

La ortodoxia metodológica en las Ciencias Sociales

[Revista del Centro de Estudios Educativos (México), vol. V, núm. 4, 1975, pp. 118-128]

Cláudio de Moura Castro,

Programa ECIEL, Río de Janeiro (Brasil)*

Nadie pone en duda que en la ciencia se deba proceder de acuerdo con reglas, cánones y reglamentos. Los que comienzan pueden proveerse de este conjunto de reglas por varios caminos.

La palabra "metodología" se ha venido usando para designar una gama excesivamente amplia de acepciones. Por un lado, vemos que el término se emplea en conexión con los problemas de la filosofía de la ciencia y de la epistemología. En tal caso se refiere a los problemas relacionados con la validación epistemológica del conocimiento científico. Se discuten cuestiones científicas que no pueden ser resueltas por la observación de los hechos, es decir, por la evidencia empírica; por ejemplo, problemas de la naturaleza misma de la explicación científica, causalidad, estructura y relación entre proposiciones científicas, etc. Por otro lado, se usa como sinónimo de lo que podría mejor llamarse "métodos y técnicas". Caen en esta categoría la técnica de preparación de cuestionarios, de entrevistas y en ocasiones la misma estadística.

En este trabajo tomamos el término en su acepción más usual, equidistante de los dos extremos. "Metodología" se refiere a cosas tales como pruebas de hipótesis, estructura formal del proceso científico, descripción de secuencias en la investigación, etcétera.

Quisiera llamar la atención hacia una falsa ortodoxia metodológica, muy común en los recién llegados a las ciencias sociales. Parece provenir de una manera peculiar de aprender las reglas del juego de la ciencia.

Cierto alumno de "Agassiz" cuenta que al iniciar sus estudios de ciencias naturales, éste le entregó un pescado muerto para que lo estudiara y lo describiera, pero con la consigna de que no leyese nada, ni conversase con persona alguna sobre el asunto. Después de un mes de trabajo se le entregó, no ya un pescado, sino un montón de pedazos de huesos. Después de examinar los pedazos, pudo reconstruir los esqueletos, así como comparar y clasificar las especies de acuerdo con las semejanzas encontradas. Una vez terminada esta fase, recibió de Agassiz algunos ejemplares de ciertos tipos de pescado. Esta vez le fue permitido al alumno, después de varios meses de estudio y ante un mutismo sepulcral de parte de Agassiz, poder leer y consultar con sus compañeros. A esta altura fue ya capaz de notar una falla importante en el esquema que Agassiz le había propuesto para la clasificación de aquella especie. Demostró con ello que había dominado los problemas básicos de la metodología biológica. Esto fue posible gracias a un entrenamiento en que a fuerza de observaciones propias el alumno redescubría los principios generales de la clasificación zoológica.

* Traducción del portugués, a cargo e A. Hernández Medina del CEE.

Lo anterior es una parábola histórica más bien que una descripción de las prácticas contemporáneas del entrenamiento en las ciencias naturales. Un estudiante, digamos, de física, entra en paciente contacto con los protocolos de su ciencia, es decir, con las observaciones y fenómenos del mundo de la física y estudia cómo los científicos trataron de describir tales hechos, generalizando y concretando ciertas regularidades. Dentro de algún tiempo, él también se aventurará primero a empresas modestas, pero irá adquiriendo progresivamente las destrezas de sus compañeros de profesión.

Comparemos esto con el noviciado de un estudiante de educación, por ejemplo. La investigación educativa es una ciencia; esto lo afirmamos con la arrogancia propia de quien afirma algo de lo que no se siente muy seguro. Esta ciencia tiene su metodología que es menester estudiar sistemáticamente. De ahí la necesidad de los cursos de metodología como algo indispensable para cualquier programa de cursos de tal naturaleza.

El químico se inicia en su oficio sin discutir problemas profundos de epistemología y sin memorizar los catecismos de las pruebas de hipótesis y de los "proyectos de investigación"; del mismo modo que un jugador gana el juego de billar sin haber antes oído hablar de Newton o del principio de conservación. Por el contrario, quien va a investigar en áreas de las ciencias sociales tiene que someterse al rito de la "iniciación metodológica".

Quisiéramos señalar algunas consecuencias poco saludables, aunque evitables, de este prefacio metodológico. Pero antes es pertinente que mostremos cómo el grado de preocupación por el método tiene mucho que ver con lo que está pasando en la ciencia.

Alguien ha sugerido que las teorías científicas presentan cíclicamente fases infantiles, maduras y seniles (Margenan, 1961: Cap. II). En la fase infantil, los científicos se enfrentan con los problemas de procedimientos, del establecimiento de paradigmas (Kuhn, 1962: Cap. II); en una palabra, con los problemas del método. En la madurez, se piensa que tales problemas han quedado sustancialmente resueltos y se avanza a cuestiones sustanciales. La fase senil comienza a perfilarse a medida que los hechos y eventos empiezan a acumularse y no es posible explicarlos según los esquemas ya aceptados. A veces tales paradigmas son rechazados y se propone entonces sustituirlos por otros, aunque sean todavía insatisfactorios. Está puesto entonces el escenario para lo que Thomas Kuhn (1962: Caps. I y II) llamó la Revolución Científica, que está marcada por un renacimiento por las preocupaciones metodológicas. Algo semejante existe en el arte: las llamadas escuelas o movimientos surgen con un pavoroso consumo de energía en los problemas de forma y búsqueda de soluciones. Viene luego un período de madurez, en que los artistas se sienten satisfechos con el lenguaje y se entregan al trabajo. Se piensa que en ese lapso crean sus mejores producciones.

Las ciencias sociales, comparadas con las físicas y matemáticas, son imberbes todavía. Y el rezago no se debe a que sus científicos sean de menor altura. Lo que pasa es que sus objetivos son de más complejidad y más huidizos; además, tienen que vérselas con dimensiones éticas e ideológicas. Con todo, dentro de las ciencias sociales existen diferencias importantes en los grados de madurez. La economía y, en menor grado, la psicología, están más provistas que la sociología, la ciencia política y la investigación educacional.¹ Por ello observamos que los economistas,

¹ La expresión "investigación educacional" la entendemos aquí como la dimensión puramente científica de la educación.

Dado el carácter cíclico del pensamiento científico, nada hay que nos garantice que nuestra afirmación será válida también en el futuro. De hecho, un grupo de "economistas radicales" de los Estados Unidos afirma que todo el aparato convencional de los economistas es irrelevante y engañoso. Dicho grupo vaticina una "revolución" en esta disciplina.

por ejemplo, se preocupan menos por las cuestiones de método. Por ello también se dan en economía menos cursos metodológicos que en otras ciencias sociales.²

A continuación presentaremos algunos ejemplos del uso espurio de “metodología”. Nuestros comentarios no deben entenderse como una toma de posición en contra de las reglas del método científico, sino solamente contra su abuso, consciente o ingenuo.

1. EI RADICALISMO DE LAS MUESTRAS ALEATORIAS

Con mucha frecuencia no es posible examinar toda la población o universo cuyos atributos estamos tratando de analizar. No podemos medir la inteligencia de toda la población del Brasil; no es posible probar la calidad de cada saco de café exportado del país; como tampoco lo es el aplicar a todos y cada uno de los estudiantes del Estado de Guanabara un cuestionario sobre su estatus socioeconómico. Se recurre entonces al arbitrio de examinar una muestra, es decir, una fracción de la población. Ya en este siglo, se demostró formalmente que si tomamos una muestra por un proceso aleatorio (en que la probabilidad de ser escogido sea la misma para cada individuo), la muestra tendrá la mayor probabilidad posible de reproducir las características de la población. Más aún, hay tablas que nos permiten cuantificar la magnitud probable del error cometido.

Dado que el tomar una muestra aleatoria implica un simple procedimiento de sortear los individuos que van a ser escogidos, no es de extrañar que las técnicas y la práctica de la muestra aleatoria se hayan difundido tanto.

Ningún radicalismo en esta área parecería justificado. Sin embargo, existe tal radicalismo y lo podemos ver nítidamente en una discusión entre alumnos sobre sus proyectos de investigación. Ante un problema cuya verificación supone un sondeo con aplicación de cuestionarios, es frecuente oírles decir que tal sondeo no es viable debido a la imposibilidad de hacer una muestra aleatoria. Por ejemplo, al tratarse de un estudio de seguimiento sobre los alumnos de una escuela, no sería posible escoger aleatoriamente a los alumnos que van a ser entrevistados, en virtud de la dificultad práctica de localizarlos. En tales circunstancias, el tema debería ser abandonado.

Adam Smith, el más famoso de los economistas, construyó su teoría de la división de trabajo y de la especialización de las funciones basándose en una fábrica de alfileres en Escocia. Marx estudió la plusvalía en una fábrica de tejidos de un amigo de Engels. Marshall, otro de los grandes del pensamiento económico, escribió basado en la observación casual del comportamiento de sus colegas, profesores de Cambridge. La obra de Freud se apoya en gran parte en sus experiencias clínicas de Viena, con personas que acudían a él con disturbios emocionales.

No existe criterio metodológico que nos dé razones terminantes para el uso de muestras aleatorias. Todo lo que la ciencia nos dice es que echemos mano de toda la evidencia disponible (Principio de evidencia total), y que la manejemos con rigor y eficiencia. No hay frontera entre lo científicamente válido y lo no válido que sea aplicable ahistóricamente, interdisciplinariamente o aun de una situación a

² Solamente referimos un hecho, decimos que está sucediendo. Por lo demás pensamos que los economistas mucho ganarían con tales cursos, con tal que no se redujeran a menos ejercicios de memorización de taxonomía o a ser vehículos disfrazados de indoctrinación ideológica.

otra. Los cálculos astronómicos que un día hicieron Newton y Galileo serían hoy del todo inaceptables aun para el más bisoño estudiante de física. Los criterios de rigor definicional que exige la física serían desastrosos para casi todo lo que se hace en las ciencias sociales.

Lo que sí es inaceptable científicamente es usar un procedimiento menos riguroso o menos eficiente del que podríamos usar en ese momento histórico o en esa situación. Respetadas las fronteras de la ciencia, que sólo incluyen aquella porción de la realidad que puede ser empíricamente examinada y aquellas proposiciones cuya validación no depende de juicios de valor o presupuestos metafísicos, cualquier problema puede ser tratado o estudiado. Lo que nos garantiza que el tratamiento es científico no es ésta o aquella regla rígida e inmutable —como la necesidad de usar muestras aleatorias—, sino el principio general de que debemos adoptar el tratamiento más riguroso que se tenga a la mano y que nuestro procedimiento sea eficiente.

Sería inadmisibles aplicar cuestionarios dentro de una escuela a una muestra que no hubiera sido obtenida aleatoriamente, ya que dicha obtención no sólo es viable sino fácil. Pero si no es posible obtener una muestra aleatoria de exalumnos, la muestra que fuere posible obtener será justificable.

La muestra aleatoria nos permite hacer inferencias respecto a la población, que no son posibles en otros tipos de muestra. Si una tercera parte de la muestra aleatoria la forman alumnos que vienen de otras comunidades, podemos afirmar, con una probabilidad de errar que es también calculable, que nuestra población contiene de 32.3% a 34.3% de inmigrantes. Tal afirmación no puede formularse con base en una muestra no-aleatoria. Es posible, de hecho, que muchos de los estudiantes hayan retornado nuevamente a su lugar de origen. Entonces, si encontramos que hay un tercio de alumnos foráneos, es probable que la población de exalumnos tenga una proporción mucho mayor de migrantes ya que los que regresaron no pudieron ser localizados.

Con todo, la muestra no-aleatoria nos puede proporcionar informaciones extremadamente útiles y valiosas. Gran parte de nuestra herencia científica tiene su base en ese tipo de muestras: pero no podemos adjudicarle las propiedades de las aleatorias.³

Siempre que usemos una muestra que no sea aleatoria, se hace necesario considerar el sesgo en que se está incurriendo. Además, el sesgo no tiene la misma importancia en todas las variables contenidas en un mismo cuestionario o entrevista.

En una investigación que hicimos entre los exalumnos de economía, la muestra consistió en intentar entrevistar a cuantos economistas que pudieran ser localizados (Moura Castro, 1970). La proporción real de los economistas que impartían clases en las facultades era ciertamente muy diferente de la proporción encontrada. La aplicación de la muestra estuvo a cargo de profesores de la propia facultad, quienes era obvio que con más facilidad localizarían a sus colegas del magisterio, dejando fuera de la muestra a muchos economistas no dedicados al magisterio. Con todo, no hay razón para creer que la preferencia por textos americanos o franceses se haya visto afectada fundamentalmente por el sesgo de la muestra.

³ Como en el ejemplo citado, cuando sabemos la dirección del sesgo o distorsión, podemos decir simplemente que "más de una tercera parte" es de fuera. En otros casos, los resultados podrían llegar a ser escasamente convincentes: si encontramos un tercio, no es de suponer que la proporción de inmigrantes sea de un sexto o de dos tercios. Lo que la muestra no-aleatoria nos da es un orden de tamaño, lo que, en ausencia de datos mejores, puede proporcionarnos una información útil, si la usamos con la debida cautela.

La importancia de la muestra aleatoria depende en buena parte del tipo de información que queremos obtener de los datos. Podemos distinguir dos situaciones diversas. Si en nuestra investigación lo que pretendemos estimar con la muestra son parámetros de la población, entonces la importancia de la muestra es enorme. En tales muestras las diferencias entre los parámetros de la muestra y de la población se deben exclusivamente al azar, es decir, a la fluctuación de la muestra. La muestra no-aleatoria produce distorsiones no conmensurables y sí incontrolables en la mayor parte de los casos. En el ejemplo citado, el número de profesores entrevistados dice relativamente poco respecto a la proporción de profesores en la población.

Con todo, la muestra aleatoria es menos importante cuando se trata de verificar las relaciones entre las variables. Este tipo de problema es mucho menos sensible a la no-aleatoriedad de la muestra. Las cuestiones más importantes de la ciencia tienen más que ver con relaciones causa-efecto o interacciones que con la medida de parámetros. En los ejemplos que hemos citado, Adam Smith, Marx y Marshall pretendían entender la regularidad del comportamiento y la asociación entre las variables; por ello su trabajo dejaba a un lado los rigores de la muestra aleatoria. Volviendo al ejemplo anterior, a pesar de que la muestra no era aleatoria, podríamos con toda legitimidad haber examinado la relación, por ejemplo, entre el salario de los economistas y su edad, la relación entre empleador y origen socioeconómico, y así otros casos.

2. LA ORTODOXIA DE LA PRUEBA DE HIPÓTESIS

Hace algún tiempo, charlaba yo sobre una investigación en que se intentaría asociar el mal comportamiento (indisciplina) con el desempeño académico y el estatus socioeconómico. Y se preguntaba qué hipótesis iba a ser probada. La respuesta de que sencillamente no existía hipótesis a probar, sonó como herejía a los oídos de alguno. ¿Cómo era posible hacer una investigación sin tener una hipótesis?

Cuando contemplamos la posibilidad de que dos variables estén funcionalmente asociadas, podemos aventurar algunas presunciones sobre la presencia y aun sobre la naturaleza de la relación. Si previamente al examen de los datos empíricos en cuestión creemos en la existencia de alguna relación específica (tal como: "Si X aumenta, Y disminuye"), decimos entonces que esa proposición constituye la hipótesis que vamos a verificar.⁴ Esta creencia en determinado resultado puede estar basada en otros datos empíricos que conocemos, en ciertas consideraciones teóricas o en lo que sea. Por eso, la epistemología de la ciencia nos exige que formulemos la hipótesis antes del examen de los datos que se van a utilizar en la prueba.⁵

Frecuentemente acontece que encontramos un problema interesante que nos sorprende totalmente impreparados para "apostar" por determinado resultado. Esto no debe impedir que exploremos el problema empíricamente, si bien en este caso haya que proceder con mucha más cautela, sobre todo en lo que se refiere a las relaciones espurias. La diferencia básica en este caso es que procedemos sin "hipótesis a verificar", y entonces la investigación misma va a generar dichas hipótesis. En lugar de comenzar la investigación empírica

⁴ Como el lector puede advertir, estamos usando el término "hipótesis por verificar" casi como sinónimo de "teoría por ser verificada", tal como lo hace Kaplan (1969: 93).

⁵ Esta precedencia en el tiempo debe ser tomada literalmente. Es inaceptable probar hipótesis que fueron establecidas a partir de un examen visual de los datos.

con una hipótesis verificada, comenzamos sin hipótesis y terminamos con una hipótesis no-verificada.

Estamos hablando en todo esto de “hipótesis a verificar”; no de hipótesis de trabajo. Estas últimas sirven para orientar y estructurar la investigación; “consisten en una idea no acerca del resultado de la investigación, sino acerca de los pasos que conviene dar enseguida. Dan una seguridad en cuanto al curso de la investigación” (Kaplan, 1969: 93).

El uso de computadoras electrónicas y la popularización de los programas de regresión “stepwise”⁶ y de aplicación iterativa de análisis de varianza, se han convertido en una tentación irresistible para ciertos investigadores. Basta introducir los datos en la computadora y ella se encarga de verificar todas las asociaciones posibles. La ausencia de hipótesis bien especificadas, aunada a la complejidad de las interacciones que se dan en las ciencias sociales, puede ser fatal para la ciencia contemporánea si el investigador se contenta con las asociaciones y correlaciones encontradas, en lugar de tomarlas sólo como punto de partida. “Las fallas ligadas al empleo de la estadística en las ciencias del comportamiento se deben atribuir, principalmente, a la tendencia a olvidar que las técnicas estadísticas son instrumentos y no sustitutos del pensamiento” (Kaplan, 1969: 263). Permitir que la computadora se convierta en una máquina de pensar, en sustitución de nuestro raciocinio e imaginación, constituye un procedimiento peligroso.

En casos más sencillos, la ausencia de “hipótesis a verificar” nos puede arrastrar a teorizar en torno de las variables que puedan eventualmente estar disponibles y que revelan, espuriamente, correlaciones significativas.⁷

Dejando a un lado posibles abusos, es importante enfatizar una vez más el hecho de que, para hacer que la ciencia avance, necesitamos de buenas preguntas, pero no necesariamente de una hipótesis segura con relación a las respuestas. Además, es menester que la hipótesis no sea meramente una corazonada. Si no tenemos razones para creer en algo, metodológicamente es indispensable que prosigamos sin hipótesis.

3. PRUEBAS ESTADÍSTICAS: ¿REGLAS DE DECISIÓN O ESCALAS DE CONFIANZA?

Los manuales de introducción a la estadística nos presentan las pruebas de hipótesis en su formulación más depurada, esto es, el “control de calidad”. Por lo demás, hay razones didácticas que nos sugieren que, siendo ese el caso más simple, debe ser el primer contacto con las reglas de decisión en la estadística. El que compra lámparas no desea recibir una partida de ellas con una vida inferior a la especificada. Y el productor no está dispuesto a ver rechazada una mercancía que, en realidad, está en buenas condiciones. Probando una muestra de esas lámparas es como hay que decidir si la entrega obedece o no a las especificaciones. La divergencia de intereses entre comprador y vendedor nos habla claramente

⁶Programas en que la máquina va experimentando cada una de las variables y va escogiendo, por sí sola, las que están asociadas significativamente con la variable dependiente.

⁷Estadísticas actuariales rusas de principios de siglo mostraron que el perjuicio causado por los incendios estaba positivamente correlacionado con el número de carros de bomberos que acudían al lugar; lo que nos podría llevar a concluir que los bomberos empeoraban, más que mejoraban, la situación. El ejemplo, no obstante ser auténtico, es grotesco; los grandes incendios que causan los mayores daños, movilizan más unidades de bomberos. Pero en situaciones menos obvias, la ausencia de hipótesis puede llevarnos a explicaciones totalmente ficticias.

de la necesidad de que se establezca, *ex ante*, una regla de decisión que tendrá validez hasta el fin, cueste lo que cueste.

Llevado al proceso científico, ese mismo esquema teórico de la prueba de hipótesis presenta perfiles menos nítidos. Trátase ahora de un investigador que está juzgando si acepta o no su propia hipótesis nula. El daño acarreado por una decisión errada es menos tangible. La tentación de abrir más un intervalo de confianza para aceptar una hipótesis marginalmente significativa es también muy grande. En todo caso, la teoría es clara: en primer lugar, tendremos la hipótesis a ser probada; en segundo, su grado de significación; finalmente, y sólo entonces, la prueba y la decisión.

Las dificultades mayores aparecen más adelante, en dos clases de situaciones:

a) existe toda una serie de medidas estadísticas en que simplemente no se requiere hipótesis previa; b) hay otros casos también en que podemos estar frente a una situación, semejante a la descrita en la sección anterior, donde no hay hipótesis a verificar.

Cuando practicamos un análisis factorial o de los componentes principales, nuestras ambiciones, en términos teóricos, son bastante modestas. Partimos de un conjunto de variables cuyo número nos parece excesivo o difícil de manejar. Nuestro objetivo se reducirá a descubrir cuáles son las variables más importantes o a derivar variables que sinteticen los conjuntos de información que serán luego desechados. No hay inferencias causales; se da simplemente un proceso algebraicamente complejo, pero epistemológicamente simple, de eliminación de variables redundantes.

En el otro caso mencionado, exploramos la posibilidad de asociación entre las variables, sin que tengamos una expectativa más nítida en cuanto a sus posibilidades. En esa situación, la configuración teórica de la "prueba de hipótesis" tórnase claramente inexplicable. No es posible probar una hipótesis que no existe. No obstante, es todavía muy conveniente la batería formal de las pruebas de hipótesis. Los valores de los parámetros que obedecen a una distribución conocida pueden ser usados para medir la probabilidad de que los resultados reflejen una asociación real entre las variables o meramente una fluctuación estocástica.

Es muy importante entender que se trata de algo fundamentalmente diferente de la prueba de hipótesis. La batería algebraica es exactamente la misma. Pero el uso es completamente diferente. En la prueba de hipótesis "apostamos" por la validez de una teoría y establecemos qué niveles de riesgo estamos dispuestos a aceptar. Desde ese momento, la computadora puede encargarse no solamente de la aritmética, sino también de la decisión: nuestra interferencia en esta fase no sólo está de más, sino que queda simplemente prohibida. Cuando no tenemos hipótesis, la comparación de nuestros parámetros con las distribuciones teóricas (puede ser una distribución normal o una *t* de *Student*) nos da una medida de la probabilidad de que los resultados se deban a fluctuaciones de la muestra y no a una asociación real entre las variables. En lugar de prueba tenemos una escala o medida de confianza en la relación encontrada.

Tal medida no pasa de ser una descripción estadística, como lo es la media, la varianza o el coeficiente de correlación.

Podemos ajustar una ecuación de regresión múltiple sin que tengamos de antemano un conjunto bien delineado de hipótesis. El parámetro *t* de *Student* nos permite entonces conocer las variables donde existe mayor confluencia en los datos, facilita la comparación del poder de explicación de cada una y nos da a conocer la probabilidad de que la asociación encontrada se deba al azar. Pero, advertimos: no estamos probando ninguna hipótesis, simplemente estamos midiendo probabilidades.

Sería injustificable en metodología ortodoxa querer aplicar la teoría nítida y perfecta de las pruebas de hipótesis a situaciones exploratorias o a métodos estadísticos que no pretenden probar hipótesis.

Finalmente, es necesario considerar que la prueba estadística no es ni el único ni necesariamente el mejor instrumento de validación. Explicación teórica y verificación empírica caminan casi paralelamente, y los progresos de una refuerzan la confianza en la otra. Se acepta que las pruebas de hipótesis validan la congruencia teoría-realidad en la misma medida en que los datos empíricos convalidan la teoría. Pero es también cierto que, en la medida en que encontremos una buena explicación teórica para un problema, los resultados estadísticos encontrados adquieren mayor confianza. En la fase inicial de la investigación, se consideran ambas direcciones. Si los datos no concuerdan con la teoría, puede suceder o que no sean datos relevantes o que el método de cuantificación sea burdo en demasía o que la muestra sea demasiado restringida. No hay manera de evitar el choque eventual entre la teoría y los hechos que la contradicen; pero, al fin de cuentas, los hechos son los que deciden. Si no se pueden conciliar y persisten las discrepancias, habrá que reformular la teoría o tendrá que desecharse. En realidad no es lo que sucede cuando intentamos, por nuestro propio esfuerzo, entender alguna cosa. Si un determinado experimento falla, es posible también que la relación entre X y Y que queremos probar se vea encubierta por la interferencia de una variable más potente, Z. O el nivel del "ruido" está demasiado alto y se hace necesario un tipo más "puro"⁸ de datos.

Otra situación común es la de la validación a través de la congruencia entre los resultados empíricos de muestras diferentes. Si se verifica una hipótesis en muestras diferentes y vemos que la naturaleza de los resultados es equivalente (misma señal, mismo formato de curvas, etc.), ese resultado puede ser tan importante como la significación de las pruebas. En una investigación deseamos probar la hipótesis de que las personas criadas en la zona rural obtienen rentas inferiores a las de origen urbano (dado el mismo nivel de educación y edad). En prácticamente todos los casos, las diferencias del rendimiento no eran estadísticamente significativas, a pesar de que en los once casos donde se hizo la prueba había una diferencia de rendimiento en la dirección prevista. Esto puede ser tan importante como validación de la teoría como lo es el criterio ortodoxo de la prueba de hipótesis. Recordemos que, con frecuencia, encontramos resultados estadísticamente significativos que confirman hipótesis diametralmente opuestas, al repetir los procedimientos en muestras sucesivas. Podríamos encontrar rentas significativamente más altas para un grupo educacional y más bajas para otro; pero eso no nos obliga en manera alguna a establecer una teoría para cada nivel educacional.

Estamos discutiendo procedimientos muy delicados que, a veces, llegan a ser conflictivos para quienes intentan establecer zonas de ortodoxia metodológica vedadas al desarrollo de las ciencias sociales. Pero si eso es el pan de cada día en la investigación, no hay por qué no discutirlo abiertamente. En realidad, ahí se da una de las pocas situaciones en que aceptamos una posición ortodoxa; no esconder a los lectores los criterios y procedimientos utilizados. Hay que decir con toda claridad si la prueba vino antes que la hipótesis, o si la teoría está apoyando la prueba. La relación entre la teoría y los hechos no es, en general, lo cristalina que desearíamos. Hay ciertas conclusiones que no podemos sacar sin dar al lector la posibilidad de valorar los criterios utilizados.

⁸ Por ejemplo, el efecto del sexo sobre el nivel de renta sólo puede ser empíricamente verificado después de controlar el efecto de la educación, la edad, etcétera.

4. HIPÓTESIS SIN PRUEBA

En la sección anterior intentamos demostrar que, si no había hipótesis, no podía haber naturalmente “prueba de hipótesis”, aun cuando la estructura formal de la inferencia estadística se pueda usar para explorar la relación entre variables. Examinaremos ahora una situación diametralmente opuesta: con frecuencia tenemos hipótesis que gustaríamos de probar, pero cuya prueba no siempre es posible. Hay dos casos importantes en que esto puede suceder.

Cuando trabajamos con una muestra y ciertas características en ella observadas las atribuimos a la población, corremos siempre el riesgo de que tales características hayan sido generadas por el proceso estocástico del muestreo. Mediante las pruebas de hipótesis, conocemos la probabilidad de estar imputando erradamente a la población características sólo observables en la muestra. Pero, según advierte Georgescu-Roegen (1969: 356-357), cuando, por el contrario, conocemos no una muestra sino toda la población, entonces las pruebas de hipótesis pierden su razón de ser.

Por ejemplo, si queremos saber si la proporción de mujeres en la población del estado de Guanabara difiere del 50%, podemos hacer una investigación por sondeo. Entonces se verifica la probabilidad de que la proporción encontrada difiere significativamente del 50%. Por lo pronto, si tenemos a mano el Censo Demográfico del Estado, éste nos muestra la proporción de hombres y mujeres existente en ese universo. En este caso, no ha habido prueba de hipótesis; grande o pequeña, la diferencia entre 50% y la proporción encontrada tienen que ser aceptadas como “significativas”.

Es mucha verdad que no siempre es clara la distinción entre lo que es la muestra y lo que es el universo. En el ejemplo citado, el estado de Guanabara es el universo. Pero podía, en otro contexto, ser definido como una muestra de los estados de la Federación. En tal caso, tendría sentido probar la hipótesis de si la proporción encontrada en Guanabara es representativa para el Brasil. Es de capital importancia delimitar bien el problema que estamos tratando de examinar; de otra forma se puede incurrir en crasos errores metodológicos.

Hay otra segunda situación en que las pruebas no son posibles. La probabilidad de errores aleatorios puede ser calculada cuando conocemos el formato analítico de la distribución considerada. Cuando la variable que estamos examinando tiene su distribución determinada por un número muy grande de factores, sin que ninguno de ellos predomine, podemos suponer que su distribución queda adecuadamente descrita por una curva normal. Si ese fuera el caso, la distribución de las medias de las muestras será también descrita por una normal. Las pruebas de hipótesis pertinentes, en ese caso, se basan en comparaciones con parámetros de la curva normal, que es ya conocida y estudiada.

Si, por el contrario, las variables que tratamos de estudiar no se distribuyen de acuerdo con la curva normal o de otras funciones conocidas, la cosa se torna muy complicada; pues la distribución de las medias de las muestras se aproxima asintóticamente a la normalidad sólo cuando el número de muestras se aproxima al infinito. Frecuentemente trabajamos con distribuciones asimétricas; como, por ejemplo, la distribución agregada de la renta o de la propiedad. A pesar de ello, es de esperar que con una distribución de renta muy desigual; la hipótesis de normalidad de la distribución de las medias de las muestras no sea suficientemente realista. Y como no conocemos las características de esta distribución, no podemos probar cosa alguna con seguridad.

Varios investigadores se han preocupado ya de este asunto, y han tratado de evaluar el nivel de imprecisión que traen consigo las distribuciones no

normales cuando son probadas por métodos que presuponen normalidad.⁹ Debemos proceder con cautela. Si la distribución no es conocida, las pruebas convencionales de hipótesis deberán ser rechazadas en principio, hasta que se demuestre empíricamente su validez. Obsérvase en este sentido cierta incuria de parte de gran número de investigadores que, por el ansia de someter a prueba alguna cosa, aplican sin detenerse pruebas normales a distribuciones desconocidas o que saben que son asimétricas, como las distribuciones de las rentas agregadas.

Las estadísticas no-paramétricas nos ofrecen con frecuencia alternativas seguras y convenientes para situaciones en que trabajamos con distribuciones “mal comportadas”. Sucede, con todo, que las pruebas no-paramétricas no son en general suficientemente conocidas. Además, son de aplicación más restringida y nos obligan muchas veces al uso de medidas estadísticas menos convenientes. La prueba de la mediana, por ejemplo, nos obliga a usar la mediana y no la media, lo cual puede ser inaceptable en el contexto de la investigación.

A fin de cuentas, ¿qué nos queda, en caso de que no podamos probar la hipótesis? La respuesta es, simplemente, que no se prueba la hipótesis. No hay mandamiento epistemológico alguno que nos constriña a probar hipótesis. Es parte de la ortodoxia metodológica que queremos subrayar en este ensayo: que existe el mito de querer probar cualquier hipótesis, cueste lo que costare. Nada podría ser más equivocado. La prueba de hipótesis no es condición *sine qua non* de la investigación científica y es una lástima que haya llegado a ser el símbolo de estatus para investigadores menores.

Posiblemente la obra más importante en economía del presente siglo fue escrita por Keynes, quien dejó también una contribución importantísima en su formulación de la teoría de las probabilidades. Sin embargo, es revelador que no haya hecho prueba de hipótesis en su Teoría General del Empleo, no obstante que ahí se proponen innumerables hipótesis osadas. La cosa era simple: no había lugar para pruebas de hipótesis con los datos que tenía a la mano.

La prueba de hipótesis es una manera formal y elegante de mostrar la confianza con que podemos recibir ciertas proposiciones. Si esa confianza puede ser medida y establecida, es injustificable omitir la prueba. Pero, cuando la naturaleza de los datos o del problema no nos permite verificar formalmente esa confianza, no es ningún desdoro, ni para la ciencia ni para el investigador, el confesarlo en el informe de su investigación.

REFERENCIAS

Georgescu-Roegen, Nicholas.

1969 “A Critique of Statistical in Relation to Social Phenomena”, en *Revue Internationale de Sociologie*, vol. V, no. 3.

Kaplan, A.

1969 *A conduta da pesquisa*. São Paulo: Editora da Universidade de São Paulo.

Kerlinger, F. N.

1964 *Foundations of Behavioural Research*. New York: Holt Rinchart and Winston.

⁹ Comentarios y bibliografía sobre este tema pueden encontrarse en Kerlinger (1964: 257-269).

Kuhn, T.

1962 *The Structure of the Scientific Revolution*. Chicago: Chicago University Press.

Margenau, H.

1961 *Open Vistas*. New York: Yale University Press.

Moura Castro, Claudio de

1970 "O que faz um economista?", en *Revista Brasileira de Economia*, no. 4.