

MATERIA  
Econometría I

unidade  
didáctica  
**3**

TITULACIÓN  
Administración e Dirección de Empresas

# Probas de hipóteses sobre os coeficientes dun MRLNC

**Juan Carlos Estévez Núñez**

Área de Economía Cuantitativa  
Departamento de Economía Cuantitativa  
Facultade de Ciencias Económicas e Empresariais

unidadesdidácticas  
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

**DESCATALOGADO**

Esta obra atópase baixo unha licenza Creative Commons BY-NC-SA 3.0. Calquera forma de reprodución, distribución, comunicación pública ou transformación desta obra non incluída na licenza Creative Commons BY-NC-SA 3.0 só pode ser realizada coa autorización expresa dos titulares, salvo excepción prevista pola lei. Pode acceder Vde. ao texto completo da licenza nesta ligazón:

<http://creativecommons.org/licenses/by-nc-sa/3.0/es/legalcode.gl>

© Universidade de Santiago de Compostela, 2015

**Deseño e maquetación**

J. M. Gairí

**Edita**

Servizo de Publicacións e intercambio científico  
da Universidade de Santiago de Compostela  
[usc.es/publicacions](http://usc.es/publicacions)

ISBN

978-84-16533-26-8

DOI

<http://dx.doi.org/10.15304/9788416533268>

**MATERIA:** Econometría I

**TITULACIÓN:** Administración e Dirección de Empresas

PROGRAMA XERAL DO CURSO

Localización da presente unidade didáctica

## **BLOQUE TEMÁTICO I. QUE É A ECONOMETRÍA?**

### **Unidade I. A Econometría**

Concepto actual e antecedentes históricos

Modelos económicos e modelos econométricos: representación formal

Clasificación dos modelos econométricos

O papel da Econometría

## **BLOQUE TEMÁTICO II. O MODELO DE REGRESIÓN LINEAR NORMAL CLÁSICO (MRLNC)**

### **Unidade II. Hipóteses básicas e estimación do MRLNC**

Introdución. Representación formal do modelo

Hipóteses que determinan que un modelo sexa un MRLNC

Métodos de estimación dun modelo de regresión

Estimación polo método de mínimos cadrados ordinarios (MCO)

Propiedades dos estimadores de MCO

Fiabilidade dos estimadores e medidas para analizar o axuste

### **Unidade III. Probas de hipóteses sobre os coeficientes dun MRLNC**

Introdución

Metodoloxía para realizar as probas de hipóteses sobre os coeficientes

Representación formal xeral e particular

Definición dos estatísticos de proba e cálculo co aplicativo GRETL

Regras de decisión e a súa interpretación

Exemplos máis representativos

### **Unidade IV. Predición, estabilidade e estimación máximo-verosímil**

Introdución

Predición e medidas avaliadoras da capacidade para predicir

Estabilidade post-mostral

Estimación máximo-verosímil

### **BLOQUE TEMÁTICO III. OUTROS TÓPICOS DUN MRLNC**

#### **Unidade V. Variables ficticias e probas de estabilidade**

Introdución de variables ficticias nun modelo econométrico

Probas de estabilidade dos coeficientes

#### **Unidade VI. Multicolineariedade e selección de regresores**

Problema da multicolineariedade e grao de presenza

Métodos para seleccionar regresores

#### **Unidade VII. Aplicativos informáticos**

LibreOffice Calc (licenza Mozilla) e Excel (de Microsoft Corporation)

GRETl (licenza GNU) e EViews (de IHS Incorporation)

PSPPP (GNU), SPSS (de IBM) e Statgraphics Centurion (StatPoint Tech. Inc.)

## ÍNDICE

---

### PRESENTACIÓN

### OBXECTIVOS

### METODOLOXÍA E ACTIVIDADES

### CONTIDOS

1. Introducción
  - 1.1. Interpretación dos coeficientes do MRLNC
  - 1.2. Tipo de cuestións que permiten responder as probas de hipóteses sobre os coeficientes
2. Metodoloxía para realizar as probas de hipóteses sobre os coeficientes
3. Representación formal xeral e particular
4. Definición dos estatísticos de proba e cálculo co aplicativo GRETL
  - 4.1. A suma dos cadrados dos erros mínimo-cuadráticos do modelo sen restricións (SCE) e do modelo restrinxido (SCER)
  - 4.2. A definición do estatístico de proba xeral e a súa distribución de probabilidade
  - 4.3. Cálculo do estatístico de proba co aplicativo GRETL
5. Regras de decisión e interpretación
6. Exemplos máis representativos
  - 6.1. Probas sobre un único parámetro
    - 6.1.1. Un caso especial: probas de nulidade dun só parámetro
    - 6.1.2. Outro caso especial: probas do signo dun parámetro
  - 6.2. Probas sobre varios parámetros
    - 6.2.1. Probas sobre subconxuntos paramétricos
    - 6.2.2. Probas sobre unha combinación linear de parámetros

### ACTIVIDADES PROPOSTAS

### AVALIACIÓN

### BIBLIOGRAFÍA

## PRESENTACIÓN

---

Nesta unidade didáctica (UD) expóñense algúns contidos fundamentais para o desenvolvemento dunha materia que se imparte no grao en Administración e Dirección de Empresas da Universidade de Santiago de Compostela (en adiante, USC), denominada *Econometría I*. Esta ensínase ao longo do primeiro cuadrimestre do terceiro curso do grao, cunha carga docente de tres horas e media semanais. Está orientada a un alumnado do que se agarda que teña fundamentos sólidos de dirección de empresas, teoría económica, estatística e matemáticas, aínda que só sexa a nivel básico.

Con esta materia preténdese instruír ao alumnado nun xeito de representar a realidade empresarial en particular, ou económica en xeral, mediante algún conxunto de hipóteses tales como as que conforman o que se denomina *modelo econométrico*. Esta idea ao noso xuízo debería desenvolverse con certos matices nos contidos dunha unidade didáctica tal como a salientada como Unidade I no apartado de localización desta.

Outra parte importante da materia *Econometría I* consiste na explicación das técnicas básicas de análise dun modelo econométrico a partir da información dispoñible de tipo cuantitativo, relativa ás magnitudes que se representan nel. Iso implica tamén advertir e xustificar cales son as técnicas máis apropiadas para cada tipo de modelo. En consecuencia, o tipo de contido suxerido para a Unidade II pode revelar os aspectos máis importantes dese proceso.

Tamén se procura expoñer os métodos apropiados para verificar mediante a evidencia empírica dispoñible, a validez das hipóteses asumidas inicialmente ao propoñer o modelo. Dada a gran diversidade deses supostos iniciais, resulta de especial interese distinguir os que fan referencia ás relacións que se establecen entre as magnitudes consideradas no modelo e todos os demais. E xusto nese senso, xorde a necesidade da presente UD para tratar polo miúdo a primeira clase deles, á que denominaremos en xeral *proba de hipóteses sobre os coeficientes*.

No desenvolvemento da citada materia tamén se clarifican os métodos básicos para determinar a capacidade dun modelo para facer predicións sobre o comportamento da magnitude de maior interese e para admitir a estabilidade dese modelo. Contidos como os indicados na Unidade IV permitirán un achegamento a eses tópicos.

E finalmente, recoméndase dedicar as tres últimas UD a estudar outros temas de especial relevancia nun curso introdutorio á Econometría, nos que se analice: como introducir información de tipo cualitativo nos modelos econométricos (Unidade V), como escoller tanto as magnitudes como as variables explicativas apropiadas que se utilizarán no modelo (Unidade VI) ou como manexar aplicativos informáticos para desenvolver as tarefas de cálculo que sexan precisas (Unidade VII), coa conseguinte interpretación dos resultados que aqueles ofrezan.

Dado entón o papel que xogan as probas de hipóteses neste esquema de investigación científica, é de vital importancia coñecer os detalles da súa posible utilización. Como resultado diso cabería agardar que o seu emprego se faga co máximo rigor, de xeito que se reforce a credibilidade deste tipo de metodoloxía. Isto último ten a súa relevancia dada a presenza de materias que esixen a utilización de

técnicas de análise cuantitativa, nos plans de estudo vixentes na maioría das universidades dos países máis desenvolvidos. E tampouco non debe perderse de vista que esta materia está obviamente relacionada coa titulada *Econometría II* que se imparte tamén no mesmo plan de estudos da USC. Ademais, a planificación académica require que esta última se imparta ao longo do segundo cuatrimestre (tamén do terceiro curso) e que os seus contidos sexan en parte continuación dos de *Econometría I*. De aí a importancia dos aquí expostos.

En canto á planificación temporal, os contidos que se presentan poderanse desenvolver en tres sesións expositivas, dunha hora de duración cada unha, nas que se explicarían os de tipo teórico, e noutras dúas sesións interactivas nas que se realizarían as correspondentes prácticas, cunha duración de dúas horas e media cada unha. Isto sumaría un total de 8 horas que representarían aproximadamente o 16% do total de horas de docencia da materia contempladas no plan de estudos.

## OBXECTIVOS

---

Con esta unidade didáctica aspírase a acadar dous tipos de metas. Por unha banda, están os obxectivos específicos da propia UD, bosquexados nos aspectos que se relacionan deseguido.

- **Obxectivo 1.** Distinguir o tipo de cuestións que permite dilucidar unha proba de hipóteses sobre os coeficientes dun MRLNC.
- **Obxectivo 2.** Seguir unha metodoloxía establecida para probar hipóteses dun modelo econométrico.
- **Obxectivo 3.** Representar formalmente toda a tipoloxía de hipóteses que se queiran probar sobre os coeficientes  $\beta_i$ .
- **Obxectivo 4.** Identificar os estatísticos de proba axeitados en cada caso.
- **Obxectivo 5.** Manexar o aplicativo GRETL para calcular o valor dos estatísticos de proba e outra información relacionada.
- **Obxectivo 6.** Tomar decisións sobre cal das hipóteses se considera máis verosímil, dada a información dispoñible.

E, por outra banda, están algúns dos **obxectivos característicos da propia materia**, dos que neste caso propóñense os seguintes.

- **Obxectivo 1.** Conseguir interpretar de forma correcta o significado dos coeficientes dun modelo econométrico básico (ao que axudará que o alumnado acade os obxectivos 1 e 3).
- **Obxectivo 2.** Utilizar un aplicativo informático que permita estimar o modelo e conseguir probas das fortalezas ou eivas deste último (ao que vai contribuír que se logren os obxectivos 4 e 5).
- **Obxectivo 3.** Recoñecer as vantaxes e as limitacións das técnicas cuantitativas de análise económica (ao que axudará que se consigan os obxectivos 1, 2 e 6).

## METODOLOXÍA

---

Para que esta UD sexa de utilidade para un eventual lector é preciso que se dean algunhas condicións. A primeira a resaltar é a necesidade de que os usuarios desta xa teñan asimilado previamente tanto unha terminoloxía coma determinados conceptos. En certo modo, contidos similares aos das Unidades I e II detalladas na parte de localización, permitirán acadar ese obxectivo.

Ademais, tendo en conta que a materia para a que se propón esta unidade didáctica ten unha certa connotación empírica, proponse utilizar unha metodoloxía de ensino que mesture a docencia expositiva coa interactiva. Isto último esixe, entre outras cousas, certo dominio tanto de conceptos estatísticos previos como no manexo de ordenadores e aplicativos informáticos. Por iso recoméndase facer unha proba sinxela de avaliación a principio de curso que inclúa estes aspectos. E se cadra, cun deseño apropiado, iso axudará a detectar as fortalezas e as limitacións que poda ter o alumnado para desenvolver os contidos da materia coa metodoloxía que aquí se propón.

Tal como se deduce do exposto na presentación, os contidos están ideados para desenvolvelos durante tres semanas do curso académico, cunha duración de tres horas e media durante as dúas primeiras e dunha hora na terceira.

Nas tres semanas, é aconsellable que se dedique a primeira hora semanal de docencia dispoñible para a materia á exposición de aspectos teóricos a un grupo de alumnos non necesariamente reducido. Isto sería mellor facelo co soporte dun encerado ou dalgún medio de presentación visual (diapositivas, transparencias, ordenador, etcétera). Nesta parte, non é necesaria a participación activa do alumnado na propia exposición máis alá do que supoña solucionar dúbidas respecto dos contidos expostos polo docente. Pero si é vital, que ao estudante lle queden moi claros os puntos de apoio irrenunciáveis da explicación. Isto é básico tanto respecto a determinados conceptos como ao propio procedemento a seguir cando se fai unha proba de hipóteses sobre os coeficientes (ou parámetros  $\beta_i$ ).

As restantes 2,5 horas da primeira semana suxírese impartilas nunha aula de informática, co alumnado das aulas expositivas dividido en grupos tan reducidos como permita a organización docente do centro. Ao longo delas, o alumnado de cada grupo haberá de estar repartido ocupando cada posto de traballo (e cada ordenador) preferentemente por unha parella de estudantes.

Recoméndase ao persoal docente que dedique a primeira destas horas a explicar a resolución dun exercicio práctico. Para iso utilizará un ordenador que dispoña dos aplicativos informáticos recomendados<sup>1</sup>, ademais dun canón vídeo-proxector ou dun encerado dixital. E presentará ao alumnado un banco de datos cargado no aplicativo GRETL, que inclúa a descrición das variables incorporadas ao mesmo, así como o ámbito mostral que se pretenda estudar.

Tamén sería oportuno presentar un ficheiro-formulario de tipo *folla de cálculo* deseñado con celas reservadas<sup>2</sup> para indicar os resultados de GRETL e os comentarios que se requiran do alumnado.

---

1 Neste caso GRETL (sigla de *Gnu Regression, Econometrics and Time-series Library*) ou Eviews.

2 O resto das celas da folla recoméndase que estean protexidas para evitar erros fortuitos.



No resto do tempo cada parella encargarase, co asesoramento puntual do docente, da realización dun exercicio propio cuxa solución haberá de entregarse ao instrutor para a súa posterior avaliación e devolución. O apropiado sería poder distribuír un banco de datos diferente a cada parella de estudantes para resolver as cuestións propostas. Este exercicio repetiríase como reforzo coa resolución doutro exercicio na segunda semana, engadindo algúns contidos novos.

Tendo en conta que é moi elevado o número de estudantes que actualmente cursan os estudos de grao en Administración e Dirección de Empresas, aconséllase propor métodos de avaliación que poidan facerse de forma eficiente cun custo de tempo razoable. Nese senso, deseñalas nun formato de tipo informático axudará tanto a que o alumnado as cumprimente, como a que o docente realice a revisión, aínda que para isto esíxese o esforzo adicional de manexar adecuadamente os aplicativos informáticos necesarios. Por iso, tamén será imprescindible poñer especial coidado en avaliar a consecución dos obxectivos específicos da UD, non confundíndoos con obxectivos que dependen das habilidades do alumnado en cuestións de tipo informático.

## CONTIDOS

---

### 1. Introducción

Tal como debера quedar claro cando se desenvolva a UD I, un *modelo económico* entraña un conxunto de hipóteses. Con elas trátase de representar de forma simplificada o xeito no que se relaciona unha magnitude que resulta de especial interese con unha ou varias das que se supón que depende, coa condición de que cando menos algunha delas teña connotacións de tipo económico. O adxectivo de *económico* fai non só unha evidente mención ao eido da economía, senón que introduce un recendo a un enfoque científico de tipo empírico.

**Exemplo:** As magnitudes relacionadas poderían ser, por un lado, o grao de interese que teñen os produtos dunha empresa para os seus consumidores reais e potenciais e, polo outro, o custo económico que ten para ela poñelos a disposición destes, o esforzo que fai a empresa para dalos a coñecer, etc.

A proposta das primeiras versións destes modelos remóntase ao período 1910-1930<sup>3</sup>, inspiradas nun espírito propio do paradigma económico neoclásico. Por iso a simplificación que implican adoitaba formalizarse utilizando unha linguaxe matemática e un enfoque de tipo analítico, nos que se distinguían os seguintes elementos:

- as *variables* como representacións das magnitudes;
- as *ecuacións* como expresións do tipo de relacións entre estas; e
- os *parámetros* como símbolos de aspectos concretos destas relacións.

---

<sup>3</sup> Cos traballos de Moore (1914), Cobb e Douglas (1928) e Schultz (1928) entre outros.

Coa adopción do concepto de regresión e dun enfoque de tipo cuantitativo á hora de afrontar a análise dos fenómenos económicos, por aquel tempo comezou a utilizarse no eido da economía o precursor do actual modelo econométrico básico. Este modelo de regresión<sup>4</sup> foi utilizado antes noutras disciplinas científicas. Actualmente matízase o seu nome engadíndolle os cualificativos de *linear*, *normal* e *clásico* para distinguilo doutro tipo de modelos similares. De aí que se utilice a sigla *MRLNC* para designalo. A súa representación formal pódese ilustrar coa expresión

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k + \varepsilon \quad (1.1)$$

onde:

- $Y$  indica a variable que se pretende estudar (ou *regresando*);
- $X_i$  sinala cada unha das variables (ou *regresores*) que se consideran máis importantes para explicar o comportamento daquela;
- $\varepsilon$  recolle todas as demais variables que poden influír no regresando pero que non se enumeran de xeito explícito; e
- $\beta_i$  denota cada un dos coeficientes que acompañan ás variables e que, nesta unidade didáctica, denomínanse *parámetros*<sup>5</sup>.

### 1.1. Interpretación dos coeficientes do MRLNC

Estes últimos elementos do modelo que nos ocupa son de especial interese na UD pois tendo en conta que

$$\frac{\partial Y}{\partial X_i} = \beta_i \quad (1.2)$$

enténdese decontado que se podan interpretar os coeficientes como indicadores das variacións marxinais do regresando ante variacións marxinais de cada regresor.

Exemplo (continuación): Supoñendo que as variables *Vendas*, *Prezo* e *Gasto en publicidade* representan ás tres magnitudes mencionadas antes de forma que a primeira é o regresando, e as outras dúas son o primeiro e o último regresor que se consideran importantes para explicar as *Vendas*, da expresión

$$\text{Vendas} = \beta_0 + \beta_1 \text{Prezo} + \dots + \beta_k \text{Gasto en publicidade} + \varepsilon \quad (1.3)$$

resulta

$$\frac{\partial \text{Vendas}}{\partial \text{Prezo}} = \beta_1 \quad (1.4)$$

Entón pódese deducir que o coeficiente  $\beta_1$  indicaría a variación marxinal das *Vendas* ante un cambio marxinal no *Prezo*.

<sup>4</sup> Galton (1886) o «inventou» para usalo no eido da bioloxía.

<sup>5</sup> Aínda que nun modelo econométrico hai outros elementos que tamén son parámetros.

Exemplo (continuación): Tamén resulta

$$\frac{\partial \text{Vendas}}{\partial \text{Gasto en publicidade}} = \beta_k \quad (1.5)$$

E o coeficiente  $\beta_k$  indicaría a variación marxinal das *Vendas* ante un cambio marxinal no *Gasto en publicidade*.

Un dos papeis máis importantes da Econometría é o de atribuír un valor que poda resultar verosímil a cada un destes parámetros  $\beta_i$ , o que se denomina *cuantificar* as relacións entre as variables. Trátase dunha tarefa ambiciosa pois non se contenta con esbozar estas relacións, senón que procura detallalas utilizando o poder simbólico do número co risco que iso supón dada a enorme simplificación que implica.

Cando se propón un modelo econométrico, algunhas das hipóteses que o definen fan referencia ao xeito en que a priori se cre que inflúen os regresores sobre o regresando. Por iso un xeito de reafirmar a súa credibilidade implica probar que ditas hipóteses son consistentes coa forma na que se comportan as variables na realidade. Isto último é debido á inspiración empírica que está no fundamento destes modelos, como se comentou con anterioridade.

## 1.2. Tipo de cuestións que permiten responder as probas de hipóteses sobre os coeficientes

Ao suxerir os posibles regresores deste tipo de modelos, non sempre é evidente o tipo de influencia que podan ter sobre a variable explicada. Así, con frecuencia acontece que unha visión pouco coídadosa parece amosar que unha magnitude é importante para explicar o comportamento doutra, pero coa análise da información dispoñible sobre as variables escollidas para representalas, iso non resulta tan evidente. Este comportamento aparentemente paradoxal pode ser debido a varias causas, das que merecen destacarse:

- unha deficiente selección de variables, regresores e regresando, para xogar ese papel representativo;
- a existencia doutras importantes magnitudes latentes nesas relacións, que non están téndose en conta ao propoñer o modelo; e
- unha incorrecta interpretación da utilidade deste sinxelo tipo de modelos econométricos, pois serve fundamentalmente para tratar de explicar a *variabilidade* do regresando.

Por esta razón, un primeiro tipo de cuestións que pode axudar a responder a realización dunha proba de hipóteses é o relativo á importancia de cada regresor para explicar o comportamento do regresando. Trátase entón de demostrar se o papel de cada un deles por separado é irrelevante dentro do modelo, ou se pode considerarse valioso. Esta situación equivale a determinar se resulta coherente dicir que o valor de cada un dos coeficientes  $\beta_j$  que acompañan a cada regresor é distinto de cero ou se non o é.

Exemplo (continuación): Polo tanto se a evidencia empírica permite soste-  
razoadamente que o coeficiente  $\beta_1$  é nulo, iso implicaría que as variacións  
do *Prezo* non suporían ningún cambio relevante nas *Vendas*, polo que  
poderíase concluír que a súa elección como regresor realmente non foi  
acertada.

Unha proba de hipóteses tamén pode ser de utilidade para comprobar se os  
datos dispoñibles sobre as variables permiten admitir que, en xeral, ao aumentar o  
valor dun regresor tamén vai aumentar o valor da variable explicada. En caso  
afirmativo iso significa que sería razoable soste- que o coeficiente que expresa esa  
relación entre as dúas deba de ter signo positivo. Pola contra, se o valor da  
variable explicada se vise en xeral reducido, o signo do coeficiente agárdase que  
sexa negativo.

Evidentemente os casos anteriores alentan de inmediato a pensar nunha  
xeneralización e mestura de ambos, que apunta a aquel carácter ambicioso da  
investigación econométrica comentado máis arriba na epígrafe 1.1, non exento de  
certos riscos. Así, outro tipo de preguntas que permite responder este tipo de  
probas é o relativo á credibilidade de soste- que un parámetro teña un valor  
concreto. Isto equivale a verificar se é admisible que un aumento dunha unidade na  
variable explicativa asociada a el, poderá provocar unha variación concreta no  
regresando (aumentándoo ou diminuíndoo), supoñendo que se manteñen estables  
o resto de variables do modelo<sup>6</sup>.

Exemplo (continuación): En consecuencia, admitir que  $\beta_1 = -2$  implicaría  
aceptar que os aumentos dunha unidade no *Prezo* (coa *cláusula ceteris*  
*paribus*) agárdase que provoquen unha diminución de 2 unidades nas  
*Vendas*.

A forma na que se utilice este tipo de razoamento pode facer dun modelo  
econométrico unha potente ferramenta de planificación para unha empresa. Pero  
tamén está latente o risco de cometer graves erros, se a metodoloxía que se utiliza  
non é consistente cos obxectivos deste tipo de modelos, ou se algunha etapa do  
proceso se fai á lixeira.

Como remate desta epígrafe cómpre dar un paso máis nesta senda de  
xeneralización, pois as probas de hipóteses das que trata esta unidade didáctica  
poden involucrar simultaneamente a varios regresores. Entón o abano de  
posibilidades concretas que se ofrece é inabarcable, aínda que cabe destacar dous  
tipos de cuestións moi relevantes. Por un lado, as que se refiren á utilidade ou  
inutilidade conxunta de varios dos regresores do modelo, o que equivale a probar  
se é admisible que os seus parámetros  $\theta$ , sexan nulos simultaneamente. E polo  
outro, as que se cuestionan posibles relacións entre os propios coeficientes, o que  
equivale a preguntarse como serán as relacións entre as propias relacións das  
variables!

---

<sup>6</sup> Suposto que se coñece como *cláusula ceteris paribus*, e que se denotará (*ccp*) en diante.

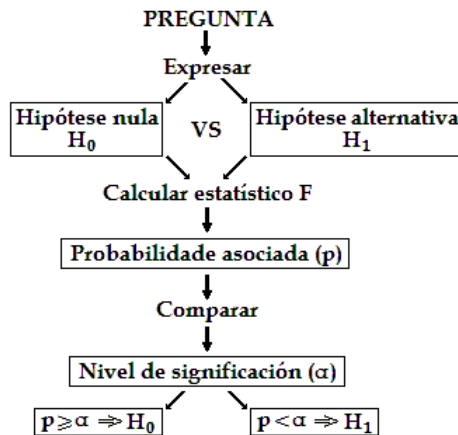
## 2. Metodoloxía para realizar as probas de hipóteses sobre os coeficientes

A continuación indícase unha proposta dunha sucesión de etapas a seguir para realizar unha proba de hipóteses sobre os parámetros  $\beta_i$  dun modelo econométrico. A grandes trazos estas poden resumirse cando menos en cinco pasos que resultan imprescindibles.

- **1º paso:** Definir claramente unha pregunta que se tratará de responder coas conclusións da proba de hipóteses. Neste punto é primordial prestar especial atención ao feito de que a cuestión estea referida ao xeito no que se relacionan os regresores co regresando, e non a outro tipo de aspectos do modelo que podan tamén interesar aínda que noutro contexto.
- **2º paso:** Expresar con notación matemática dúas hipóteses que se exclúan mutuamente e cuxa veracidade vai compararse coa axuda da proba. Isto implica ser capaces de distinguir por un lado as hipóteses que involucran a un único regresor, e que se contrastan con outras que son complementarias delas total ou parcialmente. E polo outro lado, débense considerar as que se refiren a varios regresores á vez.
- **3º paso:** Establecer o estatístico de proba que se debe utilizar en cada caso e calculalo a partir da información dispoñible sobre as variables. Unha das claves desta metodoloxía está neste paso pois introdúcese o concepto de *probabilidade*, admitindo explicitamente a posibilidade de que coa proba poida chegarse a unha conclusión errónea. Isto último dependerá da credibilidade que teña o valor do estatístico de proba calculado, e iso represéntase mediante un concepto estatístico denominado *distribución de probabilidade*.
- **4º paso:** Seguir un protocolo, é dicir unhas *regras de decisión*, que permita determinar cal das dúas hipóteses é máis verosímil. É evidente que este paso está fortemente condicionado polo anterior, pois dependerá da máxima probabilidade admisible de errar ao adoptar a decisión final.
- **5º paso:** Tomar unha decisión sendo conscientes do tipo de erro que se pode cometer ao facelo, e do grao de importancia do mesmo. Neste punto é fundamental insistir en interpretar correctamente as conclusións dunha proba de hipóteses sobre os parámetros  $\beta_i$ . Todo esforzo por marcar a diferenza entre o xeito de razoar no que a dúbida non deixa nunca de estar presente e o razoamento no que a dúbida desaparece case de milagre, pagará a pena para delimitar as formas de pensar que se consideran partícipes da ciencia contemporánea, das que non o son.

Estas etapas básicas poden representarse nunha secuencia como a indicada na figura 1.

Figura 1: Esquema da epígrafe 2



Fonte: Elaboración propia

### 3. Representación formal xeral e particular

A pesar do comentado con anterioridade na epígrafe 1.2 sobre os diferentes tipos de cuestións que permiten abordar as probas de hipóteses sobre os parámetros  $\beta_i$ , a linguaxe matemática permite representar calquera deles cunha expresión común. Para iso, o conxunto de coeficientes dun MRLNC exprésase mediante o vector

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} . \quad (3.1)$$

Chegados a este punto, convén enfocar estas hipóteses dende outro ángulo, a partir do concepto de *restrición linear*. Isto significa que as diferentes hipóteses, en troques de interpretalas coma un conxunto máis ou menos heteroxéneo e extenso de posibles valores ou combinación de valores dos coeficientes, pódense propoñer como casos particulares dun máis xeral. Logo o feito de limitar ese caso xeral a un en particular implica establecer restricións lineais sobre o xeral.

En termos matemáticos, impoñer unha das restricións equivale a establecer unha particular combinación linear dos parámetros  $\beta_i$ , cun determinado vector  $(r_n)$  de multiplicadores  $(r_{ni})$ . Ao escribir ese vector como

$$r_n = \begin{pmatrix} r_{n0} \\ r_{n1} \\ \vdots \\ r_{nk} \end{pmatrix} \quad (3.2)$$

onde o subíndice  $n$  indica que se trata da  $n$ -ésima restrición dentro dun posible conxunto  $N$  delas, poderase representar cada unha utilizando a expresión

$$\begin{pmatrix} r_{n0} & r_{n1} & \dots & r_{nk} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = v_n \quad (3.3)$$

ou expresala coa notación vectorial

$$r_n^t \beta = v_n \quad (3.4)$$

onde  $v_n$  é o valor que terá a combinación linear dos parámetros  $\beta_i$ , con  $i=0,1,\dots, k$ .

En consecuencia, supoñer un conxunto simultáneo de  $N$  restricións sobre os  $\beta_i$  convida a definir unha matriz  $R$  que ten  $N$  filas e  $k+1$  columnas, de tal forma que

$$\begin{pmatrix} r_{10} & r_{11} & \dots & r_{1k} \\ r_{20} & r_{21} & \dots & r_{2k} \\ r_{30} & r_{31} & \dots & r_{3k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{N0} & r_{N1} & \dots & r_{Nk} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_N \end{pmatrix} \quad (3.5)$$

é dicir,  $R\beta=V$ .

**Exemplo (continuación):** A hipótese de que o *Gasto en publicidade* non inflúe de forma significativa nas *Vendas (ccp)*, formularíase mediante a expresión

$$\begin{pmatrix} 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = \beta_k = 0 \quad (3.6)$$

De forma similar, a hipótese indicada na epígrafe 1.2 relativa á diminución nas *Vendas* ante un aumento dunha unidade no *Prezo (ccp)*, exprésase

$$\begin{pmatrix} 0 & 1 & \dots & 0 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = \beta_1 = (-2) \quad (3.7)$$

Por outro lado, unha hipótese que involucra a varios regresores na mesma restrición sería que o *Prezo* puidese ter unha influencia similar sobre as *Vendas (ccp)* á influencia do *Gasto en publicidade* sobre elas, o que significaría que  $\beta_1 = \beta_k$  ou tamén  $\beta_1 - \beta_k = 0$ , e podería enunciarse

$$\begin{pmatrix} 0 & 1 & \dots & -1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = 0 \quad (3.8)$$

**Exemplo (continuación):** Finalmente, unha hipótese que implica a varios regresores en varias restricións podería ser que os regresores non teñen un efecto conxunto relevante sobre as *Vendas*, o que significaría que todos os seus parámetros se anulan, é dicir  $\beta_1=0, \beta_2=0, \dots, \beta_k=0$  ou tamén

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} 0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} . \quad (3.9)$$

Para verificar calquera das hipóteses comentadas nesta epígrafe sométese cada unha delas a confrontación coa súa respectiva alternativa, polo que se utilizan as expresións *hipótese nula* e *hipótese alternativa* para referirse ás que se confrontan en cada caso. Formalmente utilízase a linguaxe matemática para representar ambas as dúas, denotándose  $H_0$  á primeira e  $H_1$  á segunda. E tendo en conta a expresión xenérica (3.5) calquera proba de hipóteses sobre os coeficientes pode indicarse por medio de

$$H_0: [R\beta = V] \quad \text{frente a} \quad H_1: [R\beta \neq V] . \quad (3.10)$$

En xeral a hipótese nula representa o cumprimento das restricións sobre os coeficientes que se pretende probar, e a alternativa o incumprimento das mesmas.

#### 4. Definición dos estatísticos de proba e cálculo co aplicativo GRETL

O enfoque da epígrafe 3 integra todas as probas de hipóteses sobre os coeficientes como casos particulares dun caso xeral. E iso vai permitir simplificar tanto a definición coma o cálculo dos estatísticos que se utilizan para resumir a información relevante, e dispoñible en forma de datos, para ditas probas. De feito, todos resultan ser casos particulares dun estatístico máis xenérico que responde a unha idea moi sinxela que se expón deseguido.

Evidentemente ao facer unha representación dunha realidade utilizando unha simplificación da mesma, cométese erros. E só fai falta un mínimo de intuición para advertir que canto menor sexa a capacidade desa simplificación para captar detalles desa realidade, maiores tenderán a ser os erros cometidos. Isto implica que un modelo econométrico, como simplificación que é, tamén responderá a esta lóxica. E cando se considera, en xeral, que un modelo ten maior grao de simplificación? A resposta no contexto que nos ocupa é obvia. Canto maior sexa o número de restricións que incorpora.

Polo tanto, a primeira idea que debe quedar absolutamente clara nesta epígrafe é que un modelo con unha ou varias restricións sobre os parámetros  $\beta_i$ , vai cometer máis erros ca un modelo sen as mesmas, pois vai simplificar máis a realidade. A clave entón para decidir se unha restrición é conveniente, estará en determinar ata que punto ese aumento nos erros cometidos cae dentro duns límites considerados tolerables para que a restrición proposta resulte admisible.



#### 4.1. A suma dos cadrados dos erros mínimo-cuadráticos do modelo sen restricións (SCE) e do modelo restrinxido ( $SCE_R$ )

Na unidade didáctica II xa debera demostrarse que logo de propoñer un modelo de regresión que se supón que cumpre as hipóteses dun MRLNC, recoméndase calcularlle os seus coeficientes aplicando o método de estimación de *mínimos cadrados ordinarios (MCO)*<sup>7</sup>. Pois tanto ese método como os estimadores a que dá lugar presentan unha serie de interesantes propiedades, se efectivamente se trata dun modelo *clásico*. Como este método arrastrará inevitablemente certos erros, defínese a noción de *suma de erros cadrados (SCE)* a partir deles.

Cando se calcula esta SCE nun modelo sen restricións lineais sobre os parámetros  $\beta$ , acostúmase a denotala como SCE, polo que ao calculala nun modelo con algunha restrición pode denotarse  $SCE_R$ . Pois ben, o estatístico xenérico para probar hipóteses sobre os coeficientes trata de medir o aumento relativo que se produce na suma de erros cadrados cando se introduce algunha ou algunhas restricións sobre os mesmos. Lembrando que se cumprirá que  $SCE_R \geq SCE$ , segundo o apuntado nos parágrafos anteriores, o incremento relativo na SCE cando se introduce algunha restrición nos parámetros  $\beta$ , dun MRLNC podería expresarse con

$$\frac{SCE_R - SCE}{SCE} \quad (4.1)$$

aínda que, como se verá, isto non abonda sen facer unha ulterior emenda.

**Exemplo (continuación):** Recompilando durante un período de varios meses, os agregados mensuais nunha empresa das variables do modelo

$$Vendas = \beta_0 + \beta_1 \text{Prezo} + \dots + \beta_k \text{Gasto en publicidade} + \varepsilon \quad (4.2)$$

xa sería posible calcular a suma de erros cadrados que cometería o método de estimación de mínimos cadrados ordinarios (MCO) aplicado a esa mostra de varios meses. Supóñase que o resultado é  $SCE = 100$ .

Cando se propón a hipótese de que quizais ningún dos regresores sexa relevante a excepción do *Prezo*, estase establecendo implicitamente un conxunto de  $k-1$  restricións de nulidade, é dicir que  $\beta_2=0, \dots, \beta_k=0$ , o que significaría que o modelo (4.2) quedaría reducido ou *restrinxido* a

$$Vendas = \beta_0^* + \beta_1^* \text{Prezo} + \xi \quad (4.3)$$

e agora ao estimar estoutro modelo co método MCO obteríase  $SCE_R$ . Supóñase que agora resulta  $SCE_R = 640$ . Entón con estes datos hipotéticos, o incremento relativo en SCE ao impoñer as  $k-1$  restricións de nulidade sería

$$\frac{(640 - 100)}{100} = 5,4$$

é dicir, a SCE aumentaría un 540% como consecuencia de restrinxir o modelo (4.2) ao obrigarse a que algúns dos seus coeficientes sexan nulos.

<sup>7</sup> Ao igual que moitos outros métodos de estimación alternativos, este trata de optimizar unha función obxectivo. E xa que se dá por descontado que o modelo vai cometer erros, a optimización deste método consistirá en tratar de minimizar unha función deses erros.

Pois ben, debe de terse en conta que este cociente estase calculando a partir de dous conceptos,  $SCE$  e  $SCE_R$ , que foron determinados con antelación utilizando un número diferente de *graos de liberdade*<sup>8</sup> para cada un. Isto esixe facer unha corrección en (4.1) que contemple esta circunstancia, polo que precísase denotar:

- $T$  para representar o número de observacións dispoñibles das variables;
- $k+1$  para indicar o número de coeficientes independentes a estimar no modelo sen restricións, co que os seus graos de liberdade son  $T-(k+1)$ ; e
- $N$  para expresar o número de restricións.

#### 4.2. A definición do estatístico de proba xeral e a súa distribución de probabilidade

Xa se está entón en condicións de dar o terceiro paso indicado na epígrafe 2 desta UD, definindo o estatístico que pode utilizarse para facer calquera proba de hipóteses relativa a restricións lineais dos parámetros  $\beta_i$  dun MRLNC. Representátese habitualmente mediante a letra  $F$  e defínese mediante

$$F = \frac{\left( \frac{SCE_R - SCE}{N} \right)}{\left( \frac{SCE}{T - k - 1} \right)} . \quad (4.4)$$

Nunha UD semellante á Unidade II indicada na sección de localización debéra explicarse en que consiste o suposto de *normalidade* das perturbacións que permite cualificar ao MRLNC como *normal*. Se o modelo proposto admite este tipo de suposición, pode demostrarse que o estatístico  $F$  se comporta como unha variable con distribución de probabilidade de tipo  $F$  de Fisher-Snedecor, no caso de que a hipótese nula sexa certa (que as restricións sexan correctas). En tal caso tería  $N$  graos de liberdade no numerador e  $(T-k-1)$  graos de liberdade no denominador.

Isto na práctica equivale a dicir que pode utilizarse a teoría da probabilidade para decidir canto teñen de cribles as restricións sobre os coeficientes. Pois pode calcularse o valor do estatístico de proba  $F$  a partir da mostra de observacións, e resulta coñecida a súa distribución de probabilidade. Con esas dúas premisas poderás asociárselle ao valor calculado para  $F$ , unha probabilidade que permitirá tomar unha decisión con certa marxe de erro, e que se denominará *probabilidade asociada* ao estatístico de proba.

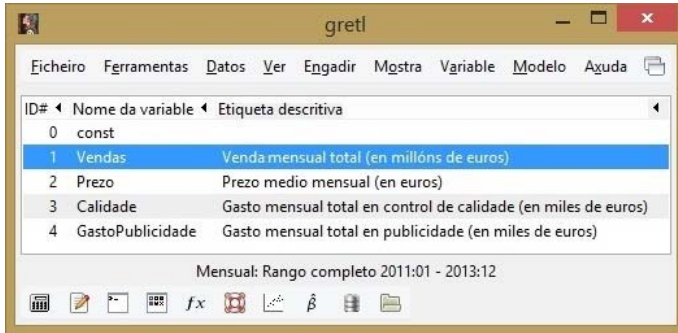
#### 4.3. Cálculo do estatístico de proba co aplicativo GRETL

Para aplicar calquera método para estimar os parámetros dun modelo econométrico, o método de mínimos cadrados ordinarios (MCO) entre eles, é preciso realizar unha serie de cálculos. E existe unha gran variedade de aplicativos informáticos especializados nesta tarefa.

<sup>8</sup> Concepto mencionado por primeira vez por R. Fisher en 1915 e que pode interpretarse no ámbito da estatística como o número de valores do cálculo dun estatístico que poden variar libremente.

O aplicativo de soporte lóxico GRETL<sup>9</sup> resulta moi útil para realizar esa tarefa coa vantaxe de ser de libre distribución. A fiestra principal do aplicativo tería o seguinte aspecto cun banco de datos mensuais que abrangueran desde xaneiro de 2011 ata Nadal de 2013:

**Figura 2:** Exemplo de fiestra principal de GRETL



Fonte: Elaboración propia

No apartado *Modelo* do menú desta fiestra pódese propoñer a estimación co método MCO, escollendo a variable dependente e as variables explicativas que se desexe utilizar (figura 3).

**Figura 3:** Proposta de estimación MCO



Fonte: Elaboración propia

9 Diferentes versións deste aplicativo para varios sistemas operativos, poden descargarse seguindo os enlaces da páxina (en español) <http://gretl.sourceforge.net/es.html>. Este é resultado da colaboración acotío dun grupo de profesores de Econometría de todo o mundo e dispón de tradución ao galego.

Tras encher o cadro e premer Aceptar, o aplicativo abre unha nova fiestra cos resultados da estimación do modelo proposto (figura 4).

Figura 4: Exemplo de modelo estimado

	Coefficiente	Desv. típica	Estatístico t	Valor p
const	-288,718	118,991	-2,426	0,0211
Prezo	-1,93959	0,187396	-10,35	9,70e-012
Calidade	2,68383	0,737698	3,638	0,0010
GastoPublicidade	0,948490	0,0594662	15,95	8,73e-017
Media var.dependente Y	47,27458	Cuasidesv. típica de Y	18,94115	
Suma de erros cadrados	863,8407	Desv.típ. regresión(S)	5,195673	
R-cadrado	0,931206	R-cadrado corrixido	0,924756	
F(3, 32)	144,3848	Valor p (de F)	1,12e-18	
Log-verosimilitude	-108,2834	Criterio de Akaike	224,5669	
Criterio de Schwarz	230,9010	Crit. de Hannan-Quinn	226,7776	
rho	0,090163	Durbin-Watson	1,817505	

Fonte: Elaboración propia

Unha vez estimado o modelo xa poderá facerse unha proba de hipóteses sobre os coeficientes, escollendo a opción *Probas* deste último tipo de fiestra. No menú que se desprega a continuación escóllese o apartado *Restricións lineais*.

Ábrese entón unha nova fiestra na que deben indicarse as restricións que se pretendan probar sobre os coeficientes. Con este fin é preciso seguir un estrito protocolo formal que obriga a respectar as letras maiúsculas e minúsculas dos nomes das variables, utilizar corchetes para recollelas e expresar todas as restricións rematadas nun signo igual, seguido dun número.

Figura 5: Proba de nulidade de tódolos coeficientes do modelo da figura 4

Especificas as restricións:  
(Consultar a Axuda para máis información)

b[const]=0  
b[Prezo]=0  
b[Calidade]=0  
b[GastoPublicidade]=0

Usar bootstrap

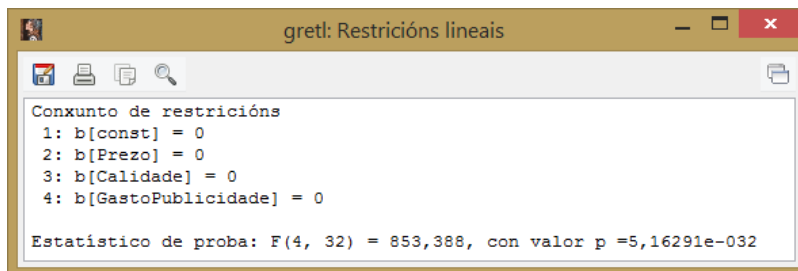
Axuda Cancelar Aceptar

Fonte: Elaboración propia

Logo de escribir as restricións e premer *Aceptar*, o aplicativo abre unha nova fiestra na que se ofrece como información máis relevante:

- o conxunto de restricións que se examinan na proba de hipóteses;
- o valor que acada o estatístico F cos graos de liberdade de numerador e denominador entre parénteses; e
- a probabilidade asociada ao mesmo (*valor p*).

**Figura 6.** Cálculo do estatístico de proba F e da súa probabilidade asociada



Fonte: Elaboración propia

E xa neste punto, abonda con seguir as regras de decisión que se explican na seguinte epígrafe desta UD, para tomar unha resolución sobre a credibilidade de cada unha das dúas hipóteses propostas. Pero todo isto sen esquecer que a decisión que se tome, sempre vai estar condicionada por algunha probabilidade de equivocarse.

## 5. Regras de decisión e interpretación

A probabilidade asociada ao estatístico F (*valor p* en GRETL) representa a probabilidade de cometer o denominado *erro tipo I* ao probar as hipóteses. Este tipo de erro é primordial pois suporía equivocarse cando se decide rexeitar a hipótese nula ( $H_0$ )<sup>10</sup>, polo que procúrase que a probabilidade de cometelo sexa razoablemente baixa.

Por iso, para decidir ata que punto é máis convincente rexeitar  $H_0$  ou o é non rexeitala, primeiro é preciso establecer un valor fronteira que exprese canto se está disposto a errar ao tomar esa decisión. Ese valor denomínase *nivel de significación* e está contido entre 0 e 1, aínda que adóitase a expresalo en porcentaxe. Habitualmente represéntase co símbolo  $\alpha$ .

Se, por exemplo, se establece que  $\alpha=0,1$  (ou 10%) asúmese que se decidirá rexeitar  $H_0$  se a probabilidade de equivocarse ao tomar esa decisión é, como máximo, do 10%. Quérese dicir que este límite marca o risco máximo de erro que se está disposto a asumir ao rexeitar  $H_0$ . En consecuencia, se a probabilidade asociada

<sup>10</sup> Nun coñecido símil que identifica a hipótese nula coa de *non culpabilidade* dun reo nun xuízo e a alternativa coa hipótese de *culpabilidade*, o *erro tipo I* representaría declarar culpable a alguén que realmente non o é.

ao estatístico F neste exemplo é menor ca 0,1 entón decídese rexeitar a hipótese nula<sup>11</sup>, o que equivale a rexeitar a validez das restricións que se propoñen na proba de hipóteses. Pola contra, se a probabilidade asociada a F fora do 0,7 por exemplo, non se rexeitaría a hipótese nula pois a probabilidade (70%) de equivocarse cando se tome esa decisión sería maior á que se estaba disposto a asumir nun principio (10%).

Polo tanto, a regra de decisión en xeral é ben sinxela:

- se resulta que (*valor p* de F)  $< \alpha \rightarrow$  rexéitase a hipótese nula e as restricións nela indicadas; e
- se resulta que (*valor p* de F)  $> \alpha \rightarrow$  a hipótese nula non debера ser rexeitada.

Do exposto no parágrafo anterior se deduce logo que a decisión sobre cal das dúas hipóteses se considera máis crible non exime da posibilidade de equivocarse. En consecuencia debe de poñerse un coidado moi especial en interpretar correctamente o resultado dunha proba de hipóteses deste tipo, procurando evitar facer afirmacións categóricas de máis ao decidirse por unha delas.

## 6. Exemplos máis representativos

Deseguido vanse expor algúns exemplos ilustrativos do xeito máis común de aplicar esta metodoloxía, utilizando o modelo estimado da figura 4 como apoio. Seguiranse as etapas indicadas na epígrafe 2 desta UD, cuxa secuencia representouse coa figura 1.

Para ligar de forma adecuada o primeiro paso co segundo, expresando correctamente a pregunta que resulte de interese en termos de dúas hipóteses que se enfronten, deberá terse moi presente a interpretación dos parámetros  $\beta_i$  do MRLNC tal como se comentou previamente na epígrafe 1.1.

Pero antes de comentar algúns deses exemplos, cabe aclarar unha cuestión. Se xa se estimaron os coeficientes dun MRLNC, para que é preciso facer probas de hipóteses sobre eles? E necesaria esta comprobación? Por exemplo, se na figura 4 xa se indica que  $b_2 = -1,93959$  (o estimador de MCO que acompaña á variable *Prezo*), ten sentido preguntarse se se cumprirá que o parámetro  $\beta_2 = -2$ ?

Pois realmente si, porque cun método de estimación como o aplicado para obter os resultados da figura 4, estímase que as *Vendas* diminuirán 1,93959 unidades ante aumentos unitarios do *Prezo* (*ccp*). Pero este cálculo está feito con unhas expresións matemáticas (os estimadores  $b_i$ ) que, por un lado, dependen da mostra de datos dispoñible e, por outro lado, conservan certa aleatoriedade<sup>12</sup>, polo que non son valores constantes senón que se asume que poden variar. Nun certo exercicio de humildade, admítase que quizais o efecto sobre as *Vendas* non sexa dunha diminución exacta de 1,93959 unidades, aínda que se agarda que sexa unha cantidade non moi diferente, quizais incluso poda ser mesmo de 2 unidades.

11 Como sucede na figura 6, onde a probabilidade asociada é aproximadamente  $5,16 \cdot 10^{-32} < 0,1$  polo que non se admite que sexan certas tódalas catro restricións desa figura.

12 Por depender das perturbacións, tal como debера explicarse nunha Unidade como a UD II.

### 6.1. Probas sobre un único parámetro

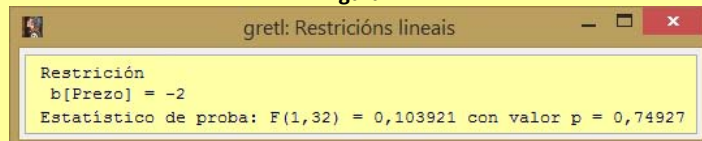
Con este tipo de probas trata de verificarse como é a influencia dun regresor sobre a variable que se trata de explicar, supoñendo que os demais regresores permanecen inalterados (*ccp*). As etapas da figura 1 correspóndense agora con isto.

- Pregunta tipo → É verosímil crer que un aumento unitario no regresor  $X_i$  ten como consecuencia unha variación de  $z$  unidades na variable explicada (*ccp*) ou esa variación non é de  $z$  unidades?
- Hipóteses →  $H_0: [\beta_i=z]$  fronte a  $H_1: [\beta_i \neq z]$ .
- Estatístico de proba e probabilidade asociada → Neste caso o estatístico de proba  $F$  ten 1 grao de liberdade no numerador pois só se propón unha restrición e  $(T-k-1)$  graos de liberdade no denominador. Con estes dous datos, unha vez achado  $F$  xa se pode obter a súa probabilidade asociada ( $p$ ).
- Regra de decisión → Tras decidir cal é o nivel de significación ( $\alpha$ ) que se está disposto a admitir, se resulta  $p \geq \alpha$  conclúese que a hipótese nula é máis crible. Noutro caso rexéitase, e considerase máis plausible a alternativa.

Exemplo (continuación): A pregunta de se é crible que un aumento unitario no *Prezo* poida provocar unha diminución de 2 unidades nas *Vendas* (véxase epígrafe 1.2 e expresión 3.7) implicaría  $H_0: [\beta_i=-2]$  fronte a  $H_1: [\beta_i \neq -2]$ .

Na fiestra de restricións de GRETl indicárase  $b[\text{Prezo}] = -2$ , co resultado:

Figura 7



Fonte: Elaboración propia

Se o nivel de significación ( $\alpha$ ) se establece no 5%, entón a hipótese nula resultaría máis crible que a alternativa, dada a mostra de datos mensuais utilizada, pois a probabilidade asociada a  $F$  é maior co nivel de significación ( $74,927\% > 5\%$ ). Por iso responderíase afirmativamente á pregunta.

#### 6.1.1. Un caso especial: probas de nulidade dun só parámetro

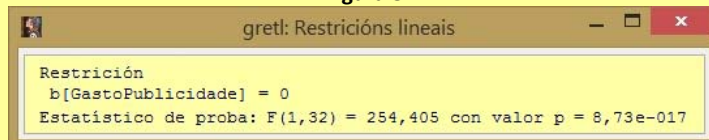
Un interesante caso deste tipo de probas é aquel no que se quere probar se o valor dun parámetro  $\beta_i$  é cero, como o comentado a raíz de definir expresión (3.6). Se iso fora certo, o regresor que o acompaña non tería influencia sobre o regresando e a especificación do modelo veríase afectada. O proceso agora é este.

- Pregunta tipo → Pode sosterse que un aumento unitario no regresor  $X_i$  non provoca ningunha variación na variable explicada (*ccp*)?
- Hipóteses →  $H_0: [\beta_i=0]$  fronte a  $H_1: [\beta_i \neq 0]$ .
- O estatístico de proba  $F$ , a súa probabilidade asociada ( $p$ ) e a regra de decisión seguirían pautas de cálculo semellantes ás do exemplo anterior.

**Exemplo (continuación):** A cuestión da expresión (3.6) sobre a falta de influencia do *Gasto en publicidade* sobre as *Vendas* levaría a ter que expresar  $H_0: [\beta_3=0]$  fronte a  $H_1: [\beta_3 \neq 0]$ .

Na fiestra de GRETL indicaríase  $b[\text{GastoPublicidade}] = 0$  e o resultado sería:

Figura 8



Fonte: Elaboración propia

Se  $\alpha = 5\%$ , entón a hipótese nula non resultaría crible dada a mostra usada, pois a probabilidade asociada ao estatístico F resulta menor ca o nivel de significación ( $8,73 \cdot 10^{-17} < 0,05$ ) polo que  $H_0$  sería rexeitada.

Conclúese así que o *Gasto en publicidade* semellaría ter unha influencia significativa sobre as *Vendas*, aínda que esta conclusión tería unha certa probabilidade de non ser correcta (do 0,000000000000000873%). Obsérvase que este dato xa estaba indicado na figura 4, na columna de valores  $p$ .

#### 6.1.2. Outro caso especial: probas do signo dun parámetro

Como consecuencia dos resultados do tipo de proba indicada na epígrafe 6.1.1, pode chegarse á conclusión de que é máis crible dicir que un regresor inflúe de xeito significativo no regresando ca dicir que non inflúe, como se deduce do exemplo da figura 8. Pero chegados a este punto, cabería preguntarse ademais se esa influencia é positiva ou negativa. E isto equivale a indagar polo signo do parámetro que acompaña ao regresor.

Para tomar unha decisión aquí coa metodoloxía das probas de hipóteses, só se modifica algo o esquema anterior, prestando moita atención á regra de decisión.

— **Pregunta tipo** → Pódese manter que os aumentos do regresor  $X_i$  provocan variacións positivas na variable explicada (*ccp*)?

— **Hipóteses** →  $H_0: [\beta_i=0]$  fronte a  $H_1: [\beta_i > 0]$ .

— O estatístico F acada xusto o mesmo valor que na epígrafe 6.1.1.

— A regra de decisión é agora diferente pois a hipótese nula só se rexeita se o signo do estimador  $b_i$  coincide co signo de  $\beta_i$  en  $H_1$  e, ademais, a metade da probabilidade que GRETL asocia ao estatístico F é inferior a  $\alpha$ .

**Exemplo (continuación):** Ao preguntarse pola credibilidade de que un aumento no *Gasto en publicidade* provoque aumentos nas *Vendas* (*ccp*) expresaríanse as hipóteses  $H_0: [\beta_3=0]$  fronte a  $H_1: [\beta_3 > 0]$ .

Dada esta última hipótese alternativa e dado que na figura 4 obtívose que  $b_3 = 0,94849 > 0$  (os signos coinciden), hai que ver se a metade do valor  $p$  de F que indica GRETL é menor que  $\alpha$  (certo pois  $[8,73 \cdot 10^{-17}/2] < 0,05$ ). Asía a  $H_0$  non se admitiría, confirmándose a relación positiva entre as variables.



Exemplo (continuación): De xeito análogo, ao pescudar se un aumento no *Prezo* provoca aumentos nas *Vendas (ccp)* expresaríase  $H_0: [\beta_1=0]$  fronte a  $H_1: [\beta_1>0]$ .

Pero dado o signo nesta última hipótese e comprobando que na figura 4 se obtivo  $b_1 = -1,93959 < 0$ , a non coincidencia de ambos os dous signos implica que a hipótese nula non pode rexeitarse.

Con este resultado aparentemente paradoxal, entón admítese que  $\beta_1=0$ ?

A resposta é negativa, por dúas razóns:

- lémbrese que preferir unha das hipóteses á outra non implica asumir que é certa pois hai unha probabilidade de equivocarse; e
- admitir unha das dúas hipóteses unicamente implica preferirla á outra, pero nada máis.

Neste último senso diríamos que preferimos a hipótese nula porque a evidencia empírica dispoñible na mostra de datos mensuais utilizada nos leva a preferir dicir que  $\beta_1=0$  antes que dicir que  $\beta_1>0$  (pois neste caso a hipótese correcta debера ser que  $\beta_1<0$ ).

Outra posibilidade consiste en preguntarse se a relación entre a variable explicativa e o regresando é inversa. Os únicos aspectos que cambian respecto do exemplo anterior son os seguintes.

- Pregunta tipo → Pódese manter que un aumento unitario no regresor  $X_1$  provoca unha diminución na variable explicada (*ccp*)?
- Hipóteses →  $H_0: [\beta_1=0]$  fronte a  $H_1: [\beta_1<0]$ .
- O valor de F sería exactamente o mesmo que no caso anterior, e as regras de decisión seguirían as mesmas pautas.

Exemplo (continuación): Para probar se pode admitirse que un aumento no *Gasto en calidade* provoca diminucións nas *Vendas (ccp)*, expresaríanse as hipóteses  $H_0: [\beta_2=0]$  fronte a  $H_1: [\beta_2<0]$ .

Pero dado o signo nesta última hipótese e comprobando que na figura 4 se obtivo  $b_2 = 2,68383 > 0$ , a non coincidencia de ambos os dous signos implicaría que sería preferible a hipótese nula antes que a alternativa.

## 6.2. Probas sobre varios parámetros

A diferenza das probas da epígrafe 6.1, nas que só tiña sentido propoñer unha restrición na hipótese nula ao estar implicado un único parámetro, estoutro tipo de probas pode implicar unha ou varias restricións. Ademais, dependendo do número de coeficientes  $\beta_i$  do modelo, as combinacións que se poden explorar son tantas que resultaría imposible detallalas. Pero si é factible deterse en dous casos particulares que teñen certa relevancia no uso decotío desta clase de modelos.

Por un lado está a posibilidade de que cada parámetro do modelo tome un valor concreto, diferente ou non aos dos demais parámetros. As probas que exploran esa situación denomínanse *probas (de varias restricións) sobre subconxuntos paramétricos*.

E por outra parte tamén pode comprobarse se unha combinación linear de varios parámetros toma un determinado valor. Estoutros casos noméanse *probas (dunha restrición) sobre unha combinación linear de parámetros*.

### 6.2.1. Probas sobre subconxuntos paramétricos

Con esta epígrafe faise referencia a aquelas situacións nas que pretende comprobarse se é convincente soste que varios parámetros toman determinados valores. O caso máis xeral deste tipo de probas sería aquel no que se considera que todos e cada un dos parámetros  $\beta_i$  do modelo toman valores concretos  $z_i$ .

Trataríase así de probar se é crible que  $\beta_0=z_0, \beta_1=z_1, \beta_2=z_2, \dots, \beta_k=z_k$  polo que a expresión matemática que recolle as restricións desa situación sería un caso particular da expresión (3.5) como se indica coa seguinte ecuación

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \\ z_k \end{pmatrix} \quad (6.1)$$

Xa que logo xorde aquí a idea de que un xeito de xustificar se ten un mínimo de validez do modelo do que se parte, consistiría en probar o caso especial no que todos os elementos do vector  $\beta$  son nulos (e tamén todo  $z_i=0$ ). Entón propónse isto.

- Pregunta tipo → É plausible soste que o modelo de partida ten un mínimo de validez? ou ben: Os regresores escollidos teñen conxuntamente influencia significativa na variable explicada?
- Hipóteses →  $H_0: [\beta=0]$  fronte a  $H_1: [\beta \neq 0]$ .
- Estatístico de proba e probabilidade asociada → Neste caso o estatístico de proba F tería tantos graos de liberdade no numerador como parámetros ten o modelo ( $k+1$ ) pois propónse ese número de restricións, e ( $T-k-1$ ) graos de liberdade no denominador. Con estes dous datos e o valor de F xa se pode obter a súa probabilidade asociada ( $p$ ).
- Regra de decisión → Abondaría de novo con verificar a relación entre o nivel  $\alpha$  e o valor  $p$  para decidir cal das dúas hipóteses resulta máis plausible.

**Exemplo (continuación):** Para verificar esta hipótese de validez do modelo, probaríase  $H_0: [\beta=0]$  fronte a  $H_1: [\beta \neq 0]$ .

A expresión que habería que utilizar con GRETL xa foi adiantada como exemplo nas figuras 5 e 6. Na derradeira compróbase que  $F= 853,388$  cun valor  $p \approx 5,163 \cdot 10^{-32} (<0,05)$ . Entón non é crible soste que os parámetros do modelo proposto sexan todos nulos (cun  $\alpha=5\%$ ). O modelo sería válido.

Dado que a constante dun MRLC só se utiliza para mellorar as propiedades matemáticas do mesmo, tamén se utiliza unha variante desta proba para verificar se os regresores son válidos en conxunto, sen prestar atención ao coeficiente  $\beta_0$ . A expresión (6.1) quedaría entón reducida á ecuación (3.9) e os graos de liberdade do numerador xa só serían  $k$ , manténdose  $T-k-1$  para o denominador. Para achar o valor do estatístico e a probabilidade asociada en GRETL abondaría con eliminar a restrición  $b[\text{const}]=0$  da fiestra de restricións.

**Exemplo (continuación):** As hipóteses serían  $H_0: [\beta_1=\beta_2=\dots=\beta_k=0]$  fronte á alternativa  $H_1: [\text{Algún } \beta_i \neq 0]$ .

Neste caso o cálculo do estatístico F e o seu valor  $p$  poderían acharse eliminando a restrición  $b[\text{const}]=0$  incluída no caso anterior. Pero en realidade xa están no cadro xeral da estimación do modelo da figura 4. Son os valores que toman  $F(3, 32)=144,3848$  e Valor  $p$  (de F)=  $1,12 \cdot 10^{-18}$ .

### 6.2.2. Probas sobre unha combinación linear de parámetros

Todas as probas que se veñen explicando na epígrafe 6 da presente UD xa se introduciron brevemente no 1.2, en cuxo parágrafo final se comentaba a posibilidade tamén de verificar se son admisibles determinadas relacións entre os coeficientes  $\beta_i$  dun modelo. Estas relacións poden expresarse mediante varias restricións, pero nesta epígrafe se comentará só o caso no que se establece unha única restrición. A representación formal da mesma xa se indicou en (3.3) coa notación desa epígrafe da UD, o que equivalía a

$$r_{10}\beta_0 + r_{11}\beta_1 + r_{12}\beta_2 + \dots + r_{1k}\beta_k = v_1 \quad (6.2)$$

onde debe de lembrarse que as constantes  $r_{10}, r_{11}, r_{12}, \dots, r_{1k}$  recóllense no vector  $r_1$  como en (3.2).

As etapas a seguir agora serían as seguintes.

- Pregunta tipo → É crible a restrición  $r_1^t \beta = v_1$  no modelo?
- Hipóteses →  $H_0: [r_1^t \beta = v_1]$  fronte a  $H_1: [r_1^t \beta \neq v_1]$ .
- Estatístico de proba e probabilidade asociada → Neste caso o estatístico de proba F terá 1 grao de liberdade no numerador porque só hai unha restrición, e  $(T-k-1)$  graos de liberdade no denominador. Considerando tamén o valor de F xa se pode calcular a súa probabilidade asociada ( $p$ ).
- Regra de decisión → Haberá que decidir cal das dúas hipóteses resulta máis crible, comparando o nivel  $\alpha$  co valor  $p$ .

O exemplo máis paradigmático deste caso é a relación que deben cumprir os coeficientes dunha función de produción de tipo Cobb-Douglas con rendementos a escala constantes. Lémbrese que esta función represéntanse coa ecuación

$$P = w T^a K^b \quad (6.3)$$

na que as variables están expresadas mediante  $P$  para a produción,  $T$  para o traballo e  $K$  para o capital, e as constantes son  $w$ ,  $a$  e  $b$ . Pero o tratamento econométrico da mesma faise mediante unha transformación logarítmica na que resulta

$$\ln(P) = \ln(w) + a \ln(T) + b \ln(K) + \varepsilon \quad (6.4)$$

o que, renomeando os elementos da ecuación, equivale formalmente a

$$Y = \beta_0 + \beta_1 \ln(T) + \beta_2 \ln(K) + \varepsilon \quad (6.5)$$

Para que efectivamente os rendementos a escala sexan constantes débese cumprir que  $\beta_1 + \beta_2 = 1$  (ou  $a + b = 1$ ), e isto é unha restrición dunha combinación linear dos parámetros

$$0\beta_0 + 1\beta_1 + 1\beta_2 = 0 \quad (6.6)$$

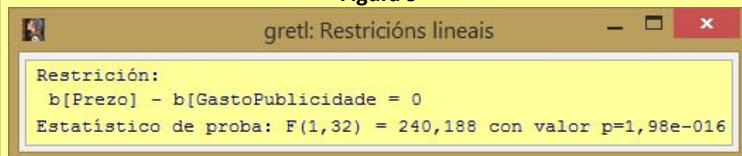
Coa información estatística que puidese proporcionar unha empresa, ou con agregados nacionais como fixeron os propios Cobb e Douglas, poden estimarse os coeficientes deste modelo e probar a hipótese de rendementos a escala constantes.

**Exemplo (continuación):** Na expresión (3.8) xa se propoñía investigar se a forma en que afecta o *Prezo* marxinalmente ás *Vendas (ccp)* é similar a como o fai o *Gasto en publicidade*. Isto aplicado ao modelo cuxa estimación se representa na figura 4 supón probar  $H_0: [\beta_1 = \beta_3]$  fronte a  $H_1: [\beta_1 \neq \beta_3]$ .

Pero no aplicativo GRETL despois do signo igual dunha restrición só se poden poñer números polo que as probas é mellor reescribilas opoñendo  $H_0: [\beta_1 - \beta_3 = 0]$  fronte a  $H_1: [\beta_1 - \beta_3 \neq 0]$ .

Indicárase na fiestra de restricións  $b[\text{Prezo}] - b[\text{GastoPublicidade}] = 0$

**Figura 9**



Fonte: Elaboración propia

Supoñendo que o nivel de significación  $\alpha = 5\%$ , a probabilidade asociada ao estatístico é menor ( $1,98 \cdot 10^{-14}\%$ ) polo que a hipótese nula non resultaría plausible, convidando a pensar que os efectos marxinais de ambos os dous regresores non son similares.

## ACTIVIDADES PROPOSTAS

Na sección de *Metodoloxía* desta UD recoméndase a exposición polo docente da metodoloxía que se representou coa figura 1 como resumo da epígrafe 2, facendo especial fincapé na interpretación dos coeficientes do MRLNC e no tipo de preguntas sobre este cuxa resposta pode depender deles. Para desenvolvela sería preciso dedicar a hora de docencia interactiva das dúas primeiras semanas, e uns 45 minutos da terceira de xeito que os 15 minutos restantes desta última poidan

utilizarse para un pequeno exercicio de avaliación escrita mediante un test. En consecuencia, unha primeira actividade que se recomenda é propoñer ao alumnado que represente formalmente por escrito diversos tipos de hipóteses sobre un modelo formulado con variables reais e poñer en común as respostas dadas. Esta actividade recoméndase desenvólvela no tramo final da segunda hora de aulas expositivas.

A segunda actividade consistiría na resolución por parte do alumnado dun exercicio similar ao explicado polo docente que ocuparía aproximadamente as 1,5 horas restantes das aulas interactivas da primeira semana, e repetiríase como reforzo na segunda semana con outro exercicio. Como aspectos máis relevantes desta actividade cabe salienta:

- a recomendación de entregar un ficheiro-formulario que debería cubrir cada parella de estudantes cos resultados, así como un banco de datos diferente para cada unha co fin de evitar o mero plaxio dos resultados finais de cada exercicio;
- a oportunidade de que o ficheiro-formulario cos resultados se envíe por correo electrónico ao profesorado ao final de cada aula interactiva, e que se comprobe a súa recepción antes de abandonala;
- a conveniencia de estimular ao alumnado para que procure axuda para as súas dúbidas entre o resto do alumnado presente na aula, promovendo así un talante cooperativo; e
- a necesidade de que o persoal docente preste continua atención e asistencia ás limitacións en habilidades informáticas do alumnado.

## AVALIACIÓN

---

Para avaliar o desenvolvemento desta unidade didáctica propóñense tarefas relacionadas coas actividades descritas na sección anterior que pretenden acadar os obxectivos indicados na sección na que estes foron descritos.

Unha das probas de avaliación se realizará nos últimos 15 minutos da última hora de aulas expositivas, como tamén se indicou na sección anterior. Consistirá na formulación por escrito das hipóteses nulas e as correspondentes alternativas, ante diferentes tipoloxías de preguntas sobre as relacións entre as variables explicativas e a explicada dun MRLNC. Con ela trataríase de comprobar ata que punto o alumnado é capaz de trasladar cada pregunta a unha representación formal adecuada en termos de parámetros.

Por outro lado corrixirase cada ficheiro-formulario enviado ao final de cada aula interactiva, para avaliar o grao de comprensión sobre como debe:

- especificarse en GRETL o cálculo dos estatísticos de cada proba;
- localizarse o resultado dos cálculos e a probabilidade asociada; e
- tomarse a decisión axeitada en cada caso.

Finalmente, no exame de avaliación final debe presentarse a estimación dun modelo a partir dun banco de datos concreto, e a partir dela formular algunhas cuestións sobre os parámetros. Na resposta, o alumnado debera establecer primeiro as hipóteses para logo buscar entre os resultados da estimación, aquela información que permita tomar unha decisión cunha marxe de erro.

## BIBLIOGRAFÍA

---

- COBB, Charles W. e Paul H. DOUGLAS (1928): *A Theory of Production*, American Economic Review, vol. 18(1), 139-172.
- DÍAZ, Montserrat e María M. LLORENTE (2013): *Econometría*, Madrid: Pirámide.
- GALTON, Francis (1886): «Family Likeness in Stature», *Proceedings of Royal Society*, London, vol. 40, 42-72.
- MOORE, Henry L. (1914): *Economic Cycles: Their Law and Cause*, New York: Macmillan.
- OTERO, José M. (1978): *Lógica y limitaciones de la Econometría*, Madrid: ICE.
- PULIDO, Antonio (2001): *Modelos econométricos*, Madrid: Pirámide.
- SCHULTZ, Henry (1928): *Statistical Laws of Demand of Supply with Special Application to Sugar*, Chicago: University of Chicago Press.



Unha colección orientada a editar materiais docentes de calidade e pensada para apoiar o traballo do profesorado e do alumnado de todas as materias e titulacións da universidade

unidadesdidácticas  
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA