

**Matemáticas  
y ciencias sociales**

# ¿Por qué probar en econometría?

---

Claramarta Adalid Diez de Urdanivia\*

*El propósito de este artículo es ofrecer una descripción de los usos de las pruebas de hipótesis en la investigación econométrica, es decir, de las pruebas acerca de las teorías económicas, de la validez del modelo y de las pruebas de diagnóstico. Se consideran particularmente los enfoques de los autores clásicos de la metodología de pruebas de hipótesis: Fisher, y Neyman y Pearson. Por otra parte, se mencionan algunos de los errores que se cometen con más frecuencia en la presentación de artículos, en el marco de modelos no anidados.*

En la bibliografía econométrica podemos encontrar expresiones frecuentes de escepticismo acerca del valor que se otorga a los procedimientos estadísticos, en particular a las denominadas pruebas de hipótesis.<sup>1</sup>

La econometría es una disciplina relacionada con la estadística y para algunos esta relación ha provocado suspicacia, a pesar de que en sus orígenes esa nueva forma de hacer economía resultaba útil para sus más entusiastas seguidores. Primero se utilizó como una herramienta descriptiva, para representar regularidades entre las va-

---

<sup>1</sup> Véase "Econometrics: alchemy or science?", en D. Hendry. *Essays in econometric methodology*: Blackwell Publishers, 1993, p. 20. En su artículo, Hendry plantea, entre otras cosas, las críticas expresadas por diversos autores acerca de los métodos estadísticos aplicados a las observaciones económicas.

\* Departamento de Política y Cultura, Universidad Autónoma Metropolitana, Unidad Xochimilco, México.

riables económicas, pero también en la medición de las mismas, que en aquella época dio como resultado la construcción de los números índice.

Con el desarrollo de la estadística matemática y sus métodos, no sólo descriptivos sino inferenciales, la econometría, y otras ciencias aplicadas como la biometría y la psicometría, pudieron ir adaptándolos a sus propias necesidades. Un avance muy importante en la metodología econométrica fue la introducción del concepto de probabilidad.

La definición para el término econometría desde la fundación de la Sociedad Econométrica, en 1930, se propuso que entre los objetivos principales de esta disciplina está el avance de la teoría económica en su relación con la estadística y la matemática.

Samuelson, Koopmans y Stone la definen como el análisis cuantitativo de fenómenos económicos del mundo real, a partir del desarrollo concurrente de la teoría y la observación, relacionados mediante métodos apropiados de inferencia.<sup>2</sup>

Para Spanos, a esta disciplina le concierne el estudio sistemático de los fenómenos económicos, para lo cual se emplean “datos que se observan”, en el marco en el que la teoría económica, así como la inferencia estadística, desempeñan un papel importante.<sup>3</sup>

Para llegar al estado actual que guarda la econometría, muchos problemas tuvieron que ser resueltos en el camino. Faltaban datos relevantes, ya que en muchos casos no existían mediciones para las variables a considerar en el modelo económico. El carácter no experimental de los datos imponía restricciones de naturaleza estadística, problema que fue teóricamente resuelto por Haavelmo en 1944 cuando introdujo el concepto de probabilidad en la práctica econométrica. Haavelmo argumentaba que los economistas no están en posición de aislar, controlar y manipular las condiciones económicas: no pueden realizar experimentos. En vez de eso, deben improvisar con observaciones pasivas (aquellas de los experimentos de la naturaleza) que se ven influidas por una gran cantidad de factores no tomados en cuenta por la teoría.<sup>4</sup>

De este modo, la econometría, en un principio, establece las relaciones entre variables económicas como cantidades y precios, ingresos y ventas, etc., abstrayendo

<sup>2</sup> Carl Christ. *Modelos y métodos econométricos*: México, Limusa, 1976.

<sup>3</sup> A. Spanos. *Statistical foundations of econometric modelling*: Cambridge University Press, Cambridge, 1986.

<sup>4</sup> M. Morgan. *The history of econometric ideas*: Cambridge University Press, Cambridge, 1990.

el fenómeno que se considera y estableciendo teorías en forma matemática. La utilidad empírica de los modelos resultantes es evaluada con los distintos métodos estadísticos. Las teorías económicas, en general, mejoran considerablemente en calidad y credibilidad cuando se contrastan con los datos económicos.<sup>5</sup> Podemos decir que una razón muy importante por la que se confrontan modelos y datos es que éste es uno de los procedimientos del método científico.

Así pues, usar pruebas estadísticas en la práctica económica forma parte relevante del método econométrico que usa las herramientas del análisis estadístico, no sólo descriptivo, sino principalmente inferencial. Entre los métodos inferenciales, las pruebas de hipótesis han desempeñado un papel fundamental, pues establecen el puente entre la teoría y el mundo real.

### **Las pruebas de hipótesis estadísticas**

Las pruebas de hipótesis estadísticas están íntimamente relacionadas con el tema de la estimación, que forma parte de la teoría estadística matemática. Su estudio se remonta a 1738, cuando en un ensayo escrito por Bernoulli se menciona que calculó una estadística de prueba para ensayar su hipótesis en el campo de la astronomía. Pero la formulación, como es conocida actualmente, fue desarrollada entre 1915 y 1933 por tres hombres: Ronald Fisher, Jerzy Neyman y Egon Pearson.<sup>6</sup>

La teoría moderna de las pruebas de hipótesis comienza con Gosset (cuyo seudónimo era Student) y su descubrimiento de la *prueba t* en 1908. A partir de entonces Fisher desarrolló el tema y realizó una serie de artículos que culminan en 1925 con la publicación de su libro *Métodos estadísticos para investigadores* (*Statistical methods for research workers*). Desde entonces, el tema de pruebas de hipótesis es considerado uno de los métodos cuantitativos más utilizados en casi todas las disciplinas.

Si bien es cierto que los estudiantes de estadística aprenden a probar hipótesis de acuerdo con cierta metodología, ésta no es una teoría unificada sino una amalgama de los resultados encontrados por Ronald Fisher, por una parte, y Neyman y

---

<sup>5</sup> D. Hendry. *Essays in econometric methodology*: Blackwell Publishers, 1993.

<sup>6</sup> W. Guenther. "Hypothesis testing", en Samuel Kotz y Norman L. Johnson (eds.). *Encyclopedia of statistical sciences*, vol. 8: Wiley, Nueva York, 1988.

Pearson, por la otra. Las teorías desarrolladas por estos autores se denominaron de manera distinta como: pruebas de significancia para el enfoque de Fisher y pruebas de hipótesis para el de Neyman-Pearson.<sup>7</sup>

Fisher propone la prueba de una hipótesis nula  $H_0$ :  $\theta = \theta_0$  usando el valor probabilístico o valor  $p$  para decidir acerca de la fuerza de la estadística de prueba: “si el valor  $p$  es pequeño, esto implica que, o la realización observada de la estadística de prueba constituye un evento raro, o que la hipótesis nula postulada es inválida”. Cuando no se rechaza la hipótesis nula sólo significa que es aceptada por el momento sobre una base provisional. Éste es, en esencia, el argumento de las pruebas de significancia: “Fisher consideraba que la estadística era la ciencia de la incertidumbre inferencial capaz de ofrecer la clave que permitiría resolver el largamente debatido problema de inducción”.<sup>8</sup> Sin embargo, ¿cómo escoger entre las distintas estadísticas de prueba?, ¿por qué unas y no otras?

Para resolver esta pregunta, Neyman y Pearson desarrollaron la teoría para calcular y derivar pruebas de significancia óptimas. Según la teoría de Neyman y Pearson, al establecimiento de la hipótesis nula  $H_0$ :  $\theta = \theta_0$  se le agrega la hipótesis alternativa, que en realidad es un conjunto de posibles alternativas,  $H_1$ :  $\theta = \theta_1 = \theta - \theta_0$ . La introducción de lo que conocemos como la hipótesis alternativa da por resultado el establecimiento de los errores de tipo I, falso rechazo, y de tipo II, falsa aceptación; así como la elección de una estadística de prueba óptima aparece como el resultado de maximizar la función potencia en el espacio de parámetros  $\theta_1$ . En esta teoría se reemplaza el valor  $p$  con una regla de decisión apoyada en la noción del nivel de significancia (o tamaño)  $\alpha$  de la prueba.

Neyman y Pearson resolvieron, el caso de la prueba de una hipótesis simple contra otra hipótesis simple con el lema de Neyman-Pearson, que en versión moderna establece: “*La prueba de la razón de verosimilitud es la prueba más poderosa de la hipótesis  $H_0: \theta = \theta_0$  en contra de la posibilidad simple  $H_1: \theta = \theta_1$* ”.<sup>9</sup>

<sup>7</sup> Un buen desarrollo de las diferencias y los usos de estos dos enfoques se puede encontrar en H. Roberts. “For what use are tests and hypotheses and tests of significance”, en O. Kempthorne y J. Pratt. *Communications in statistics*, 1976, VA5 N 7.9.

<sup>8</sup> E. L. Lehman. “The Fisher, Neyman-Pearson theories of testing hypotheses: one theory or two?”, en *Journal of the American Statistical Association*, vol. 88, núm. 424 (Theory and Methods), 1993.

<sup>9</sup> H. D. Brunk. *Introducción a la estadística matemática*: Trillas, México, 1979, p. 415.

Entonces, el principal objetivo en la perspectiva de Fisher es utilizar los datos como evidencia para fortalecer la validez de la hipótesis nula. Es decir, el enfoque es inferencial, mientras que Neyman y Pearson opinan que es necesario *tomar una decisión* de aceptar o rechazar la hipótesis nula.

La aproximación de Neyman-Pearson a las pruebas de hipótesis, método que fue popularizado en econometría por Haavelmo en 1944, entra en el enfoque hipotético deductivo,<sup>10</sup> mientras que la aproximación de Fisher está orientada a la inferencia inductiva, a partir de muestras. Por otro lado, Neyman y Pearson interpretan las pruebas como reglas de aceptación en el contexto del muestreo repetido; Fisher estaría de acuerdo en aceptarlo como una interpretación en problemas comerciales o tecnológicos, pero no para corroborar hipótesis científicas, ya que en estos casos el muestreo repetido resulta en una confusión y no existe un problema de decisión bien definido.<sup>11</sup>

La postura de los autores acerca de sus diferentes teorías no era conciliadora. Fisher fue, en todo momento, un crítico del trabajo de Neyman y Pearson, pues creía firmemente que las pruebas de hipótesis no tenían como objetivo decisiones finales e irrevocables, como se dejaba ver en los ejemplos presentados por éstos.<sup>12</sup> Las diferencias tanto filosóficas como prácticas fueron ferozmente defendidas por Fisher y Neyman como “una batalla que tuvo un efecto muy destructivo en la profesión estadística”.<sup>13</sup> Esta discusión duró mientras Fisher estuvo vivo.

La forma en que actualmente se presenta la teoría de pruebas de hipótesis, en la gran mayoría de los textos, es una amalgama de ambas aproximaciones y Gigerenzer describe la situación como sigue:

Aunque el debate [Fisher *versus* Neyman-Pearson] sigue presente entre los estadísticos, fue resuelto, silenciosamente, en los “libros de cocina” escritos entre los cuarenta y los sesenta, principalmente por no estadísticos, para enseñar a estudiantes de ciencias sociales las “reglas de la estadística”. La teoría de pruebas de

<sup>10</sup> Morgan, *ibidem*, 1990.

<sup>11</sup> A. Spanos. *Probability theory and statistical inference*: Cambridge University Press, Cambridge, 1995.

<sup>12</sup> En un interesante artículo, los autores ofrecen una revisión de las pruebas de hipótesis. T. W. Nix y B. Jackson. “The data analysis dilemma: ban or abandon”, en *Research in the Schools*, vol. 5, núm. 2, 1998, pp. 3-14.

<sup>13</sup> S. L. Zabell, “R. A. Fisher and the fiducial argument”, en *Statistical Science*: 7, 1992, pp. 369-387, citado en Lehman, 1993, *ibidem*.

significancia de Fisher, que históricamente apareció primero, fue mezclada con los conceptos de la teoría de Neyman y Pearson y enseñada como la estadística *per se*. Le llamamos a esto “teoría híbrida” de la inferencia estadística y esto continúa aun sin mencionar siquiera que ni Fisher ni Neyman-Pearson hubieran visto con buenos ojos el florecimiento de este matrimonio forzado.<sup>14</sup>

Pero para otros autores, como Lehman, aun aceptando el hecho de que hay diferencias entre ambas teorías, en algunas circunstancias pueden combinarse en una aproximación unificada. Este enfoque “combinado” es el que aparece en la gran mayoría de los textos, no sólo de econometría, sino de estadística. Las pruebas de hipótesis aparecen como uno de los métodos de inferencia estadística más utilizados hasta el momento.

Entonces, y siguiendo el razonamiento de Spanos, algunas hipótesis en el terreno econométrico deberán probarse desde la perspectiva de Fisher, como las pruebas de especificación incorrecta del modelo, en las que se prueba lo adecuado de las suposiciones probabilísticas en el marco de las observaciones.

Es decir, el modelo es:

$$X_k = E(X_k) + \epsilon_k \quad k = 1, \dots, N$$

El modelo probabilístico propuesto es:

$$f(x; \theta) = \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2} \right] \quad , x \in R$$

$$\mu = E(X_k) \quad \sigma^2 = \text{Var}(X_k)$$

los supuestos probabilísticos a probar son:

<sup>14</sup> G. Gigerenzer. *The empire of chance: How probability changed science and everyday life*: Cambridge University Press, Cambridge, 1987, en Spanos, 1995, *ibidem*, p. 727.

$$X_k \sim N(\mu, \sigma^2) \quad k = 1, 2, \dots$$

$(X_1, X_2, \dots, X_n)$  es una muestra independiente<sup>15</sup>  
 $(X_1, X_2, \dots, X_n)$  es una muestra idénticamente distribuida

### Los diferentes tipos de pruebas

Son diferentes los propósitos que nos llevan a utilizar la metodología de pruebas de hipótesis en una investigación econométrica. De acuerdo con Keuzenkamp y Magnus,<sup>16</sup> se propone una clasificación de las diversas formas en que son utilizadas las pruebas de hipótesis y las de uso más frecuente en investigación. Y aunque esta clasificación no es única ni exhaustiva pueden distinguirse cuatro tipos principales:

#### *Las pruebas acerca de las teorías económicas*

En algunas ocasiones, el interés del investigador radica en la confrontación de alguna especificación particular de la teoría con los hechos, que en no pocas ocasiones representa un problema, principalmente por el carácter no experimental de los datos económicos. Pero también, porque puede no conocerse la teoría. Ya Tinbergen menciona las dificultades que había encontrado para confrontar las teorías disponibles acerca de los ciclos económicos y decidir cuáles eran los correctos.

Esto fue percibido claramente por algunos autores que realizaron críticas al trabajo econométrico. Koopmans,<sup>17</sup> que en aquel momento era un investigador distinguido de la Comisión Cowles, en un artículo titulado “Medición sin teoría” (Measurement without theory. Debate), critica a los economistas Burns y Mitchell, quienes trataron de medir los ciclos económicos sin contar con la teoría de cómo funcionan éstos.

<sup>15</sup> Spanos, capítulo 15, *ibidem*, 1995.

<sup>16</sup> J. Keuzenkamp, y R. Magnus. “The significance of testing in econometrics”, en *Journal of Econometrics*, vol. 67, núm. 1, 1995, pp. 5-24.

<sup>17</sup> Este artículo aparece reproducido en D. Hendry y M. Morgan. *The foundations of econometric analysis*: Cambridge University Press, Cambridge, 1995, p. 43.

Con el trabajo de diversos autores, entre los que sobresalen Tinbergen y Haavelmo, estos problemas fueron identificados y resueltos. Tinbergen, por un lado, a partir de su investigación acerca de los ciclos económicos, publicada en dos volúmenes y titulada *Pruebas estadísticas de las teorías de los ciclos económicos*, plantea que las teorías deben establecerse como relaciones capaces de ser confrontadas estadísticamente, y Haavelmo, por el otro, inicia lo que se denominó “estrategia tradicional”, en donde los teóricos proporcionan los modelos y los econométricos los estiman y prueban. Esta forma de trabajo ha resultado exitosa para cierto tipo de problemas en los que la teoría puede expresarse en forma de relaciones lineales, o no lineales, entre las variables.

En este entorno, la herramienta más utilizada es el modelo de regresión y las pruebas se limitan a probar la significancia de los coeficientes de regresión y a verificar que sus signos sean los correctos. El modelo de regresión permite, también, considerar formas funcionales distintas, usar variables rezagadas que ajustan el proceso, variables aproximadas (*proxies*) para aquellas no observables, etcétera.

A partir de los años setenta esta estrategia ha tenido diversas críticas. En primer lugar, había evidencia creciente de que los modelos no representaban a los datos, y en segundo, porque los economistas teóricos insistían en que los modelos no representaban a la teoría. Además, los tomadores de decisiones argüían que los modelos eran poco efectivos para pronosticar o realizar un análisis político.<sup>18</sup>

Las respuestas a estas críticas fueron de diversa índole, algunas tan extremas como priorizar la representación de los datos restándole importancia a la teoría. Esta propuesta tuvo como resultado los modelos de vectores autorregresivos en donde:  $y_t = [y_{1t}, y_{2t}, \dots, y_{kt}]'$  es un vector de  $k$  variables diferentes que se expresa en valores pasados del mismo vector. El resultado es un vector autorregresivo o VAR. El proceso VAR( $p$ ) es

$$y_t = m + A_1 y_{t-1} + A_2 y_{t-2} + \dots + A_p y_t + \epsilon_t$$

con  $A_j$  matrices de coeficientes  $k \times k$ ,  $m$  es un vector  $k \times 1$  de constantes y  $\epsilon_t$  es un vector de ruido blanco con las siguientes propiedades:

<sup>18</sup> M. H. Pesaran y R. Smith. “The role of theory in econometrics”, en *Journal of Econometrics*, 1995.

$$E\left(\begin{matrix} \\ \\ \\ \end{matrix} \right) = 0 \text{ para toda } t; \quad E\left(\begin{matrix} & & s = t \\ & & \\ 0 & s & t \end{matrix} \right) =$$

y es la matriz de covarianza que se supone positiva definida.<sup>19</sup>

Otros modelos que resultaron de la crítica a la estrategia tradicional fueron aquellos que introducían un “mecanismo de corrección de errores” con el que intentan reparar algún desequilibrio de largo plazo.

Estos dos procedimientos se combinaron para obtener el análisis de *cointegración*, el cual sostiene que, en economía, la gran mayoría de las series de tiempo están sujetas a determinada tendencia, es decir, son no estacionarias. Aquellas series de observaciones que aunque sean no estacionarias puedan representarse como una combinación lineal en una sola serie estacionaria se denominan *series cointegradas*.<sup>20</sup> Otros procedimientos han probado ser más o menos exitosos, como el uso de procedimientos de optimización.

### *Las pruebas acerca de la validez del modelo*

Este tipo de pruebas son muy utilizadas y sirven para conocer si el modelo está bien especificado. Sea:

$$y = X\beta + u$$

un modelo de regresión lineal.

La especificación de un modelo lineal está centrado alrededor del vector de perturbaciones  $u$  y de la matriz de observaciones  $X$ . Los que se supone cumplen con lo siguiente:

$$\begin{aligned} &u_i \text{ son independientes e idénticamente distribuidos } N(0, \sigma^2) \quad i = 1, \dots, n \\ &E(X_i' u) = 0 \text{ para toda } i = 1, \dots, k \text{ y } s = 1, \dots, n \\ &X \text{ es no estocástica de rango completo } k. \end{aligned}$$

Entre las distintas posibilidades que existen para violar estos supuestos se men-

<sup>19</sup> J. Johnston y J. Dinardo. *Econometric methods*, 4a. ed.: McGraw-Hill, 1997.

<sup>20</sup> W. Charemza y D. F. Dreadman. *New directions in econometric practice*, 2a. ed.: Edward Elgar, EUA.

cionan las más comunes, aunque debe indicarse que este tema es mucho más complejo de lo que aquí se presenta.

Problemas con los errores:

- 1) Si los errores son independientes e idénticamente distribuidos, pero no tienen una distribución normal, los procedimientos inferenciales son válidos sólo en forma asintótica.
- 2) La presencia de errores autocorrelacionados

$$E(u_t u_{t-s}) = 0, \quad s \neq 0$$

que se verifica con las diferentes versiones de la prueba de Durbin-Watson, la prueba Breusch-Godfrey y la estadística Box-Pierce-Ljung, entre otras.

Cuando esto sucede, es evidencia de alguna mala especificación; por ejemplo, que algunas variables relevantes han sido omitidas del modelo de regresión, pero puede indicar también que el modelo es autorregresivo de algún orden, de rezagos distribuidos, con factores comunes.

- 3) La matriz de varianzas covarianzas del vector  $u$  presenta varianzas diferentes en la diagonal principal. Es decir:

$$E(uu) = \text{diag} \left[ \begin{matrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \dots & \\ & & & \sigma_n^2 \end{matrix} \right]$$

Cuando esto sucede se presenta un problema de heteroscedasticidad que es más frecuente en estudios de sección cruzada. Es muy importante probar que este supuesto se cumple ya que, si no es así, los estimadores de mínimos cuadrados dejan de ser eficientes. Entre las pruebas de diagnóstico posibles están la prueba de White, las distintas versiones de la Goldfeld-Quandt y otras.

En un artículo más o menos reciente, Engle<sup>21</sup> planteó la posibilidad de que este

<sup>21</sup> R. Engle. "Autorregresive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation", en *Econometría*, vol. 50, núm. 1, 1982.

problema se presentara también en el contexto de las series de tiempo y particularmente en el pronóstico de series de tipo financiero, como precios de acciones, tasas de inflación, etc. Engle desarrolló el modelo autorregresivo de heteroscedasticidad condicional (ARCH), en donde la idea central es que la varianza del error en el tiempo  $t$ , depende del tamaño del error al cuadrado en el tiempo  $t-1$ , apoyando la noción de que el pasado reciente puede ofrecer información acerca de la varianza condicional del error:

$$\sigma_t^2 = \omega + \alpha_1 u_{t-1}^2 + \dots + \alpha_p u_{t-p}^2$$

donde la impresión es que la varianza está correlacionada.

Una varianza como la anterior puede provenir de una perturbación definida por:

$$u_t = \epsilon_t \left[ \omega + \alpha_1 u_{t-1}^2 + \dots + \alpha_p u_{t-p}^2 \right]^{1/2}$$

donde  $\{\epsilon_t\}$  es un proceso de ruido blanco con varianza uno. Este modelo se denomina ARCH( $p$ ).

### *Las pruebas de diagnóstico del modelo*

En el enfoque tradicional se considera que en un análisis econométrico se deben estudiar los siguientes aspectos:

- a) La teoría económica.
- b) Los datos estadísticos.
- c) La formulación del modelo sobre la base de la teoría o de investigaciones econométricas recientes.
- d) El cálculo de los parámetros del modelo usando información estadística relevante.
- e) La inspección de los resultados estadísticos obtenidos para juzgar cuán adecuado es el modelo.

En este contexto, la inspección consiste en medir la bondad del ajuste del modelo, que en forma general se hace a través del coeficiente de determinación  $R^2$ : verificar que los estimadores de los coeficientes de regresión sean significativos, con

los signos esperados *a priori*, y que el modelo tenga poder de predicción, *dentro* y *fuera* de la muestra dada. Entre las pruebas que existen para detectar errores de especificación están la prueba de Durbin, la RESET de Ramsey y la del multiplicador de Lagrange.

Si después de aplicar la batería de pruebas el modelo se considera satisfactorio, entonces se tiene en cuenta y se usa en las predicciones. Pero si el modelo se considera insatisfactorio, entonces el investigador debe encontrar la mejor especificación, probando quizá múltiples formulaciones con docenas de corridas y pruebas en la computadora, en un afán por dar con la ecuación "satisfactoria". Este proceso de búsqueda se ha denominado, peyorativamente, *data-mining*.

Frente a esta posición existen opiniones encontradas. Algunos autores, como Darnell,<sup>22</sup> consideran que si se modifica un poco la metodología tradicional, puede continuar siendo la metodología preferida por los investigadores.

La otra posición es la que desarrolla Hendry y es conocida como metodología de lo *general a lo particular*; en ella se parte de un modelo general y con la ayuda de una batería de pruebas de diagnóstico se reduce a un modelo particular. No se piense que esta búsqueda es automática: el investigador requiere tener una idea clara de la forma específica del modelo, que es el hilo conductor de la investigación. Si no es así, modelar de lo general a lo particular puede convertirse en un método de simplificación de modelos, es decir, "un método de descubrimiento más que de confirmación".<sup>23</sup>

### *Confirmación de la prueba*

Finalmente, se menciona que un objetivo de una prueba es realizar alguna toma de decisiones, como en los procesos de control de calidad, pero también con fines políticos, en los que se puede considerar la variable independiente como variable de control y la dependiente como variable objetivo.<sup>24</sup>

Si el modelo es capaz de resistir el cúmulo de pruebas y salir bien librado, podemos decir que la hipótesis o la teoría en consideración se ha confirmado, y entonces puede utilizarse para pronóstico o con fines de control o de política.

<sup>22</sup> A. Darnell y L. Evans. *The limits of econometrics*: Edward Elgar Publishing, 1990.

<sup>23</sup> W. Charemza y D. Dreadman. *New directions in econometric practice*, 2a. ed.: Edward Elgar Publishing, EUA, 1993, p. 77.

<sup>24</sup> D. Gujarati. *Econometría*, 3a ed.: McGraw-Hill, México, 1997.

### ¿Por qué entonces el escepticismo de algunos autores?

Seguramente este escepticismo no es exclusivo de la práctica econométrica y es casi seguro que podremos encontrar algunas críticas a esta metodología en todas las disciplinas en donde se lleven a cabo pruebas estadísticas. Las críticas van desde lo que pareciera muy trivial, como “¿cuál es el sentido de rechazar una hipótesis que se sabe que es falsa desde el comienzo?”,<sup>25</sup> hasta el argumento de Spanos, que cuestiona la relevancia misma de la argumentación estadística: “Ninguna teoría económica ha sido descartada aunque no demostrara su validez bajo la evidencia empírica, ni existió una clara distinción entre teorías en competencia, porque la evidencia empírica era inadecuada para dicho propósito”.<sup>26</sup>

Aun cuando podemos aceptar el hecho de que cierto grado de *data-mining* es inevitable, ya que refleja el problema de la imposibilidad de conducir experimentos controlados,<sup>27</sup> en algunas investigaciones este proceso es llevado a extremos: “corriendo” un gran número de regresiones con el mismo conjunto de datos para seleccionar la “mejor” de acuerdo con el coeficiente de determinación más alto, sin mostrar el proceso con el que se alcanzaron los resultados publicados.

McAleer,<sup>28</sup> en un llamativo artículo, revisa la bibliografía econométrica en diversas revistas especializadas, considerando los modelos no anidados, que son muy utilizados en investigación, y evalúa, en el contexto de modelos de regresión lineal, log lineales y no lineales, la importancia que tiene probar este tipo de modelos. Bajo el subtítulo de deficiencias en publicaciones empíricas encuentra, entre otras cosas:

- 1) Que existe una descripción insuficiente de los datos utilizados, especialmente si son cuatrimestrales, ajustados o no para estacionalidad.
- 2) Una falta sorprendente de pruebas de diagnóstico.
- 3) Inferencias equivocadas que se obtenían cuando los investigadores no se percataban

---

<sup>25</sup> G. R. Loftus. “On the tyranny of hypothesis testing in the social sciences”, en *Contemporary Psychology*, vol. 36, núm. 2, 1991, p. 102. En este artículo se hace una crítica, en general, al uso de las pruebas de hipótesis y, en particular, a los conceptos usados en éstas.

<sup>26</sup> Spanos. “On theory tTesting in econometrics modelling with non experimental data”, en *Journal of Econometrics*, vol. 67, 1995, p. 190.

<sup>27</sup> Charemza *et al.*

<sup>28</sup> M. McAleer. “The significance of testing empirical non-nested models”, en *Journal of Econometrics*, vol. 67, 1995, pp. 149-171.

del tipo de pruebas no anidadas que estaban considerando, y una mala interpretación de los resultados en las pruebas estadísticas (como decir que un valor que resultaba estadísticamente significativo se interpretaba como “insignificante”).

- 4) No considerar las propiedades de muestra finitas para algunas pruebas. (Encontró doce artículos en los que la muestra era menor a 25, y dos en que las muestras eran de 15 y 16 observaciones.)
- 5) El uso de modelos inadecuados que no han pasado algunas pruebas, como las de correlación serial.

Por último, hace notar que los temas de no estacionalidad y cointegración no se han examinado seriamente en la investigación de modelos no anidados.

La utilización de la metodología de las pruebas de hipótesis ha probado su importancia en diversas formas: es el puente que enlaza la teoría con las observaciones en las ciencias aplicadas, en particular en la economía, donde el análisis cuantitativo de los fenómenos económicos ha tenido un gran desarrollo en los últimos cincuenta años. Para su utilización ha sido necesario sortear algunos problemas: como el que en ocasiones no existieran mediciones para las variables que se consideraban, y como el hecho de que los datos en economía son de carácter no experimental —este problema tuvo que esperar a ser resuelto por Haavelmo, cuando introdujo la noción de probabilidad en la práctica econométrica.

Sin embargo, aun cuando cada vez es mayor la utilización de los modelos de regresión en la investigación económica y, en consecuencia, el uso de las pruebas de hipótesis, sigue presente el escepticismo en algunos autores. Esto se debe a diversos factores, entre los que podemos mencionar: la descripción insuficiente de los datos que se utilizan, si están —o no— ajustados para estacionalidad, o si han sido tomados en forma trimestral o cuatrimestral, etc.; la falta de pruebas de diagnóstico; obtener inferencias equivocadas y otros.

A pesar de los problemas que pueden presentarse, las pruebas de hipótesis son, hasta el momento, uno de los métodos más utilizados en la inferencia estadística, no sólo en disciplinas de las ciencias más cuantitativas, como la física o la biología, sino también en las sociales, principalmente en la economía.

Así, podemos estar de acuerdo con el economista Hendry,<sup>29</sup> de que en econometría, para pasar de la alquimia a la ciencia, debemos *probar, probar y probar*:

<sup>29</sup> D. Hendry. “Econometrics: alchemy or science?”, *ibidem*, p. 27.