

IV Encuentro Latinoamericano de Metodología de las Ciencias Sociales –ELMeCS

“La investigación social ante desafíos transnacionales: procesos globales, problemáticas emergentes y perspectivas de integración regional”

Heredia, Costa Rica, 27 al 29 de agosto de 2014

MESA 5 - Problemas de enseñanza-aprendizaje de la metodología en el grado y postgrado universitario

El arte de enseñar a crear: reflexiones sobre la enseñanza del proceso de análisis sociológico.

Gonzalo Seid (Universidad de Buenos Aires, Argentina) gonzaloseid@gmail.com

Jessica Malegarie (Universidad de Buenos Aires, Argentina) jessi.malegarie@gmail.com

Patricia Fernández (Universidad de Buenos Aires, Argentina) pfernandez4@yahoo.com.ar

Introducción

El propósito de esta ponencia es sistematizar un conjunto de reflexiones sobre el proceso de enseñanza-aprendizaje del análisis de datos dentro de un enfoque cuantitativo de investigación. La misma se encuadra en nuestra experiencia como docentes de la materia Metodología de Investigación Social de la carrera de Sociología de la Universidad de Buenos Aires.

Dentro del campo de las ciencias sociales, en el proceso de análisis, surgen algunos obstáculos y se ponen de manifiesto ciertas complejidades que hacen de dicho proceso una tarea para nada mecánica ni lineal. Por el contrario, se trata de una actividad que requiere la toma de un conjunto de decisiones teóricas y metodológicas de cara a los desafíos que se van presentando. Es por esta razón que enseñar cómo transitar esta etapa del proceso de investigación requiere, por parte de los docentes, un entrenamiento en pos de construir una

serie de herramientas pedagógicas y didácticas para dar acompañamiento a los estudiantes en los primeros pasos de su formación en investigación. No se trata éste, tampoco, de un proceso mecánico ni lineal de transmisión de fórmulas establecidas sino, por el contrario, de una tarea continua de reflexión y de adaptación a las dificultades que se van presentando en el espacio de aprendizaje dentro del aula.

En este mismo sentido, el contenido que presenta el documento invita a reflexionar conjuntamente sobre cómo enseñar la lógica del análisis cuantitativo poniendo por delante en todo momento el rol fundamental que ocupa la teoría en el conjunto de decisiones a tomar. Se abordarán algunos de los obstáculos más habituales que se presentan en el proceso de enseñanza de esta temática: la vinculación entre variables, el armado del plan de análisis, la distinción entre la significación estadística y la relevancia teórica, la interpretación de valores de coeficientes, la lectura porcentual guiada por la hipótesis, entre otros aspectos. Entendemos que *“durante esta etapa, las dificultades más frecuentes que surgen entre los estudiantes de sociología en el aprendizaje son de lo más variados y comprenden desde la correcta construcción y porcentualización de las tablas, hasta errores como lecturas univariadas que no contrastan las hipótesis, falta de vinculación de las frecuencias condicionales con las marginales, dificultades relativas a la incorporación de variables de control y la interpretación de los resultados del cruce de tres o más variables”* (Fernández y Malegarie, 2011).

En cada apartado se trabajará un aspecto particular del análisis de datos, reforzando los elementos que componen su enseñanza y aquellos que presentan mayor dificultad de aprendizaje. La primera sección profundizará sobre dónde y de qué manera la teoría está presente al momento del análisis. La segunda abordará la relación entre significación teórica y significación estadística. El tercer apartado analizará el rol de los coeficientes de asociación y su interpretación. En cuarto lugar, se ahondará sobre la lectura porcentual guiada por la hipótesis. Por último, la quinta sección refiere al análisis multivariado y a los criterios por los cuales se selecciona una determinada variable de control y se procede a su lectura e interpretación.

La teoría en el momento del análisis

Durante una investigación, la etapa del análisis constituye el momento en que se procura dar cumplimiento a los objetivos de investigación y responder a las preguntas que dieron origen a la misma. En este proceso se busca colaborar con la acumulación del conocimiento en determinada temática y contribuir con hallazgos empíricos a un determinado posicionamiento teórico, a la vez que esta mirada conceptual permite echar luz para interpretar la realidad.

Cuando se trabaja con un enfoque cuantitativo, una de las herramientas más utilizadas es la lectura de tablas bivariadas y multivariadas, a partir de las cuales se puede analizar el comportamiento de una determinada población a partir de la relación de dos o más variables. Es importante señalar, en el sentido que venimos planteando, que toda relación entre variables pautada a partir de una hipótesis –y plasmada en una tabla de contingencia– encierra finalmente una relación entre conceptos. Dichos conceptos surgen y forman parte del marco conceptual. Es por esta razón que podemos afirmar que no es posible soslayar la teoría en el momento del análisis y que toda construcción de un cuadro y su posterior interpretación implica un procedimiento que es tanto teórico como metodológico. Precisamente la elección del entrecruzamiento entre esas variables –y no otras– descansa en la idea de que se espera desde la teoría que dichos conceptos estén de alguna manera vinculados en la realidad. Esta explicitación sobre el rol de la teoría –que en definitiva atraviesa todo el proceso de investigación– es necesaria ante las dificultades de los estudiantes para abordar el análisis teniendo siempre presente el problema y los objetivos desde un determinado posicionamiento teórico. A continuación, se expondrán algunos aspectos que pueden conducir a errores, simplificaciones o confusiones en algunos de los recodos y vericuetos del laberinto de la investigación. En todos los casos es la reflexión teórica la que funcionará como hilo conductor para proseguir en el recorrido de este laberinto.

A. Una dificultad que aparece muchas veces al momento de realizar un análisis adecuado de las tablas es que los estudiantes no siempre tienen claro para qué se construyen. En estos casos seleccionan variables sin situar la lectura de cuadros dentro del proceso de investigación que se ha emprendido y sin basar la construcción de la tabla en un aporte teórico específico. Dicho en otros términos, cuando se decide observar la relación entre dos o más variables los

estudiantes muchas veces no logran considerar la relevancia teórica de dicha relación y la necesaria existencia de alguna hipótesis de trabajo que la sustente.

Un primer recurso pedagógico en estos casos consiste en transmitir que estas hipótesis no necesariamente deben ser planteadas al inicio del proceso en tanto respuestas anticipadas a la pregunta-problema, sino que son hipótesis más puntuales que pueden formularse en el transcurso de la investigación, en vinculación directa con las tablas que se elaboran. A su vez es importante señalar que las hipótesis, basadas en una referencia teórica, serán las que darán un marco y sostén a lo que se plantea analizar, le dará fundamento al análisis y se disminuirán los riesgos de plantear relaciones que carezcan de sentido. Basar el análisis de las relaciones entre variables en una hipótesis con un contenido teórico es uno de los elementos que marcan la diferencia entre una elucubración del sentido común y un análisis sociológico.

Habiendo superado este obstáculo, otra dificultad que surge es que, aún analizando la relación entre variables desde una elección basada en la teoría, en ocasiones se olvida que puede ser necesario un reprocesamiento para lograr mayor relevancia en el análisis. En este sentido un aporte didáctico importante es el de señalar las distintas posibilidades de agrupar y reagrupar las categorías de las variables en distintos sistemas de categorías. Esto implica variaciones en las distribuciones de casos en los espacios de propiedades generados y, por lo tanto, la hipótesis puede verse reforzada e incluso verificada cuando se trabaja con cierto sistema de categorías y no con otros. Aquí la decisión de unir categorías puede fundamentarse en un criterio de relevancia, agrupando los valores de una variable según determinada referencia teórica, o puede descansar en algún criterio empírico que busque reducir la cantidad de espacios de propiedades.

Para definir qué variables y categorías han de ser puestas en relación en las hipótesis (y por lo tanto en las tablas de contingencia) se propone trabajar con los estudiantes profundamente en el cómo y el por qué de la construcción de un plan de análisis. Esto implica una labor de previsión y reflexión sobre las estrategias de análisis y el detalle de procedimientos que deben llevarse a cabo para cumplir los objetivos. Este procedimiento descansará en tres pilares: las variables que hayan sido incluidas en el instrumento de recolección de la información; de éstas, aquellas variables que contengan una frecuencia de respuestas que permita una fragmentación a través de la inclusión de otro/s criterio/s como son la variable independiente y de control; y a su vez el criterio teórico que permita establecer dichas relaciones. El plan de análisis cuantitativo contiene un esquema sobre el conjunto de cuadros esperados, esto es, las

decisiones relativas a todas las tablas que se confeccionarán, formalizando las hipótesis en torno a las cuales se han de interpretar los datos producidos.

Hasta aquí lo remarcado refiere a que cada tabla que se elaborará tiene su razón de ser en que hay un supuesto teórico acerca de lo que espera observarse. Si en lugar de darle formato de procesamiento a las hipótesis se procediera observando todos los cruces posibles de variables, ésta sería una tarea no sólo innecesaria y engorrosa, sino ante todo carente de relevancia teórica, que tendría como resultado datos desconectados de la teoría, que por sí solos no pueden constituir un avance en la producción de conocimiento científico. Así por ejemplo, los programas informáticos como SPSS, Excel o Access permiten sin grandes dificultades “cruzar todo con todo” y eventualmente podrían servir para explorar relaciones impensadas en los datos, pero debe advertirse que de ninguna manera se estaría produciendo conocimiento científico si se realizara una lectura de datos desprovista de la teoría que le da sustento. Como contrapartida, debe advertirse que tampoco es posible forzar la lectura teórica y “estirar” los conceptos para que digan lo “queremos que digan”. La común unión entre teoría y empiria es la que garantiza un adecuado proceso de investigación.

Por todo lo señalado es que se trabaja dentro del aula para desterrar la idea de que han de ponerse todas las variables en relación, rompiendo con una lógica empirista de investigación. Es así que, al momento de enseñar la etapa de análisis, la construcción del plan de cuadros como expresión de hipótesis resulta imprescindible. En dicha actividad se visualiza que existe una decisión teórica tras la construcción de todo cuadro.

B. Al examinar qué ha ocurrido con cada hipótesis sustantiva y los resultados de un cuadro, se está en la etapa del análisis propiamente dicho. Aquí se busca que los estudiantes logren comprender que la contrastación de cada hipótesis permite avanzar en la producción de conocimiento, pero no agota en sí misma los objetivos de investigación, sino que es el conjunto de las hipótesis de trabajo contrastadas en las tablas que se producen, con la articulación teórica y la ilación lógica que le confieren quienes investigan, lo que puede dar respuesta a los interrogantes o no (es también un aprendizaje interesante el de aceptar que ciertas preguntas de investigación puedan no ser total o parcialmente respondidas en un proceso de investigación y que se requiera plantear un nuevo diseño para intentar responderlas).

La lectura de un cuadro no debe ser entendida entonces como un fin en sí mismo sino como un medio para el análisis e interpretación de los resultados. Nunca debe perderse de vista el panorama general del plan de análisis, de modo que los avances parciales en el conocimiento a partir de la lectura de cada cuadro puedan ser articulados de un modo coherente y enfocado en dar cumplimiento a los objetivos.

La importancia de enfatizar el para qué de la lectura de cuadros se debe a que cuando se trabaja con abordajes cuantitativos, se utilizan habitualmente herramientas estadísticas, que en no pocas ocasiones son erróneamente consideradas como si fueran utilizables de manera mecánica mediante meros procedimientos técnicos e impersonales. Para evitar este error, el rol del docente se centra en distinguir análisis cuantitativo de análisis estadístico, teniendo presente que el dato es siempre una construcción teórica y que no puede ser interpretado desde una fórmula sino desde la teoría y desde un procedimiento teórico metodológico determinado. Esta propuesta de enseñanza considera el análisis estadístico como una importante herramienta de trabajo, que desde la metodología se incorpora y complementa con un abordaje teórico que sustenta el análisis de los resultados.

Blalock (1986: 19) recuerda que la estadística no comprende problemas como los de medición y que las consideraciones estadísticas intervienen en decisiones relativas a la construcción de la muestra y a la manipulación de cifras, dando por supuesto el cumplimiento de requisitos en la medición. El hecho de que la estadística sólo intervenga en algunas decisiones técnicas del muestreo y análisis no debe entenderse como si el investigador pudiera llevar a cabo su trabajo sin conocimiento de estadística y delegar en un especialista el análisis estadístico, sino que los problemas del análisis han de anticiparse, influyendo las consideraciones estadísticas a lo largo de todo el proceso de investigación. En un sentido similar se ha afirmado que “no hay nada más técnico (en el sentido de conocimiento objetivado, impersonal, basado en procedimientos normados, recurrentemente aplicables y ampliamente aceptados) en la investigación estándar que sus herramientas de análisis (estadístico) (...) Esto no significa que las técnicas estadísticas resuelvan por sí solas los múltiples y complejos aspectos que encierra la mediación entre un conjunto de informaciones empíricamente construidas y algún tipo de interpretación de éstas” (Marradi, Archenti y Piovani, 2010: 264).

En definitiva, dentro del proceso de enseñanza aprendizaje se remarca fuertemente que el investigador se vale de la estadística pero no puede dejar de tomar decisiones teórico-metodológicas indelegables si quiere producir conocimiento relevante para llevar a buen puerto la investigación. De lo contrario, el riesgo es caer en lo que Pierre Bourdieu (2011)

denominaba la autonomización del método, que ocurre cuando el desarrollo de técnicas estadísticas y sofisticados modelos matemáticos se convierten en el objeto mismo de las investigaciones, quedando los procesos sociales reducidos a relaciones estadísticas entre variables. Teoría, metodología y técnicas de análisis deben articularse reflexivamente, atendiendo a la coherencia y ejerciendo vigilancia epistemológica.

C. Otro aspecto de importancia para la reflexión teórica y la vigilancia epistemológica es lo relativo a la construcción de las variables a partir de la información relevada. Si bien suele resultar innecesario reiterar las limitaciones de una fuente de datos cada vez que se analiza un cuadro, puede ser provechoso explicitarlas toda vez que su lectura produce la impresión de que si se hubieran construido las variables de otra manera o con otra técnica los datos serían otros. Al reflexionar se puede dilucidar si esa impresión se debe a un prejuicio porque los datos no se ajustan a las expectativas o si debe cuestionarse la construcción de ese dato. Puesto que “los datos no hablan” es necesario mantener una actitud de vigilancia epistemológica constante respecto a la evidencia construida, controlando las prenociones del investigador y la de los sujetos estudiados o informantes.

Cuando se utiliza el método de encuesta, es necesario interrogarse sobre las propias preguntas de cuestionario y cómo estas afectan la construcción del objeto que se realiza. Bourdieu, Chamboredon y Passeron (2011) aluden al señalamiento de Goldthorpe y Lockwood acerca de que las respuestas a una pregunta varían según sea una pregunta abierta o cerrada, y según los términos elegidos para brindar como opciones. Ejemplifican que un trabajador manual que se declara de clase media puede significar que se considera por encima de la infraclass, que se considera en una posición media al interior de la clase obrera, que se siente en el mismo nivel económico que empleados de cuello blanco y pequeños comerciantes, que aspira a un estilo de vida de clase media o bien que su origen familiar es de clase media.

En este sentido, cuando se analizan variables construidas a partir de lo que los encuestados han informado, siempre es posible matizar la aparente objetividad del dato con consideraciones en torno a lo que la etnometodología ha denominado la indexicalidad interminable de las expresiones¹. Considerando que el lenguaje natural es un recurso obligado

¹ Para esta corriente, resulta irremediable la dependencia contextual de lo que se dice. No siendo posible establecer un sentido invariable, las expresiones del lenguaje ordinario deben analizarse conforme a la situación particular en que fueron producidas. El significado cabal de las palabras y expresiones depende del contexto de intercambio lingüístico en que aparecen, de la biografía e intención inmediata del locutor, de la relación con el oyente, de conversaciones anteriores, etc.

de cualquier indagación sociológica, la necesidad de tomar en cuenta las propiedades indexicales del discurso práctico conlleva un cuestionamiento al tratamiento que los sociólogos le dan a los cuestionarios que utilizan suponiendo cierta homogeneidad semántica de las palabras. Tal como plantea Aaron Cicourel (1982), es necesario indagar en los sentidos típicos que se atribuyen a las preguntas, los motivos y roles que se presuponen. El análisis de las nociones que entran en juego en los actos cotidianos es la forma de aproximarse a medidas más rigurosas, sabiendo que nunca podrán ser exactas, puesto que no es legítimo suponer que el investigador y el actor comparten las mismas estructuras de sentido subjetivo para atribuir significaciones. Cuando se trabaja con variables construidas a partir de información proveniente de encuestas, no puede soslayarse la atención al razonamiento y la comprensión de los sujetos respondentes para entender cómo dichas encuestas reconstruyen el conocimiento de los actores sobre lo social.

Significación estadística y relevancia teórica

Luego de haber elaborado las hipótesis a partir de la teoría y de efectuado el procesamiento del que resulta una distribución de casos en los espacios de propiedades, se está en el momento del análisis de una tabla. Una estrategia posible es comenzar realizando una prueba de hipótesis. Cuando se utilizan estas pruebas debe evitarse una aplicación mecánica y metodológicamente irreflexiva; es necesario que se evalúe el alcance de las mismas a los fines de la investigación. Tomemos el ejemplo del ji cuadrado, que es una de las pruebas de hipótesis empleadas más habitualmente en el análisis de tablas. El ji cuadrado pone a prueba la hipótesis nula, es decir, la hipótesis que niega la relación entre las variables analizadas. Se plantea así una hipótesis que sostiene que la relación que esperamos teórica y empíricamente no tendrá lugar. Esta prueba de hipótesis, de acuerdo a sus resultados, define el escenario estadístico en que se trabaja, por eso puede constituir un buen primer paso en el análisis.

Si se rechaza la hipótesis nula, se descarta asumiendo determinado riesgo que se trate de una situación de independencia estadística y se acepta que hay alguna asociación estadística entre las variables de la hipótesis de trabajo, con determinado nivel de confianza. En cambio, cuando la hipótesis nula no puede ser rechazada, la hipótesis de trabajo no puede ser aceptada como estadísticamente significativa, para determinado nivel de confianza. Sin embargo, el

hecho de que una hipótesis no sea estadísticamente significativa para determinado nivel de confianza no implica necesariamente que se acepte la hipótesis nula ni que deba darse por finalizado el análisis de la relación.

La ausencia de significación estadística puede deberse a diversos motivos, tales como el modo de construir las variables, una insuficiente cantidad de casos, la amplitud del sistema de categorías de cada variable (en términos estadísticos, los grados de libertad del cuadro), las características de la estrategia de medición, etc. Por ello, la ausencia de significación estadística en la relación entre las variables analizadas puede ser tan relevante como la determinación de su existencia, siendo posible avanzar en la producción de conocimiento cuando se trata de una situación de “significación” teórica y sociológica. A los fines sociológicos, la situación en que el χ^2 no asume significación estadística no menoscaba la relevancia teórica de una hipótesis y, en contrapartida, el hecho de que una relación entre variables sea estadísticamente significativa no es suficiente por sí mismo para producir conocimiento científico.

Uno de los principales obstáculos al aprendizaje por parte de los estudiantes se centra en la dificultad real de asumir que la relación no debe ser descartada, aún no pudiendo rechazar la hipótesis nula. Es por esta razón que en la enseñanza de este eje el trabajo fundamental del docente se ubica en dos aspectos. Por un lado, el constante énfasis en que el motor que guía el análisis es la búsqueda de nueva información que aporte conocimiento sobre una problemática social específica y no el establecimiento de un conjunto de varas de aceptación-rechazo de carácter estadístico, vacías de contenido. Por otro lado, se remarca nuevamente el papel relevante de la teoría al momento del análisis. Se recupera así la noción de conocimiento acumulado, donde la ciencia busca permanentemente la ampliación del saber sobre una determinada temática más allá de su validación estadística. Por último, y no por ello menos importante, debe recordarse permanentemente que lo que se está realizando es un proceso de contrastación teórico-empírica, lo que se está poniendo a prueba es una relación teórico metodológica y no un procesamiento estadístico.

Cabe aclarar también que la advertencia anteriormente explicitada debe tenerse en cuenta también en la enseñanza de los tipos de errores estadísticos. En el error de tipo I se acepta la hipótesis de trabajo cuando la relación no es estadísticamente significativa. La situación inversa es el error de tipo II, que es la probabilidad de no rechazar la hipótesis nula siendo ésta falsa y de rechazar la hipótesis de trabajo que sí era estadísticamente significativa. En ambas situaciones no se trata de errores en tanto decisiones desacertadas por la impericia del

investigador, sino de errores estadísticos, probabilidades que quienes investigan asumen como riesgos propios de los instrumentos estadísticos que se utilizan.

Interpretación de valores de los coeficientes a partir de la comparación

Las pruebas de hipótesis son un paso a realizarse antes, después o en paralelo con el uso de coeficientes, al que referiremos en el presente apartado. A modo organizativo se presenta de esta manera, sin perjuicio de que pueda asumir otro orden en la práctica de investigación.

En el presente apartado se analiza el rol que asumen los coeficientes de asociación en la interpretación de la relación entre dos variables. Si para determinar la existencia o no de asociación estadísticamente significativa se utilizan pruebas de hipótesis, para conocer la intensidad de una asociación entre las variables de una tabla es posible valerse de coeficientes de asociación. Sin embargo, para interpretar adecuadamente el valor de un coeficiente es necesario tener ciertos recaudos.

En primer lugar, seleccionar adecuadamente el tipo de coeficiente que se utiliza, lo cual dependerá no solo de las características de la variable que se ha construido (su nivel de medición y sistema de categorías) sino también de la hipótesis planteada. Si la hipótesis es rinconal, dependiendo del tipo de variables involucradas, se utilizan los coeficientes curvilineales como Q de Yule o Gamma. En la interpretación de los valores de estos coeficientes, debe tenerse presente que pueden asumir altos valores puesto que los mismos se basan en una definición menos restrictiva de asociación perfecta, relativa a una hipótesis rinconal que puede cumplirse a la perfección poniendo una sola condición -de modo que para que haya asociación perfecta en una tabla de dos por dos bastaría con que un solo espacio tenga cero frecuencias-. Inversamente, los coeficientes lineales como V de Cramer tienden a subestimar el grado de relación por suponer una definición exigente de asociación perfecta, lo cual debe ser tenido en cuenta en la interpretación del valor de este coeficiente.

La definición del tipo de hipótesis (rinconal o diagonal) suele ser un aspecto que genera dificultad para los estudiantes en la comprensión del tipo de coeficiente a utilizar. Por un lado, en ocasiones no resulta fácil la identificación del tipo de hipótesis de que se trata en cada caso y, por otro lado, resulta compleja la comprensión de que el tipo de hipótesis planteada es

independiente de lo observado en el cuadro. Frente a estas dificultades una estrategia muy útil es trabajar reiteradamente sobre diferentes formulaciones de hipótesis asociadas a una misma tabla de contingencia, lo que permite observar que pueden corroborarse o refutarse distintas hipótesis frente a una misma distribución porcentual observada en un cuadro bivariado.

Otro aspecto importante a destacar en la enseñanza de coeficientes de asociación se relaciona con la interpretación de su valor resultante. El valor que arroje un coeficiente de asociación ha de ser entendido según cuánto se aproxime al 0 y al 1 (o también -1 en algunos coeficientes). A medida que los coeficientes comienzan a alejarse de cero están expresando algún tipo de relación entre las variables. No obstante, un importante recaudo consiste en que el valor no sea interpretado aisladamente como alto, medio, bajo, etc., sino que su interpretación esté fundada en la realización de comparaciones, para las cuales se puede tomar como referencia el valor del mismo coeficiente en otros casos o para otra hipótesis. Es posible interpretar el valor de un coeficiente a la luz del valor que el mismo adquiere para la relación de la misma variable dependiente con otra variable independiente. Para que las comparaciones respecto a la fuerza de incidencia de distintas variables independientes sean correctas es necesario realizarlas con el mismo coeficiente (frecuentemente deben usarse aquellos que son para variables nominales si hay por lo menos una de dicho nivel de medición).

También el conocimiento teórico acumulado sobre el comportamiento del fenómeno contribuirá a determinar cuán alto o bajo es el resultado, en relación con expectativas teóricas o derivadas de la experiencia previa. En cualquier caso, la comparación es un paso nodal en el análisis, que posibilita que los recursos técnicos sean subordinados a las decisiones teórico-metodológicas, contribuyendo a la producción de conocimiento teóricamente relevante.

En no pocas ocasiones los estudiantes tienden a hacer un uso mecánico de los coeficientes y una lectura demasiado simple de una relación bivariada como si fuera en sí misma "fuerte o débil", sin comparar esa supuesta fortaleza de asociación con la de otra relación bivariada. Esto reduce las posibilidades de conocer qué variable independiente describe mejor la asociación con una determinada variable dependiente. Reconociendo que la comparación es esencial en toda estrategia de análisis, interpretar individualmente cada coeficiente en una escala de mayor o menor fuerza de asociación acorta la comprensión global de un fenómeno a la luz del impacto de diferentes variables. En suma, resulta imprescindible la estrategia de comparar los resultados del mismo coeficiente frente a distintas situaciones que pongan en relación una variable dependiente con distintas variables independientes, observando cuál de ellas contiene una mayor asociación con la variable que queremos analizar.

Una de las estrategias pedagógicas utilizadas refiere a indagar siempre sobre el para qué de la elección de determinado coeficiente y el para qué de su aplicación. Es respondiendo a estos interrogantes que su utilización estará fundamentada en un aporte al análisis y no en una “receta” que se sigue acríticamente. Si al aplicar un determinado coeficiente, y su comparación con otros, se arroja luz al análisis y se puede tener mayor claridad sobre la intensidad de distintas relaciones, su utilización tendrá un para qué con sentido dentro del proceso de contrastación teórico empírica.

Lectura porcentual guiada por la hipótesis

El recurso central, que se complementa con las pruebas de hipótesis y los coeficientes, es la lectura porcentual de las tablas. Como se ha señalado en reiteradas ocasiones, toda construcción de una tabla de contingencia y el sentido porcentual otorgado a dicha relación descansa en un supuesto teórico expresado a través de una hipótesis. Mientras que las pruebas de hipótesis y los coeficientes permiten un conocimiento sintético acerca de la relación entre variables, la lectura porcentual de tablas de contingencia -generadas a partir de la tabulación simultánea de dos o más variables en función de las hipótesis-, permite un análisis cuantitativo refinado, analítico y “artesanal”.

Una vez que se tiene una tabla con las frecuencias absolutas distribuidas en los espacios de propiedades, es necesario calcular porcentajes para poder realizar comparaciones válidas con mayor facilidad. Porcentualizar es una manera de estandarizar las frecuencias condicionales para liberarlas de los efectos de las diferencias de los marginales y hacer comparable una parte respecto al todo (Baranger, 1999). No obstante, es necesario saber sobre qué total de casos han sido calculados los porcentajes, ya sea para poder re-porcentualizar agrupando categorías o, más importante, para tener los debidos recaudos en la interpretación de porcentajes calculados sobre pocos casos. La cantidad de decimales con la cual puede trabajarse depende de la cantidad de casos, lo cual debe tenerse en cuenta para evitar que los porcentajes aparenten un nivel de exactitud que no pueden tener cuando la muestra es pequeña. Para facilitar el análisis se comenta a los estudiantes que generalmente los porcentajes suelen presentarse sin decimales o con un decimal.

Cuando se trata de un cuadro de dos categorías entrecruzadas con dos categorías -usualmente llamados cuadros de dos por dos- bastará con comparar dos porcentajes para leer el cuadro por completo, puesto que habiendo leído los dos porcentajes de una categoría de la variable dependiente, los de la otra categoría de la variable dependiente son complementarios. La diferencia entre los dos porcentajes comparados será la diferencia porcentual o épsilon (ϵ) que en este caso sintetiza todo el cuadro, indicando la fuerza de la relación.

En los cuadros de más de dos por dos, ya no habrá una única diferencia porcentual sino varias y por ende ninguna de ellas sintetizará en una única medida la fuerza la relación entre las variables. Cuantas más sean más las categorías de las variables, mayores son las posibilidades de comparaciones que se abren. Probablemente no sea necesario leer todos los porcentajes cuando haya muchas categorías de las variables, siendo siempre la hipótesis lo que orientará la lectura indicando qué espacios de propiedades son los más relevantes según las expectativas teóricas -lo que no significa que la lectura deba restringirse a la de las celdas verificadoras de la hipótesis-.

Todo tipo de situaciones intermedias, variaciones sutiles y matices respecto al modo de relación entre las variables pueden ser descriptos pormenorizadamente mediante la lectura porcentual, lo cual permite analizar en profundidad la información resultante de un cuadro, algo que no es posible realizar sólo con los coeficientes de asociación.

Las dificultades de los estudiantes en lo que respecta a la lectura y análisis del cuadro abarcan desde el sentido en el cual calculan los porcentajes hasta la realización de lecturas erróneas respecto a dichas porcentualizaciones. Podemos decir que de todos los aspectos antes mencionados, aquel que conlleva mayor complejidad para los estudiantes es la correcta construcción de un cuadro en función de una hipótesis y su ulterior lectura e interpretación. Y en este punto la estrategia no implica ni más ni menos que dedicar el mayor tiempo que la currícula permita a realizar distintos tipos de ejercitaciones. Las mismas abarcan instancias individuales, conjuntas, verbales, por escrito y puestas en común, para superar las distintas limitaciones que surgen en el aprendizaje: análisis realizado correctamente en forma verbal pero con dificultad de plasmarlo por escrito, pormenorizado análisis escrito que pierde el sentido orientador de la lectura basado en la hipótesis, lecturas incorrectas que llevan a conclusiones equivocadas, por poner algunos ejemplos. Lo que se busca en definitiva es la apropiación real del uso de una técnica que será seguramente la que con más frecuencia utilicen los estudiantes a lo largo de la carrera y de su desempeño profesional.

Antes de pasar al apartado siguiente, nos parece importante destacar un aspecto adicional del aprendizaje del análisis de tablas, más allá de las cuestiones técnicas, vinculado a aprender a percibir y dotar de sentido lo que “dice” el cuadro, toda la información que el cuadro tiene para proporcionar. Seguramente algunas sutilezas sólo podrían resultar sugerentes para quienes cuentan con mayor experiencia realizando investigaciones, pero vale la pena tener en cuenta que muchas veces los aspectos más interesantes del proceso de producción de conocimiento pueden surgir de las maneras menos esperadas. Por este motivo, aunque las tablas se elaboren habitualmente para verificar hipótesis, también debe considerarse la posibilidad de que el cuadro pueda ser utilizado como elemento sugerente de nuevas hipótesis “más allá de lo que dice el cuadro”, con procesos o variables no relevadas, que contribuyen a la interpretación de los datos. Una distribución o alguna celda que se aleja de lo esperado, acompañada de una mirada aguda, suele disparar hipótesis de por qué ocurre, “interpretaciones” que suelen emerger casi espontáneamente de un sentido común sociológico o del conocimiento teórico acumulado. Pueden resultar particularmente fructíferas las ideas que emergen en las discusiones de un grupo de investigación, pues pueden ser fuente de hipótesis o bien tratarse de preconcepciones a ser explicitadas y en torno a las que deba reflexionarse con la actitud de vigilancia epistemológica.

Interrogarse permanentemente cómo se vinculan las variables

Habiendo señalado muy sucintamente el proceso del cual se puede servir un estudiante para analizar relaciones bivariadas, podemos ahora introducirnos en el análisis de relaciones multivariadas. Cuando se corrobora que existe relación entre dos variables, y bajo una lógica explicativa, es habitual la decisión de introducir variables de control para examinar si esa asociación inicial persiste, se modifica o desaparece con la presencia de otras variables. Con la introducción de sucesivas variables de control es posible aproximarse a la explicación de un fenómeno, lo cual no implica necesariamente la causalidad. En cualquier caso es necesaria la reflexión teórica acerca de cómo se vinculan ambas variables. Si las hipótesis han sido adecuadamente planteadas a partir de la teoría, es probable que ya se haya conjeturado algún mecanismo que vincule las variables, el cual puede estar explicitado y clarificado en mayor o menor medida. Nuevamente, se refuerza la idea que la elección de la variable de control es

una decisión teórica que desprende una hipótesis de asociación, ahora de carácter multivariado.

Los problemas que suelen surgir entre el estudiantado se refieren a los ya señalados sobre la inconsistencia teórica en la selección de la tercer variable y la correspondiente construcción de hipótesis alternativas. Se procede en este caso a la incorporación de cualquier variable “cruzando por cruzar”, sin sustento conceptual. En estos casos se enfatiza en el rol de la teoría y el sustento que ésta ofrece al momento de elaborar una conclusión.

En otros casos los estudiantes no saben cómo proceder durante el análisis una vez que han elegido la variable a incorporar y han construido la hipótesis alternativa. En algunos casos, los estudiantes confunden el lugar donde debe estar ubicada la tercera variable y por lo tanto se les dificulta la lectura correspondiente y la interpretación global del rol que ocupa la variable introducida en la relación original. Se remarca entonces que el nuevo “estímulo” debe fragmentar la lectura de la relación anterior, es decir que la relación bivariada debe ser repetida y analizada para cada una de las categorías de la nueva variable.

Luego se señala que una de las tareas a realizar será comparar los resultados que se obtuvieron en la tabulación cruzada bivariada con las que se obtengan al agregar sistemáticamente terceras variables. Se compararán entonces los coeficientes y las distintas lecturas porcentuales y se evaluará si la relación original fue afectada o no por la introducción de una nueva variable.

Se transmiten a los estudiantes las posibilidades más conocidas respecto a lo que puede ocurrir con la introducción de la variable de control, aclarando siempre que se trata de situaciones límite. La hipótesis original podría ser descartada si la introducción de la nueva variable en el análisis diluye lo que se había establecido por el momento (las diferencias entre porcentajes que aparecen en los espacios de propiedades de la relación bivariada decrecen notablemente al introducir una nueva variable). Esto no implica que la relación original haya sido falsa, sino que en dicha covariación estaba incidiendo de algún modo la variable de control. Otra de las situaciones que pueden suceder es que la introducción de la nueva variable no modifique notablemente la relación bivariada y se sostenga la hipótesis original. La tercera opción es que la relación bivariada se mantenga o fortalezca para una de las categorías de la variable de control y no para otras, es decir, la hipótesis se especifica (se ven modificaciones en los coeficientes y porcentajes entre categorías de la variable de control).

Para interpretar cada una de estas posibilidades es preciso reflexionar en torno a los mecanismos que vinculan las variables, confrontándolos con la nueva evidencia.

Un mecanismo es “el modo en que procede un proceso, es un proceso esencial capaz de producir o impedir algún cambio en un sistema concreto” (Bunge, 2001). La mayoría de los mecanismos están ocultos, son imperceptibles, tienen que conjeturarse y luego contrastarse empíricamente. De este modo es posible responder detalladamente cómo funciona algo, yendo más allá de la mera descripción y de las denominadas “teorías de caja negra” que sólo incluyen variables externas observables, sin ser capaces de dar cuenta del funcionamiento, de cómo un fenómeno condiciona, influye, determina de algún modo o causa otro fenómeno.

Aunque se descubra una asociación entre variables, no se debe aceptar como explicación válida si no se encuentra para ella algún mecanismo plausible. Si no hay algún mecanismo en la realidad, no se podrá realizar una explicación plausible y por lo tanto no se podrá comprender satisfactoriamente. La regla metodológica al respecto consiste en buscar el o los mecanismos que subyacen a toda asociación o cambio constante. Por ejemplo, en la sociología de la movilidad social habitualmente se considera a la clase social de una persona como variable dependiente de la clase social de su padre. Cuando se analizan los datos, se corrobora que existe relación de modo tal que se concentran en la diagonal de la tabla una proporción de casos mayor que la esperable bajo el supuesto de independencia estadística. Esto significa que un determinado porcentaje tiende a reproducir la posición de clase de su familia de origen, mientras que entre aquellos que hubo movilidad las probabilidades no son las mismas de acceder a cualquier clase social cuando provienen de distintos orígenes. Más allá de la necesidad de introducir en el análisis variables de control, -algunas de las cuales podría jugar papeles cruciales como variables intervinientes como la educación-, existen múltiples procesos como la socialización, la transmisión de capitales y recursos, la herencia patrimonial, las dinámicas familiares, etc. cuya complejidad los vuelve difícilmente reductibles a un conjunto de variables relevadas en una investigación, pero que resultan insoslayables para una comprensión cabal de los procesos que se manifiestan en la asociación entre variables analizada, amén de lo cual deben tomarse en cuenta para contextualizarla e interpretarla cabalmente, pues constituyen precisamente los mecanismos que vinculan la variable independiente “clase social del padre” con la variable dependiente “clase social del encuestado”.

Breves conclusiones

Como cierre nos interesa explicitar dos aspectos que asumen importancia en la enseñanza durante la etapa de análisis. Son, a nuestro entender, elementos que deben remarcarse permanentemente en el proceso de enseñanza, evitando así que los estudiantes conciban la existencia de fórmulas pre-establecidas a la hora de realizar el análisis de datos.

Por un lado, el rol fundamental que posee la teoría durante todo el proceso de investigación, tanto en la delimitación de la pregunta problema, en la definición conceptual, en la operacionalización de variables, en la construcción del instrumento y por supuesto también en la decisión de qué y cómo interpretar los resultados de una investigación. La teoría es entonces aquello que enmarca y establece los límites posibles de una investigación, definiendo también nuevos interrogantes resultantes que darán paso a posibles nuevos estudios. Este aspecto fue repetidamente remarcado en la ponencia con el fin de explicitar el énfasis que ponemos sobre este aspecto en nuestra tarea docente dentro del aula.

Por otro lado, nos parece importante remarcar que todo abordaje metodológico implica un diseño único, que es propio y particular a cada problema de investigación. De esta manera, si bien existen reglas generales en relación a cómo y en qué momentos se aplican una serie de instrumentos, herramientas y técnicas en el proceso de investigación y en la etapa de análisis, las mismas solo toman forma y contenido si son aplicadas a un particular problema de investigación. En este marco la tarea del docente implica remarcar que toda decisión es una decisión teórica y metodológica y una decisión que sólo puede ser evaluada en términos del problema que se está investigando y los objetivos propios de esa investigación.

Referencias bibliográficas

- Baranger, D. (1999) *Construcción y Análisis de Datos. Introducción al uso de Técnicas Cuantitativas en la Investigación Social*. Posadas: Editorial Universitaria de la Universidad Nacional de Misiones.
- Blalock, H. (1986) *Estadística social*. México: Fondo de Cultura Económica.
- Bourdieu, P.; Chamboredon, J.C. y Passeron, J.C. (2011) *El oficio del sociólogo. Presupuestos epistemológicos*. Buenos Aires: Siglo XXI.
- Bunge Mario (2001) *La relación entre la Sociología y la Filosofía*. Madrid: EDAF.
- Cicourel, A. (1982) *El método y la medida en sociología*. Madrid: Editora Nacional.
- Fernández. P. y Malegarie, J. (2011) “Lectura de cuadros. Desafíos del proceso de enseñanza aprendizaje”. Ponencia presentada en el XXVIII Congreso Internacional de la Asociación Latinoamericana de Sociología, UFPE, Recife-PE.
- Marradi, A.; Archenti, N. y Piovani, J. (2010) *Metodología de las Ciencias Sociales*. Buenos Aires: Cengage Learning.