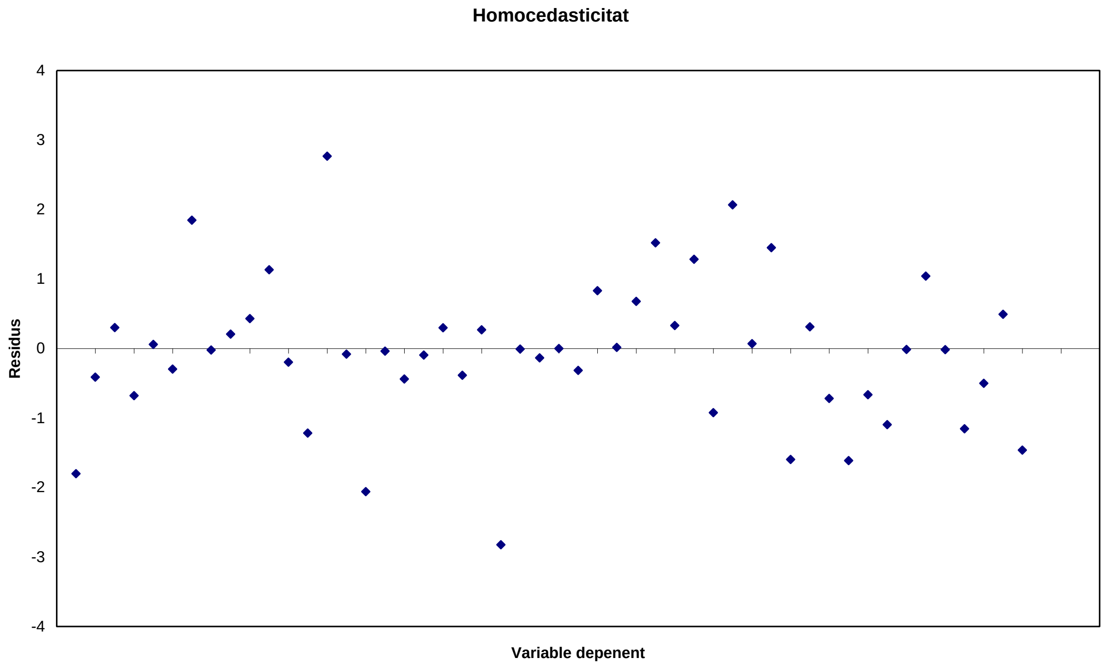
4) Millores en la qualitat de la informació disponible: En sèries temporals, pot ser que en el temps es vagin millorant les tècniques de recol·lecció de dades i això provoqui que els errors de mesura es vagin fent més petits.

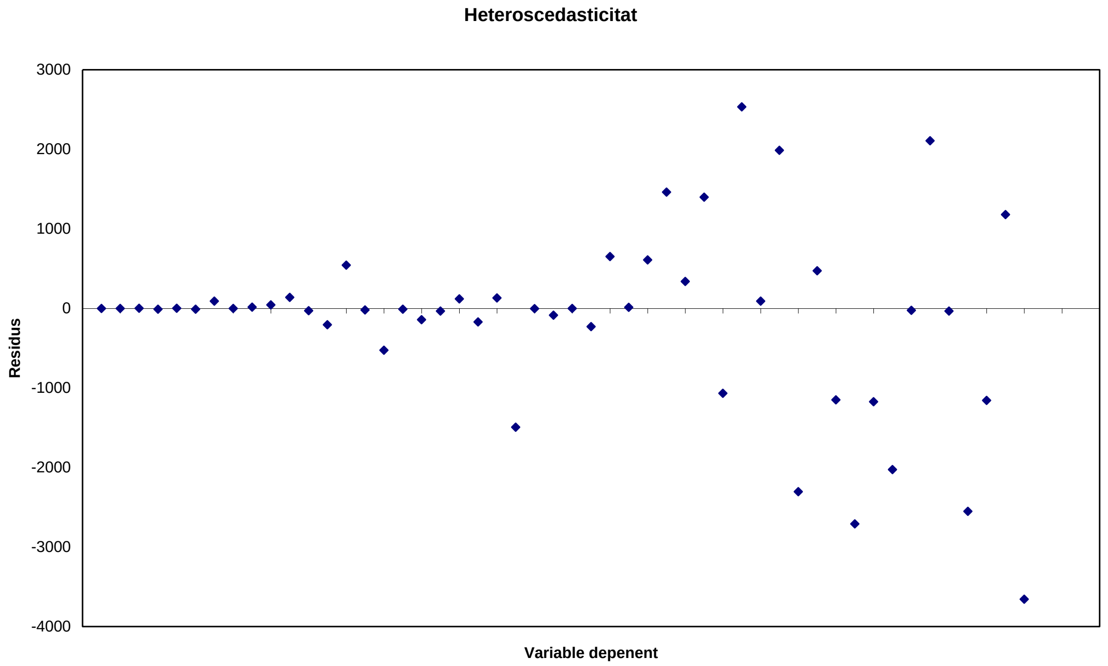
3.2.3. Tests per detectar la presència d'heteroscedasticitat

Hi ha diferents formes de detectar l'heteroscedasticitat.

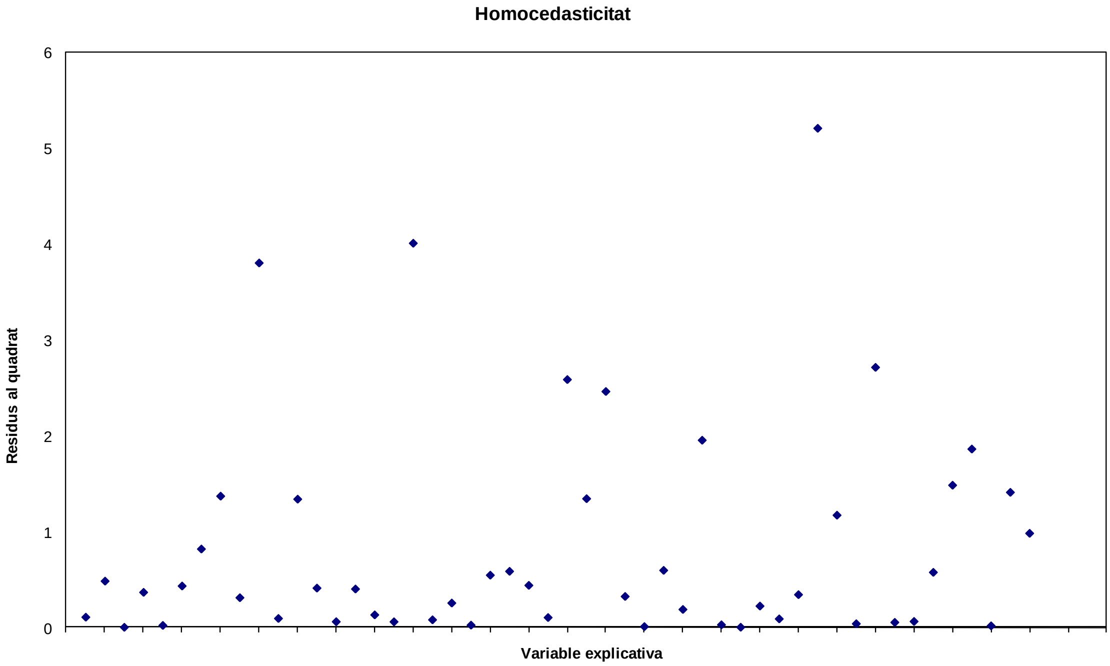
- **Natura del problema**: Utilitzant informació i resultats d'estudis preliminars del mateix problema, podem esperar trobar heteroscedasticitat a la mostra. De fet, a les dades de tall transversal, l'heteroscedasticitat és més la regla que l'excepció.

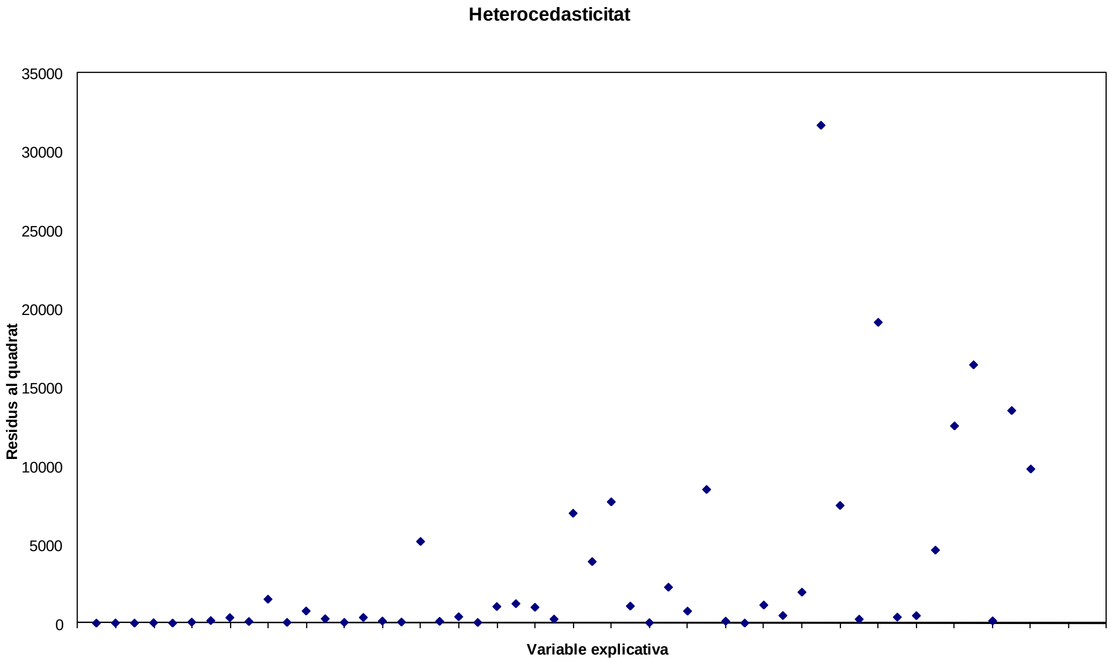
- **Inspecció visual dels residus**: Si la dispersió dels residus no és constant és indicador de la presència d'heteroscedasticitat.





També es pot graficar els residus al quadrat contra la variable explicativa:





-**Test de Park**: Es suggereix l'existència d'una relació funcional entre la variància de la variable aleatòria del model i la variable explicativa X_j de la següent forma:

$$\sigma_i^2 = \sigma^2 \cdot x_{ji}^{\beta} \cdot e^{\nu_i} \qquad \forall i = 1, 2, \ldots, n$$

o equivalentment, prenent logaritmes

$$\ln \sigma_i^2 = \ln \sigma^2 + \beta \cdot \ln x_{ji} + \nu_i \qquad \forall i = 1, 2, \ldots, n$$

Com els valors de σ_i^2 són desconeguts, s'utilitza per fer les regressions els valors dels residus, per tant:

$$\ln \hat{u}_i^2 = \ln \sigma^2 + \beta \cdot \ln x_{ji} + \nu_i = \alpha + \beta \cdot \ln x_{ji} + \nu_i \qquad \forall i = 1, 2, \ldots, n$$

La hipòtesi de no heteroscedasticitat és la següent:

$H_0 : \beta = 0$ (pertorbacions homocedàstiques)

S'accepta la hipòtesi de homocedasticitat quan el coeficient de regressió β no sigui estadísticament significatiu utilitzant la prova t-student.

-**Test de Glejser**: És similar al test de Park, però en aquest cas es fa la regressió dels valors absoluts dels residus sobre les variables explicatives utilitzant diverses relacions funcionals:

$$|\hat{u}_i| = \alpha + \beta \cdot x_{ji} + \nu_i \qquad \forall i = 1, 2, \dots, n$$

$$|\hat{u}_i| = \alpha + \beta \cdot \sqrt{x_{ji}} + \nu_i \qquad \forall i = 1, 2, \dots, n$$

$$|\hat{u}_i| = \alpha + \beta \cdot \frac{1}{x_{ji}} + \nu_i \qquad \forall i = 1, 2, \dots, n$$

- **Test de White**: La idea al darrera és la següent: Si les pertorbacions són homocedàstiques, llavors les x_i o funcions de x_i no haurien d'ajudar a explicar $E(u_i^2)$.
Per tant, seria el següent:

$$y_i = x'_i \beta + u_i \qquad i = 1, \dots, n$$
$$\operatorname{var}(u_i) = \sigma_i^2$$
$$\sigma_i^2 = f(z_{1i}, z_{2i}, \dots.)$$

sota homocedasticitat, f no hauria d'explicar σ_i^2
En concret, White proposa:

$$\sigma_i^2 = \alpha_0 + \alpha_1 x_{2i} + \alpha_2 x_{3i} + \alpha_3 x_{2i}^2 + \alpha_4 x_{3i}^2 + \dots + \alpha_p x_{k-1,i} x_{k,i}$$

$$\sigma_i^2 = \alpha_0 + z'_i \alpha \quad \text{on} \quad z'_i = \left(x_{2i}, x_{2i}^2, x_{3i}, x_{3i}^2, \dots, \textit{termes creuats}\right)$$

o equivalentment,

$$\hat{u}_i^2 = \alpha_0 + z'_i \alpha + \nu_i \quad \text{on} \quad E(\nu_i) = 0$$
$H_0 : \alpha = 0$ (pertorbacions homocedàstiques)

Estadístic de White: $n\tilde{R}^2 \underset{sota\,H_o}{\overset{a}{\sim}} \chi^2(p)$ p = Nombre de coeficients dins H_0

on $\tilde{R}^2$ és el coeficient de determinació de la regressió auxiliar de White.

- **Altres tests** per detectar l'heteroscedasticitat són:

- Test de Levene
- Test de Goldfeld-Quandt
- Test de Breusch-Pagan-Godfrey
- Test de Peak
- Test de classificació de la correlació de Spearman

3.2.4. Estimació per MQG sota heteroscedasticitat

Model original:

$$Y = X\beta + u$$
$$u \sim (0, \Sigma) \qquad on\ \Sigma\ diagonal$$

Estimador MQG:

$$\hat{\beta}_{MQG} = \left(X'\Sigma^{-1}X\right)^{-1} X'\Sigma^{-1}Y$$

$$\Sigma^{-1} = \begin{bmatrix} 1/\sigma_1^2 & 0 & 0 & . & 0 \\ 0 & 1/\sigma_2^2 & 0 & . & 0 \\ 0 & 0 & 1/\sigma_3^2 & . & 0 \\ . & . & 0 & . & 0 \\ 0 & . & . & 0 & 1/\sigma_n^2 \end{bmatrix}$$

A la pràctica, per trobar l'estimador MQG es pot fer via MQO al model transformat. Per trobar el model transformat necessitem una matriu P que compleixi $P'P = \Sigma^{-1}$

La matriu de transformació P és:

$$P = \begin{bmatrix} 1/\sigma_1 & 0 & 0 & . & 0 \\ 0 & 1/\sigma_2 & 0 & . & 0 \\ 0 & 0 & 1/\sigma_3 & . & 0 \\ . & . & 0 & . & 0 \\ 0 & . & . & 0 & 1/\sigma_n \end{bmatrix}$$

per tant, el model transformat queda:

$$Y^* = X^*\beta + u^*$$
$$u^* \sim (0, I)$$

$$Y^* = PY = \begin{bmatrix} 1/\sigma_1 & 0 & 0 & . & 0 \\ 0 & 1/\sigma_2 & 0 & . & 0 \\ 0 & 0 & 1/\sigma_3 & . & 0 \\ . & . & 0 & . & 0 \\ 0 & . & . & 0 & 1/\sigma_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix} = \begin{bmatrix} y_1/\sigma_1 \\ y_2/\sigma_2 \\ . \\ . \\ y_n/\sigma_n \end{bmatrix}$$

$$X^* = PX = \begin{bmatrix} 1/\sigma_1 & 0 & 0 & . & 0 \\ 0 & 1/\sigma_2 & 0 & . & 0 \\ 0 & 0 & 1/\sigma_3 & . & 0 \\ . & . & 0 & . & 0 \\ 0 & . & . & 0 & 1/\sigma_n \end{bmatrix} \begin{bmatrix} 1 & x_{21} & . & . & x_{k1} \\ 1 & x_{22} & . & . & x_{k2} \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & x_{2n} & . & . & x_{kn} \end{bmatrix} =$$

$$= \begin{bmatrix} 1/\sigma_1 & x_{21}/\sigma_1 & . & . & x_{k1}/\sigma_1 \\ 1/\sigma_2 & x_{22}/\sigma_2 & . & . & x_{k2}/\sigma_2 \\ . & . & . & . & . \\ . & . & . & . & . \\ 1/\sigma_n & x_{2n}/\sigma_n & . & . & x_{kn}/\sigma_n \end{bmatrix}$$

o alternativament,

$$\frac{y_i}{\sigma_i} = \beta_1 \frac{1}{\sigma_i} + \beta_2 \frac{x_{2i}}{\sigma_i} + \ldots + \beta_k \frac{x_{ki}}{\sigma_i} + \frac{u_i}{\sigma_i}$$

Aleshores, s’aplica MQO a aquest model per trobar l’estimador MQG.

3.2.5. Estimació factible davant la presència d’heteroscedasticitat

Per estimar els paràmetres d’un model per MQGF, primer cal estimar els paràmetres de la matriu Σ

Per estimar Σ:

(1) Estimar Σ implica estimar N paràmetres i per tant necessitem supòsits per reduir el nombre de paràmetres a estimar

$$\sigma_i^2 = f(z_i, \theta)$$

(2) Si podem estimar θ de forma consistent, llavors podem obtenir un estimador consistent de les σ_i^2, i per tant, un estimador consistent de Σ:

$\hat{\sigma}_i^2 = f(z_i, \hat{\theta})$ on $\hat{\theta}$ estimador consistent de θ

Un cop estimades les variàncies es poden seguir dos camins:

(1) Construir la matriu $\hat{\Sigma}$ i aplicar $\hat{\beta}_{MQGF} = \left(X'\hat{\Sigma}^{-1}X\right)^{-1} X'\hat{\Sigma}^{-1}Y$

(2) Aplicar MQO al model transformat: $\frac{y_i}{\hat{\sigma}_i} = \frac{\beta_1}{\hat{\sigma}_i} + \beta_2 \frac{x_{2i}}{\hat{\sigma}_i} + \ldots + \beta_k \frac{x_{ki}}{\hat{\sigma}_i} + u_i^* \quad i = 1, \ldots, n$

- Exemples de supòsits sobre el comportament de σ_i^2:

(1) $E(u_i^2) = \sigma^2 X_{ji}^2 \quad \forall i = 1, 2, \ldots, n$

En aquest cas es divideix el model per X_{ji}

$$\frac{y_i}{x_{ji}} = \frac{\beta_1}{x_{ji}} + \beta_2 \frac{x_{2i}}{x_{ji}} + \ldots + \beta_k \frac{x_{ki}}{x_{ji}} + \frac{u_i}{x_{ji}} \quad i = 1,\ldots,n$$

(2) $E\left(u_i^2\right) = \sigma^2 X_{ji} \quad \forall i = 1,2,\ldots,n$

En aquest cas la transformació té que ser

$$\frac{y_i}{\sqrt{x_{ji}}} = \frac{\beta_1}{\sqrt{x_{ji}}} + \beta_2 \frac{x_{2i}}{\sqrt{x_{ji}}} + \ldots + \beta_k \frac{x_{ki}}{\sqrt{x_{ji}}} + \frac{u_i}{\sqrt{x_{ji}}} \quad i = 1,\ldots,n$$

3.2.6. Reducció de l'heteroscedasticitat especificant un nou model

A vegades es pot reduir l'heteroscedasticitat fent una transformació logarítmica del model. D'aquesta forma no fa falta especular sobre la forma que tindrà l'heteroscedasticitat com hem fet a l'apartat anterior.

Si estimem

$$\ln y_i = \alpha + \beta \cdot \ln x_i + u_i$$

El problema de l'heteroscedasticitat pot ser menys greu amb aquesta transformació, ja que la transformació logarítmica comprimeix les escales en les que es mesuren les variables.

3.3. Autocorrelació

3.3.1. Presentació

El model amb pertorbacions autocorrelacionades és el següent:

$$y_t = x'_t \beta + u_t \quad t = 1,\ldots,T$$
$$u_t \sim \left(0, \sigma^2\right)$$
$$\operatorname{cov}\left(u_t, u_{t-s}\right) = \sigma_{t,t-s}$$

o alternativament, en forma matricial

$$\begin{array}{l} Y = X\beta + u \\ u \sim \left(0, \Sigma\right) \end{array} \qquad \text{on} \quad \Sigma = \begin{bmatrix} \sigma^2 & \sigma_{12} & . & . & \sigma_{1T} \\ \sigma_{21} & \sigma^2 & . & . & . \\ . & . & . & . & . \\ . & . & . & . & \sigma_{T-1,T} \\ \sigma_{T1} & . & . & \sigma_{T,T-1} & \sigma^2 \end{bmatrix} \neq \sigma^2 I$$

Els estimadors MQO aplicats sobre models amb autocorrelació tenen les següents propietats:

1) Els estimadors MQO segueixen sent lineals
2) Segueixen no tenint biaix
3) Ja no són eficients, perquè ja no tenen variància mínima
4) Les fòrmules habituals per estimar les variàncies dels estimadors MQO donen variàncies esbiaixades
5) Les proves de contrastació d'hipòtesis i els intervals de confiança de les distribucions t i F ja no són fiables

Els estimadors MQO, MQG i MQGF són:

- MQO: $\hat{\beta} = (X'X)^{-1}X'Y$
- MQG: $\hat{\beta}_{MQG} = \left(X^*{}'X^*\right)^{-1}X^*{}'Y^* = \left(X'\Sigma^{-1}X\right)^{-1}X'\Sigma^{-1}Y$
- MQGF: $\hat{\beta}_{MQGF} = \left(X'\hat{\Sigma}^{-1}X\right)^{-1}X'\hat{\Sigma}^{-1}Y$

Els tipus bàsics d'autocorrelació en les pertorbacions són els següents:

- AR(p): $u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + ... + \rho_p u_{t-p} + \varepsilon_t$
- MA(q): $u_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q}$

Com a casos particulars d'aquests es poden estudiar els models AR(1) i MA(1):

- Model amb pertorbacions AR(1) (estacionari):

$$y_t = x'_t \beta + u_t$$
$$u_t = \rho u_{t-1} + \varepsilon_t \qquad |\rho| < 1 \qquad \varepsilon_t \sim i.i.d\left(0, \sigma_\varepsilon^2\right) \qquad E\left(u_{t-s}\varepsilon_t\right) = 0 \quad s \geq 1$$

Aleshores, per substitució succesiva:

$$u_t = \sum_{j=0}^{\infty} \rho^j \varepsilon_{t-j}$$

Així: $\sigma_u^2 = \mathrm{var}(u_t) = E\left(u_t^2\right) = \sum_{j=0}^{\infty} \rho^{2j} E\left(\varepsilon_{t-j}\right)^2 + termes\ creuats = \dfrac{\sigma_\varepsilon^2}{1-\rho^2}$

$$\sigma_{t,t-1} = \mathrm{cov}(u_t, u_{t-1}) = E(u_t u_{t-1}) = E\left((\rho u_{t-1} + \varepsilon_t)u_{t-1}\right) = \rho E\left(u_{t-1}^2\right) = \rho\sigma_u^2$$
$$\sigma_{t,t-2} = \mathrm{cov}(u_t, u_{t-2}) = E(u_t u_{t-2}) = E\left((\rho u_{t-1} + \varepsilon_t)u_{t-2}\right) = \rho^2\sigma_u^2$$
…
$$\sigma_{t,t-s} = \mathrm{cov}(u_t, u_{t-s}) = ... = \rho^s \sigma_u^2$$

Per tant, sota pertorbacions AR(1):

$$\mathrm{var}(u) = \Sigma = \sigma_u^2 \begin{bmatrix} 1 & \rho & \rho^2 & . & \rho^{T-1} \\ \rho & 1 & \rho & . & . \\ \rho^2 & \rho & 1 & . & \rho^2 \\ . & . & . & . & \rho \\ \rho^{T-1} & . & \rho^2 & \rho & 1 \end{bmatrix}$$

Fixem-nos que el coeficient d'autocorrelació d'ordre s és ρ^s. Per tant ρ és el coeficient d'autocorrelació d'ordre 1.

- Model amb pertorbacions MA(1):

$$y_t = x'_t \beta + u_t$$
$$u_t = \varepsilon_t + \varphi\varepsilon_{t-1} \qquad \varepsilon_t \sim i.i.d\left(0, \sigma_\varepsilon^2\right) \qquad E\left(u_{t-s}\varepsilon_t\right) = 0 \quad s \geq 1$$

En aquest cas tenim,

$$\sigma_u^2 = \mathrm{var}(u_t) = E(u_t^2) = E(\varepsilon_t + \varphi\varepsilon_{t-1})^2 = \sigma_\varepsilon^2(1+\varphi^2)$$

$$\sigma_{t,t-s} = \mathrm{cov}(u_t, u_{t-s}) = E(u_t u_{t-s}) = E[(\varepsilon_t + \varphi\varepsilon_{t-1})(\varepsilon_{t-s} + \varphi\varepsilon_{t-s-1})] =$$
$$= termes\ creuats + \begin{cases} \varphi\sigma_\varepsilon^2 & si\ \ s = 1 \\ 0 & si\ \ s \geq 2 \end{cases}$$

Per tant ρ_s el coeficient d'autocorrelació d'ordre s $= \dfrac{\mathrm{cov}(u_t, u_{t-s})}{\mathrm{var}\, u_t} = \begin{cases} \dfrac{\varphi}{1+\varphi^2} & si\ s = 1 \\ 0 & si\ s \geq 2 \end{cases}$

Així,

$$\mathrm{var}(u) = \Sigma = \sigma_\varepsilon^2 \begin{bmatrix} 1+\varphi^2 & \varphi & 0 & . & 0 \\ \varphi & 1+\varphi^2 & \varphi & . & . \\ 0 & . & . & . & . \\ . & . & . & . & \varphi \\ 0 & . & 0 & \varphi & 1+\varphi^2 \end{bmatrix}$$

3.3.2. Causes de l'autocorrelació

L'autocorrelació es presenta principalment a les sèries temporals. Les causes més comuns són les següents:

1) Inèrcia de la sèrie: Les series temporals macro com el PIB, la producció o els índexs de preus mostren cicles econòmics. Quan hi ha una pujada, aquesta pujada continua fins que passa alguna cosa (canvi d'impostos, taxes d'interès, ...). Per tant les successives observacions són interdependents o estan correlacionades.

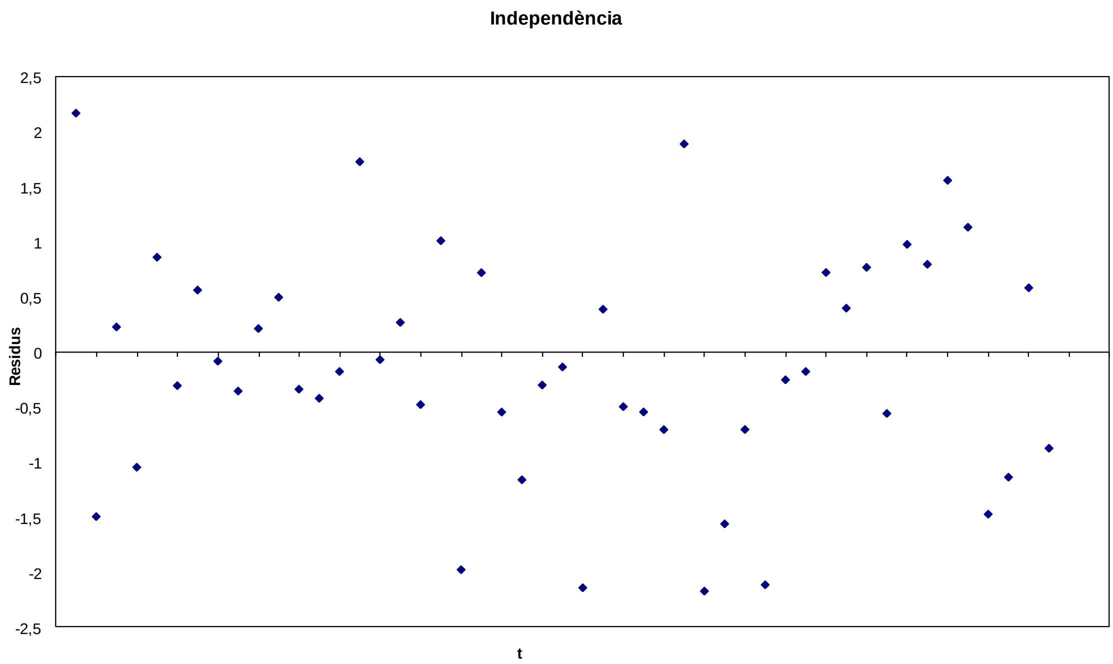
2) Temps d'ajust: Un fet succeït en un període pot tenir impacte en diferents períodes de temps degut a que els agents econòmics necessiten temps per reaccionar al canvi.

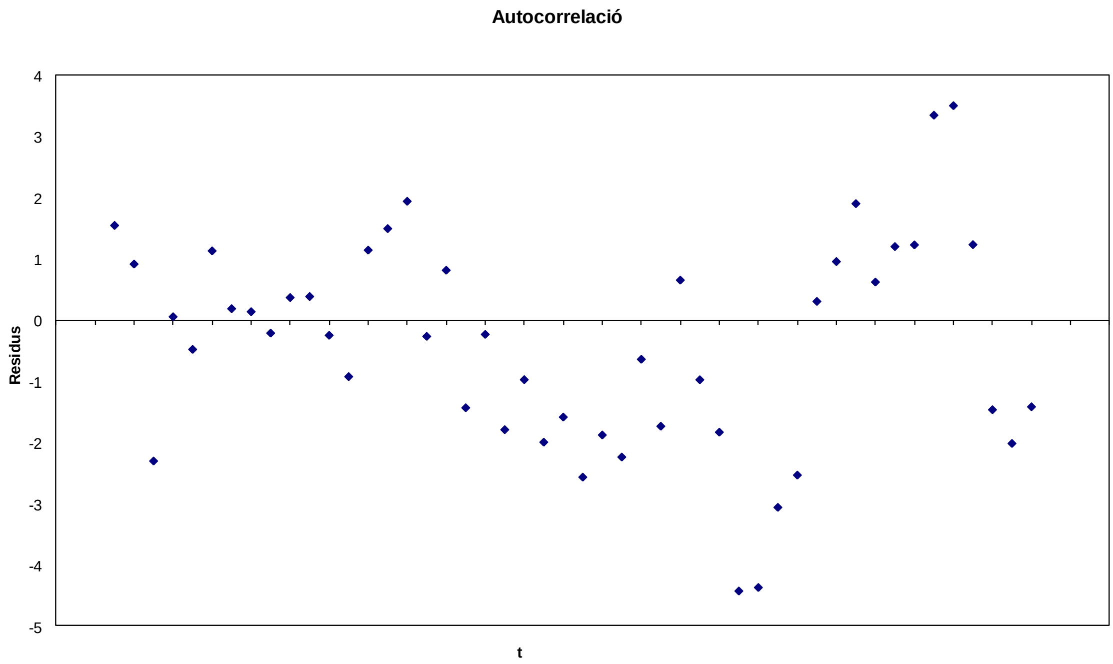
3) Errors d'especificació: L'elecció d'una forma funcional incorrecta pot provocar l'aparició d'un comportament sistemàtic al terme d'error.

4) Manipulació de dades: Per exemple quan es fan regressions amb dades trimestrals que han estat obtingudes a partir de dades mensuals. Això suavitza la sèrie, eliminant soroll, però introdueix autocorrelació en les dades.

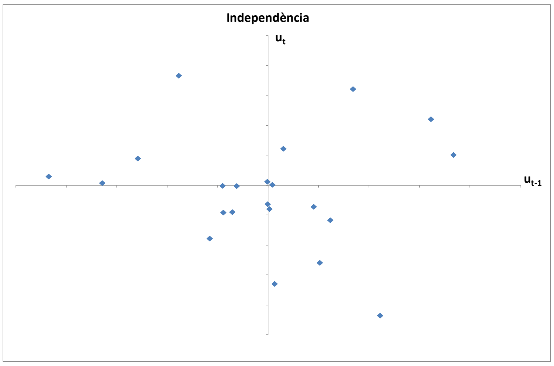
3.3.3. Tests per detectar la presència d'autocorrelació

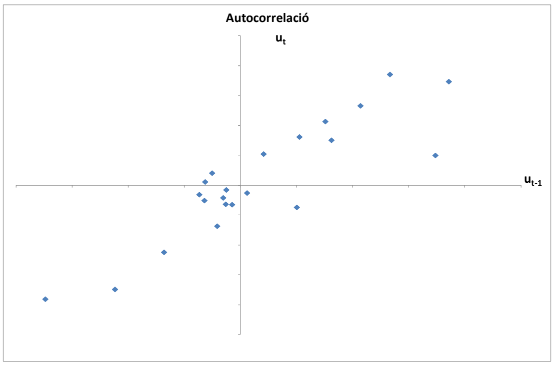
-**Inspecció visual dels residus**: La inspecció visual dels residus dóna una primera idea de si hi ha presència d'autocorrelació. Per exemple, es poden graficar els residus respecte el temps:





També es poden graficar els residus en t respecte els residus en t-1:





- **Test de Durbin-Watson**

$H_0 : \rho_1 = 0$ (coeficient d'autocorrelació d'ordre 1 = 0)

Es defineix l'estadístic de Durbin-Watson com:

$d \equiv \frac{\sum_{t=2}^{T} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{T} \hat{u}_t^2} \approx 2(1-r)$ on $r = \frac{\sum_{t=2}^{T} \hat{u}_t \hat{u}_{t-1}}{\sum_{t=1}^{T} \hat{u}_t^2}$ és el coeficient de correlació simple mostral entre $\hat{u}_t$ i $\hat{u}_{t-1}$

Donat que $r \xrightarrow[p]{} \rho_1$, llavors:

- Si pertorbacions no autocorrelacionades: $d \xrightarrow[p]{} 2$
- Si $u_t \sim AR(1)$ amb $\rho = 1$: $d \xrightarrow[p]{} 0$
- Si $u_t \sim AR(1)$ amb $\rho = -1$: $d \xrightarrow[p]{} 4$
- 0<d<4

Sobre la distribució de d sota H_0 es pot dir el següent:

- Sota H_0: $d \xrightarrow[p]{} 2$
- Simètrica entorn de 2
- Var(d) depèn de la matriu de regressors.

Els valors de d que difereixen significativament de 2 suggereixen autocorrelació. Les regles de decisió del test de Durbin són les següents:

H_0: No hi ha autocorrelació positiva Es rebutja H_0 si $d < d_L$

H_0: No hi ha autocorrelació negativa Es rebutja H_0 si $d > 4-d_L$

H_0: No hi ha autocorrelació positiva ni negativa No es rebutja H_0 si $d_U < d < 4-d_U$

A continuació es mostra una taula amb els valor tabulats de d_L i d_U per diferents nombres de regressors i mides mostrals.

Table Critical Values for the Durbin-Watson Test: 5% Significance Level[a]

	K = 2		K = 3		K = 4		K = 5		K = 6		K = 7		K = 8		K = 9		K = 10		K = 11	
T	d_L^*	d_U^*	d_L^*	d_U^*	d_L^*	d_U^*	d_L^*	d_U^*	d_L^*	d_U^*	d_L^*	d_U^*	d_L^*	d_U^*	d_L^*	d_U^*	d_L^*	d_U^*	d_L^*	d_U^*
6	0.610	1.400																		
7	0.700	1.356	0.467	1.896																
8	0.763	1.332	0.559	1.777	0.368	2.287														
9	0.824	1.320	0.629	1.699	0.455	2.128	0.296	2.588												
10	0.879	1.320	0.697	1.641	0.525	2.016	0.376	2.414	0.243	2.822										
11	0.927	1.324	0.758	1.604	0.595	1.928	0.444	2.283	0.316	2.645	0.203	3.005								
12	0.971	1.331	0.812	1.579	0.658	1.864	0.512	2.177	0.379	2.506	0.268	2.832	0.171	3.149						
13	1.010	1.340	0.861	1.562	0.715	1.816	0.574	2.094	0.445	2.390	0.328	2.692	0.230	2.985	0.147	3.266				
14	1.045	1.350	0.905	1.551	0.767	1.779	0.632	2.030	0.505	2.296	0.389	2.572	0.286	2.848	0.200	3.111	0.127	3.360		
15	1.077	1.361	0.946	1.543	0.814	1.750	0.685	1.977	0.562	2.220	0.447	2.472	0.343	2.727	0.251	2.979	0.175	3.216	0.111	3.438
16	1.106	1.371	0.982	1.539	0.857	1.728	0.734	1.935	0.615	2.157	0.502	2.388	0.398	2.624	0.304	2.860	0.222	3.090	0.155	3.304
17	1.133	1.381	1.015	1.536	0.897	1.710	0.779	1.900	0.664	2.104	0.554	2.318	0.451	2.537	0.356	2.757	0.272	2.975	0.198	3.184
18	1.158	1.391	1.046	1.535	0.933	1.696	0.820	1.872	0.710	2.060	0.603	2.257	0.502	2.461	0.407	2.667	0.321	2.873	0.244	3.073
19	1.180	1.401	1.074	1.536	0.967	1.685	0.859	1.848	0.752	2.023	0.649	2.206	0.549	2.396	0.456	2.589	0.369	2.783	0.290	2.974
20	1.201	1.411	1.100	1.537	0.998	1.676	0.894	1.828	0.792	1.991	0.692	2.162	0.595	2.339	0.502	2.521	0.416	2.704	0.336	2.885
21	1.221	1.420	1.125	1.538	1.026	1.669	0.927	1.812	0.829	1.964	0.732	2.124	0.637	2.290	0.547	2.460	0.461	2.633	0.380	2.806
22	1.239	1.429	1.147	1.541	1.053	1.664	0.958	1.797	0.863	1.940	0.769	2.090	0.677	2.246	0.588	2.407	0.504	2.571	0.424	2.734
23	1.257	1.437	1.168	1.543	1.078	1.660	0.986	1.785	0.895	1.920	0.804	2.061	0.715	2.208	0.628	2.360	0.545	2.514	0.465	2.670
24	1.273	1.446	1.188	1.546	1.101	1.656	1.013	1.775	0.925	1.902	0.837	2.035	0.751	2.174	0.666	2.318	0.584	2.464	0.506	2.613
25	1.288	1.454	1.206	1.550	1.123	1.654	1.038	1.767	0.953	1.886	0.868	2.012	0.784	2.144	0.702	2.280	0.621	2.419	0.544	2.560
26	1.302	1.461	1.224	1.553	1.143	1.652	1.062	1.759	0.979	1.873	0.897	1.992	0.816	2.117	0.735	2.246	0.657	2.379	0.581	2.513
27	1.316	1.469	1.240	1.556	1.162	1.651	1.084	1.753	1.004	1.861	0.925	1.974	0.845	2.093	0.767	2.216	0.691	2.342	0.616	2.470

S'han de tenir en compte les limitacions del test de Durbin-Watson:

1. El model de regressió inclou punt de tall. El test no es pot utilitzar en models de regressió lineal que passen per l'origen.
2. Els regressors X són no estocàstics.
3. Les pertorbacions u_t són AR(1), és a dir, venen generades pel següent mecanisme:

$$u_t = \rho u_{t-1} + \varepsilon_t \qquad |\rho| < 1$$

4. La regressió no inclou els valors retardats de la variable dependent com una de les variables explicatives.

- **Estadístic de Breusch-Godfrey**:

La idea al darrera és la següent: Si no hi ha autocorrelació, llavors els residus en períodes anteriors a t no han d'ajudar a explicar la magnitud del residu en el moment t.

La regressió de Breusch-Godfrey és la següent:

$$\hat{u}_t = x'_t\,\delta + \gamma_1 \hat{u}_{t-1} + \gamma_2 \hat{u}_{t-2} + \ldots + \gamma_p \hat{u}_{t-p} + \nu_t$$

La hipòtesis nul.la de no autocorrelació es pot expressar com:

$$H_0 : \gamma_1 = \gamma_2 = \ldots = \gamma_p = 0$$

L'estadístic de Breusch-Godfrey és $(T-p)R^2 \overset{sota\,H_0}{\underset{a}{\sim}} \chi^2(p)$ on R^2 és el coeficient de determinació de la regressió de B-G.

- **Altres tests** per detectar l'autocorrelació són:

- Prova h de Durbin

- Proves de Box-Pierce i Ljung-Box

3.3.4. Estimació per MQG del model sota el supòsit de pertorbacions AR(1):

Model:

$$y_t = x'_t \beta + u_t$$
$$u_t \sim AR(1)$$

o alternativament, en forma matricial

$$Y = X\beta + u \qquad u \sim (0, \Sigma) \qquad \text{on} \quad \Sigma = \sigma_u^2 \begin{bmatrix} 1 & \rho & \rho^2 & . & \rho^{T-1} \\ \rho & 1 & \rho & . & . \\ \rho^2 & . & . & . & . \\ . & . & . & . & \rho \\ \rho^{T-1} & . & \rho^2 & \rho & 1 \end{bmatrix} = \sigma_u^2 \Omega$$

L'estimador MQG es pot escriure com:

$$\hat{\beta}_{MQG} = \left(X'\Sigma^{-1}X\right)^{-1} X'\Sigma^{-1}Y = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}Y$$

$$\text{on} \quad \Omega^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1 & -\rho & 0 & . & . & 0 \\ -\rho & 1+\rho^2 & . & . & . & . \\ 0 & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & 1+\rho^2 & -\rho \\ 0 & . & . & 0 & -\rho & 1 \end{bmatrix}$$

A la pràctica, una forma equivalent de trobar l'estimador MQG és aplicant MQO al model transformat. Per trobar el model transformat mecessitem la matriu de transformació P / $P'P = \Omega^{-1}$

$$\hat{\beta}_{MQG} = \left(X'\Omega^{-1}X\right)^{-1} X'\Omega^{-1}Y = \left(X^{*}{}'X^{*}\right)^{-1} X^{*}{}'Y^{*} \quad \text{on} \quad X^{*} = PX, \; Y^{*} = PY$$

La matriu de transformació P:

$$P \equiv \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & . & . & 0 \\ -\rho & 1 & 0 & . & . & 0 \\ 0 & -\rho & 1 & 0 & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & 1 & 0 \\ 0 & . & . & 0 & -\rho & 1 \end{bmatrix}$$

Així $P'P = \Omega^{-1}$ (excepte pel terme $\dfrac{1}{1-\rho^2}$)

Per tant, el model transformat en forma matricial queda:

$$Y^* = PY = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & . & 0 & 0 \\ -\rho & 1 & 0 & . & . & 0 \\ 0 & -\rho & 1 & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & -\rho & 1 \end{bmatrix} . \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_{T-1} \\ y_T \end{bmatrix} = \begin{bmatrix} \sqrt{1-\rho^2}\, y_1 \\ y_2 - \rho y_1 \\ y_3 - \rho y_2 \\ . \\ . \\ y_T - \rho y_{T-1} \end{bmatrix}$$

$$X^* = PX = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & . & 0 & 0 \\ -\rho & 1 & 0 & . & . & 0 \\ 0 & -\rho & 1 & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & -\rho & 1 \end{bmatrix} \begin{bmatrix} 1 & x_{21} & . & x_{k1} \\ 1 & x_{22} & . & x_{k2} \\ 1 & x_{23} & . & x_{k3} \\ . & . & . & . \\ . & . & . & . \\ 1 & x_{2T} & . & x_{kT} \end{bmatrix} =$$

$$= \begin{bmatrix} \sqrt{1-\rho^2} & x_{21}\sqrt{1-\rho^2} & . & x_{k1}\sqrt{1-\rho^2} \\ 1-\rho & x_{22} - \rho x_{21} & . & x_{k2} - \rho x_{k1} \\ 1-\rho & x_{23} - \rho x_{22} & . & x_{k3} - \rho x_{k2} \\ . & . & . & . \\ . & . & . & . \\ 1-\rho & x_{2T} - \rho x_{2,T-1} & . & x_{kT} - \rho x_{k,T-1} \end{bmatrix}$$

i en forma d'equacions:

$$y_t\sqrt{1-\rho^2} = \beta_1\sqrt{1-\rho^2} + \beta_2 x_{2t}\sqrt{1-\rho^2} + \ldots + \beta_k x_{kt}\sqrt{1-\rho^2} + u_t^* \quad t = 1$$
$$y_t - \rho y_{t-1} = \beta_1(1-\rho) + \beta_2\left(x_{2t} - \rho x_{2,t-1}\right) + \ldots + \beta_k\left(x_{kt} - \rho x_{k,t-1}\right) + u_t^* \quad t = 2,\ldots,T$$

3.3.5. Estimació per MQGF del model sota el supòsit de pertorbacions AR(1):

A la pràctica quan el valor de la correlació ρ és desconegut es pot aplicar un dels mètodes següents:

- **Mètode de Cochrane-Orcutt**:

Primer, s'estima per MQO el model original:

$$y_t = \beta_1 + \beta_2 \cdot x_{2t} + \cdots + \beta_k \cdot x_{kt} + u_t \qquad t = 1, \ldots, T$$

El segon pas consisteix en trobar un estimador consistent de ρ, utilitzant els residus MQO de la regressió anterior:

$$\hat{u}_t = \rho \hat{u}_{t-1} + \varepsilon_t \qquad \Rightarrow \quad \hat{\rho}$$

L'últim pas consisteix en utilitzar l'estimador consistent de ρ, per construir l'estimador MQGF:

$$\hat{\beta}_{MQGF} = \left(X'\hat{\Omega}^{-1}X\right)^{-1} X'\hat{\Omega}^{-1}Y$$

on $\hat{\Omega} = \begin{bmatrix} 1 & \hat{\rho} & \hat{\rho}^2 & . & \hat{\rho}^{T-1} \\ \hat{\rho} & 1 & \hat{\rho} & . & . \\ \hat{\rho}^2 & \hat{\rho} & 1 & . & . \\ . & . & . & 1 & \hat{\rho} \\ \hat{\rho}^{T-1} & . & \hat{\rho}^2 & \hat{\rho} & 1 \end{bmatrix}$

A la pràctica pot ser més fàcil aplicar MQO al model transformat:

$$y_t\sqrt{1-\hat{\rho}^2} = \beta_1\sqrt{1-\hat{\rho}^2} + \beta_2 x_{2t}\sqrt{1-\hat{\rho}^2} + ... + \beta_k x_{kt}\sqrt{1-\hat{\rho}^2} + u_t^* \quad t=1$$
$$y_t - \hat{\rho}y_{t-1} = \beta_1(1-\hat{\rho}) + \beta_2\left(x_{2t} - \hat{\rho}x_{2,t-1}\right) + ... + \beta_k\left(x_{kt} - \hat{\rho}x_{k,t-1}\right) + u_t^* \quad t = 2,...,T$$

Si s'ignora la primera observació s'anomena el mètode de Cochrane-Orcutt. Si no es descarta la primera observació s'anomena el mètode de Praïs-Winsten.

- **Mètode de Hildreth-Lu:**

És un mètode de rastreig. El recorregut possible de ρ, l' interval (-1,1) es particiona en una sèrie de punts ρ_i normalment equidistants. Per cada ρ_i s'estima mitjançant MQO el model amb les variables transformades tal com s'ha fet amb el mètode de Cochrane-Orcutt:

$$y_t - \hat{\rho}y_{t-1} = \beta_1(1-\hat{\rho}) + \beta_2\left(x_{2t} - \hat{\rho}x_{2,t-1}\right) + \cdots + \beta_k\left(x_{kt} - \hat{\rho}x_{k,t-1}\right) + u_t^* \qquad t = 2, \ldots, T$$

Es calcula per cada ρ_i la suma de quadrats residuals SQR, i s'escull l'estimació $(\hat{\rho}, \hat{\beta})$ que tingui la SQR mínima.

4. Temes avançats d'econometria

4.1. Models econòmics dinàmics

4.1.1. Presentació

Els models econòmics dinàmics són models que impliquen canvis al llarg del temps, degut a que l'efecte d'una variació unitària en les variables explicatives perdura durant un cert número de períodes.

Un model de retards distribuïts en k períodes amb una variable explicativa es pot escriure com:

$$y_t = \alpha + \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + ... + \beta_k x_{t-k} + u_t$$

β_0 es coneix com multiplicador d'impacte (multiplicador a curt plaç)

$\sum_{i=0}^{k}\beta_i$ es coneix com multiplicador total (multiplicador a llarg plaç)

4.1.2. Estimació dels models de retards distribuïts

Es pot utilitzar MQO ja que X no és estocàstica, però poden surgir els següents problemes pràctics:

1) S’ha de determinar quans valors retardats s’han d’introduïr, ja que moltes vegades la teoria no ho diu.

2) Si s’introdueixen molts retards es perden graus de llibertat, i per tant la inferència estadística és cada vegada menys fiable.

3) Ens podem trobar amb el problema de la col.linealitat, ja que els valors succesius de moltes variables econòmiques tendeixen a estar correlacionats.

4.1.3. Models de Koyck, d’expectatives adaptatives i d’ajustos parcials

Aquests models permeten reduir el nombre de termes retardats i el problema de la col.linealitat. Els models distribuïts es poden escriure com:

$$y_t = \gamma_1 + \gamma_2 x_t + \gamma_3 y_{t-1} + v_t$$

Aquest model es pot dir que és autorregressiu ja que el valor retardat de la variable dependent apareix com a variable explicativa al costat de la dreta.

Ara ja no es pot dir que la matriu de regressors sigui no estocàstica. Per tant per utilitzar MQO, ens tenim que assegurar que el terme d’error v_t i la variable retardada y_{t-1} no estan correlacionades. Si estan correlacionades, els estimadors MQO tenen biaix i no són consistents.

Si v_t i y_{t-1} no estan correlacionades, els estimadors MQO tenen biaix per mostres petites, però aquest biaix tendeix a desaparèixer a mesura que la mida de la mostra creix, per tant els estimadors MQO són consistents.

El paràmetre γ_2 mostra l’efecte a curt plaç d’una variació unitària de x_t sobre y_t mig i $\frac{\gamma_2}{1-\gamma_3}$ mostra l’efecte a llarg plaç d’una variació unitària de x_t sobre y_t mig.

4.2. No estacionarietat, regressions espúries i cointegració

4.2.1. Presentació

Un procés estocàstic és estacionari si la seva mitjana i variància són constants en el temps i el valor de la covariància entre dos períodes temporals depèn únicament del retard entre els dos períodes temporals, i no del moment del temps pel qual es calcula la covariància.

Per tant, si Y_t és una sèrie temporal estocàstica, diem que és estacionària si es compleix:

Mitjana: $E(Y_t) = \mu$

Variància: $E(Y_t - \mu)^2 = \sigma^2$

Covariància: $\gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)]$

Si no es compleixen aquestes condicions diem que la sèrie és no estacionària

4.2.2. Exemple de procés no estacionari: El passeig aleatori

El model de passeig aleatori amb deriva és:

$Y_t = d + Y_{t-1} + u_t$ on u_t és el terme d’error aleatori amb mitjana 0 i variància constant σ^2

L'esperança i variància d'aquest model és:

$$E(Y_t) = Y_0 + t \cdot d$$
$$\operatorname{var}(Y_t) = t \cdot \sigma^2$$

La mitjana i la variància augmenten contínuament, per tant és un procés aleatori no estacionari en mitjana i en variància.

Una variable aleatòria que la seva mitjana i variància depenen del temps es diu que segueix una tendència estocàstica.

4.2.3. Regressions espúries

Les regressions amb dades de sèries temporals ofereixen algunes vegades resultats espuris, o de valor dubtós, en el sentit de que els resultats semblen, a primera vista, bons, però una investigació més pormenoritzada revela que són sospitosos.

El fenòmen de les regressions espúries surgeix quan es fa una regressió d'una sèrie temporal no estacionària sobre una altra sèrie temporal no estacionària. Uns valors elevats de R^2 i t poden fer pensar que s'ha trobat una relació significativa entre les variables. De fet, un R^2 alt pot reflexar únicament que les sèries comparteixen tendències comuns, i per tant que no hi ha cap relació entre les dues.

Les regressions espúries presenten aquestes característiques:

- ✓ Els estimadors són estadísticament significatius amb valors de t i F elevats
- ✓ El valor de R^2 és proper a 1 indicant que el model és adecuat
- ✓ El valor d de Durbin-Watson és petit

Un $R^2 > d$ és una bona regla intuitiva per sospitar que la regressió és espúria.

4.2.4. Tests d'estacionarietat

S'han desenvolupat diversos tests per saber si una sèrie és estacionària. A continuació es presenta el test de l'arrel unitària:

Si Y_t és la sèrie que ens interessa s'estima la següent regressió:

$$\Delta Y_t = \alpha_1 + \alpha_2 t + \alpha_3 Y_{t-1} + u_t$$

La hipòtesis nul.la del test és

$H_0 : \alpha_3 = 0$ (la sèrie Y_t és no estacionària) (hipòtesi de l'arrel unitària)

No és pot utilitzar el test t perquè aquest només és vàlid si les sèries subjacents són estacionàries.

S'ha d'utilitzar el test tau de Dickey-Fuller. Si el valor tau (es calcula com la t) és més petit en valor absolut als valors crítics de Dickey-Fuller no rebutgem la hipòtesi nul.la i en conseqüència la sèrie temporal és no estacionària.

4.2.5. Sèries temporals cointegrades

Encara que les sèries temporals sobre les que fem una regressió no siguin estacionàries, pot ser que existeixi una relació d'equilibri o estable (a llarg plaç) entre elles. En aquest cas, diem que aquestes estan cointegrades.

Per saber si 2 sèries temporals no estacionàries estan cointegrades es fa un test d'estacionarietat sobre les discrepàncies de la regressió entre elles. Si la sèrie de discrepàncies és estacionària, aleshores vol dir que les dues sèries temporals estan cointegrades i per tant, la regressió entre elles no és espúria i té significació econòmica.

És a dir, X_t i Y_t són sèries no estacionàries i volem saber si la regressió $Y_t = \beta_1 + \beta_2 X_t + u_t$ és espúria. Calculem els residus $\hat{u}_t = Y_t - \hat{\beta}_1 - \hat{\beta}_2 X_t$ i fem el test d'estacionarietat $\Delta\hat{u}_t = \gamma \cdot \hat{u}_{t-1} + v_t$. Si el paràmetre γ és significatiu vol dir que les sèries X_t i Y_t estan cointegrades i per tant, la regressió entre elles no és espúria.

I. Annexe – El model de Regressió clàssic

a) *Supòsits del model*

El model de regressió lineal clàssic es basa en els següents supòsits:

- Linealitat:

El model és una funció lineal del vector de paràmetres β:

$$y_t = x_t^{'}\beta + u_t$$

o en forma matricial

$$Y = X\beta + u$$

on Y és un vector nx1, X és una matriu nxK , u és un vector nx1, n és el número d'observacions i K és el número de regressors.

- Errors amb mitja zero, independents i idènticament distribuits (IID):

$$E(u) = 0$$
$$Var(u) = E(uu^{'}) = \sigma_0^2 I_n$$

- Regressors linealment independents i no estocàstics:

La matriu X té rang K i és no estocàstica. És a dir, la matriu X té les columnes linealment independents, on K és el número de regressors. A més els regressors no són aleatoris. També s'ha de complir:

$$\lim_{n\to\infty} \frac{1}{n} X'X = Q_{XX} \quad matriu\ definida\ \ positiva$$

- Normalitat:

El vector dels errors té distribució normal, $u \sim N(0, \sigma^2 I_n)$

b) *Distribució de la variable dependent*

La distribució de la variable dependent y està molt lligada a la distribució de la variable aleatòria u.

$$\left.\begin{array}{r} E(u_i)=0 \\ y_i = x_i'\beta + u_i \end{array}\right\} \Rightarrow E(y_i) = x_i'\beta$$

$$\left.\begin{array}{r} Var(u_i)=\sigma^2 \\ y_i = x_i'\beta + u_i \end{array}\right\} \Rightarrow Var(y_i) = \sigma^2$$

$$u_i \text{ , } u_j \text{ } independents \Rightarrow y_i \text{ } y_j \text{ } independents$$

$$u_i \sim N(0,\sigma^2) \Rightarrow y_i \sim N(x_i'\beta, \sigma^2)$$

c) Estimador MQO

L'estimador MQO és el que minimitza la suma dels errors al quadrat, és a dir

$$\underset{\beta}{Min}\, SQR = \underset{\beta}{Min} \|Y - X\beta\|^2 \Rightarrow \hat{\beta} = (X'X)^{-1} X'Y$$

En el cas particular del model de regressió lineal simple, la solució és:

$$\hat{\beta} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \qquad \hat{\alpha} = \bar{y} - \hat{\beta} \cdot \bar{x}$$

d) Estimació de σ^2

$$X\hat{\beta} = \hat{Y}$$

$\hat{u} = Y - \hat{Y}$ són els residus MQO

L'estimador MQO de σ^2 és $\hat{\sigma}^2 = \dfrac{\hat{u}'\hat{u}}{n-k}$

e) Bondat de l'ajust

Mesura fins a quin punt l'anàlisi fet permet explicar la variabilitat observada en la mostra de la variable dependent. La desviació de la variable dependent respecte la seva mitjana es pot descomposar de la següent manera:

$$y_i - \bar{y} = \hat{y}_i - \bar{y} + y_i - \hat{y}_i$$

Es defineix:

$$SQT = \sum_{i=1}^{n} (y_i - \bar{y})^2 \qquad SQE = \sum (\hat{y}_i - \bar{y})^2 \qquad SQR = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \hat{u}_i^2$$

Una mesura per la bondat de l'ajust és el coeficient de determinació que indica la proporció de variabilitat observada en la mostra que sabem explicar amb el model:

$$R^2 = \frac{SQE}{SQT} = 1 - \frac{SQR}{SQT}$$

Per poder comparar la bondat de l'ajust de regressions amb diferent nombre de regressors s'utilitza el coeficient de determinació corregit,

$$\overline{R}^2 = 1 - \frac{\frac{SQR}{n-k}}{\frac{SQT}{n-1}}$$

f) Distribució i propietats del estimador MQO

Suposem que tenim una mostra de mida n i $\hat{\theta}_n$ és un estimador de θ. Definim les propietats següents:

- Diem que $\hat{\theta}_n$ és un estimador sense biaix de θ si E($\hat{\theta}_n$)=θ .
- Diem que $\hat{\theta}_n$ és un estimador eficient de θ, si entre tots els estimadors sense biaix de θ és el de mínima variància, és a dir, Var($\hat{\theta}_n$) < Var($\tilde{\theta}$) per tot $\tilde{\theta}$ estimador sense biaix de θ

$$\hat{\beta} = (X'X)^{-1}X'Y = (X'X)^{-1}X'(X\beta + u) = \beta + (X'X)^{-1}X'u$$ per tant,

$$E(\hat{\beta}) = \beta \qquad Var(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

Per tant com sabem que u té distribució normal,

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

Propietats:

-- $\hat{\beta}$ no té biaix $E(\hat{\beta}) = \beta$

- $\hat{\sigma}^2$ no té biaix $E(\hat{\sigma}^2) = \sigma^2$
- $\hat{\beta}$ és eficient, és el millor estimador lineal sense biaix (**Teorema de Gauss-Markov**): Sota els supòsits clàssics, l'estimador MQO és l'estimador lineal sense biaix amb variància més petita.

g) Contrast d'hipòtesi sobre els paràmetres d'estimació

Donat el model $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + ... + \beta_k x_{ki} + u_i$
Es fa inferència sobre una d'aquestes betes:

$$H_0: \quad \beta_j = \beta_{j0}$$

Com no es coneix el valor de σ s'ha d'utilitzar el valor de la seva estimació $\hat{\sigma}$. Al substituir el valor de σ pel seu estimador $\hat{\sigma}$, la normal s'ha de substituir per la distribució t-student

$$\frac{\hat{\beta}_j - \beta_{j0}}{\sqrt{\frac{\hat{u}'\hat{u}}{n-k}\left(X'X\right)^{-1}_{jj}}} = \frac{\hat{\beta}_j - \beta_{j0}}{\sqrt{\hat{\sigma}^2\left(X'X\right)^{-1}_{jj}}} = \frac{\hat{\beta}_j - \beta_{j0}}{\sqrt{\hat{\operatorname{var}}\,\hat{\beta}_j}} = \frac{\hat{\beta}_j - \beta_{j0}}{\hat{\sigma}_{\hat{\beta}_j}} \underset{H_0}{\sim} t_{n-k}$$

Com a cas particular està el Test de significació individual . S'utilitza per veure si una variable és significativa o no. La hipòtesis nul.la en aquest cas és $H_0: \quad \beta_j = 0$ i l'estadístic de contrast

és $t = \frac{\hat{\beta}_j}{\hat{\sigma}_{\hat{\beta}_j}}$

També es pot utilitzar l'estadístic t per contrastar una combinació lineal de parámetres, com per exemple:

$$H_0: \quad \beta_1 + \beta_2 = 3$$

$$t = \frac{\hat{\beta}_1 + \hat{\beta}_2 - 3}{\sqrt{\hat{\operatorname{var}}\left(\hat{\beta}_1 + \hat{\beta}_2\right)}} = \frac{\hat{\beta}_1 + \hat{\beta}_2 - 3}{\sqrt{\hat{\operatorname{var}}\,\hat{\beta}_1 + \hat{\operatorname{var}}\,\hat{\beta}_2 + 2\,\hat{\operatorname{cov}}\left(\hat{\beta}_1, \hat{\beta}_2\right)}}$$

h) Contrast per σ^2

La hipòtesis nul.la és:

$$H_0: \quad \sigma^2 = \sigma_0^2$$

Com $u \sim N\left(0, \sigma^2 I\right) \quad \Rightarrow \quad \frac{\hat{u}'\hat{u}}{\sigma^2} \sim \chi^2_{n-k}$

Per tant, sota H_0 ,

$$\frac{\hat{\sigma}^2(n-k)}{\sigma_0^2} \sim \chi^2_{n-k}$$

i) Estimació en forma d'interval

L'estimació en forma d'interval ens serveix per veure la capacitat que tenim de precisar on està el paràmetre.

Sabem $\frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}_{\hat{\beta}_j}} \sim t_{n-k}$

Per tant,

$$\Pr ob\left\{-t_{1-\alpha/2,n-k} < \frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}_{\hat{\beta}_j}} < t_{1-\alpha/2,n-k}\right\} = 1-\alpha$$

$$\Pr ob\left\{\hat{\beta}_j - t_{1-\alpha/2,n-k}\hat{\sigma}_{\hat{\beta}_j} < \beta_j < \hat{\beta}_j + t_{1-\alpha/2,n-k}\hat{\sigma}_{\hat{\beta}_j}\right\} = 1-\alpha$$

De l'expressió anterior es troba l'estimació en forma d'interval:

$$\hat{\beta}_j \pm t_{1-\alpha/2,n-k}\hat{\sigma}_{\hat{\beta}_j}$$

j) Predicció

Donat el model

$$\left.\begin{aligned} y_i &= x'_i \beta + u_i \\ u_i &\sim i.i.N\left(0,\sigma^2\right) \end{aligned}\right\} \quad i = 1,..,n$$

s'intenta fer predicció sobre el comportament de les y per un valor concret de les variables explicatives $x'_p = \left[1, x_{2p}, ..., x_{kp}\right]$

Es considera $y_p = x'_p \beta + u_p$ per qualsevol x_p

S'ha de distingir entre predicció sobre la mitjana i predicció individual:

Predicció sobre la mitjana del valor de la y:

$E\left(y_p / x_p = x_0\right) = E\left(y_0\right) = x'_0 \beta$. Per tant $\hat{E}\left(y_0\right) = x'_0 \hat{\beta}$

Predicció individual:

$\left(y_p / x_p = x_0\right) = y_0 = x'_0 \beta + u_0$. Per tant $\hat{y}_0 = x'_0 \hat{\beta}$

Encara que les prediccions individuals i sobre la mitjana coincideixen tenen errors diferents:

Error Predicció sobre la mitjana:

$$\hat{E}\left(y_o\right) = x'_0 \hat{\beta}$$

$$error = \hat{E}\left(y_0\right) - E\left(y_0\right) = x'_0 \hat{\beta} - x'_0 \beta = x'_0 \left(\hat{\beta} - \beta\right)$$

Per tant si es millora l'estimació de β es redueix l'error de predicció.

Error Predicció individual:

$$\hat{y}_0 = x'_0 \hat{\beta}$$

$$error = \hat{y}_0 - y_0 = x'_0 \hat{\beta} - \left(x'_0 \beta + u_0\right) = x'_0 \left(\hat{\beta} - \beta\right) - u_0$$

En aquest cas l'error depèn de la diferència en l'estimació de la β i a la naturalesa del component aleatori.

Per trobar les prediccions en forma d'interval es necessita coneixer la distribució del error:

Predicció en forma d'interval sobre la mitjana:

$$\hat{E}\left(y_0\right) = x'_0 \hat{\beta}$$

$$error = x'_0 \left(\hat{\beta} - \beta\right)$$

Com $\hat{\beta} \sim N\left(\beta, \sigma^2 \left(X'X\right)^{-1}\right) \quad \Rightarrow \quad E\left(error\right) = 0$

$$\operatorname{var}\left(error\right) = \operatorname{var}\left(x'_0 \left(\hat{\beta} - \beta\right)\right) = x'_0 \operatorname{var}\left(\hat{\beta} - \beta\right) x_0 = x'_0 \operatorname{var}\left(\hat{\beta}\right) x_0 = x'_0 \sigma^2 \left(X'X\right)^{-1} x_0$$

L'interval de confiança per $E\left(y_0\right)$

$$\left[\hat{E}\left(y_0\right) \pm t_{1-\alpha/2,n-k} \sqrt{\hat{\sigma}^2 x'_0 \left(X'X\right)^{-1} x_0}\right]$$

Predicció en forma d'interval individual:

$$\hat{y}_0 = x'_0\,\hat{\beta}$$
$$error = x'_0\left(\hat{\beta} - \beta\right) - u_0$$

Per tant la distribució de l'error en aquest cas és $error \sim N\left(0, \sigma^2\left(1 + x'_0\,(X'X)^{-1}x_0\right)\right)$
L'interval de confiança és en aquest cas

$$\left[\hat{y}_0 \pm t_{1-\alpha/2,n-k}\sqrt{\hat{\sigma}^2\left(1 + x'_0\,(X'X)^{-1}x_0\right)}\right]$$

Fórmules de predicció en el cas particular del model simple

$$\left.\begin{array}{l} y_i = \beta_1 + \beta_2 x_{2i} + u_i \\ \quad u_i \sim i.i.N\left(0, \sigma^2\right) \end{array}\right\}$$

Els intervals de predicció en aquest cas són

mitjana $\quad E(y_0):\quad \hat{y}_0 \pm t_{1-\alpha/2,n-k}\sqrt{\hat{\sigma}^2\left\{\frac{1}{n} + \frac{(x_{20} - \bar{x}_2)^2}{\sum (x_{2i} - \bar{x}_2)^2}\right\}}$

individual $\quad y_0:\quad \hat{y}_0 \pm t_{1-\alpha/2,n-k}\sqrt{\hat{\sigma}^2\left\{1 + \frac{1}{n} + \frac{(x_{20} - \bar{x}_2)^2}{\sum (x_{2i} - \bar{x}_2)^2}\right\}}$

Bibliografia

- **Econometría** - Montserrat Díaz Fernández, María del Mar Llorente Marrón – Pirámide

- **Análisis Econométrico** – William H. Greene – Prentice Hall

- **Principios de Econometría** – Damodar N. Gujarati -- McGrawHill

- **Introduction to the Theory and Practice of Econometrics** – Judge, Hill, Griffiths, Lütkephol, Lee – Wiley

-**The Econometric Modelling of Financial Time Series** – Terence C. Mills – Cambridge University Press

-**The Econometrics of Financial Markets** – John Campbell – Princeton University Press