UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Definición sistemática del significado de los parámetros en modelos de regresión lineal con variables cualitativas

Apellidos, nombre	Chirivella González, Vicente (vchirive@eio.upv.es)
Departamento	Estadística e Investigación Operativa Aplicadas y Calidad
Centro	Facultad de Administración y Dirección de Empresas

1 Resumen de las ideas clave

En este artículo docente te presento los pasos que debemos dar para definir, de forma sistemática, los parámetros de un modelo de regresión que tenga variables explicativas de naturaleza cualitativa. Introducimos variables cualitativas en el modelo (en forma de variables ficticias o *dummy*) porque deseamos cuantificar diferencias entre grupos de datos, cuantificación de diferencias que llevamos a cabo mediante los parámetros del modelo. Es obvio que conocer el significado de los parámetros será una cuestión fundamental, y me temo que también será complicada, si no procedemos de una forma sistemática a definir y a entender su significado.

2 Introducción

En el modelo de Keynes, que explica el consumo de un producto mediante la renta de los consumidores y el precio del propio producto, parece claro que dicha explicación pueda ser diferente dependiendo de si estamos en un momento de crisis económica o de si no lo estamos. Y estar, o no estar en crisis, no es una variable cuantitativa que pueda formar parte (sin más) del modelo propuesto. Más bien al contrario, es una variable cualitativa, una variable que me identifica dos (posibles) grupos diferentes, *crisis – no crisis*.

Además, tal vez no nos interese tanto el valor de la renta de los consumidores como el agruparlos en ciertas categorías representativas: mísero, pobre, medio, rico y podrido de dinero. Éstos podrían ser, por ejemplo, los posibles niveles de la renta de los consumidores que nos interesen estudiar en nuestro análisis del consumo del producto. Habrá que asignar los valores de renta que correspondan a cada intervalo, claro, pero de esta forma convertiríamos una variable cuantitativa en una cualitativa.

En esencia, con una variable cualitativa estaríamos identificando diferentes grupos en el modelo de regresión, con lo que podríamos establecer y cuantificar las diferencias existentes entre dichos grupos.

En estos casos recurriremos a las llamadas variables ficticias, unas variables que indican y codifican los niveles de la variable cualitativa para así poder utilizarlas en los modelos de regresión. Estas variables nos permitirán describir y cuantificar diferencias en efectos temporales, cuando los grupos se refieren a distintos períodos de tiempo; espaciales, cuando se refieren a regiones o países; industriales, cuando hacen referencia al sector industrial, etc.

3 Objetivos

Los objetivos marcados para este artículo docente son:

1. Entender el concepto de variable ficticia y su uso en los modelos de regresión.
2. Definir de forma sistemática el significado de los parámetros de un modelo que contenga variables ficticias, para saber qué diferencias estamos midiendo.
3. Interpretar las estimaciones de los parámetros desde el punto de vista económico mediante las definiciones realizadas.

4 Variables cualitativas

Comenzaremos presentando un ejemplo que nos ayudará a entender mejor los pasos para la definición sistemática de los parámetros.

Se desea establecer la relación que existe entre el gasto realizado por los consumidores mediante terminales punto de venta (TPV) y el número de tarjetas a débito existentes. Para ello se dispone del IMPORTE del gasto realizado, en miles de millones de euros, y el número de tarjetas de débito, NTARJETAS, en miles de unidades. Al representar gráficamente ambas variables se encuentran ***dos puntos de ruptura en su relación****, uno en el* ***año 2002****, que coincide con la expansión de la burbuja económica de España, y otro en el* ***año 2008*** *cuando comienza a tener efecto la crisis actual en el consumo. También se tiene que el número total de tarjetas de débito existentes aumentan hasta el año 2002, y a partir de esa fecha se produce una disminución, moderada hasta 2007, y muy importante desde ese año. Con todo ello se propone el siguiente modelo, donde al número de tarjetas total se le ha restado la cantidad de 32700 miles de tarjetas (el número de tarjetas existentes en el año 2002), y se introducen dos variables ficticias,* ***A1996*** *que corresponde al* ***periodo 1996-2001*** *y* ***A2008*** *corresponde al* ***periodo 2008-2012****.*

$$IMPORTE = \beta_0 + \beta_1 A1996 + \beta_2 A2008 + \beta_3 (NTARJETAS - 32700)$$
$$+\beta_4 A1996(NTARJETAS - 32700) + \beta_5 A2008(NTARJETAS - 32700) + U$$

Ecuación 1: Modelo para explicar el importe del gasto

¿Por qué he propuesto este modelo? Pues porque es el modelo más general posible, al introducir todas las variables ficticias existentes (A1996 y A2008), todas las variables explicativas cuantitativas existentes (NTARJETAS), y el producto de todas las variables ficticias por todas las variables explicativas cuantitativas existentes (A1996*NTARJETAS y A2008*NTARJETAS). Añadimos un parámetro beta a cada una de ellas, el término constante, y ya tenemos el modelo. ¿Por qué he restado 32700 miles de tarjetas al número de tarjetas? Pues porque en caso contrario ciertos parámetros no tendrían sentido (lo vemos luego). Vale,... y ¿por qué exactamente 32700 miles de tarjetas? Pues porque es el número de tarjetas del año 2002, el momento de un cambio, y quiero cuantificar lo que ocurre en ese cambio.

4.1 Variables ficticias

La forma en que introducimos las variables cualitativas en los modelos de regresión son las variables ficticias. Estas variables sólo toman los valores 0 ó 1, y las utilizamos como variables indicadoras de los niveles de la variable cualitativa.

Así, el intervalo de tiempo considerado en nuestro ejemplo, desde el año 1996 al año 2012, se divide en tres etapas: de 1996 a 2001, de 2002 a 2007, y de 2008 a 2012. Tenemos, por tanto, tres posibles niveles, **periodo previo**, **periodo de burbuja** y **periodo de crisis**. Podemos crear, por lo tanto, tres variables ficticias, una para indicar cada nivel o periodo de tiempo:

A1996	1	si el dato es del periodo 1996 a 2001
	0	si no lo es
A2002	1	si el dato es del periodo 2002 a 2007
	0	si no lo es
A2008	1	si el dato es del periodo 2008 a 2012
	0	si no lo es

Tabla 1: Valores de las variables ficticias

Sin embargo, no introducimos todas las posibles variables ficticias en el modelo de regresión. De forma general, **si una variable cualitativa tiene *k* niveles (valores), debemos crear *k-1* variables ficticias para introducidas en el modelo de regresión**. Uno de los niveles queda fuera del modelo, y lo utilizamos como nivel de referencia.

Los parámetros que acompañen a las variables ficticias medirán algo (depende del modelo) que se referirá a ese grupo (nivel) de referencia, o bien medirá diferencias de un grupo con respecto al grupo de referencia.

¿Cuál es el nivel ausente o de referencia en el problema de los importes de las tarjetas de débito? ¿Por qué es buena referencia?

4.2 Definición sistemática de parámetros

Si queremos definir el significado de los parámetros del modelo, lo único que debemos hacer es asignar valores adecuados a las variables explicativas, de forma que nos quede únicamente en el modelo aquél parámetro que nos interesa definir. O tal vez puede que nos quede algún parámetro más, siempre y cuando ya lo hayamos definido. Cuando esto no sea posible, entonces transformaremos adecuadamente el modelo (por ejemplo tomando incrementos) y asignamos valores a las variables explicativas.

Esta es una tarea laboriosa, que necesita de un tratamiento sistemático si queremos definirlos correctamente y sin demasiada complicación. Por ese motivo propongo seguir los siguientes pasos:

1.- **Identificamos las poblaciones o grupos en estudio**, anotando los valores que le corresponden a cada variable ficticia en el mismo.

2.- **Separamos el modelo original** en un modelo para cada grupo, sustituyendo en el mismo los valores correspondientes, 0 ó 1, de cada variable ficticia.

3.- **Determinamos el significado de los parámetros del modelo más simple** de ellos, los del modelo con menos parámetros a definir, y que corresponde al grupo escogido como referencia.

4.- **Determinamos el significado del resto de los parámetros en los restantes modelos**. Para ello debemos observar que, en dichos modelos y entre paréntesis, tenemos juntos a parámetros que no hemos definido con parámetros que si lo hemos hecho. Para definir el parámetro desconocido escribimos la palabra "*diferencia*", luego copiamos la definición del parámetro al que acompaña en el paréntesis (ya definido), y cambiamos en dicha definición al "*grupo de referencia*" por la expresión "*del grupo actual respecto del grupo de referencia*".

5.- **Repasamos las definiciones**, por si un mismo parámetro aparece en solitario en más de un modelo, lo cual significaría que es el mismo parámetro para más de un nivel y cambiaría su significado original.

Pero será mejor que apliquemos todo esto al ejemplo que nos ocupa para que quede más claro…

4.3 Aplicando la definición sistemática al ejemplo

Vamos a aplicar los pasos de la definición sistemática a los parámetros del modelo que explica el gasto realizado mediante tarjeta de débito.

4.3.1 Definición de los parámetros

1.- Identificar las poblaciones o grupos en estudio.

Tenemos que identificar los grupos. Ya hemos visto en la Tabla 1 que son tres, de **1996 a 2001**, de **2002 a 2007**, y de **2008 a 2012**. Reservamos un espacio suficiente en nuestra hoja de respuesta (ver el siguiente punto) y escribimos los nombres de los tres grupos, poniendo al lado el valor que le corresponde a cada variable ficticia. El primer grupo es **1996 a 2001** al que le corresponde los valores A1996=1 por ser de dicho intervalo de tiempo, y A2008=0 porque no corresponde. El segundo es **2002 a 2007**, con A1996=0 y A2008=0 porque en ambos casos no corresponden al intervalo de tiempo, y finalmente el tercer grupo es **2008 a 2012**, con A1996=0 porque no corresponde al intervalo y A2008=1 porque si lo hace.

2.- Separar el modelo original en un modelo para cada grupo.

Debemos separar nuestro modelo original (Ecuación 1) en tres modelos, uno para cada grupo, asignado los valores 0 ó 1 que hemos visto tienen las variables ficticias, según corresponda, sustituyendo y simplificando:

1996 a 2001 (A1996=1, A2008=0) **Periodo previo**

$$IMPORTE = (\beta_0 + \beta_1) + (\beta_3 + \beta_4)(NTARJETAS - 32700) + U$$

2002 a 2007 (A1996=0, A2008=0) **Periodo de burbuja económica**

$$IMPORTE = \beta_0 + \beta_3(NTARJETAS - 32700) + U$$

2008 a 2012 (A1996=0, A2008=1) **Periodo de crisis**

$$IMPORTE = (\beta_0 + \beta_2) + (\beta_3 + \beta_5)(NTARJETAS - 32700) + U$$

Observa que hemos añadido algunos paréntesis, que no eran necesarios desde el punto de vista matemático, pero que agrupan convenientemente a ciertos parámetros a efectos de definirlos posteriormente.

3.- Determinar el significado de los parámetros del modelo más simple.

El modelo más simple de los tres que hemos obtenido es el del periodo de burbuja, y lo es porque tiene menos parámetros que definir. Determinaremos entonces el significado de sus parámetros, β_0 y β_3. Aquí no hay variables ficticias, así que podemos definir los parámetros sin problemas. Coloreamos la definición de estos parámetros para resaltar así ciertas partes y compararlas con las definiciones de otros parámetros más adelante:

$$IMPORTE = \beta_0 + \beta_3(NTARJETAS - 32700) + U$$

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

β_0 – es el promedio del importe del gasto cuando el número de tarjetas es de 32,7 millones, en el periodo de burbuja económica

β_3 – es el incremento del promedio del importe del gasto por cada mil tarjetas disponibles, en el periodo de burbuja económica

¿A qué parámetro afecta la cantidad de 32700 miles de tarjetas?

Si no hubiésemos restado esta cantidad, el parámetro β_0 sería el promedio del importe del gasto cuando no hay tarjetas de débito, en el periodo de burbuja económica. Si no hay tarjetas, ¿qué gasto puede haber? Así que es mejor asegurar un parámetro con sentido restando esta cierta cantidad.

4.- Determinar el significado del resto de los parámetros en los restantes modelos.

Vamos con los restantes modelos. Allí encontramos entre paréntesis a parejas de parámetros, uno de ellos todavía no definido y el otro que si lo está. Para definir el parámetro todavía no definido escribimos la palabra "diferencia", luego copiamos la definición del parámetro al que acompaña en el paréntesis (ya definido), cambiando en dicha definición al "grupo de referencia" por la expresión "del grupo actual respecto del grupo de referencia".

Comencemos por el primer modelo, el del periodo previo:

$$IMPORTE = (\beta_0 + \beta_1) + (\beta_3 + \beta_4)(NTARJETAS - 32700) + U$$

El parámetro β_1 no está definido, pero acompaña a otro parámetro que si lo está, β_0. Escribimos la palabra "diferencia", luego copiamos la definición de β_0, "promedio del importe del gasto cuando el número de tarjetas es de 32,7 millones, en el periodo de burbuja económica", pero donde pone "en el periodo de burbuja económica", que es la referencia, ponemos "en el periodo previo respecto del periodo de burbuja", puesto que el periodo de este modelo es el previo, y la referencia es la burbuja.

β_1 – es la diferencia del promedio del importe del gasto cuando el número de tarjetas es de 32,7 millones, en el periodo previo respecto del periodo de burbuja

y como ya conocemos la mecánica, definimos el parámetro β_4, que acompaña a β_3 en el paréntesis:

β_4 – es la diferencia del incremento del promedio del importe del gasto por cada mil tarjetas disponibles, en el periodo previo respecto del periodo de burbuja

Es de suponer que en este momento es sencillo definir los parámetros restantes, en el tercer modelo, el de periodo de crisis:

$$IMPORTE = (\beta_0 + \beta_2) + (\beta_3 + \beta_5)(NTARJETAS - 32700) + U$$

β_2 – es la diferencia del promedio del importe del gasto cuando el número de tarjetas es de 32,7 millones, en el periodo de crisis respecto del periodo de burbuja

β_5 – es la diferencia del incremento del promedio del importe del gasto por cada mil tarjetas disponibles, en el periodo de crisis respecto del periodo de burbuja

Por cierto, ¿has comparado, palabra por palabra, las definiciones de β_1 y de β_2? ¿Y las de β_4 y β_5? ¿Qué nueva regla podemos extraer de la comparación?

5.- Repasar las definiciones

No es el caso, puesto que no vuelve a aparecer en solitario un parámetro que ya ha sido definido, así que no daríamos éste paso en el ejemplo. Pero bueno, supongamos que β_2=0. En el tercer modelo aparecería en solitario (de nuevo) el parámetro β_0.

$$IMPORTE = \beta_0 + (\beta_3 + \beta_5)(NTARJETAS - 32700) + U$$

¿Qué implicaría esto en las definiciones dadas anteriormente? Pues hay algunos cambios:

- Este parámetro β_2 cuantificaba diferencias entre el periodo de crisis y el de burbuja, y como la diferencia es cero, esto quiere decir que no hay diferencia. Por lo tanto, del promedio del importe del gasto cuando el número de tarjetas es de 32,7 millones es el mismo en el periodo de crisis que en el periodo de burbuja.

- El parámetro β_0 cambia su significado, porque ya no se refiere únicamente a la burbuja, sino que ahora se refiere también a la crisis

  β_0 – es el promedio del importe del gasto cuando el número de tarjetas es de 32,7 millones, tanto en el periodo de burbuja económica como en el de posterior crisis

4.3.2 (Re)Interpretación de los parámetros

Volviendo al ejemplo, y con los datos correspondientes realizamos el ajuste del modelo analizado mediante Statgraphics, y obtenemos el resultado de la Tabla 2.

Multiple Regression - IMPORTE

Parameter	*Estimate*	*Standard Error*	*T Statistic*	*P-Value*
CONSTANT	28,8384	0,865588	33,3166	0,0000
A1996	-15,0519	1,59475	-9,4384	0,0000
A2008	11,1528	1,93165	5,77373	0,0001
NTARJETAS-32700	-0,00734078	0,00110456	-6,64588	0,0000
A1996*(NTARJETAS-32700)	0,0080961	0,00112196	7,21606	0,0000
A2008*(NTARJETAS-32700)	0,00646634	0,00118522	5,45581	0,0002

Tabla 2: Estimación de parámetros del modelo

Donde todas las estimaciones de parámetros son significativas (significativamente distintas de cero) por tener `P-Value` inferiores a un 5%, y los valores calculados por el programa (columna `Estimate`) serían aceptables.

Ahora que sabes cuánto valen, ¿podrías volver a definir los parámetros del modelo pero ahora con los valores numéricos? Por ejemplo β_0. El promedio del importe del gasto cuando el número de tarjetas es de 32,7 millones y en el periodo de burbuja económica es de 28,8384 millones de euros. ¿Te atreves a interpretar el resto de valores? ¡Venga!

Examina el Gráfico 1 y trata de situar en el mismo los parámetros del modelo. Fíjate que el periodo de burbuja es una buena referencia, pues permite comparar con los estados anterior y posterior. No hay parámetros que comparen el periodo de crisis y el periodo previo, pero al estar el de burbuja en medio, también podemos extraer conclusiones en la comparación de periodos previo - crisis, aunque no existan parámetros y pruebas de hipótesis para ello.

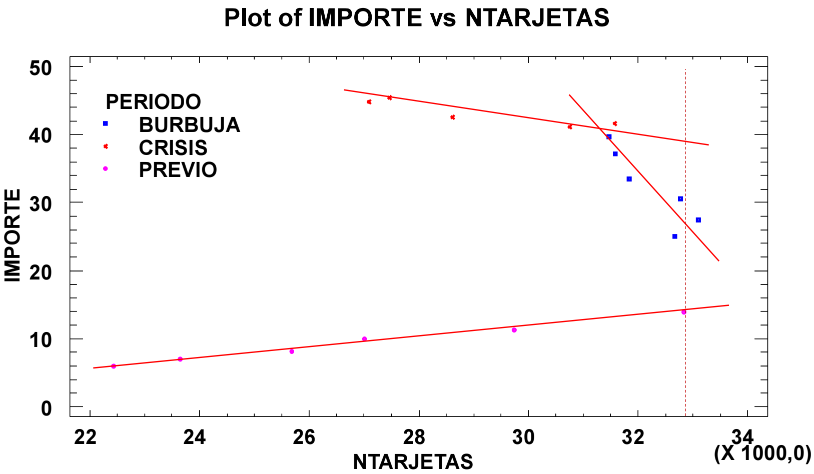

Gráfico 1: Relación entre el importe y el número de tarjetas, con las rectas ajustadas

4.4 Ejercicios de aplicación

Aplica la definición sistemática de parámetros a los siguientes modelos. Las soluciones las encontrarás en el siguiente apartado. Sigue los pasos, de forma ordenada, y llegarás al final con tus definiciones.

4.4.1 Ejercicios propuestos

Ejercicio 1 *Para explicar la PRODUCCIÓN DE ALGODÓN en España se propone utilizar la superficie cultivada y el método de hacerlo, en secano o en regadío. Se dispone de la PRODUCCIÓN de algodón, en Qm, la SUPERFICIE cultivada, en Ha, a la cual se resta su propio valor medio en el modelo, y se define la variable ficticia SECANO que indica si el dato corresponde a una producción en secano. Determina el significado de los parámetros del modelo:*

$$PRODUCCION = \beta_0 + \beta_1 SECANO + \beta_2 (SUPERFICIE - \overline{SUPERFICIE}) + \beta_3 SECANO(SUPERFICIE - \overline{SUPERFICIE}) + U$$

Ejercicio 2 *Se desea explicar la PRODUCCIÓN DE CAVA en España según su destino mediante la producción total del mismo. Para ello se dispone de la producción de cava con destino al consumo interior, con destino a consumo en el resto de la UE, y con destino a terceros países, además de la producción total (PROD.CAVA), medidas todas ellas en millones de botellas. Como se desea explicar una hipotética distribución de los destinos cuando se fabrican 200 millones de botellas, se restan 200 a la producción total en el modelo. Se han considerado las variables ficticias INTERNO y TERCEROS que identifican al consumo interno y al de terceros países diferentes del resto de la Unión Europea, respectivamente. Determina el significado de los parámetros y el reparto de los 200 millones de botellas según el destino.*

$$PROD.CAVA_{DESTINO} = \beta_0 + \beta_1 INTERNO + \beta_2 (PROD.CAVA - 200) + \beta_3 INTERNO(PROD.CAVA - 200) + \beta_4 TERCEROS(PROD.CAVA - 200) + U$$

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Ejercicio 3 *En el análisis de la tasa de actividad en función del nivel de estudios, cabría esperar que a medida que aumente el nivel de estudios aumente la tasa de actividad, es decir, aumente la probabilidad de tener trabajo. La variable nivel de estudios está compuesta por los niveles: analfabetos, primarios, medios, técnicos superiores, 1ciclo, 2ciclo, 3ciclo, y cada nivel de estudios es superior al nivel anterior, de forma que hay un orden natural (salvo técnicos superiores). Se definen las variables ficticias ANALF, EPRIM, EMED, ETP, E1C, E2C, E3C, que se corresponden con los respectivos niveles de estudio mencionados. La variable ficticia Nivel(i) vale 1 si el nivel de estudios considerado es de Nivel(i) o superior, y 0 en otro caso. Por ejemplo, si se tiene un dato de una persona con estudios de primer ciclo, las variables ficticias toman los valores, ANALF=1, EPRIM=1, EMED=1, ETP=1, E1C=1, porque el nivel actual, estudios de primer ciclo, es igual o superior a analfabetos, primarios, medios, técnicos y primer ciclo, pero E2C=0 y E3C=0 porque el nivel de estudios actual, primer ciclo, no es igual o superior a segundo y tercer ciclo. Determina el significado de los parámetros del modelo:*

$$TASACT = \beta_0 + \beta_1 EPRIM + \beta_2 EMED + \beta_3 ETP + \beta_4 E1C + \beta_5 E2C + \beta_6 E3C + U$$

4.4.2 Soluciones

<table>
<tr><td colspan="2">Ejercicio1</td></tr>
<tr><td>β_0</td><td>Promedio de la producción de algodón en regadío cuando la superficie cultivada es la superficie media.</td></tr>
<tr><td>β_1</td><td>Diferencia del promedio de la producción de algodón, en secano respecto de regadío, cuando la superficie cultivada es la superficie media.</td></tr>
<tr><td>β_2</td><td>Incremento del promedio de la producción de algodón en regadío por cada hectárea cultivada.</td></tr>
<tr><td>β_3</td><td>Diferencia del incremento del promedio de la producción de algodón, en secano respecto de regadío, por cada hectárea cultivada.</td></tr>
<tr><td colspan="2">Ejercicio2</td></tr>
<tr><td>β_0</td><td>Promedio de la producción de cava con destino a la UE y a terceros países cuando la producción total es de 200 millones de botellas.</td></tr>
<tr><td>β_1</td><td>Diferencia del promedio de la producción de cava con destino a consumo interno respecto del de la UE y de terceros países cuando la producción total es de 200 millones de botellas.</td></tr>
<tr><td>β_2</td><td>Incremento del promedio de la producción de cava con destino a la UE por cada millón de botellas que se produce.
También es el porcentaje de botellas dedicado a exportar a la UE de ese millón de botellas que se ha producido.</td></tr>
<tr><td>β_3</td><td>Diferencia del incremento del promedio de la producción de cava, con destino al consumo interno respecto del de la UE, por cada millón de botellas que se produce.</td></tr>
<tr><td>β_4</td><td>Diferencia del incremento del promedio de la producción de cava, con destino a terceros respecto del de la UE, por cada millón de botellas que se produce.</td></tr>
<tr><td colspan="2">De los 200 millones de botellas, la cantidad $\beta_0+\beta_1$ se dedica al consumo interno, la cantidad β_0 se dedica al consumo en la UE, y la cantidad β_0 se dedica al consumo en terceros países, con lo que su suma deberían ser los 200 millones, $3\beta_0+\beta_1=200$</td></tr>
</table>

Ejercicio3	
β_0	Promedio de la tasa de actividad de los analfabetos
β_1	Diferencia del promedio de la tasa de actividad de los que tienen estudios primarios respecto de los analfabetos
β_2	Diferencia del promedio de la tasa de actividad de los que tienen estudios medios respecto de los que tienen estudios primarios
β_3	Diferencia del promedio de la tasa de actividad de los que tienen estudios técnicos profesionales respecto de los que tienen estudios medios
β_4	Diferencia del promedio de la tasa de actividad de los que tienen estudios de primer ciclo respecto de los que tienen estudios técnicos profesionales
β_5	Diferencia del promedio de la tasa de actividad de los que tienen estudios de segundo ciclo respecto de los que tienen estudios de primer ciclo
β_6	Diferencia del promedio de la tasa de actividad de los que tienen estudios de tercer ciclo respecto de los que tienen estudios de segundo ciclo

5 Cierre

Las variables cualitativas enumeran los posibles valores (grupos) de una característica no cuantificable observada en un individuo. Si deseamos tener en cuenta una variable cualitativa en un modelo de regresión es porque queremos encontrar diferencias entre los grupos que define dicha variable.

La manera correcta de introducir las variables cualitativas en el modelo de regresión es mediante las variables ficticias. Una variable cualitativa tiene un cierto número de valores o niveles, y las variables ficticias se definen para indicar los niveles de dicha variable. Las variables ficticias sólo toman dos valores, 0 ó 1, hay tantas como niveles tenga la variable cualitativa, y el valor 1 sirve para indicar uno de esos niveles.

En el modelo de regresión no pueden aparecer todas las posibles variables ficticias (todos los niveles), una de ellas debe quedar fuera para ser utilizado como nivel de referencia. Los parámetros del modelo cuantificarán un concepto para ese nivel de referencia, o medirán diferencias respecto de ese nivel de referencia.

Definir el significado de los parámetros en un modelo con variables ficticias es complicado, pero es posible hacerlo si se siguen los pasos propuestos en este documento.

6 Bibliografía

D. Gujarati : "Basic Econometrics", Ed. McGrawHill - 4ª edition, páginas 297-334.

D. Peña: "Estadística: Modelos y Métodos. (Vol.2) Modelos lineales y Series temporales", Ed. Alianza Universidad-Textos, páginas 307-548.