1. La ecuación (5.2.1) no dice que la probabilidad de que β_2 se encuentre entre los límites dados sea $1 - \alpha$. Puesto que se supone que β_2, aún siendo desconocido, es un número fijo, se dice que está o no está dentro del intervalo. La ecuación (5.2.1) establece que, al utilizar el método descrito en este capítulo, la probabilidad de construir un intervalo que contenga β_2 es $1 - \alpha$.

2. El intervalo (5.2.1) es un **intervalo aleatorio**; es decir, variará de una muestra a la siguiente debido a que está basado en $\hat{\beta}_2$, el cual es aleatorio. (¿Por qué?)

3. Puesto que el intervalo de confianza es aleatorio, los enunciados probabilísticos que le corresponden deben ser entendidos en un sentido de largo plazo, es decir, para muestreo repetido. Más específicamente, (5.2.1) significa: Si se construyen intervalos de confianza como el anterior con base probabilística de $1 - \alpha$, entonces, a largo plazo, en promedio, tales intervalos contendrán, en $1 - \alpha$ de los casos, el valor verdadero del parámetro.

4. Como se mencionó en 2, el intervalo (5.2.1) es aleatorio siempre y cuando $\hat{\beta}_2$ sea desconocido. Pero una vez que se tenga una muestra específica y se obtenga un valor numérico específico de $\hat{\beta}_2$, el intervalo (5.2.1) deja de ser aleatorio, quedando entonces fijo. En este caso, **no se puede** hacer la afirmación probabilística (5.2.1); así, no se puede decir que la probabilidad de que un intervalo *fijo* dado incluya el verdadero β_2 sea $(1 - \alpha)$. En esta situación, β_2 está en el intervalo fijo, o por fuera de éste. Por consiguiente, la probabilidad será 1 o 0. Por tanto, en nuestro ejemplo hipotético consumo-ingreso, si el intervalo de confianza al 95% fuera obtenido $(0.4268 \leq \beta_2 \leq 0.5914)$, se mostrará en (5.3.9) que **no se puede** decir que la probabilidad de que este intervalo incluya el verdadero β_2 sea de 95%. Esa probabilidad es 1 o 0.

¿Cómo se construyen los intervalos de confianza? De la exposición anterior se puede esperar que si se conocen las **distribuciones muestrales o de probabilidad** de los estimadores, se puedan hacer afirmaciones sobre intervalos de confianza tales como (5.2.1). En el capítulo 4 se vio que bajo el supuesto de normalidad de las perturbaciones u_i, los estimadores MCO $\hat{\beta}_1$ y $\hat{\beta}_2$ están también normalmente distribuidos y que el estimador MCO, $\hat{\sigma}^2$, está relacionado con la distribución χ^2 (ji-cuadrada). Entonces, parecería que la labor de construir intervalos de confianza es muy sencilla. ¡Y, de hecho, lo es!

5.3 INTERVALOS DE CONFIANZA PARA LOS COEFICIENTES DE REGRESIÓN β_1 Y β_2

Intervalo de confianza para β_2

En el capítulo 4, sección 4.3, se mostró que bajo el supuesto de normalidad de u_i, los estimadores MCO $\hat{\beta}_1$ y $\hat{\beta}_2$ son en sí mismos normalmente distribuidos con medias y varianzas allí establecidas. Por consiguiente, por ejemplo, la variable

$$\begin{aligned} Z &= \frac{\hat{\beta}_2 - \beta_2}{\text{ee}(\hat{\beta}_2)} \\ &= \frac{(\hat{\beta}_2 - \beta_2)\sqrt{\sum x_i^2}}{\sigma} \end{aligned} \tag{5.3.1}$$

como se anotó en (4.3.6), es una variable normal estandarizada. Por consiguiente, parece que se puede utilizar la distribución normal para hacer afirmaciones probabilísticas sobre β_2 siempre que se conozca la verdadera varianza poblacional σ^2. Si σ^2 se conoce, una propiedad importante de una variable normalmente distribuida con media μ y varianza σ^2 es que el área bajo la curva normal entre $\mu \pm \sigma$ esté cercana al 68%, que entre $\mu \pm 2\sigma$ esté alrededor del 95% y que entre los límites $\mu \pm 3\sigma$ el área se acerque al 99.7%.

Pero σ^2 raramente es conocida y, en la práctica, está determinada por el estimador insesgado $\hat{\sigma}^2$. Si se reemplaza σ por $\hat{\sigma}$, (5.3.1) puede escribirse así

$$
\begin{aligned}
t &= \frac{\hat{\beta}_2 - \beta_2}{\text{ee}(\hat{\beta}_2)} = \frac{\text{estimador} - \text{parámetro}}{\text{error estándar estimado del estimador}} \\
&= \frac{(\hat{\beta}_2 - \beta_2)\sqrt{\sum x_i^2}}{\hat{\sigma}}
\end{aligned}
\tag{5.3.2}
$$

donde ee($\hat{\beta}_2$) se refiere ahora al error estándar estimado. Puede demostrarse (véase el apéndice 5A, sección 5A.2) que la variable t, así definida, sigue la distribución t con $n - 2$ g de l [*Nótese* la diferencia entre (5.3.1) y (5.3.2).] Por consiguiente, en lugar de utilizar la distribución normal, se puede utilizar la distribución t para construir un intervalo de confianza para β_2 de la siguiente forma:

$$
\Pr(-t_{\alpha/2} \le t \le t_{\alpha/2}) = 1 - \alpha \tag{5.3.3}
$$

donde el valor t en el centro de esta doble desigualdad es el valor t dado por (5.3.2) y donde $t_{a/2}$ es el valor de la variable t obtenida de la distribución t para un nivel de significancia de $\alpha/2$ y $n - 2$ g de l; frecuentemente es llamado el valor **crítico** t a un nivel de significancia $\alpha/2$. Sustituyendo (5.3.2) en (5.3.3), se obtiene

$$
\Pr\left[-t_{\alpha/2} \le \frac{\hat{\beta}_2 - \beta_2}{\text{ee}(\hat{\beta}_2)} \le t_{\alpha/2}\right] = 1 - \alpha \tag{5.3.4}
$$

Reorganizando (5.3.4), se obtiene

$$
\Pr[\hat{\beta}_2 - t_{\alpha/2}\text{ee}(\hat{\beta}_2) \le \beta_2 \le \hat{\beta}_2 + t_{\alpha/2}\text{ee}(\hat{\beta}_2)] = 1 - \alpha \tag{5.3.5}
$$
[3]

La ecuación (5.3.5) proporciona un **intervalo de confianza** para β_2 al 100 $(1 - \alpha)$%, el cual puede ser escrito en forma más compacta como

Intervalo de confianza para β_2 al $100(1 - \alpha)$%:

$$
\hat{\beta}_2 \pm t_{\alpha/2}\text{ee}(\hat{\beta}_2) \tag{5.3.6}
$$

---

[3] Algunos autores prefieren escribir (5.3.5) con los g de l indicados explícitamente. Por tanto, ellos escribirían

$$
\Pr[\hat{\beta}_2 - t_{(n-2),\alpha/2}\text{ee}(\hat{\beta}_2) \le \beta_2 \le \hat{\beta}_2 + t_{(n-2)\alpha/2}\text{ee}(\hat{\beta}_2)] = 1 - \alpha
$$

Pero, por simplicidad, utilizaremos nuestra notación; el contexto clarifica los g de l apropiados.

Mediante argumentación análoga y utilizando (4.3.1) y (4.3.2), se puede escribir:

$$\Pr\left[\hat{\beta}_1 - t_{\alpha/2}\,\text{ee}\,(\hat{\beta}_1) \le \beta_1 \le \hat{\beta}_1 + t_{\alpha/2}\,\text{ee}\,(\hat{\beta}_1)\right] = 1 - \alpha \tag{5.3.7}$$

o, en forma más compacta,

Intervalo de confianza para β_1 al $100(1 - \alpha)\%$:

$$\hat{\beta}_1 \pm t_{\alpha/2}\,\text{ee}\,(\hat{\beta}_1) \tag{5.3.8}$$

Obsérvese un rasgo importante de los intervalos de confianza dados en (5.3.6) y (5.3.8): En ambos casos *la amplitud del intervalo de confianza es proporcional al error estándar del estimador*. Es decir, entre más grande sea el error estándar, más amplio será el intervalo de confianza. Expresado de otra forma, entre más grande sea el error estándar del estimador, mayor será la incertidumbre de estimar el verdadero valor del parámetro desconocido. Así, el error estándar de un estimador es descrito frecuentemente como una medida de la **precisión** del estimador, es decir, qué tan preciso mide el estimador al verdadero valor poblacional.

Volviendo al ejemplo ilustrativo consumo-ingreso, en el capítulo 3 (sección 3.6), se encuentra que $\hat{\beta}_2 = 0.5091$, ee$(\hat{\beta}_2) = 0.0357$, y g de l = 8. Si se supone que $\alpha = 5\%$, es decir, un coeficiente de confianza de 95%, entonces la tabla t muestra que para 8 g de l el **valor crítico** $t_{\alpha/2} = t_{0.025} = 2.306$. Sustituyendo estos valores en (5.3.5), el lector debe verificar que el intervalo de confianza para β_2 al 95% es el siguiente:

$$0.4268 \le \beta_2 \le 0.5914 \tag{5.3.9}$$

O, utilizando (5.3.6), es

$$0.5091 \pm 2.306(0.0357)$$

es decir,

$$0.5091 \pm 0.0823 \tag{5.3.10}$$

La interpretación de este intervalo de confianza es: Dado el coeficiente de confianza de 95%, a largo plazo, en 95 de cada 100 casos, intervalos como (0.4268 a 0.5914) contendrán el verdadero β_2. Pero, como se advirtió antes, obsérvese que no se puede decir que la probabilidad de que el intervalo específico (0.4268 a 0.5914) contenga el verdadero β_2 sea de 95% porque este intervalo es ahora fijo y no aleatorio; por consiguiente, β_2 se encontrará o no dentro de él: la probabilidad de que el intervalo específicamente fijado incluya el verdadero β_2 es por consiguiente 1 o 0.

Intervalo de confianza para β_1

Siguiendo (5.3.7), el lector puede verificar fácilmente que para el ejemplo consumo-ingreso, el intervalo de confianza para β_1 al 95% es:

$$9.6643 \leq \beta_1 \leq 39.2448 \tag{5.3.11}$$

O, utilizando (5.3.8), se encuentra que es

$$24.4545 \pm 2.306(6.4138)$$

es decir,

$$24.4545 \pm 14.7902 \tag{5.3.12}$$

Nuevamente, se debe ser cauteloso al interpretar este intervalo de confianza. A largo plazo, en 95 de cada 100 casos, intervalos como (5.3.11) contendrán el verdadero β_1; la probabilidad de que este intervalo fijo incluya el verdadero β_1 es 1 o 0.

Intervalo de confianza para β_1 y β_2 simultáneamente

Hay ocasiones en las cuales se necesita construir un *intervalo de confianza conjunto* para β_1 y β_2 tal que, para un coeficiente de confianza $(1 - \alpha)$, digamos del 95%, ambos β_1 y β_2 caigan simultáneamente dentro de ese intervalo. Puesto que este tema es complejo, el lector quizá desee consultar referencias.[4] (Estudiaremos brevemente este tema en los capítulos 8 y 10).

5.4 INTERVALO DE CONFIANZA PARA σ^2

Como se señaló en el capítulo 4, sección 4.3, bajo el supuesto de normalidad, la variable

$$\chi^2 = (n-2)\frac{\hat{\sigma}^2}{\sigma^2} \tag{5.4.1}$$

sigue la distribución χ^2 con $n - 2$ g de l.[5] Por consiguiente, podemos utilizar la distribución χ^2 para establecer el intervalo de confianza para σ^2

$$\Pr\left(\chi^2_{1-\alpha/2} \leq \chi^2 \leq \chi^2_{\alpha/2}\right) = 1 - \alpha \tag{5.4.2}$$

donde el valor de χ^2 en medio de esta doble desigualdad es igual a la planteada en (5.4.1) y donde $\chi^2_{1-\alpha/2}$ y $\chi^2_{\alpha/2}$ son dos valores de χ^2 (los valores **críticos** χ^2)

---

[4] Para una exposición clara, véase John Neter, William Wasserman y Michael H. Kutner, *Applied Linear Regression Models*, Richard D. Irwin, Homewood, Ill., 1983, capítulo 5.

[5] Para una demostración, véase Robert V. Hogg y Allen T. Craig, *Introduction to Mathematical Statistics*, 2a. ed., Macmillan, Nueva York, 1965, p. 144.

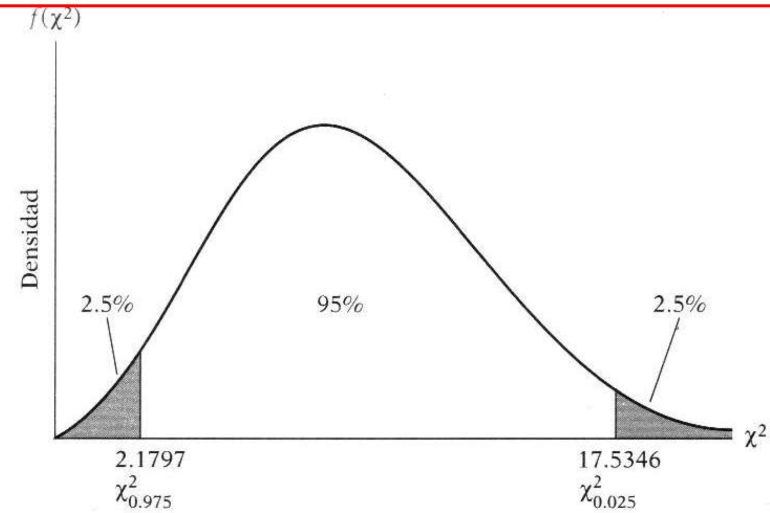

FIGURA 5.1 Intervalo de confianza del 95% para χ^2 (8 g de l).

obtenidos de la tabla ji cuadrada para $n - 2$ g de l, de tal manera, que ellos cortan $100(\alpha/2)\%$ de las áreas de las colas de la distribución χ^2 como se muestra en la figura 5.1.

Sustituyendo χ^2 de (5.4.1) en (5.4.2) y reorganizando los términos, se obtiene

$$\Pr\left[(n-2)\frac{\hat{\sigma}^2}{\chi^2_{\alpha/2}} \leq \sigma^2 \leq (n-2)\frac{\hat{\sigma}^2}{\chi^2_{1-\alpha/2}}\right] = 1 - \alpha \tag{5.4.3}$$

que da el intervalo de confianza para σ^2 de $100(1 - \alpha)\%$.

Para ilustrar, considérese este ejemplo. Del capítulo 3, sección 3.6, se obtuvo $\hat{\sigma}^2 = 42.1591$ y g de l = 8. Si se escoge α igual al 5%, la tabla ji cuadrada para 8 g de l da los siguientes valores críticos: $\chi^2_{0.025} = 17.5346$ y $\chi^2_{0.975} = 2.1797$. Estos valores muestran que la probabilidad de un valor ji cuadrada que exceda 17.5346 es 2.5% y el de 2.1797 es 97.5%. Por consiguiente, el intervalo entre estos dos valores es el intervalo de confianza para χ^2 al 95%, como se muestra en el diagrama de la figura 5.1. (Obsérvese la característica asimétrica de la distribución ji-cuadrada.)

Sustituyendo los datos del ejemplo en (5.4.3), el lector debe verificar que el intervalo de confianza para σ^2 al 95% es el siguiente:

$$19.2347 \leq \sigma^2 \leq 154.7336 \tag{5.4.4}$$

La interpretación de este intervalo es la siguiente: Si establecemos límites de confianza al 95% sobre σ^2 y si afirmamos *a priori* que entre estos límites caerá el verdadero σ^2, se acertará a largo plazo el 95% de las veces.

5.5 PRUEBA DE HIPÓTESIS: COMENTARIOS GENERALES

Habiendo estudiado el problema de la estimación puntual y de intervalos, se considerará ahora el tema de la prueba de hipótesis. En esta sección se analizarán brevemente algunos aspectos generales de este tema; el **apéndice A** proporciona algunos detalles adicionales.

El problema de la prueba de hipótesis estadística puede plantearse sencillamente de la siguiente manera: *¿Es compatible o incompatible una observación dada o un hallazgo, con algunas hipótesis planteadas?* La palabra "compatible", se utiliza aquí en el sentido de que la observación está lo "suficientemente" cercana al valor hipotético, de tal forma que no se rechaza la hipótesis planteada. Así, si alguna teoría o experiencia previa lleva a creer que el verdadero coeficiente de la pendiente β_2 en el ejemplo consumo-ingreso es la unidad, ¿es el $\hat{\beta}_2 = 0.5091$ obtenido de la muestra de la tabla 3.2 consistente con la hipótesis planteada? De ser así, no se rechaza la hipótesis; de lo contrario, se puede rechazar.

En el lenguaje de estadística, la hipótesis planteada es conocida como **hipótesis nula** y está denotada por el símbolo H_0. La hipótesis nula es usualmente probada frente a una **hipótesis alternativa** (también conocida como **hipótesis mantenida**) denotada por H_1, que puede plantear, por ejemplo, que el verdadero β_2 es diferente a la unidad. La hipótesis alterna puede ser **simple** o **compuesta**.[6] Por ejemplo, H_1: $\beta_2 = 1.5$ es una hipótesis simple, pero H_1: $\beta_2 \neq 1.5$ es una hipótesis compuesta.

La teoría de prueba de hipótesis se preocupa por el diseño de reglas o procedimientos que permitan decidir si se rechaza o no la hipótesis nula. Hay dos métodos *mutuamente complementarios* para diseñar tales reglas, a saber: el **intervalo de confianza** y la **prueba de significancia**. Estos dos enfoques plantean que la variable (el estadístico o estimador) bajo consideración sigue alguna distribución de probabilidad y que la prueba de hipótesis establece afirmaciones sobre el (los) valor(es) del (los) parámetro(s) de tal distribución. Por ejemplo, se sabe que con el supuesto de normalidad $\hat{\beta}_2$ está normalmente distribuida con media igual a β_2 y varianza dada por (4.3.5). Si formulamos la hipótesis de que $\beta_2 = 1$, se está haciendo una afirmación sobre uno de los parámetros de la distribución normal, por ejemplo, la media. La mayoría de las hipótesis estadísticas que se encuentran en este texto serán de este tipo, haciendo afirmaciones sobre uno o más valores de los parámetros de algunas distribuciones de probabilidad supuestas, tales como la normal, F, t o χ^2. En las dos secciones siguientes se estudia la forma como se logra esto.

5.6 PRUEBA DE HIPÓTESIS: MÉTODO DEL INTERVALO DE CONFIANZA

Prueba de dos lados o dos colas

Para ilustrar el método del intervalo de confianza, una vez más se hace referencia al ejemplo consumo-ingreso. Como se sabe, la propensión marginal a consumir estimada (PMC), $\hat{\beta}_2$, es 0.5091.

Supóngase que se postula que

$$H_0: \beta_2 = 0.3$$
$$H_1: \beta_2 \neq 0.3$$

[6] Se dice que una hipótesis estadística se denomina **hipótesis simple** si especifica el (los) valor(es) preciso(s) del (los) parámetro(s) de una función de densidad de probabilidad (fdp); de lo contrario, es llamada **hipótesis compuesta**. Por ejemplo, en la fdp normal $(1/\sigma\sqrt{2\pi})\exp\left\{-\frac{1}{2}[(X-\mu)/\sigma]^2\right\}$, si afirmamos que H_1: $\mu = 15$ y $\sigma = 2$, se trata de una hipótesis simple; pero si H_1: $\mu = 15$ y $\sigma > 15$, es una hipótesis compuesta, porque la desviación estándar no tiene un valor específico.

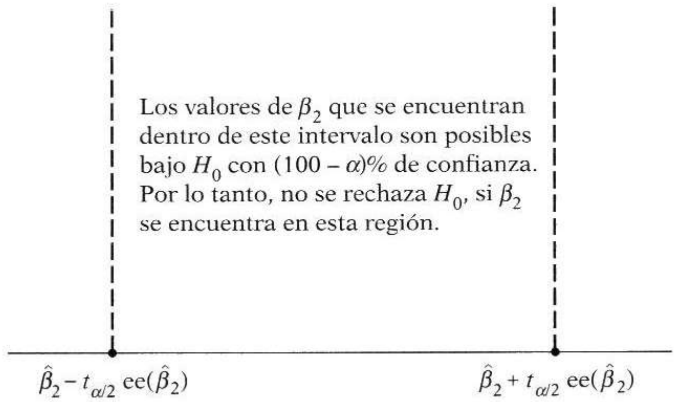

FIGURA 5.2 Intervalo de confianza para β_2 al $(100 - \alpha)\%$.

es decir, bajo la hipótesis nula, la verdadera PMC es 0.3 pero, según la hipótesis alterna, su valor es menor o mayor de 0.3. La hipótesis nula es una hipótesis simple, mientras que la hipótesis alterna es compuesta; y, en la práctica, se conoce como **hipótesis de dos lados o de dos colas**. Muy frecuentemente dicha hipótesis alterna de dos lados refleja el hecho de que no se tiene una expectativa fuerte *a priori* o teórica sobre la dirección en la cual debe moverse la hipótesis alterna con respecto a la hipótesis nula.

¿Es el $\hat{\beta}_2$ compatible con H_0? Para responder a esta pregunta, se hace referencia al intervalo de confianza (5.3.9). Se sabe que, a largo plazo, intervalos como (0.4268, 0.5914) contendrán el verdadero β_2 con una probabilidad del 95%.

En consecuencia, a largo plazo (es decir, en muestreo repetido) tales intervalos proporcionan un recorrido o límites dentro de los cuales puede encontrarse el verdadero β_2 con un coeficiente de confianza de, digamos, 95%. Por tanto, el intervalo de confianza proporciona un conjunto de hipótesis nulas posibles. Por consiguiente, si β_2 bajo H_0 se encuentra dentro del intervalo de confianza $100(1 - \alpha)\%$, no se rechaza la hipótesis nula; si éste se encuentra por fuera del intervalo, se puede rechazar.[7] Este rango se ilustra esquemáticamente en la figura 5.2.

> **Regla de decisión:** Constrúyase un intervalo de confianza para β_2 al $100(1 - \alpha)\%$. Si β_2 bajo H_0 se encuentra dentro de este intervalo de confianza, no se rechace H_0, pero si está por fuera del intervalo, rechace H_0.

Siguiendo esta regla, para el ejemplo hipotético, H_0: $\beta_2 = 0.3$ es claro que éste se encuentra por fuera del intervalo de confianza al 95% dado en (5.3.9). Por consiguiente, se puede rechazar la hipótesis de que la verdadera PMC sea 0.3, con 95% de confianza. Si la hipótesis nula fuera cierta, la probabilidad de obtener por casualidad un valor de PMC igual a 0.5091 es, como máximo de 5%, una probabilidad pequeña.

[7] Siempre téngase en mente que hay una posibilidad de $100\alpha\%$ de que el intervalo de confianza no contenga a β_2 bajo H_0 aunque la hipótesis esté correcta. En pocas palabras hay una posibilidad de $100\alpha\%$ de cometer un **error tipo I**. Así, si $\alpha = 0.05$, hay una posibilidad de 5% de que se pueda rechazar la hipótesis nula aun cuando ésta sea verdadera.

En estadística, cuando se rechaza la hipótesis nula, se dice que el hallazgo es **estadísticamente significativo**. Por otra parte, cuando no se hace, se dice que el hallazgo **no es estadísticamente significativo**.

Algunos autores utilizan frases como "altamente significativo desde un punto de vista estadístico". Con este término, generalmente quieren decir que cuando ellos rechazan la hipótesis nula, la probabilidad de cometer un error tipo I (por ejemplo, α) es un número pequeño, usualmente 1%. Pero, como lo demostrará el análisis del **valor p** en la sección 5.8, es mejor dejar que el investigador califique el hallazgo estadístico como "significativo" , "moderadamente significativo" o "altamente significativo".

Prueba de un lado o de una cola

Algunas veces tenemos una gran expectativa *a priori* o teórica (o existen expectativas basadas en algún trabajo empírico previo) de que la hipótesis alterna es de un lado o de una dirección, en lugar de ser de dos lados o dos colas, como se acaba de analizar. Así, para el ejemplo consumo-ingreso, se puede postular que

$$H_0\colon \beta_2 \le 0.3 \qquad \text{y} \qquad H_1\colon \beta_2 > 0.3$$

Puede ser que la teoría económica o el trabajo empírico previo sugieran que la propensión marginal a consumir es mayor de 0.3. Aunque el procedimiento para probar esta hipótesis puede derivarse fácilmente de (5.3.5), el mecanismo real está mejor explicado en términos del método de prueba de significancia analizado a continuación.[8]

5.7 PRUEBA DE HIPÓTESIS: MÉTODO DE PRUEBA DE SIGNIFICANCIA

Prueba de significancia de los coeficientes de regresión: la prueba *t*

Un *método alternativo, pero complementario* al de intervalos de confianza para probar hipótesis estadísticas es el **método de la prueba de significancia** desarrollado en forma independiente por R.A. Fisher y conjuntamente por Neyman y Pearson.[9] **En términos generales, una prueba de significancia es un procedimiento mediante el cual se utilizan los resultados muestrales para verificar la verdad o falsedad de una hipótesis nula**. La idea básica detrás de las pruebas de significancia es la de un **estadístico de prueba** (un estimador) y su distribución muestral bajo la hipótesis nula. La decisión de aceptar o rechazar H_0 se lleva a cabo con base en el valor del estadístico de prueba obtenido a partir de los datos disponibles.

[8] Si se desea utilizar el método de intervalos de confianza, constrúyase un intervalo de confianza para β_2 al $(100 - \alpha)\%$ de *un lado* o *de una cola*. ¿Por qué?

[9] Se pueden encontrar detalles en E.L. Lehman, *Testing Statistical Hypotheses*, John Wiley & Sons, Nueva York, 1959.

Como ilustración, recuérdese que, bajo el supuesto de normalidad, la variable

$$t = \frac{\hat{\beta}_2 - \beta_2}{\text{ee}(\hat{\beta}_2)} = \frac{(\hat{\beta}_2 - \beta_2)\sqrt{\sum x_i^2}}{\hat{\sigma}} \tag{5.3.2}$$

sigue la distribución t con $n - 2$ g de l. Si el valor del verdadero β_2 es especificado bajo la hipótesis nula, el valor t de (5.3.2) puede ser calculado fácilmente a partir de la muestra disponible y, por consiguiente, puede servir como un estadístico de prueba. Debido a que este estadístico de prueba sigue una distribución t, pueden hacerse afirmaciones sobre los intervalos de confianza como la siguiente:

$$\Pr\left[-t_{\alpha/2} \le \frac{\hat{\beta}_2 - \beta_2^*}{\text{ee}(\hat{\beta}_2)} \le t_{\alpha/2}\right] = 1 - \alpha \tag{5.7.1}$$

donde β_2^* es el valor de β_2 bajo H_0 y donde $-t_{\alpha/2}$ y $t_{\alpha/2}$ son los valores de t (los valores **críticos** de t) obtenidos de la tabla t para un nivel de significancia ($\alpha/2$) y $n - 2$ g de l [cf. (5.3.4)]. La tabla t está en el **apéndice D**.

Reorganizando (5.7.1), se obtiene

$$\Pr\,[\hat{\beta}_2 - t_{\alpha/2}\,\text{ee}(\hat{\beta}_2) \le \hat{\beta}_2 \le \beta_2^* + t_{\alpha/2}\,\text{ee}(\hat{\beta}_2)] = 1 - \alpha \tag{5.7.2}$$

que da el intervalo en el cual se encontrará $\hat{\beta}_2$ con probabilidad $1 - \alpha$, dado $\beta_2 = \beta_2^*$. En el lenguaje de prueba de hipótesis, el intervalo de confianza al 100 $(1 - \alpha)\%$ establecido en (5.7.2) es conocido como la **región de aceptación** (de la hipótesis nula) y la(s) *región(es)* que queda(n) por fuera del intervalo de confianza es (son) llamada(s) la(s) **región(es) de rechazo** (de la H_0) o la(s) **región(es) crítica(s)**. Como se anotó previamente, los límites de confianza dados por los puntos extremos del intervalo de confianza son llamados también **valores críticos**.

La estrecha conexión entre los enfoques de intervalo de confianza y prueba de significancia para realizar la prueba de hipótesis puede verse ahora comparando (5.3.5) con (5.7.2). En el procedimiento de intervalo de confianza se trata de establecer un rango o intervalo que tenga una probabilidad determinada de contener el verdadero, aunque desconocido β_2, mientras que en el método de prueba de significancia se somete a hipótesis algún valor de β_2 y se trata de ver si el $\hat{\beta}_2$ calculado se encuentra dentro de límites (de confianza) razonables alrededor del valor sometido a hipótesis.

Considérese una vez más el ejemplo de consumo-ingreso. Se sabe que $\hat{\beta}_2 =$ 0.5091, ee $(\hat{\beta}_2) = 0.0357$ y g de l = 8. Si se supone $\alpha = 5\%$, $t_{\alpha/2} = 2.306$. Si se plantea que H_0: $\beta_2 = \beta_2^* = 0.3$ y H_1: $\beta_2 \neq 0.3$, (5.7.2) se convierte en

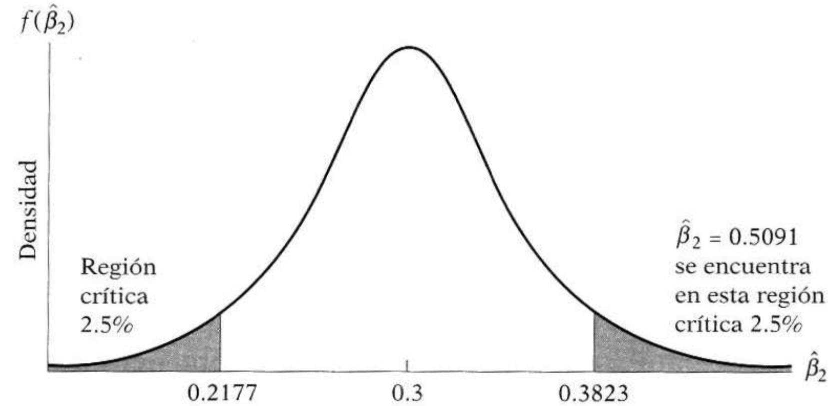

FIGURA 5.3 Intervalo de confianza al 95% para $\hat{\beta}_2$ bajo la hipótesis de que $\beta_2 = 0.3$.

$$\Pr(0.2177 \leq \hat{\beta}_2 \leq 0.3823) = 0.95 \tag{5.7.3}$$ [10]

como se muestra en el diagrama de la figura 5.3. Puesto que $\hat{\beta}_2$ se encuentra en la región crítica, se rechaza la hipótesis nula de que el verdadero $\beta_2 = 0.3$.

En la práctica, no hay necesidad de estimar (5.7.2) explícitamente. Se puede calcular el valor de t del centro de la doble desigualdad dada en (5.7.1) y ver si ésta cae entre los valores críticos t o por fuera de éstos. Para el ejemplo,

$$t = \frac{0.5091 - 0.3}{0.0357} = 5.86 \tag{5.7.4}$$

es claro que este valor se encuentra en la región crítica de la figura 5.4. La conclusión se mantiene; es decir, rechazamos H_0.

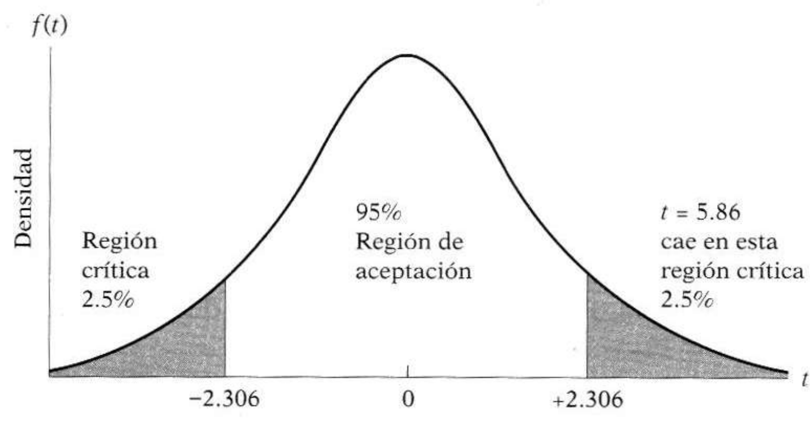

FIGURA 5.4 Intervalo de confianza del 95% para t(8 g de l).

[10] En la sección 5.2, punto 4, se afirma que *no se puede* decir que la probabilidad de que el intervalo fijo (0.4268, 0.5914) incluya el verdadero β_2, sea del 95%. Pero podemos hacer la afirmación probabilística dada en (5.7.3) porque $\hat{\beta}_2$, siendo un estimador, es una variable aleatoria.

Obsérvese que si el $\beta_2(=\hat{\beta}_2)$ estimado es igual al β_2 hipotético, el valor t en (5.7.4) será cero. Sin embargo, en la medida en que el valor de β_2 estimado se aleje del valor hipotético de β_2, el $|t|$ (es decir, el valor absoluto de t. *Nota:* t puede ser positivo o negativo) será cada vez mayor. *Por consiguiente, un valor "grande" de* $|t|$ *será evidencia en contra de la hipótesis nula*. Siempre se puede utilizar la tabla t para determinar si un valor t particular es grande o pequeño; la respuesta, como se sabe, depende de los grados de libertad igual que de la probabilidad del error tipo I que se esté dispuesto a aceptar. Como se puede observar en la tabla t dada en el **apéndice D**, para cualquier valor dado de g de l, la probabilidad de obtener un valor de $|t|$ cada vez mayor se hace progresivamente menor. Por tanto, para 20 g de l, la probabilidad de obtener un valor $|t|$ mayor o igual a 1.725 es 0.10 o 10%, pero para los mismos g de l, la probabilidad de obtener un valor $|t|$ mayor o igual a 3.552 es tan sólo 0.002 o 0.2%.

Puesto que se utiliza la distribución t, el anterior procedimiento de prueba es llamado apropiadamente la **prueba t**. **En el lenguaje de las pruebas de significancia, se dice que un estadístico es estadísticamente significativo si el valor del estadístico de prueba cae en la región crítica. En este caso, la hipótesis nula se rechaza. De la misma manera, se dice que una prueba es estadísticamente no significativa si el valor del estadístico de prueba cae en la región de aceptación.** En esta situación, la hipótesis nula no se rechaza. En nuestro ejemplo, la prueba t es significativa y por tanto se rechaza la hipótesis nula.

Antes de concluir la exposición de pruebas de hipótesis, obsérvese que el procedimiento de prueba presentado se conoce como el procedimiento de las pruebas de significancia de **dos lados, o dos colas**, ya que se consideran las dos colas extremas de la distribución de probabilidad relevante, como regiones de rechazo, y se rechaza la hipótesis nula si cae en cualquiera de ellas. Esto sucede porque la H_1 era una hipótesis compuesta de dos lados; $\beta_2 \neq 0.3$ significa que β_2 es mayor que o menor que 0.3. Supóngase que la experiencia sugiere que la PMC sea mayor que 0.3. En este caso se tiene: H_0: $\beta_2 \leq 0.3$ y H_1: $\beta_2 > 0.3$. Aunque H_1 es aún una hipótesis compuesta, tiene ahora tan sólo un lado. Para probar esta hipótesis, se utiliza una **prueba de una cola** (la cola derecha), como se observa en la figura 5.5. (Véase también el análisis en la sección 5.6.)

El procedimiento de prueba es similar al anterior excepto que el límite de confianza superior o valor crítico corresponde ahora a $t_\alpha = t_{.05}$, es decir, al nivel del 5%. Como lo indica la figura 5.5, en este caso no es preciso considerar la cola inferior de la distribución t. La utilización de una prueba de significancia de una o dos colas dependerá de la forma como esté formulada la hipótesis alterna, la cual, a su vez, puede depender de algunas consideraciones *a priori* o de experiencia empírica previa. (Otras consideraciones se dan en la sección 5.8.)

Un resumen del método de la prueba t de significancia para la prueba de hipótesis se presenta en la tabla 5.1.

Prueba de significancia para σ^2: la prueba χ^2

Como otro ejemplo de la metodología de las pruebas de significancia, considérese la siguiente variable:

$$\chi^2 = (n-2)\frac{\hat{\sigma}^2}{\sigma^2} \tag{5.4.1}$$

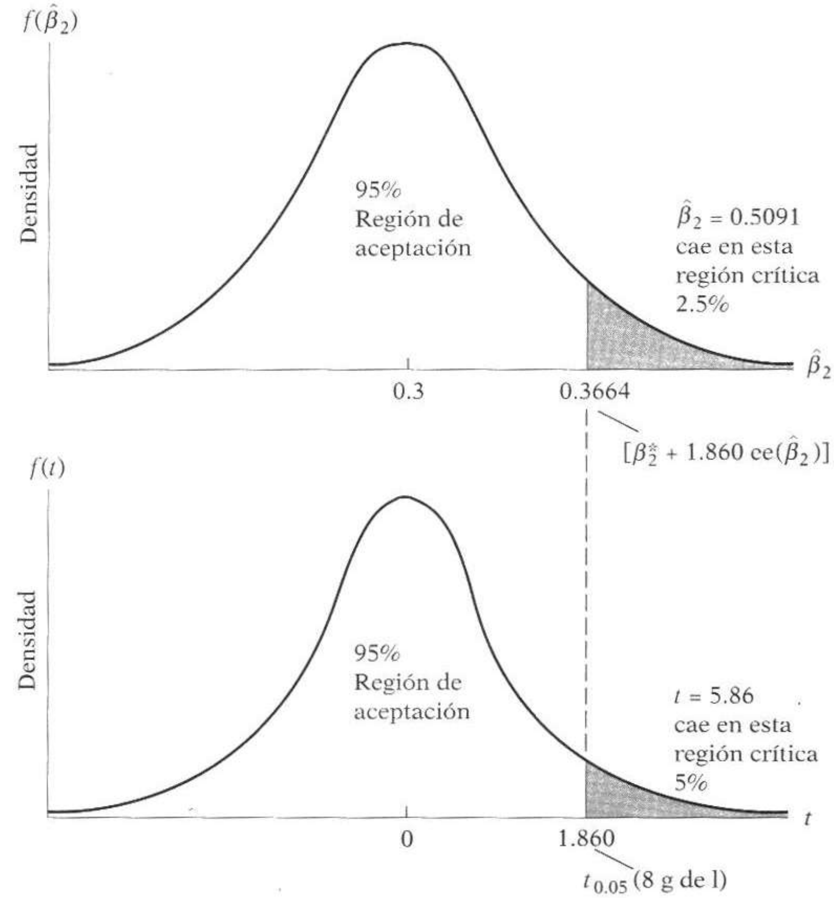

FIGURA 5.5 Prueba de significancia de una cola.

la cual, como se anotó previamente, sigue una distribución χ^2 con $n - 2$ g de l. Para el ejemplo hipotético, $\hat{\sigma}^2 = 42.1591$ y g de l = 8. Si se postula que H_0: $\sigma^2 = 85$ *vs.* H_1: $\sigma^2 \neq 85$, la ecuación (5.4.1) proporciona el estadístico de prueba para H_0. Sustituyendo los valores apropiados en (5.4.1), puede encontrarse que bajo H_0, $\chi^2 = 3.97$. Si se supone $\alpha = 5\%$, los valores críticos χ^2 son 2.1797 y 17.5346. Puesto que el χ^2 calculado cae dentro de estos límites, los datos apoyan la hipó-

TABLA 5.1 LA PRUEBA *t* DE SIGNIFICANCIA: REGLAS DE DECISIÓN

Tipo de hipótesis	H_0: la hipótesis nula	H_1: la hipótesis alterna	Regla de decisión: rechazar H_0 si
Dos colas	$\beta_2 = \beta_2^*$	$\beta_2 \neq \beta_2^*$	$\lvert t\rvert > t_{\alpha/2,\text{ g de l}}$
Cola derecha	$\beta_2 \leq \beta_2^*$	$\beta_2 > \beta_2^*$	$t > t_{\alpha,\text{ g de l}}$
Cola izquierda	$\beta_2 \geq \beta_2^*$	$\beta_2 < \beta_2^*$	$t < -t_{\alpha,\text{ g de l}}$

Notas: β_2^* es el valor numérico de β_2 hipotético.
$\lvert t\rvert$ significa el valor absoluto de t.
t_α o $t_{\alpha/2}$ significa el valor crítico de t al nivel de significancia α o $\alpha/2$.
g de l: grados de libertad, $(n - 2)$ para el modelo de dos variables, $(n - 3)$ para el modelo de tres variables y así sucesivamente.
Para probar hipótesis sobre β_1 se sigue un procedimiento similar.

TABLA 5.2 RESUMEN DE LA PRUEBA χ^2

H_0: la hipótesis nula	H_1: la hipótesis alterna	Región crítica: rechazar H_0 si
$\sigma^2 = \sigma_0^2$	$\sigma^2 > \sigma_0^2$	$\dfrac{\text{g de l}(\hat\sigma^2)}{\sigma_0^2} > \chi^2_{\alpha,\ \text{g de l}}$
$\sigma^2 = \sigma_0^2$	$\sigma^2 < \sigma_0^2$	$\dfrac{\text{g de l}(\hat\sigma^2)}{\sigma_0^2} > \chi^2_{(1-\alpha),\ \text{g de l}}$
$\sigma^2 = \sigma_0^2$	$\sigma^2 \neq \sigma_0^2$	$\dfrac{\text{g de l}(\hat\sigma^2)}{\sigma_0^2} > \chi^2_{\alpha/2,\ \text{g de l}}$ o $< \chi^2_{(1-\alpha/2),\ \text{g de l}}$

Nota: σ_0^2 es el valor de σ^2 bajo la hipótesis nula. El primer subíndice asociado a χ^2 en la última columna es el nivel de significancia, en tanto que el segundo indica los grados de libertad. Éstos son los valores críticos ji-cuadrada. Obsérvese que si el modelo de regresión es de dos variables, los g de l son $(n-2)$, si el modelo de regresión es de tres variables son $(n-3)$ y así sucesivamente.

tesis nula y no se la rechaza. (*Véase* figura 5.1.) Este procedimiento de prueba se denomina la **prueba de significancia ji cuadrada**. El enfoque de la prueba de significancia χ^2 para la prueba de hipótesis se resume en la tabla 5.2.

5.8 PRUEBA DE HIPÓTESIS: ALGUNOS ASPECTOS PRÁCTICOS

El significado de "aceptar" o "rechazar" una hipótesis

Si, con base en una prueba de significancia, por ejemplo, la prueba t, se decide "aceptar" la hipótesis nula, todo lo que se está diciendo es que con base en la evidencia dada por la muestra, no existe razón para rechazarla; no se está diciendo que la hipótesis nula sea verdadera con absoluta certeza. ¿Por qué? Para responder esto, téngase en cuenta el ejemplo consumo-ingreso y supóngase que H_0: β_2 (PMC) = 0.50. Ahora, el valor estimado de la PMC es $\hat\beta_2 = 0.5091$ con un ee $(\hat\beta_2) = 0.0357$. Entonces, con base en la prueba t, se encuentra que $t = (0.5091 - 0.50)/0.0357 = 0.25$, que es no significativo, es decir, para un $\alpha = 5\%$. Por consiguiente, se dice que "aceptamos" H_0. Pero ahora supóngase H_0: $\beta_2 = 0.48$. Aplicando la prueba t, se obtiene $t = (0.5091 - 0.48)/0.0357 = 0.82$, el cual tampoco es estadísticamente significativo. Entonces, se dice ahora que "se acepta" esta H_0. ¿Cuál de estas dos hipótesis nulas es la "verdadera"? No se sabe. Por consiguiente, en la "aceptación" de una hipótesis nula siempre se debe tener presente que puede existir otra hipótesis nula igualmente compatible con los datos. Es preferible, por tanto, decir que se puede aceptar la hipótesis nula en lugar de decir que se la acepta. Mejor aún,

> ...de la misma manera que una corte pronuncia un veredicto de "no culpable" en lugar de decir "inocente", así la conclusión de un estadístico de prueba es la de "no rechazar" en lugar de "aceptar".[11]

[11] Jan Kmenta, *Elements of Econometrics*, Macmillan, Nueva York, 1971, p. 114.

Hipótesis nula o "cero" y regla práctica "2-*t*"

La hipótesis nula que es objeto frecuente de prueba en el trabajo empírico es H_0: $\beta_2 = 0$, es decir, el coeficiente de la pendiente es cero. Esta hipótesis nula de "cero" es un mecanismo para establecer si Y tiene relación con X, la variable explicativa. Si, para empezar, no existe relación entre Y y X, entonces la prueba de hipótesis tal como $\beta_2 = 0.3$ o cualquier otro valor no tiene significado.

Esta hipótesis nula puede probarse fácilmente mediante los enfoques de intervalos de confianza o prueba t estudiados en las secciones anteriores. Pero, muy frecuentemente, tales pruebas formales pueden abreviarse adoptando la regla de significancia "2-t" que puede expresarse así:

Regla práctica "2-*t*". Si el número de grados de libertad es 20 y si α, el nivel de significancia, se fija en 0.05, entonces la hipótesis nula $\beta_2 = 0$ puede ser rechazada si el valor de $t\,[=\hat{\beta}_2/\text{ee}(\hat{\beta}_2)]$ calculado a partir de (5.3.2) excede a 2 en valor absoluto.

El razonamiento para esta regla no es muy difícil de entender. De (5.7.1) se sabe que se rechazará H_0: $\beta_2 = 0$ si

$$t = \hat{\beta}_2/\text{ee}(\hat{\beta}_2) > t_{\alpha/2} \qquad \text{cuando } \hat{\beta}_2 > 0$$

o

$$t = \hat{\beta}_2/\text{ee}(\hat{\beta}_2) < -t_{\alpha/2} \qquad \text{cuando } \hat{\beta}_2 < 0$$

o cuando

$$|t| = \left|\frac{\hat{\beta}_2}{\text{ee}(\hat{\beta}_2)}\right| > t_{\alpha/2} \tag{5.8.1}$$

para los grados de libertad apropiados.

Ahora, si se examina la tabla t dada en el **apéndice D**, se ve que para g de l alrededor de 20 o más, un valor calculado t mayor que 2 (en términos absolutos), es decir, 2.1, es estadísticamente significativo al nivel del 5%, lo cual implica rechazo de la hipótesis nula. Por consiguiente, si se encuentra que para 20 o más g de l el valor t calculado es, digamos, 2.5 o 3, ni siquiera se tiene que consultar la tabla t para asegurar la significancia del coeficiente de la pendiente estimado. Por supuesto, siempre puede referirse a la tabla t para obtener el nivel preciso de significancia. Sin embargo, esto debe hacerse siempre que los g de l sean inferiores, por ejemplo, a 20.

A propósito, obsérvese que si se está probando la hipótesis de un lado $\beta_2 = 0$ *vs*. $\beta_2 > 0$ o $\beta_2 < 0$, entonces se debe rechazar la hipótesis nula si

$$|t| = \left|\frac{\hat{\beta}_2}{\text{ee}(\hat{\beta}_2)}\right| > t_{\alpha} \tag{5.8.2}$$

Si se fija α en 0.05, entonces en la tabla t se observa que, para 20 o más g de l, un valor t mayor que 1.73 es estadísticamente significativo al nivel de significancia del 5% (de una cola). Por lo tanto, siempre que un valor t exceda, por ejemplo, 1.8 (en términos absolutos) y los g de l sean 20 o más, no es necesario consultar la tabla t para la significancia estadística del coeficiente observado. Es claro que, si se escoge α igual a 0.01 o cualquier otro nivel, se tendrá que decidir sobre el valor apropiado de t como valor crítico de referencia; el lector deberá ser capaz de hacer eso.

Formación de las hipótesis nula y alterna[12]

Dadas las hipótesis nula y alterna, probar su significancia estadística no debe seguir siendo un misterio. Pero, ¿cómo se formulan estas hipótesis? No existen reglas específicas. Muy frecuentemente el fenómeno bajo estudio sugerirá la forma de las hipótesis nula y alterna. Por ejemplo, se pide estimar la línea del mercado de capitales (LMC) de la teoría de portafolio, que postula que $E_i = \beta_1 + \beta_2\sigma_i$, donde E = ganancia esperada sobre el portafolio y σ = la desviación estándar de la ganancia, una medida de riesgo. Puesto que se espera que la ganancia y el riesgo estén relacionados positivamente, entre mayor sea el riesgo, más alto será la ganancia; la hipótesis alterna natural a la hipótesis nula, $\beta_2 = 0$, sería $\beta_2 > 0$. Es decir, no se considerarán valores de β_2 menores de cero.

Pero considérese el caso de la demanda de dinero. Como se demostrará más adelante, uno de los determinantes importantes de la demanda de dinero es el ingreso. Estudios anteriores de las funciones de demanda de dinero han mostrado que la elasticidad ingreso de la demanda de dinero (el cambio porcentual en la demanda de dinero por un cambio porcentual de 1% en el ingreso) ha estado típicamente dentro de un rango de 0.7 a 1.3. Por consiguiente, en un nuevo estudio de la demanda de dinero, si se postula que el coeficiente β_2 de la elasticidad ingreso es 1, la hipótesis alterna podría ser que $\beta_2 \neq 1$, una hipótesis alterna de dos lados.

Por tanto, las expectativas teóricas o el trabajo empírico previo o ambos pueden ser la base para la formulación de hipótesis. Sin embargo, sin importar la forma como se postulen las hipótesis, *es extremadamente importante que el investigador plantee estas hipótesis antes de llevar a cabo la investigación empírica*. De lo contrario, él o ella serán culpables de razonamientos circulares o de profecías autocumplidas. Es decir, si se formulara la hipótesis después de examinar los resultados empíricos, podría presentarse la tentación de formular la hipótesis de tal manera que justifique los resultados obtenidos. Una práctica así debe ser evitada a cualquier costo, al menos para salvar la objetividad científica. ¡Recuérdese la cita de Stigler que se presenta al principio de este capítulo!

Selección del nivel de significancia α

Del análisis adelantado hasta el momento, debe tenerse claro que el hecho de rechazar o no una hipótesis nula depende en forma crítica de α, el nivel de significancia o *probabilidad de cometer un error **tipo I***, o sea, la probabilidad de rechazar la hipótesis cuando es verdadera. En el **apéndice A** analizamos en detalle

[12] Para una exposición interesante sobre la formulación de hipótesis, véase J. Bradford De Long y Kevin Lang, "Are All Economic Hypotheses False?", *Journal of Political Economy*, vol 100, núm. 6, 1992, pp. 1257-1272.

la naturaleza del llamado error tipo I, su relación con el error ***tipo II*** (la probabilidad de aceptar la hipótesis cuando es falsa) y la razón por la cual la estadística clásica se centra generalmente en el error tipo I. ¿Por qué generalmente se considera α como un valor del 1%, 5%, o máximo del 10% para α? De hecho, no hay nada sagrado acerca de estos valores; cualquier otro valor sería igualmente apropiado.

En un libro introductorio como éste, no es posible analizar a fondo la razón por la cual se escogen los niveles de significancia 1, 5 o 10%, ya que esto nos llevaría al campo de la toma de decisiones estadísticas, que de por sí es una disciplina completa. Sin embargo, puede ofrecerse un breve resumen. Como se estudió en el **apéndice A**, para un tamaño de muestra dada, si tratamos de reducir un *error tipo I*, aumenta un *error tipo II* y viceversa. Es decir, dado el tamaño de la muestra, si tratamos de reducir la probabilidad de rechazar la hipótesis cuando es verdadera, se puede aumentar al mismo tiempo la probabilidad de aceptarla cuando es falsa. Por tanto, dado el tamaño de la muestra, existe una conexión de intercambio entre estos dos tipos de error. Ahora, la única forma en la cual se puede decidir sobre esta conexión es encontrar los costos relativos de los dos tipos de error. Entonces,

> Si el error de rechazar la hipótesis nula cuando es verdadera (error tipo I) es costoso en comparación con el error de no rechazar la hipótesis nula cuando es falsa (error tipo II), será razonable fijar la probabilidad de ocurrencia del primer tipo de error a niveles bajos. Si, por otra parte, el costo de incurrir en el error tipo I es bajo comparado con el costo de cometer el error tipo II, se justificará que la probabilidad del primer tipo de error sea alta (rebajando así la posibilidad de incurrir en el segundo tipo de error).[13]

Desde luego, pocas veces se conocen los costos de cometer los dos tipos de error. Por tanto, los econometristas tienen por costumbre fijar el valor de α a niveles del 1, 5 o 10% como máximo y escogen un estadístico de prueba que minimice la probabilidad de cometer un error tipo II. Puesto que uno menos la probabilidad de cometer un error tipo II se conoce como la **potencia de la prueba**, este procedimiento equivale a maximizar esa potencia de la prueba. (Véase en el **apéndice A** un análisis de la potencia de una prueba.)

Pero todo este problema relacionado con la selección del valor apropiado de α puede ser evitado si se utiliza lo que se conoce como el *valor p* o "***P-value***" del estadístico de prueba, que se analiza a continuación.

Nivel exacto de significancia: valor *p* o "*P-value*"

Como recién se anotó, el talón de Aquiles del método clásico de la prueba de hipótesis es su arbitrariedad en la selección de α. Una vez que se ha obtenido un estadístico de prueba (es decir, el estadístico t) en un ejemplo dado, ¿por qué no consultar sencillamente la tabla estadística apropiada y encontrar la probabilidad real de obtener un valor del estadístico de prueba tan grande o mayor que el obtenido en el ejemplo? Esta probabilidad se denomina el **valor *p*** (es decir, **el valor de probabilidad**), también conocido como el **nivel observado o exacto de significancia** o la **probabilidad exacta de cometer un error tipo I**. Más

[13] Jan Kmenta, *Elements of Econometrics*, Macmillan, Nueva York, 1971, pp. 126-127.

técnicamente, el valor p está definido como el **nivel de significancia más bajo al cual puede rechazarse una hipótesis nula**.

Para ilustrar, recuérdese el ejemplo consumo-ingreso. Dada la hipótesis nula de que la verdadera PMC es 0.3, se obtuvo un valor t de 5.86 en (5.7.4). ¿Cual es el valor p o "p-value" de obtener un valor t igual o superior a 5.86? En la tabla t del **apéndice D**, se observa que para 8 g de l la probabilidad de obtener tal valor t debe estar muy por debajo de 0.001 (una cola) o 0.002 (dos colas). Mediante el uso de la computadora puede mostrarse que la probabilidad de obtener un valor t mayor o igual a 5.86 (8 g de l) es alrededor de 0.000189.[14] Éste es el valor p del estadístico t observado. Este nivel de significancia observado o exacto del estadístico t es mucho menor que los niveles de significancia del 1%, 5% o 10% fijados convencional y arbitrariamente. De hecho, si fuéramos a utilizar el valor p recién calculado y rechazar la hipótesis nula de que la verdadera PMC es 0.3, la probabilidad de que se cometa un error tipo I es sólo de cerca de 0.02%, es decir, ¡solamente 2 en 10 000!

Como se anotó anteriormente, si los datos no apoyan la hipótesis nula, el $|t|$ obtenido bajo tal hipótesis nula será "grande" y, por consiguiente, el valor p de obtener tal $|t|$ será "pequeño". En otras palabras, para un tamaño de muestra dado, a medida que aumenta $|t|$, el valor p se reduce y se puede, por consiguiente, rechazar la hipótesis nula con mayor confianza.

¿Cuál es la relación entre el valor p y el nivel de significancia α? Si se adquiere el hábito de fijar α igual al valor p de un estadístico de prueba (es decir, el estadístico t), entonces no hay conflicto entre estos dos valores. Expresado en otros términos, **es mejor no fijar de forma arbitraria α a algún nivel sino escoger simplemente el valor p del estadístico de prueba**. Es preferible dejar que el lector decida si debe rechazar la hipótesis nula al valor p dado. Si, en una aplicación, el valor p de un estadístico de prueba resulta ser, por ejemplo, 0.145 o 14.5% y si el lector desea rechazar la hipótesis nula a este nivel (exacto) de significancia, entonces lo puede hacer. No está mal tomar el riesgo de equivocarse un 14.5% de las veces si se rechaza la hipótesis nula verdadera. De manera similar, como en el ejemplo de consumo-ingreso, no está mal si el investigador desea escoger un valor p cercano al 0.02% y no tomar el riesgo de equivocarse en más de ¡2 veces de cada 10 000! Después de todo, ¡algunos investigadores pueden ser amantes del riesgo y otros opuestos a él!

En el resto de este texto, se citará generalmente el valor p de un estadístico de prueba dado. Algunos lectores pueden desear fijar α a algún nivel y rechazar la hipótesis nula si el valor p es menor que α. Es su opción.

Significancia estadística *versus* significancia práctica

Recuérdese el ejemplo consumo-ingreso y ahora plantéese que la verdadera PMC es 0.61 (H_0: $\beta_2 = 0.61$). Basados en el resultado muestral de $\hat{\beta}_2 = 0.5091$, se obtuvo el intervalo (0.4268, 0.5914) al 95% de confianza. Puesto que este intervalo no incluye 0.61, es posible decir, con un 95% de confianza, que el valor estimado es estadísticamente significativo, es decir, significativamente diferente de 0.61.

[14] Es posible obtener el valor p utilizando tablas estadísticas electrónicas con varios lugares decimales. Desafortunadamente, las tablas estadísticas convencionales, debido a falta de espacio, no pueden ser así de refinadas. Existen paquetes estadísticos como Micro TSR SHAZAM y ET entre muchos otros que imprimen, hoy en día, los valores p o "p-values".

Pero, ¿cuál es el significado práctico o real del hallazgo? Es decir, ¿qué diferencia existe entre asignar a la PMC un valor de 0.61 o uno de 0.5091? ¿Es la diferencia de 0.1009 entre las dos PMC así de importante en la práctica?

La respuesta a esta pregunta depende de lo que en realidad se haga con estos estimados. Por ejemplo, de la macroeconomía se sabe que el multiplicador del ingreso es 1/(1 – PMC). Por tanto, si la PMC es 0.5091, el multiplicador es 2.04, pero será 2.56 si la PMC es igual a 0.61. Esto es, si el gobierno fuera a incrementar su gasto en US$1 para sacar la economía de una recesión, el ingreso aumentaría en ese caso en US$2.04, si la PMC es 0.5091 pero lo hará en US$2.56 si la PMC es 0.61. Y esa diferencia podría ser crucial para reactivar la economía.

El punto de toda esta exposición es que no se debe confundir la significancia estadística con la significancia práctica o económica. Como lo afirma Goldberger:

> Cuando una hipótesis nula, digamos $\beta_j = 1$, se especifica, lo que se busca es que β_j esté *cercano* a 1, tan cerca que para todos los propósitos prácticos pudiera ser tratada *como si fuera* 1. Pero el que 1.1 sea "prácticamente lo mismo que" 1.0 es un asunto de economía, no de estadística. No se puede resolver el asunto apoyándose en una prueba de hipótesis porque el estadístico de prueba $[t =](b_j - 1)/\hat{\sigma}_{bj}$ mide el coeficiente estimado en unidades de errores estándar, las cuales no tienen significado para medir el parámetro económico $\beta_j - 1$. Puede ser una buena idea reservar el término "significancia" para el concepto estadístico, adoptando la palabra "sustancial" para el concepto económico.[15]

El punto expresado por Goldberger es importante. A medida que el tamaño de la muestra se hace muy grande, la importancia de los temas relacionados con significancia estadística se hace mucho menor pero los temas de significancia económica adquieren importancia crítica. En efecto, puesto que con muestras grandes casi todas las hipótesis nulas serán rechazadas, puede haber estudios en los cuales la magnitud de los valores estimados puntuales pueda ser lo único importante.

Selección entre los métodos del intervalo de confianza y la prueba de significancia en la prueba de hipótesis

En la mayor parte de los análisis económicos aplicados, la hipótesis nula postulada hace las veces de comodín, y el objetivo del trabajo empírico es tumbarlo, es decir, rechazar la hipótesis nula. Por tanto, en nuestro ejemplo consumo-ingreso, la hipótesis nula de que la PMC, $\beta_2 = 0$ es claramente absurda, pero con frecuencia la utilizamos para ejemplificar los resultados empíricos. Parece ser que los editores de publicaciones especializadas de renombre no encuentran emocionante publicar un trabajo empírico que no rechace la hipótesis nula. De alguna manera, como noticia es más novedoso el hallazgo de que la PMC sea estadísticamente diferente de cero que el hallazgo de que sea igual a, digamos, ¡0.7!

[15] Arthur S. Goldberger, *A Course in Econometrics*, Harvard University Press, Cambridge, Massachusetts, 1991, p. 240. *Obsérvese* que b_j es el estimador MCO de β_j y $\hat{\sigma}_{bj}$ es su error estándar. Este enfoque está corroborado en D.N. McCloskey, "The Loss Function Has Been Mislaid: The Rhetoric of Significance Tests", *American Economic Review*, vol. 75, 1985, pp. 201-205. Véase también D.N. McCloskey y S.T. Ziliak, "The Standard Error of Regression," *Journal of Economic Literature*, vol. 37, 1996, pp. 97-114.

Por tanto, J. Bradford De Long y Kevin Lang sostienen que es mejor para los economistas

> ...concentrarse en las magnitudes de los coeficientes y dar informes sobre los niveles de confianza y no sobre las pruebas de significancia. Si todas, o casi todas, las hipótesis nulas son falsas, tiene poco sentido concentrarse en averiguar si un estimado es o no indistinguible de su valor predicho bajo la hipótesis nula. En lugar de esto, deseamos averiguar cuáles modelos son buenas aproximaciones, para lo cual es necesario que conozcamos los intervalos de los valores de los parámetros excluidos por los estimados empíricos.[16]

En resumen, estos autores prefieren el método del intervalo de confianza al de la prueba de significancia. El lector puede desear tener este consejo en mente.[17]

5.9 ANÁLISIS DE REGRESIÓN Y ANÁLISIS DE VARIANZA

En esta sección se estudiará el análisis de regresión desde el punto de vista del análisis de varianza y se introducirá al lector hacia una forma complementaria de mirar el problema de la inferencia estadística.

En el capítulo 3, sección 3.5, se desarrolló la siguiente identidad:

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum \hat{u}_i^2 = \hat{\beta}_2^2 \sum x_i^2 + \sum \hat{u}_i^2 \tag{3.5.2}$$

es decir, STC = SEC + SRC, la cual descompone la suma total de cuadrados (STC) en dos componentes: la suma explicada de cuadrados (SEC) y la suma de residuales al cuadrado (SRC). Un estudio de estos componentes de STC se conoce como el **análisis de varianza** (ANOVA) desde el punto de vista de la regresión.

Asociado con toda suma de cuadrados están sus g de l, es decir, el número de observaciones independientes sobre las cuales está basada. La STC tiene $n - 1$ g de l porque se pierde 1 g de l en el cálculo de la media muestral $\bar{Y}$. La SRC tiene $n - 2$ g de l. (¿Por qué?) (*Nota*: Esto es cierto solamente para el modelo de regresión con dos variables con presencia del intercepto β_1.) SEC tiene 1 g de l (de nuevo, esto es cierto solamente para el caso de dos variables), lo cual se deduce del hecho de que SEC $= \hat{\beta}_2^2 \sum x_i^2$ es una función de $\hat{\beta}_2$ sólo si $\sum x_i^2$ es conocida.

Reorganícense las sumas de cuadrados y sus g de l asociados en la tabla 5.3, que es la forma estándar de la tabla AOV denominada algunas veces la **tabla ANOVA**. Dada la información de la tabla 5.3, considérese ahora la siguiente variable:

$$\begin{aligned} F &= \frac{\text{SPC de SEC}}{\text{SPC de SRC}} \\ &= \frac{\hat{\beta}_2^2 \Sigma x_i^2}{\Sigma \hat{u}_i^2/(n-2)} \\ &= \frac{\hat{\beta}_2^2 \Sigma x_i^2}{\hat{\sigma}^2} \end{aligned} \tag{5.9.1}$$

[16] Véase su artículo citado en la nota de pie de página 12, p. 1271.

[17] Para alguna otra perspectiva diferente, véase Carter Hill, William Griffiths y George Judge, *Undergraduate Econometrics*, Wiley & Sons, Nueva York, 2001, p. 108.