



EL ANEXO A DE LA PRIMERA EDICIÓN DE *THE CIVIC CULTURE*

Rubén HERNÁNDEZ CID*

SUMARIO: I. *Introducción*. II. *Las ciencias sociales y la estadística, una relación complicada*. III. *Estadística y ciencias sociales en acción, el anexo A*. IV. *Comentarios de siete párrafos seleccionados del Apéndice A*. V. *Bibliografía*.

I. INTRODUCCIÓN

Las ciencias sociales y la estadística han tenido, desde siempre, una relación intensa, productiva y creciente que, sin embargo, con frecuencia ha sido motivo de sospechas para los más ortodoxos practicantes de cada uno de estos dominios de estudio. Hoy, a cincuenta años de la publicación de *The Civic Culture*, las dudas acerca de que el pensamiento estadístico pueda estar en el centro de una investigación social conviven aún con las reservas de que los estudios sociales pueden resultar excelentes oportunidades de investigación en estadística matemática. Así pues, en este trabajo se expondrá una reflexión acerca de la importancia del contenido del Anexo A de la primera edición de tal obra desde la perspectiva de la historia de la Estadística en la segunda mitad del siglo XX.

II. LAS CIENCIAS SOCIALES Y LA ESTADÍSTICA, UNA RELACIÓN COMPLICADA

Hasta hace apenas unos pocos años era relativamente frecuente que algunos libros de texto de estadística, sobre todo en el nivel universitario, incluyeran una nota histórica acerca de los orígenes de tal disciplina. Desde el muy célebre *Statistics Methods for research workers* de Fisher (1946), que incluye un breve esbozo con nombres tales como Bayes, Laplace, Gauss, Pearson y “Student”, hasta obras mucho más recientes como Peña, D. y Romo,

* Departamento de Estadística, ITAM (hernande@itam.mx).

J. (2003) que cita las mismas fuentes e incorpora los nombres —entre otros— del propio Fisher, Galton, Pearson (Karl y Egon), así como de Neyman. Este hecho que se puede evocar como un dato curioso, puede ser visto también como un elemento para que el lector pueda convencerse de que esa obra que tiene frente a sí trata, verdaderamente, de un conjunto de conocimientos que pueden serle de utilidad en su propio espacio de investigación. Los grandes nombres serían, en esta perspectiva, una suerte de argumento adicional frente a la, por fortuna cada vez menos frecuente, desconfianza de los futuros investigadores sociales hacia el pensamiento cuantitativo.

Si bien los nombres evocados en tales obras son, en efecto, los de muy notables académicos, llama la atención el que en tales notas históricas no se mencionen nombres tales como el de Durkheim. Este autor representa, sin duda, uno de los más importantes promotores del uso del “pensamiento estadístico” en la investigación social.¹ “Pensamiento estadístico” que no debe ser confundido con el concepto generalizado de que los métodos estadísticos son, en las Ciencias Sociales, algo como una caja de herramientas tal que, en las manos de un buen artesano-estadístico puede ayudar a “tratar” conjuntos de datos recabados de alguna forma más o menos metódica.

En este contexto, el aporte de *The Civic Culture* resulta importante en la cultura estadística. El proyecto de esta investigación conjuga la Ciencia Política con la Estadística en un mismo fin donde cada vertiente tiene su sustento en la otra. Este estudio reposa, en última instancia, en un diseño que toma en cuenta los aspectos teóricos de la Ciencia Política sobre los que se sustentan las hipótesis estadísticas fundamentales que tendrán de ser probadas a partir de muestras independientes, con cuestionarios equivalentes y utilizando argumentos probabilísticos apropiados para detectar diferencias significativas entre los países analizados. Esto, que puede parecer en nuestros días la tarea rutinaria de un investigador, tuvo otra connotación para los profesores Almond y Verba, dado que en aquel entonces lo que hoy se llama la “inferencia estadística simultánea” estaba en pleno proceso de creación y era apenas objeto de discusiones en revistas especializadas.² Si los objetivos del trabajo, desde el punto de vista de la Ciencia Política, eran tan ambiciosos como complicados de lograr,

¹ En particular *Le Suicide* (1960) resulta fundamental en esta perspectiva. Si el aporte de esta obra en la historia de la sociología resultó capital, también es un hecho que su influencia en el uso y en la investigación de nuevos resultados en el análisis estadístico multivariado es relevante.

² El primer texto dedicado exclusivamente a este tema (*Simultaneous Statistical Inference* de Rupert Miller) aparece en 1996.

también era necesario justificar y preparar a los futuros lectores acerca de los resultados que estarían basados en razonamientos estadísticos que en aquel entonces serían de vanguardia.

Por otra parte, y con objeto de comprender mejor la talla del esfuerzo que fue necesario imaginar y realizar, conviene saber que hace cincuenta años el pensar, planear y, sobre todo, hacer análisis estadístico a partir de grandes bases de datos era prácticamente una labor heroica. Si bien los programas especializados en el análisis estadístico aparecen, de forma comercial, hacia 1970,³ la irrupción de los equipos personales de cómputo comienza, por su parte, hasta mediados de los años ochenta. Con estos datos puede resultar complicado para el investigador moderno imaginar siquiera la compleja labor que resultó el tratamiento de aquellas encuestas hace más de cincuenta años.

III. ESTADÍSTICA Y CIENCIAS SOCIALES EN ACCIÓN: EL ANEXO A

El Anexo A de *The Civic Culture* es, en esta perspectiva, una pieza fundamental para entender los alcances, pero también los posibles riesgos, de un trabajo como el emprendido por sus autores. De manera abreviada, se puede decir que el aporte fundamental de tal Anexo consiste en sensibilizar al lector acerca del alcance de las conclusiones para prevenir posibles errores en la interpretación de los mismos. No son abundantes las investigaciones que, basadas en muestras aleatorias, prevengan al lector acerca de los eventuales excesos en cuanto a las interpretaciones. Como un dato adicional es pertinente mencionar que este Anexo no se incluyó en las subsecuentes ediciones de esta obra. Sabemos que esta omisión atañe exclusivamente a los editores,⁴ pero es posible que, con los años, los avances de la inferencia estadística simultánea hayan estado al alcance de un número mayor de lectores y el sentido didáctico de ese texto haya desaparecido. El Anexo está dividido en dos partes claramente diferenciadas: la primera se refiere a los aspectos relevantes de las muestras obtenidas en cada uno de los cinco países y la segunda —mucho más breve— trata sobre el error muestral.

En lo que corresponde a la muestra, los autores detallan el esquema de muestra aleatoria a ser empleado en cada país (similar en el diseño, estra-

³ El *Statistical Package for the Social Sciences* (SPSS) se hace público a partir de 1969.

⁴ Entrevista con Sidney Verba, Ciudad de México, 22 de octubre 2009.

tificado, polietapico) y subrayan lo que consideran más importante: que cada familia, en cada nación tenga la misma probabilidad de que uno de sus miembros de dieciocho años o más, sea seleccionado.⁵ En el caso mexicano, aclaran los autores, esto no fue posible por consideraciones técnicas y de costos, por lo que tuvieron que concentrarse en las ciudades de más de diez mil habitantes. Conviene subrayar el énfasis empleado al subrayar que, en todo caso, la selección de cada familia y del entrevistado en tal familia era de tipo “estrictamente aleatorio y en consecuencia fuera del control del entrevistador”.⁶ Los equipos responsables que realizaron el trabajo de campo en cada país eran reconocidos especialistas en ese tipo de estudios empírico.

La segunda parte del Anexo A con sus siete párrafos, dos fórmulas y dos pies de página representa el centro de atención de estos comentarios tanto por su aspecto didáctico como por su valiosa discusión del tema de las comparaciones múltiples en la historia de la estadística social. En efecto, en cuanto a su valor didáctico, el texto es una notable lección acerca de cómo realizar la interpretación de los resultados de pruebas de hipótesis estadísticas simultáneas cuando los datos provienen, como ya se ha mencionado, de un conjunto de muestras aleatorias independientes. Con respecto a los aspectos relativos al uso de las comparaciones múltiples, baste mencionar que ese tema representaba todavía un campo de confrontación entre dos de las más sólidas posiciones de la estadística de los inicios de la década de los años sesenta del siglo XX.

Con objeto de destacar de mejor manera las dos facetas aludidas, se incluye a continuación una lectura comentada de los párrafos mencionados.⁷ La versión consultada es la primera edición de *The Civic Culture* de la Princeton University Press. Con estos comentarios se pretende compartir a los nuevos lectores este sutil aspecto que, de otra manera, podría pasar desapercibido entre las otras, muy trascendentales, aportaciones.

⁵ Es un hecho muy conocido que para esas fechas, el estado del arte de los diseños de muestras puede ser considerado como muy avanzado, pero debido a los problemas de estimación numérica, se tendía a preferir diseños con este tipo de restricciones. En otras palabras, los diseños no autoponderados eran ya utilizados pero, en general, para estudios de menor talla que el de este caso. Para mayores referencias acerca de los diseños muestrales usados en aquellos tiempos, consultar Gray, P. G. y T. Corlett (1950).

⁶ Para el caso mexicano conviene referirse a los comentarios del capítulo IX de *The Civic Culture Revisited* (1989), en el que Craig y Cornelius incluyen, entre otros temas, una crítica de dicha muestra.

⁷ Se trata de una traducción que, sin duda, puede ser mejorada pero que pone el énfasis en los aspectos estrictamente estadísticos.

IV. COMENTARIOS DE SIETE PÁRRAFOS SELECCIONADOS DEL APÉNDICE A⁸

APÉNDICE A ERROR MUESTRAL

Todas las muestras aleatorias están sujetas a error muestral. El error muestral es la diferencia entre los atributos de la muestra y los atributos de la población de la que es extraída la muestra. Tales errores no pueden ser eliminados pero con el muestreo aleatorio es posible estimar la probable magnitud de dichos errores. Tales estimaciones son usualmente obtenidas a través del uso de pruebas de significancia estadística. Con estas pruebas podemos decir qué tan cercana parece estar la verdadera proporción de un atributo en la población de la correspondiente proporción muestral (por ejemplo, 95% de las veces), o cuáles son las chances⁹ de que una particular diferencia entre dos grupos pueda ser debida al error muestral. Como una convención, diferencias tan grandes que puedan ser debidas al error muestral sólo en una de veinte ocasiones (al nivel del 95%) son generalmente aceptadas como estadísticamente significativas. La aplicación de tales tests de significancia al material discutido en este libro debe ser tomada con precaución.

Comentario: Entre las décadas de los cuarenta y los cincuenta del siglo pasado, la comunidad mundial de estadísticos asistió a una fuerte controversia acerca de la interpretación que debería darse a las pruebas de hipótesis. Por un lado, la escuela clásica fundada por Fisher veía en el nivel de significancia una medida de la evidencia aportada por la muestra en favor de la hipótesis que se deseaba probar. Esta medida es precisamente la probabilidad estimada de que el resultado obtenido se deba exclusivamente al error muestral. De esta manera un nivel de significancia pequeño querría decir que era muy poco probable que lo observado fuese producto del azar y, por lo tanto, debería pensarse que la hipótesis propuesta (por ejemplo, que no existe una diferencia en un par de proporciones de algún atributo en dos subpoblaciones) era apropiada. Frente a esta posición, estadísticos tales como Neyman,¹⁰ ubicaban las pruebas de hipótesis desde un contexto falsa-

⁸ Con objeto de facilitar la lectura, se presenta la traducción en tres segmentos seguidos, en cada caso, de los comentarios que resultan pertinentes. La traducción se presenta en negritas e itálicas.

⁹ Se conserva en esta traducción el término *chance* y no el de *probabilidad* debido a que esto forma parte de la discusión conceptual que se comenta más adelante.

¹⁰ El texto de Neyman y Pearson (1932) es un clásico en este campo aun cuando la discusión de los conceptos expresados ahí no fueron inmediatamente aceptados por toda la co-

cionista en el que la hipótesis estadística nula es la negación de la hipótesis de trabajo y el concepto de significancia (*valor p*, en la posición antagónica) era la probabilidad estimada de cometer el error del tipo I (rechazar la hipótesis nula cuando en realidad es verdadera).

De esta manera, la lectura de estas primeras líneas parece estar influenciada por una visión más *fisheriana* que *neymaniana*.¹¹ Si bien esta última posición ha sido la que ha terminado por imponerse en la literatura estadística no deja de ser interesante el interés y la descripción de los argumentos planteados.

Los tests de significancia estadística pueden exagerar la importancia de los hallazgos de diferentes maneras. En primer lugar, por supuesto, la significancia estadística no es una medida de la significancia sustantiva del hallazgo. Esto simplemente indica que una particular diferencia no parece ser debida a errores aleatorios en la muestra. Una diferencia de uno por ciento en las proporciones de dos grupos teniendo un atributo particular puede ser estadísticamente significante si la muestra es suficientemente grande pero, puede ser no sustantivamente significante. De manera similar, los tests de significancia estadística no son tests de asociación. No se puede inferir un nivel alto de asociación a partir de un nivel alto de significancia.

Aún más importante, quizás: tales pruebas de significancia son más apropiadas para el tipo de investigación en la que algunas hipótesis rigurosamente definidas están siendo probadas y son menos apropiadas para investigaciones más exploratorias en las cuales muchas de las relaciones sobre las que se comenta “emergen” a partir de una inspección *a posteriori* de los datos. Con un número suficientemente grande de diferencias a inspeccionar, el investigador está seguro de encontrar algunas que sean estadísticamente significantes únicamente por azar, aun si en realidad no existen diferencias en la población muestreada.

Comentario: Notar lo actual que resulta la lectura de estas líneas, en las que se destaca la importancia de la planeación detallada y evitar los “hallazgos de último momento” que, por desgracia, se suelen encontrar de manera frecuente. La irrupción de los modernos sistemas de cómputo estadístico ha mostrado que su uso descuidado y sobre todo su abuso, puede llevar a conclusiones erróneas. La idea acerca de que los datos pueden, después de ser sometidos a los cada vez más numerosos “análisis estadísticos”, revelar

munidad estadística. Lejos de eso, este artículo, junto con algunos otros, formó parte de la discusión aludida en el párrafo comentado.

¹¹ Para tener una mejor idea del carácter de las discusiones a las que se alude en el comentario, consultar, por ejemplo, a Selvin, H. (1957).

algún resultado interesante debería ser ponderada con ejemplos esclarecedores en los cursos para los futuros investigadores sociales.

Este problema podría ser aún más serio en nuestros datos cuando, por ejemplo, las diferencias en las proporciones teniendo un atributo particular fueran comparadas. Dado que hay cinco naciones, existen diez posibles pares que pueden ser comparados y una diferencia estadísticamente significante entre uno de los pares es bastante posible de suceder únicamente por azar. Es sólo cuando la diferencia particular es predicha que puede ser considerada estadísticamente significante. Puesto que este estudio fue exploratorio, no hemos dudado en reportar los resultados interesantes que emergen de los datos, aunque en muchos casos las diferencias que son reportadas y enfatizadas son aquellas que eran esperadas al inicio del estudio. Sin embargo, esta limitación sobre el uso de los tests estadísticos debe tenerse en mente.

Por último, los tests estadísticos debido a que tratan únicamente con los errores aleatorios del muestreo pueden exagerar la validez de un hallazgo. Éstos no toman en consideración otras fuentes de error —en la entrevista, la codificación, traducción, etcétera— que pueden no variar de manera aleatoria. En general, entonces, uno debería usar tests de significancia estadística con una gran cautela por temor a que nos den un falso sentido de seguridad científica.

[...]

La conclusión general de esta discusión parece ser que el valor de los tests de significancia estadística para estudios de este tipo son dudosos. Las críticas que se han hecho en la literatura a tales tests parecen ser aplicables de manera aún más poderosa a nuestro estudio de cinco naciones.

Comentario: Como puede observarse, la lectura de este texto sería de gran utilidad no solo para jóvenes estudiantes que están interesados en la investigación basada en muestras aleatorias, sino para un buen número de practicantes confirmados que, en ocasiones, tienden a privilegiar la naturaleza insospechada del supuesto hallazgo a la nitidez metodológica de la que es necesario conocer apropiadamente sus límites.

Finalmente, el hecho de que hayan pasado ya más de cincuenta años de la publicación de este texto no significa que las críticas y las dudas acerca de la relación entre la ciencia política (y en general, las ciencias sociales) con la estadística hayan concluido. Algunos temas han cambiado, otros permanecen en el centro del debate pero es de destacar que el profesor Verba continúa teniendo una parte activa en él.¹² Sería justo esperar que el nombre *The Civic Culture* y los de sus autores pudieran ser citados en los

¹² Ver por ejemplo McKeown, Timothy (1999).

textos básicos de estadística de los futuros investigadores sociales como un ejemplo de esa asociación provechosa entre las ciencias sociales y la estadística.

V. BIBLIOGRAFÍA

- ALMOND, Gabriel A. y VERBA, Sydney, *The Civic Culture*, Princeton, Princeton University Press, 1969.
- DURKHEIM, Émile, *Le Suicide (nouvelle Edition)*, París, Presses Universitaires de France, 1960.
- FISHER, Ronald A., *Statistics Methods for Research Workers*, Edinburgh, Oliver and Boyd, 1946.
- GRAY P. G. y CORLETT, T., “Sampling for the Social Survey”, *Journal of the Royal Statistical Society. Series A (General)*, vol. 113, núm. 2, 1950.
- MCKEOWN, Timothy J., “Review of King, Keohane, and Verba’s Designing Social Inquiry: Scientific Inference in Qualitative Research”, *International Organization*, vol. 53, núm. 1, Winter 1999.
- MILLER, Rupert G. Jr., *Simultaneous Statistical Inference*, Nueva York, Springer, 1981.
- NEYMAN, J. y PEARSON, Egon S., “On the Problem of the Most Efficient Tests of Statistical Hypotheses”, *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 231, 1993.
- PEÑA, Daniel y ROMO, Juan, *Introducción a la estadística para las ciencias sociales*, Madrid, McGraw-Hill, 2003.
- SELVIN, Hanan C., “A Critique of Tests of Significance in Survey Research”, *American Sociological Review*, vol. 22, núm. 5, octubre de 1957.