



De los conceptos a la construcción de los tipos sociales agrarios. Una mirada sobre distintos modelos y las estrategias metodológicas

From concepts to the construction of agricultural social types. A look on different models and methodological strategies

Guillermo De Martinelli

Resumen.

El presente trabajo se propone avanzar en la reflexión metodológica respecto a las diferentes etapas y supuestos que articulan el desarrollo de tipologías de sujetos agrarios, en tanto instrumentos de análisis social. Para ello revisamos los alcances y limitaciones de distintas estrategias metodológicas para el armado de tipologías y la construcción de tipos sociales agrarios correspondientes a la región pampeana argentina. El artículo se ocupa de explorar el uso de las denominadas tablas de contingencia y luego las denominadas técnicas multivariadas, focalizando la atención en el análisis de conglomerados. Particularmente interesa avanzar en el debate acerca de las características que asumen los sujetos agrarios pero desde la perspectiva que provee la reflexión metodológica sobre las herramientas que suelen utilizarse en el proceso de identificación y clasificación de los sujetos agrarios.

Palabras claves: tipologías; sujetos agrarios; técnicas multivariadas; tablas de contingencia; análisis de cluster.

Abstract.

This paper intends to advance a methodological reflection on the different stages and assumptions that underlie the development of typologies of agrarian subjects, as an instrument of social analysis. In order to do that, we analyze the scope and limitations of different methodological strategies for the assembly of typologies and the construction of agrarian social types in the Pampas region of Argentina. The paper focuses on exploring the use of so-called contingency tables and multivariate techniques, paying special attention to cluster analysis. We aim to contribute particularly to advancing the debate about the features assumed by different agrarian subject, from the perspective provided by methodological reflection on the tools commonly used in the process of identification and classification of such social subjects.

Keywords: typologies; agricultural subjects; multivariate techniques; contingency tables; cluster analysis.

I. Introducción

El presente trabajo se propone avanzar en la reflexión metodológica respecto a las diferentes etapas y supuestos que articulan el desarrollo de tipologías de sujetos agrarios, en tanto instrumentos de análisis social. Para ello hemos ordenado la exposición a través de tres planos de reflexión. En el primero, detendremos la atención en el análisis de la construcción de las tipologías como instrumento analítico, revisando cuestiones centrales respecto al armado de los indicadores y el tipo de fuentes utilizadas. El segundo plano de reflexión se referirá a las distintas estrategias adoptadas para construir tipologías, considerando sus alcances y limitaciones. Por último, cerraremos el trabajo discutiendo las implicancias que se derivan de este tipo de construcción analítica.

II.1 Tipologías y construcción de tipos sociales

Si asumimos que los fenómenos sociales son por definición multidimensionales (Samaja, 1993), deberíamos comenzar por interrogarnos sobre cuáles son las operaciones que permiten que dicha multidimensionalidad logre reducirse y hacerse, en alguna medida, comprensible.

A través de la construcción de tipos y modelos se ha buscado resolver, aunque sea en parte, esta tensión, propiciando un acercamiento a los fenómenos sociales a través de la clasificación, simplificación y reducción de la multidimensionalidad que los procesos sociales poseen.

En este sentido, la construcción de tipos sociales tomó una dirección precisa a partir de los desarrollos realizados por Weber. La orientación se dio entonces básicamente a través de categorías ideales, en tanto las mismas eran elaboradas a partir de la acentuación unilateral de uno o algunos puntos de vista y mediante la reunión de gran cantidad de fenómenos concretos individuales, existentes de un modo difuso. Es decir, fenómenos que se amoldan a aquellos puntos de vista destacados de modo unilateral hasta constituir una formación conceptual unitaria.

Desde los trabajos de Becker (1956) se ha propuesto sustituir el "tipo ideal" weberiano, de concepción más intuitiva y subjetiva, por el "tipo estructurado" o "tipo construido", elaborado con el apoyo de la estadística y las diferentes formas de análisis. De modo que el valor de las tipologías fue modificándose para avanzar a través de su capacidad de explicar la realidad, ordenando los hechos observados de manera tal que facilitara su comparación.

La utilización de "tipologías" para la clasificación, simplificación y reducción de la dimensionalidad social permite contar con un elemento heurístico que intenta establecer un puente entre la necesidad de generalización que tiene la ciencia y los fenómenos particulares que se confrontan en la realidad.¹

Los criterios de ordenamiento adquieren en el proceso de clasificación un rol fundamental, dado que su elección e inclusión dentro del modelo determinan las características y significación que asumen los grupos resultantes. Pero, la centralidad que toman los criterios de clasificación en la construcción de las tipologías, no siempre son debidamente ponderados, como tampoco suelen considerarse los problemas inherentes a su diseño.²

¹ Como señala Ander-Egg, la operación de construir tipos, consiste en seleccionar y acentuar una o varias características de un fenómeno social a partir de un número significativo de datos empíricos con el fin de simplificar y explicar lo existente (Ander-Egg, 1995: 82).

² También es necesario incluir en el origen de la diversidad de tipologías, los diferentes objetivos que guían su construcción, dado que no es lo mismo la preocupación académica por explorar algún tipo de teoría sobre el desarrollo del capitalismo, que el interés de ciertos organismos oficiales, orientados por el diseño de políticas sectoriales. Como señalan Tort y Román (2005: 36), para el caso de la explotaciones familiares "Desde trabajos académicos, preocupados por establecer el rol del campesinado en la evolución de la sociedad y el

En relación al campo de los estudios agrarios, la identificación de los límites precisos, que permiten diferenciar un tipo de explotación de otra, han sido señalado asimismo como un problema sustancial por varios autores preocupados por el tema (Peón, 1992; Tort y Román, 2005).

La diversidad de fuentes utilizadas en la construcción de clasificaciones también suele mencionarse como un problema, en particular, cuando se busca replicar o extender las propuestas de clasificación a otros contextos y períodos.³

Pero sin duda, es la construcción conceptual de los tipos sociales agrarios y su definición lo que ha generado el mayor nivel de disidencia. Los múltiples enfoques y el peso diferencial otorgado a las distintas dimensiones utilizadas han dado por resultado una gran cantidad de categorías analíticas.⁴

Asimismo, la *cosificación* de estas categorías, ha generado como resultado la invisibilidad de las dimensiones que le dieron origen, y más aún, de los indicadores que la expresan empíricamente, alejándonos del principio que señala que todo tipo social es por definición una construcción teórica y una forma parcial de aproximarse al fenómeno que intentamos abordar (McKinney, 1968).

Parece apropiado entonces desagregar las dimensiones y los indicadores fundidos en las tipologías para lograr una perspectiva que permita conjugar mejor los conceptos (por definición, generalizadores) y captar las particularidades de un determinado proceso histórico. En este sentido, debemos tener presente que toda tipología es una conceptualización y conceptualizar es generalizar hasta un cierto grado (McKinney, 1968).

La reducción del espacio de propiedades y dimensiones que posee una entidad social, como las explotaciones agropecuarias, a unos pocos criterios de clasificación, implica avanzar a través de una serie de decisiones teóricas sobre el modo en que pensamos la relación entre las distintas dimensiones e indicadores, y el objeto que finalmente buscamos construir.

El modo en que pensamos la operacionalización de dimensiones como el tamaño de las explotaciones, el nivel de capitalización, la presencia o no de asalariados, el régimen de tenencia de la tierra y el uso del suelo y los valores que toman sus posibles indicadores, aportan una gran cantidad de combinaciones, que necesariamente deben considerarse bajo el marco de algún criterio teórico o metodológico de referencia.

Pero estas teorías o criterios de referencia no son neutros, ni ajenos a los procesos sociales que intentan explicar. Es necesario señalar entonces que el peso diferencial otorgado a ciertas dimensiones e indicadores influye decisivamente en la estimación de la significación de los tipos sociales resultantes.⁵

desarrollo del capitalismo, hasta programas y políticas de gobierno, destinado a productores familiares o con especial preocupación por asignar los recursos entre estos tipos de productores. Estos diferentes objetivos hacen que el resultado de la construcción conceptual sea desigual y, en ocasiones, no comparable".

³ La necesidad de contar con una estrategia de clasificación común que nos permita observar comparativamente diferentes agrupamientos se convierte al mismo tiempo en un desafío que debe saldarse, si se pretende intervenir en este campo de conocimientos.

⁴ A modo de ejemplo, y sin ánimo de ser exhaustivo, podemos mencionar, entre otras, las siguientes definiciones de tipos sociales agrarios vinculados a la región pampeana: "Productor subfamiliar", "Productor familiar", "Productor multifamiliar mediano" y "Productor multifamiliar grande" (Comité Interamericano de Desarrollo Agrícola, 1965); "Pequeña cuenta propia no propietario", "Pequeño productor propietario", "Mediano productor propietario con escaso capital", "Cuenta propia capitalizada sin tierra", "Mediano productor propietario capitalizado", "productor propietario mediano - grande capitalizado" (Tort, 1983); "Explotación familiar combinada", "Explotación agraria capitalista", "Gran empresa diversificada", "Empresa mediana chica", "Empresa agrícola especializada" (Solá, 1985); "Tipo social minifundista", "Tipo social familiar capitalizado", "Tipo social empresario agropecuario" (Caracciolo de Basco y Neiman, 1979); "Organización familiar", "Organización familiar contractual", "Organización familiar empresarial", "Organización empresarial" (Martínez, Cloquell, y otros, 1982); "Farmer", "Farmer con asalariados", "Terrateniente - capitalista", "Arrendatario familiar", "Arrendatario familiar con asalariados", "Arrendatario empresarial" (Balsa, 2006).

⁵ La conceptualización de un agro dominado por "grandes y medianas explotaciones capitalizadas", o por el contrario, la visión de un agro en donde la "explotación familiar" es predominante, puede deberse a diferentes operacionalizaciones de determinadas dimensiones de interés.

II.II Los indicadores en el diseño de tipologías

La definición de las dimensiones de interés supone asimismo el desarrollo de procedimientos mediante los cuales dichas dimensiones, se vuelven observables. Esto implica la definición teórica y metodológica de estrategias y diseños que se aproximen a las propiedades más relevantes que permitan una adecuada clasificación. La medición de las dimensiones relevantes es un proceso complejo, que implica un conjunto de aspectos y definiciones teóricas y metodológicas, que deben ser adecuadamente resueltas.

Pero esta operación reviste, como lo menciona Lazarsfeld (1973: 39), algunas dificultades. La primera de ellas se refiere al tránsito entre la imagen teórica que devuelven nuestras representaciones sobre el fenómeno que interesa investigar y sus posibles indicadores. Cuando se afirma que una explotación es "familiar" se suele sostener que dicha explotación adopta un cierto número de propiedades o atributos característicos de la familia, es decir, sus integrantes se encuentran comprometidos con la supervivencia de la unidad, se observa la transmisión del patrimonio de generación a generación, perduran ciertas prácticas de organización familiar, etcétera. El término "familiar" es entonces una forma más o menos típica de expresar de un modo abstracto un rasgo común a estas unidades productivas.

Pero la determinación de los criterios que limitan la elección de los indicadores constituye un problema delicado. En consecuencia, este proceso tiene que ser especificado para que la relación entre los indicadores seleccionados y las unidades de interés se ajusten a determinadas definiciones. Es decir, para que una explotación sea considerada "familiar" no es necesario que dicha unidad incorpore todas las características que podrían definir una unidad productiva de este tipo. Es suficiente que sea probable que realice y/o asuma determinadas características de una unidad familiar.

La relación entre cada indicador y el concepto fundamental queda definida entonces en términos de probabilidad y no de certeza. Por tal motivo, como señala Lazarsfeld, es absolutamente necesario utilizar en la medida de lo posible, un gran número de indicadores.

Sobre este punto cabe una aclaración en relación a cuáles son las características que se consideran relevantes y el peso o ponderación que se le asigna a cada una de ellas. La teoría juega un rol central en esta cuestión, dado que a partir de ella se seleccionarán los atributos que se considerarán más influyentes en su definición.

En este proceso se incluye como etapa necesaria el desarrollo de los procedimientos mediante los cuales obtendremos los datos, a partir del recorte de la realidad que hemos desarrollado a través de los indicadores seleccionados. Los procedimientos que resultan en el proceso de medición proporcionan los elementos necesarios para la observación de las dimensiones de interés, que en principio son inobservables, tal como lo plantea Samaja (1993).

La complejidad de hacer observables dimensiones o conceptos se expresa entonces en cada una de las etapas que deben transitarse en el proceso que lleva a la obtención del o los indicadores de la dimensión en cuestión. Cada operación realizada implica la presencia de procesos de validación teórica y empírica que responden, en mayor o menor medida, a las formas en que representamos socialmente los conceptos y las dimensiones.

Si bien los objetos o elementos clasificados son en su naturaleza objetos constantes y por lo tanto no son medibles en sí mismos, sí son mensurables sus propiedades (Bar, 2000: 5). La elección de estas propiedades depende, como señalábamos más arriba, del marco teórico con el que estamos trabajando pero también de una serie de decisiones referidas a las posibilidades técnicas de medición (Korn, 1971: 11).

De modo que en la elección y construcción de los indicadores, se incluyen diferentes opciones mediante las cuales los tipos sociales toman una entidad definida y específica.

Las diferencias en las escalas y niveles de medición también influyen en el modo en que se conceptualiza la estructura social agraria. El establecimiento de ciertos principios ordenadores y de determinadas formas de conceptualizar y definir las unidades de análisis, y el modo en que las mismas se constituyen en objetos con el propósito de ser agrupados o diferenciados, implica una tarea de definición teórica en donde el proceso de medición asume un rol significativo.

Pero esta tarea no es sencilla, dado que debe existir correspondencia entre los significados nominales de los conceptos del modelo y el significado operacional de las mediciones del sistema operacional (Willer, 2001: 139).

Este problema de la relación entre el plano de la teoría y el de la investigación ha sido frecuentemente señalado por la bibliografía. Blalock plantea que al respecto "parece existir una brecha intrínseca entre los lenguajes de la teoría y la investigación (...) Uno piensa en términos de un lenguaje teórico (...) pero lleva a cabo las pruebas en términos de covariaciones, operaciones y lecturas de puntajes" (Blalock, 1991).

El debate sobre el alcance de la medición en ciencias sociales vuelve a renovarse cada vez que enfrentamos un fenómeno dinámico que requiere el desarrollo de nuevos conceptos, y la definición de nuevas relaciones que surgen como expresión de las transformaciones socioeconómicas experimentadas por la estructura social, como las ocurridas en las últimas décadas.

El tradicional empleo de ciertas medidas e indicadores, sobre aspectos que se renuevan o modifican con el paso del tiempo, y la ausencia de una crítica teórica y operacional de las mismas, hace que dichos elementos tengan escasa relación con el significado nominal que promueve su comprensión.

El uso de las clásicas dimensiones vinculadas al capital, el trabajo y la tierra, merecen una reformulación y una nueva redefinición, ya sea porque el desarrollo del capitalismo agrario las vuelve más complejas en su articulación, o bien porque los fenómenos de los que dan cuenta los indicadores asociados a dichas dimensiones adquieren nuevas articulaciones. El examen de los indicadores que surgen de las diferentes dimensiones necesita entonces una nueva exploración que permita conocer tanto los aspectos positivos, como los rasgos negativos que las definen y limitan en su capacidad para dar cuenta de la realidad social agraria.

Estas dimensiones (trabajo - tierra - capital - uso del suelo) merecen una definición teórica precisa que indique cuáles son los fenómenos relevantes a los que deben referirse las observaciones que realicemos.

Por ejemplo, la definición y operacionalización de la dimensión *trabajo*, en términos de "organización social del trabajo" conduce a pensar en una serie de indicadores que presentan un sentido preciso y que dejan por fuera otro tipo de propiedades o atributos.⁶ Es decir, recortamos un posible espacio de propiedades, de los múltiples espacios teóricos posibles. El grupo de conceptos "organización social del trabajo" es teóricamente diferente entonces al concepto de "trabajo", concebido sólo por su carácter de mercancía o bien de cambio.

Este mismo ejercicio de reflexión deberíamos extenderlo al resto de las dimensiones consideradas. No basta saber que ha de medirse una dimensión o fenómeno particular, denominado tierra o capital, también es necesario saber cuál de sus propiedades es medida (Willer, 2001: 156). Como señala Blalock "las mediciones suelen aclarar nuestro pensamiento teórico y sugerirnos la consideración de nuevas variables" (Blalock, 1994: 98). Es recomendable entonces avanzar en la "ruptura" de los preconceptos que llevan a asociaciones que permanecen cosificadas en el sentido común y que conducen a mantener imágenes estáticas de la sociedad.

De modo que las definiciones operacionales y teóricas desempeñan un rol fundamental, dado que la construcción de los tipos o colectivos sociales, debe incluir tanto los aspectos lógicos que permiten una definición precisa de los agentes involucrados, como también una definición sustancial, surgida de las propiedades que permitan considerar distintas explotaciones agropecuarias como agentes pertenecientes a un mismo conjunto social. Es decir, conjuntos de actores sociológica e históricamente ubicados en una determinada posición social.

La construcción de los tipos sociales demanda entonces una reflexión sobre la forma en que se clasifican los elementos de interés y el modo en que se construyen las diferencias y similitudes

⁶ En el siguiente capítulo buscaremos construir medidas o indicadores que nos permitan observar fenómenos vinculados a la forma en que el trabajo es desarrollado en la explotación, considerando sus posibles derivaciones en términos de las formas sociales que promueve y la capacidad de generar diferentes tipos y categorías. Lo mismo intentaremos realizar con las otras dos dimensiones señaladas (tierra y capital).

entre las unidades relevadas. La clasificación implica la revisión de la forma en que un elemento es reconocido como similar a otro, pero al mismo tiempo, distinto de aquellos con los cuales mantiene diferencias en las dimensiones observadas.

El problema de la clasificación presenta al menos tres aspectos a considerar: el primero de ellos se refiere a las dimensiones e indicadores seleccionados como criterios para la diferenciación de los elementos, y junto con ello, todo el proceso que deriva en la construcción de dichos indicadores, los que buscarán dar cuenta de la dimensión observada.⁷ En segundo lugar, debe considerarse la metodología empleada en la agrupación de las unidades de análisis, es decir, a través de qué estrategias se intenta ordenar y clasificar los elementos. Un tercer aspecto a considerar se vincula con la forma en que se determina la proximidad o lejanía de un elemento respecto a otro, y los límites que diferencian a las distintas unidades.⁸ Sobre estos dos últimos aspectos avanzaremos en el próximo apartado.

II.III Algunos comentarios sobre las fuentes de datos y su construcción

Frecuentemente la descripción de la estructura social y la caracterización de los sujetos sociales agrarios se ha sustentado en la exploración de la información provista básicamente por censos, y en menor medida, encuestas. Ambas fuentes, más allá de sus características técnicas –que las vuelven más o menos confiables– se encuentran mediadas desde su propia concepción por operaciones de recorte y selección. Es importante tener presente que las fuentes consultadas (censos o encuestas) contienen y expresan determinadas perspectivas y/o enfoques sobre la realidad social a la que buscan aproximarse, y que por definición es multidimensional y compleja.

Las fuentes de información no son equiparables a una cantera en la que los datos se encuentran en estado de pureza y neutralidad, y a la espera de ser utilizados por los investigadores para que éstos, a su vez, puedan desarrollar sus análisis. Las fuentes de datos son construcciones sociales, y por tal motivo debería ser importante no perder de vista esta afirmación.

Como señala Samaja, el término "fuente de datos" se asocia a la idea de un "cierto lugar" en donde "ocasionalmente" están los datos. La palabra "fuente de datos" resume en una representación cosificada, es decir, lo que una cierta sociedad obra sobre una cierta realidad, transformándola así en una realidad intersubjetiva y, consecuentemente, en objeto de conocimiento (Samaja, 1993: 256). Entonces, toda fuente de datos constituye una cierta configuración práctica instituida socialmente.

El uso de este tipo de información para caracterizar y mensurar distintos tipos de productores requiere estar atento a algunas dificultades que emanan tanto de los cambios en las metodologías de los relevamientos, como de cierta ilusión de realismo que los datos numéricos parecieran transmitir.

La primera advertencia que habría que formular es la simple recordación de que los censos no son *la realidad*. Son sólo un discurso sobre esa realidad (más allá de su formato numérico) que se basa en la producción de datos de un determinado modo posible, entre muchos otros (Balsa, 2007: 216).

La segunda cuestión que deberíamos tener presente se vincula con la tensión que se genera entre el dinamismo que poseen los sistemas sociales y el modo en que se intenta dar cuenta del mismo. La construcción de la información censal se asocia con una mirada puntual y sincrónica de los fenómenos sociales. Entonces suelen confundirse los alcances que inherentemente poseen las fuentes, generándose interpretaciones que dichas fuentes no pueden sustentar. En este sentido, las posibilidades de observar y seguir ciertos procesos de gran relevancia e impacto (como la concentración de la producción, la agriculturización, la centralización de la tierra, etc.), a partir de

⁷ Este aspecto será abordado en profundidad en el siguiente apartado, aunque en esta parte del trabajo nos referiremos a él desde una perspectiva centrada en la reflexión epistemológica.

⁸ Esta consideración es tan relevante como las anteriores ya que la forma en que se calculan las distancias entre las unidades de interés puede modificar sustancialmente la clasificación de dichos elementos. Sin embargo se carece de estudios en donde se discuta este tipo de aspectos.

registros puntuales y no sistemáticos, es una limitación que no siempre se tiene presente y que podría distorsionar las miradas sobre las características que asumen dichos procesos.

Asimismo, los cambios en las metodologías y las redefiniciones de las variables en los diferentes momentos, es un aspecto crítico a considerar. Por lo tanto, el análisis de la definición operativa que adoptan las variables es sumamente necesario cada vez que se utilizan distintas fuentes de información.

El repaso de estas cuestiones es particularmente relevante para nuestro trabajo, dado que el mismo se construye a partir del empleo de información proveniente del Censo Nacional Agropecuario. Específicamente, el análisis que hemos desarrollado buscó extraer el máximo potencial que la información censal permite, teniendo en cuenta la disponibilidad de datos existentes.⁹

En ese sentido, el fuerte sesgo productivista que poseen los últimos censos nacionales agropecuarios condiciona en cierta medida las posibilidades que la exploración teórica permite.¹⁰ Sin embargo, estas limitaciones no deberían justificar la ausencia del análisis y la evaluación teórica de los indicadores considerados como criterios para el armado de los tipos sociales.

En relación con esto, en el siguiente apartado avanzaremos en la exploración de los criterios que han guiado la conformación de una serie de indicadores con el propósito de conocer cómo han sido construidos, y qué elementos se han priorizado en su construcción. Este análisis permitirá avanzar en la etapa de definición de los indicadores que utilizaremos para desarrollar nuestra tipología de explotaciones agropecuarias.

III. Estrategias metodológicas en la construcción de tipologías

El desarrollo de tipologías, como veíamos más arriba, se ha vinculado fuertemente a la noción de tipos ideales o cuasi ideales. Los modelos de clasificación se han ido construyendo con la intención de generar un "juicio de imputación", ordenando de un modo más o menos preciso los diferentes elementos, con el propósito de alcanzar como resultado un cuadro aprehensible del fenómeno en cuestión.

Para el caso de la construcción de los tipos sociales agrarios podríamos agrupar las propuestas en, básicamente, tres esquemas de clasificación. El primero de ellos consiste en la utilización de esquemas univariados. En este caso los tipos agrarios se conforman a partir de la identificación de determinados "cortes" o estratificaciones al interior de la variable seleccionada.

El segundo modelo se corresponde con el uso de esquemas basados en el empleo de tablas de contingencia que suponen la combinación de dos o más variables. Por último, el tercer esquema se vincularía con la utilización de técnicas de análisis multivariadas, en particular aquellas que tienen por objetivo la reducción de la dimensionalidad y la clasificación de los elementos de interés. A continuación revisaremos algunos de los aspectos centrales que intervienen en el empleo de cada uno de estos esquemas.

III.1 Los esquemas univariados y las tablas de contingencia

Los esquemas de clasificación univariados han sido ampliamente utilizados en la construcción de los tipos sociales agrarios. La ventaja de su aplicación se vincula con la posibilidad de establecer, de un modo relativamente sencillo y práctico, diferentes grupos en función del comportamiento asumido por la variable considerada en el modelo. Pero los resultados obtenidos a partir del empleo de este tipo de esquemas presentan varias limitaciones, entre ellas, la elección de

⁹ La disponibilidad de información sólo podría ampliarse realizando un relevamiento más específico, por medio de encuestas, o bien incorporando al Censo Agropecuario un mayor número de preguntas sobre dimensiones sociales, como se puede observar en los censos que se realizan en otros países, como en el caso norteamericano.

¹⁰ Es necesario aclarar que el Censo Agropecuario Experimental, realizado en el partido de Pergamino en el año 1999, relevaba ciertas dimensiones sociales, que finalmente no fueron incorporadas a la cedula del CNA 2002.

los criterios de corte que determinan los límites entre las diferentes categorías, una reducción del nivel de medición de las variables, y la posibilidad de que las variables consideradas no posean un nivel apropiado de discriminación en términos de diferenciar distintos tipos de explotaciones agropecuarias.

De este modo, variables que originalmente se expresan en niveles de medición de intervalo, como por ejemplo la superficie que opera la explotación, la cantidad de asalariados que contratan, la cantidad de maquinaria o el número de existencias ganaderas que poseen las unidades, suelen ser redefinidas a niveles ordinales (por segmentos o estratos), o definidas en términos de pequeñas, medianas o grandes, sin plantear en la mayoría de los casos una justificación teórica que las sustente y que permita asegurar que las características que asumen las explotaciones pequeñas son diferentes de aquellas que se encuentran agrupadas en los segmentos medios o grandes.

Pero además este tipo de aproximaciones, basadas en un único indicador, puede generar clasificaciones poco prácticas o demasiado generales, sobre todo cuando las variables consideradas poseen una elevada asimetría o bien las condiciones y el contexto modifican las clásicas segmentaciones.¹¹

Las limitaciones observadas en esta forma básica de clasificación, se han tratado de superar a través del empleo de esquemas bivariados o multivariados, basados en el uso de tablas de contingencia o de múltiples entradas. Estas aplicaciones buscaron captar de un modo más completo las diferencias asumidas por las distintas unidades, generando agrupamientos más sólidos y consistentes que los registrados a través de las clasificaciones univariadas, aunque en algunos puntos se sigue arrastrando con ciertas limitaciones que se observan en los modelos basados en una única variable.

Pero si bien la cantidad de variables que pueden incluirse no implicaría una limitación en sí misma, como tampoco lo sería la cantidad de combinaciones posibles, sí en cambio resulta poco práctica su aplicación.

Veamos como ejemplo el caso de una clasificación de explotaciones agropecuarias generada a través de una tabla construida a partir de dos variables: "superficie" y "presencia de asalariados", en donde cada una de ellas asume dos valores posibles (entre 1 y 199 hectáreas y más de 200 hectáreas, en el caso de la variable "superficie"; y "sí" y "no", en el caso de la variable "presencia de asalariados").¹²

En este ejemplo, son cuatro los tipos de explotaciones que podríamos identificar a partir de la combinación de los valores de las dos variables consideradas, como puede observarse en el esquema N° 1.

Esquema N° 1. Tabla de contingencia, con cuatro tipos

		Superficie	
		1 a 199 ha	Más de 200 ha
Asalariados	Sí	Tipo I	Tipo II
	No	Tipo III	Tipo IV

¹¹El proceso de estratificación, cualquiera sea el reagrupamiento seleccionado, implicará necesariamente una pérdida de información (entre las diferencias de los valores dentro de cada estrato o segmento) y, con independencia del cuidado que se ponga en justificar la pertinencia de esa decisión, estaremos optando siempre por un "mal menor", por un "poco más o menos" siempre sospechoso. Aunque sólo fuera para permitir la comparación con otras investigaciones, estos estratos suelen ser casi siempre los mismos y se encuentran, de alguna manera, naturalizados (Combessie, Jean-Claude. 2003. *El método en sociología*, Córdoba, Ferreyra Editor, pp.89).

¹² La reducción de todos los valores posibles a sólo dos valores implica una fuerte simplificación, que se vincula con los problemas mencionados para el caso de las clasificaciones basadas en una única variable. Un ejemplo de este tipo de problema puede observarse en: Ramos, M (2010) "Tipología de explotaciones agropecuarias arroceras de la provincia de Corrientes", en VII Jornadas de *Investigación y Debate Conflictos rurales en la Argentina del Bicentenario. Significados, Alcances y Proyecciones*, Universidad Nacional de Quilmes.

Ahora bien, si sumáramos una nueva variable, como por ejemplo la contratación de servicios de labores a la tipología que estamos construyendo, y asumiendo que dicha variable –al igual que las anteriores– posee solo dos valores (si contrata y no contrata), nuestra tipología pasaría a tener ocho tipos agrarios, tal como se refleja en el esquema N° 2.

Esquema N° 2. Tabla de contingencia, con ocho tipos

Superficie							
1 a 199 ha				1 a 199 ha			
		Contratismo				Contratismo	
		Sí	No			Sí	No
Asalariados	Sí	Tipo I	Tipo II	Asalariados	Sí	Tipo V	Tipo VI
	No	Tipo III	Tipo IV		No	Tipo VII	Tipo VIII

La posibilidad de sumar nuevas variables, desagregar en más categorías cualquiera de las variables consideradas, o bien, optar por una combinación de ambas alternativas (más variables y más categorías) es una opción factible y muy utilizada en los estudios agrarios.

Sin embargo, este tipo de estrategia presenta limitaciones cuando avanzamos en esta dirección. Un ejemplo lo encontramos en la clasificación realizada por Balsa (2002). En dicho trabajo el autor considera cuatro dimensiones, algunas de ellas con varias subdimensiones. En el caso de la tenencia, Balsa distingue cuatro formas de no-propiedad, y al mismo tiempo, aconseja discriminar las formas mixtas de tenencia del suelo (lo que arroja 16 situaciones). En la organización social del trabajo, el autor propone considerar además de la composición de la mano de obra (siete posiciones), la tercerización de las labores (tres opciones), y la forma de administración (dos tipos). A su vez la dimensión *tamaño* se divide en cuatro estratos y finalmente, la racionalidad de los productores en dos: de la combinatoria de todas las labores surgen "8.064" tipos de productores posibles (Balsa, 2002: 47).

Obviamente es imposible trabajar con una cantidad de tipos agrarios tan elevada. Al respecto, Balsa señala la necesidad de reducir este enorme espacio de propiedades retomando las sugerencias de Barton (1973). Este "exceso de discriminación" es abordado siguiendo como estrategia la reducción de las posibles categorías, pero manteniendo las variables de interés consideradas inicialmente.¹³

El ejercicio final resulta en la elaboración de 288 tipos agrarios, que sigue siendo una cantidad demasiado grande para resultar práctica. Si bien muchas de las posibles posiciones que expresa esta combinación no presentan casos el número total de tipos agrarios continúa siendo elevado. La contrastación empírica entonces reduce de hecho la cantidad de tipos pero esta operación no alcanza para que la tipología propuesta sea operativa, por lo que el autor debe a su vez emplear otros procedimientos para construir la tipología propuesta. Uno de los procedimientos empleados es el de resaltar alguno de todos los tipos, optando en primer término por la identificación de una serie de tipos "ideales" de productores, mientras que por otra parte se seleccionan aquellos tipos que presentaban mayor relevancia numérica, independientemente de su similitud con los tipos ideales. Finalmente Balsa utiliza una técnica de análisis estadístico para tablas de contingencia que le permitirá reducir a 9 tipos agrarios sus 8.064 tipos sociales iniciales (Balsa, 2002: 48-49).

¹³ Las cuatro dimensiones consideradas fueron: tenencia, reducida a una única variable que da cuenta de la proporción de la superficie en propiedad que tenía cada explotación, pero categorizada en cuatro posiciones posibles. La dimensión tamaño, se operacionalizó en función de la superficie total operada, discriminando cuatro categorías (pequeñas, medianas - chicas; medianas - grandes; y grandes). En la dimensión organización social del trabajo se consideraron dos indicadores: uno referido a la conformación de la mano de obra, compuesta por cuatro categorías (sin asalariados; 1-2 asalariados; 3-4 asalariados y 5 y más asalariados) y otra vinculada a la presencia de tercerización en las labores, con tres categorías (organización directa de las labores por el productor; tercerización y organización mixta). Por último la racionalidad de los productores fue estimada de un modo indirecto a través del modo de vida, en particular en función del lugar de residencia con dos posibles valores: reside en la explotación o reside fuera de ella (Balsa, 2002: 42).

Ahora bien, este punto es en sí mismo muy interesante de examinar, dado que el problema reconoce varios planos de análisis. En principio, observamos una primera cuestión que tiene que ver con el desplazamiento del plano de las variables al de las categorías, abriéndose paso a múltiples combinaciones de categorías, que es necesario comparar entre sí, para observar la existencia de relaciones entre las mismas.¹⁴

Por otra parte, se plantea el problema de la practicidad u operatividad de una clasificación que genera una enorme cantidad de tipos. Finalmente, y partir de esta situación, se presenta el interrogante referido a cómo reducir la cantidad de posiciones resultante de la combinatoria de las distintas categorías.

El modelo de tablas de contingencia contiene entonces, además de una pérdida de información (por la reducción en el nivel de medición de las variables), las dificultades asociadas más arriba, sobre todo cuando se incorporan nuevas variables y/o categorías.

Frente a estas limitaciones observadas en la metodología de tablas de contingencia, una alternativa posible es avanzar a través de análisis multivariados más complejos. Particularmente, una opción que supera las dificultades antes apuntadas lo constituyen la técnica de conglomerados o cluster.

III.11 Los esquemas multivariados y el análisis cluster

El empleo de técnicas multivariadas constituye una alternativa a las propuestas que hemos comentado en el apartado anterior. En particular la técnica de análisis de conglomerados o análisis cluster se han convertido en una opción válida para avanzar en la identificación de distintos tipos sociales agrarios, intentando superar las limitaciones que impone la generación de tipologías constituidas por una gran cantidad de tipos.

Los procedimientos multivariados tienen como punto de referencia la construcción de una matriz de distancias o proximidades entre pares de elementos (en nuestro caso, las explotaciones agropecuarias), o variables que permite cuantificar su grado de similitud o semejanza en el caso de las proximidades, o su grado de disimilitud o desemejanza en el caso de las distancias.

El análisis cluster es una técnica eminentemente exploratoria dado que la mayor parte de las veces no utiliza ningún tipo de modelo estadístico para llevar a cabo el proceso de clasificación. En este sentido, resulta una técnica adecuada para extraer información de un conjunto de datos sin imponer restricciones previas en forma de modelos estadísticos, al menos de forma explícita, y por ello puede llegar a ser muy útil como herramienta para la elaboración de hipótesis (Figueras, 2001: 1).

Esta técnica presenta la ventaja de mantener el nivel de medición original de las variables, al tiempo que es posible incluir muchas variables sin que ello genere como resultado una enorme cantidad de tipos o conglomerados, como sí ocurría en el caso de los esquemas de tablas de contingencia. De modo que se gana en precisión y detalle, al no bajar los niveles de medición, pero también al no encontrarse limitada la cantidad de variables a incluir en el análisis.

En el gráfico N° 3.4.2 se presenta la distribución hipotética de dos variables. Como podemos observar los elementos del plano se agrupan en cinco conglomerados en función de su cercanía relativa. La cercanía de los puntos se determina a través de la presencia de características similares en relación a las variables consideradas.

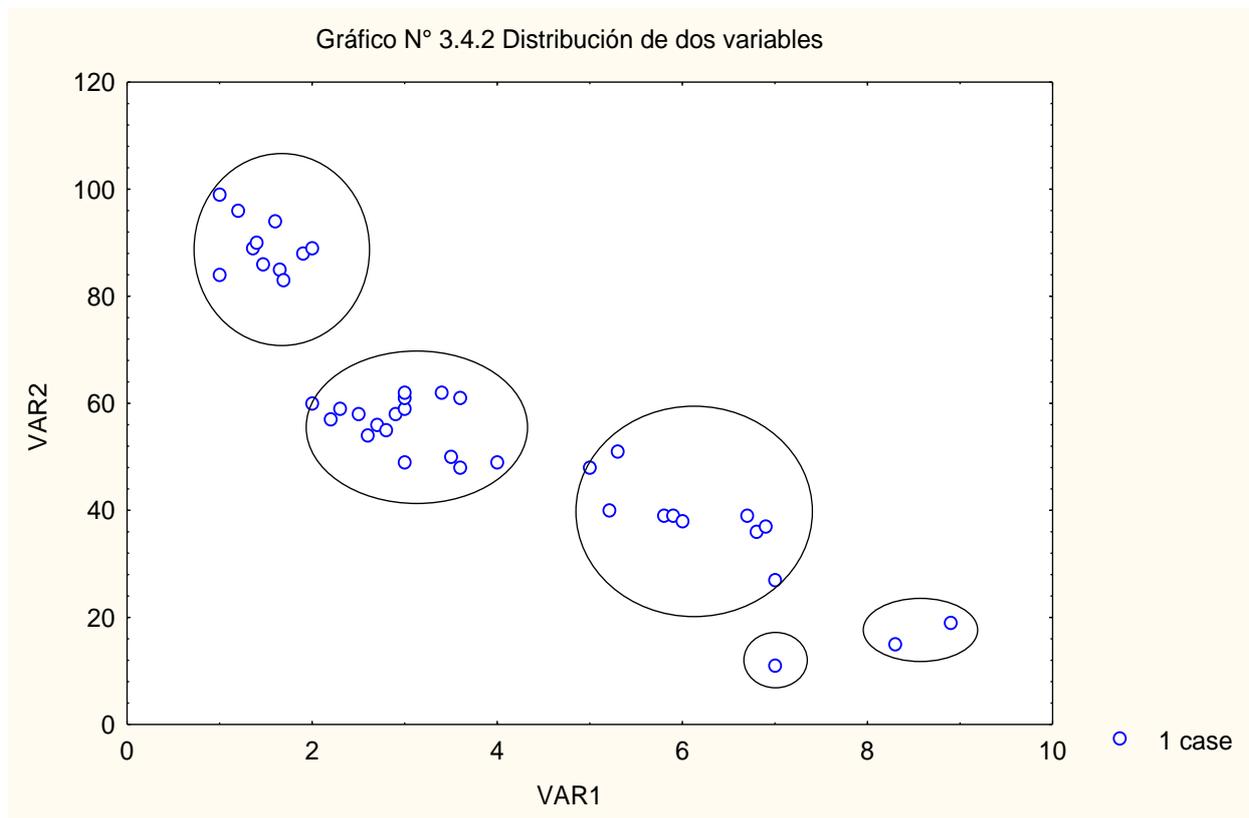
En este ejemplo, sólo se representan gráficamente dos variables, pero el esquema puede ampliarse, teóricamente, a más variables con la salvedad de que no podremos graficar la relación de más de tres variables.

De modo que la similitud entre elementos determina en este esquema su inclusión en un grupo, que entonces se caracteriza por manifestar rasgos similares entre sus componentes, al

¹⁴ El avance sobre el plano de la relación entre categorías nos lleva a la consideración de las diferentes medidas que suelen utilizarse para dar cuenta de su nivel de relación. En este sentido, es apropiado traer a cuenta la reflexión realizada por Sánchez Carrión acerca de la necesidad de olvidarnos de la posibilidad de encontrar un único número que resuma la relación entre categorías (Sanchez Carrión, 1992: 2)

tiempo que se diferencian de otros elementos que constituyen, por su parte, otros agrupamientos. La lógica del proceso de conglomerados se basa en la noción de homogeneidad entre los elementos y de heterogeneidad entre los grupos.

Gráfico N° 1. Relación entre variables.¹⁵



Sin embargo, y al igual que en el caso de las tablas de contingencia, la técnica de análisis de cluster también presenta una serie de cuestiones que debemos analizar. En primer término tenemos que señalar que las variables consideradas y su forma de inclusión, determinan la construcción de determinados tipos de grupos y no de otros.

Es decir, las variables consideradas intervienen directamente en la construcción de los grupos, y por lo tanto, es necesario que las mismas se seleccionen y justifiquen de acuerdo a los criterios teóricos que guían su utilización, y no sólo en función de su potencial poder discriminante en términos estadísticos.

Es notable observar trabajos en donde esta técnica se emplea de un modo mecánico, desplazando a un segundo plano la reflexión teórica sobre las variables utilizadas. Esto suele ocurrir particularmente cuando se construyen grupos o conglomerados a partir de la complementación en el uso de otras técnicas multivariadas, como el análisis factorial (AF).¹⁶

El uso de esta técnica expresa permanentemente la tensión entre el carácter teórico y la potencia empírica que poseen como herramientas de análisis.

En segundo término, debemos señalar que en el caso de las técnicas de cluster, existen distintos procedimientos de clasificación. Entre ellos, los más difundidos son el procedimiento de

¹⁵ Elaboración propia: *Relaciones bivariadas*. Ficha de la Cátedra de Metodología I, Carrera de Historia, Universidad Nacional de La Plata, 2010.

¹⁶ Al respecto, es interesante retomar el planteo de Weber cuando señala que las tipologías proceden de una elección. Estos puntos de vista (...) no son en modo alguno los únicos posibles. Al igual que ocurre con cada fenómeno histórico, otros puntos de vista –otras variables o criterios– harían aparecer otros rasgos como "esenciales" (Weber, Max, 1984).

análisis definido mediante la formación de "*conglomerados jerárquicos*" y el denominado "*análisis de conglomerados de K medias*".

La elección de uno u otro tipo de procedimiento tampoco es neutral en sus resultados, dado que en el caso del procedimiento de conglomeración jerárquica no se determina a priori la cantidad de grupos a formarse, mientras que esto sí ocurre en el caso del método de k-medias. En este último procedimiento, la determinación previa de la cantidad de grupos posibles suele forzar la inclusión de los elementos en los grupos definidos a priori, con el riesgo de generar agrupamientos poco homogéneos.

Este problema no ocurre en el caso del procedimiento de conglomeración jerárquica. Sin embargo, este procedimiento al permitir la combinación de distintos métodos de conglomeración y formas de calcular las distancias entre los elementos, genera como resultado diferencias en los posibles agrupamientos en función del tipo de método y distancia que se utilice. Su definición es, por lo tanto, determinante aunque no siempre es debidamente evaluada en sus aplicaciones.

Ahora bien, además de las cuestiones referidas a los criterios que guían la introducción de las diferentes variables en el modelo (cuestión que también es fundamental en la construcción de tipologías mediante tablas de contingencia), debemos estar especialmente atentos a tres aspectos centrales que forman parte del proceso de construcción de los conglomerados.

El primero de ellos se vincula con la forma en que se mide la distancia entre los elementos a clasificar. La distancia entre los diferentes elementos se puede calcular a través de distintas medidas y no existe consenso sobre cuál es el mejor modo de calcular las diferencias entre los casos (o variables).¹⁷

La segunda cuestión se refiere al método de aglomeración que se utiliza en el armado de los grupos. Al igual que en el caso de las distancias, existen diferentes alternativas, que determinan distintos resultados en relación al armado de los posibles grupos. Por último, el tercer aspecto a considerar en la utilización de la técnica de conglomerados se vincula con la métrica y la variabilidad que poseen las variables que se incluyen en el modelo. Este aspecto también es clave cuando las variables consideradas poseen unidades de medición diferentes y/o expresan grandes variaciones.¹⁸

III.III Las técnicas de conglomeración: métodos y tipos de medidas

Las técnicas de clasificación se utilizan para generar agrupamientos de objetos o individuos, en función del comportamiento que asumen dichos elementos en un cierto número de variables o características. El objetivo de esta técnica es obtener grupos de elementos de forma que, por un lado, los objetos pertenecientes a un mismo grupo sean semejantes entre sí, es decir, que el grupo esté cohesionado internamente y, por el otro, los objetos pertenecientes a grupos diferentes tengan un comportamiento distinto con respecto a las variables analizadas, es decir, que cada grupo se diferencie de los demás.

Específicamente, el análisis de conglomerados jerárquicos se construye a través de una secuencia de pasos que finalizan cuando se logra obtener un único conglomerado formado por el conjunto de los elementos considerados. El primer paso entonces se inicia con el cálculo de la matriz de distancias entre los elementos relevados. Esta matriz contiene las distancias existentes entre cada elemento y las restantes unidades.

El siguiente paso consiste en identificar los dos elementos más próximos, en términos de la distancia existente entre ellos, para conformar el primer conglomerado. La importancia de

¹⁷ Frecuentemente, los trabajos en donde se utiliza la técnica de cluster, no suelen incluir ningún comentario o reflexión acerca de la posibilidad de obtener diferentes agrupamientos a partir del uso de distintos métodos de conglomeración y medidas de distancia. Algunos ejemplos en este sentido pueden encontrarse en: Natera Rivas (2005); Fariña Gómez y otros (2004).

¹⁸ Es el caso de las variables referidas a la dotación de ciertos factores o la posesión de determinados recursos. Las variables relacionadas con el tamaño de la explotación, en términos de superficie, o el nivel de capitalización que posee la misma, presentan un elevado nivel de variabilidad, como así también unidades de medición diferentes.

considerar distintos tipos de distancias radica justamente en que diferentes formas de cálculo pueden arrojar agrupamientos diferentes entre los elementos considerados.

Dicho conglomerado es indivisible a partir de ese momento, dando sentido a la denominación de "jerárquico". Luego se irán agrupando, en sucesivos pasos, el resto de los elementos en distintos grupos hasta conformar finalmente un único conglomerado que incluirá a todas las explotaciones.¹⁹

En relación a su aplicación práctica como herramienta para la clasificación de unidades agropecuarias pampeanas ha sido escasamente utilizada.²⁰ Sí, en cambio, ha tenido un uso relativamente difundido en la clasificación de unidades espaciales agregadas, como los departamentos provinciales,²¹ los conglomerados urbanos,²² como así también en la clasificación de unidades productivas de espacios extrapampeanos.²³

Desde un punto de vista teórico, la decisión de optar por una reducción de la dimensionalidad que presenta el fenómeno estudiado a través de la extracción de factores, hace que se desplace del centro de interés la reflexión teórica sobre la potencialidad explicativa de los indicadores incluidos, por una elección guiada mediante otro tipo de lógica, basada en el cálculo de correlaciones.

En algunos análisis este problema se vuelve preocupante, sobre todo cuando las dimensiones latentes se conforman a partir de indicadores que teóricamente no expresan diferencias sustanciales o directamente no aportan elementos a la teoría que la sustenta. Sólo en situaciones en donde la exploración de los datos se hace sin un conocimiento previo de sus propiedades y covarianzas, es recomendable la aplicación de esta técnica. Pero siempre que exista algún conocimiento sobre el fenómeno que se intenta estudiar, la aplicación del análisis factorial debería cumplir un estricto rol de control, y de ningún modo, reemplazar el análisis teórico de los indicadores que se seleccionan.

III.IV Variables, medidas, distancias y métodos utilizados

En relación a la aplicación de esta técnica debemos especificar una serie de cuestiones metodológicas, que inciden de forma directa en la determinación de los conglomerados.

En primer término, las variables que se incluyen en el modelo de clasificación no sólo se deben justificar teóricamente, sino que también es recomendable que posean un adecuado nivel de variación. En relación con el método de conglomerados jerárquicos, las variables que mejor discriminan son aquellas que presentan una elevada heterogeneidad.

Ahora bien, para avanzar en el proceso de agrupamiento es necesario calcular la distancia que existe entre los elementos a partir de las variables consideradas. Este es un aspecto clave en el armado de los conglomerados. El procedimiento "análisis de conglomerados jerárquicos" que seguiremos a continuación permite elegir entre una amplia variedad de distancias, en función del

¹⁹ En cada paso del proceso pueden agruparse elementos individuales, conglomerados previamente formados o un caso individual con un conglomerado previamente formado.

²⁰ Cfr: Truffer y otros (2003); Djenderedjian (2001).

²¹ Cfr: Lago Martínez, Mauro y Alvarez (2000); y Natali (s/f).

²² Cfr: Oliva (1997)

²³ Cfr: Paz (1994); Coronel de Renolfi y Ortuño Pérez (2005).

nivel de similaridad²⁴ o disimilaridad²⁵ que deseamos observar, en nuestro caso de análisis, entre las explotaciones agropecuarias que conforman nuestra población de interés.²⁶

Por otra parte, otro de los aspectos relevantes a considerar se vincula con la inclusión de las dimensiones de interés y las características que las mismas buscan describir. Específicamente, el aspecto a tener presente se refiere a la incidencia que poseen los indicadores utilizados en la construcción de los conglomerados. Deberemos estar atentos a la posibilidad de asignar a ciertas variables una ponderación diferencial, de modo que las mismas no sobredeterminen la conformación de los grupos.

Asimismo, es pertinente considerar el problema vinculado a las magnitudes que expresan las distintas variables. Es decir, cuando las variables poseen diferentes unidades de medida, es conveniente unificar las magnitudes y transformar sus valores en alguna medida homogénea. Las medidas de distancia no son invariantes respecto a la métrica que poseen las variables. De modo que las diferencias existentes entre las variables con valores muy altos pueden anular las diferencias existentes entre las variables con valores bajos.

Frecuentemente, los diferentes indicadores utilizados no sólo presentan distintas unidades de medida, sino que también poseen amplitudes muy variables. Si tomamos indicadores con unidades de medida diferentes, al calcular la distancia entre dos elementos, las diferencias en la variable con mayor variación tendrán mucha más influencia que el impacto de la diferencia en la variable con menor variación relativa.

Para controlar este problema se debe emplear algún tipo de transformación de los valores originales en puntuaciones de un mismo rango de variación. Entre las transformaciones posibles encontramos el rango 0 - 1,²⁷ las "puntuaciones Z",²⁸ el rango "-1 a 1",²⁹ la Magnitud máxima de 1,³⁰ la Media 1³¹ y la Desviación Típica 1, habitualmente denominada "estandarización".³²

Por último, un aspecto que también debemos tener en cuenta se refiere al modo en que se agrupan los elementos. El análisis de conglomerados jerárquico avanza "etapa a etapa", uniendo en cada paso los dos elementos de la matriz de distancias que se encuentran más próximos entre sí. En cada etapa se funden dos elementos o grupos. Una vez calculada la matriz de distancias, los dos elementos más próximos (los más similares o menos distantes) son fundidos en un mismo conglomerado. Estos dos casos, que forman el primer conglomerado constituyen una unidad que como tal, posee su propia distancia respecto al resto de los elementos de la matriz de distancias. La

²⁴ Las medidas de similaridad miden el grado de proximidad entre los elementos. De modo que los valores más altos indican mayor proximidad entre los elementos comparados. Cuando dos elementos se encuentran juntos, el valor de las medidas es máximo.

²⁵ Las medidas de disimilaridad en cambio, se basan en la diferencia o lejanía existente entre elementos. Los valores más altos indican en este caso mayor diferencia o lejanía entre los objetos comparados. Si dos elementos se hallan juntos la distancia entre los mismos es nula.

²⁶ Entre las diferentes formas de calcular la distancia entre las explotaciones agropecuarias, se encuentran la Distancia Euclídea al Cuadrado, la Distancia Euclídea, Coseno, Correlación de Pearson, Chevychev, Bloques, Minkowsky perteneciente al grupo de las medidas de disimilitud.

²⁷ A cada valor se le resta el valor más pequeño de la variable y esa diferencia se divide entre el rango o amplitud de dicha variable. Para calcular estos valores se toma la siguiente expresión: $X_i - X_{\min} / \text{amplitud}$.

²⁸ En el caso de la transformación a valores Z, a cada valor se le resta la media de la variable y esa diferencia se divide por la desviación típica de la variable, obteniéndose valores estandarizados con media 0 y desviación típica 1.

²⁹ Para el rango -1 a 1, el procedimiento consiste en dividir cada valor por el rango o amplitud de la variable, obteniéndose valores estandarizados con amplitud 2 en una escala cuya unidad de medida es el rango o amplitud de la variable.

³⁰ Con respecto a la "Magnitud máxima de 1", la transformación se realiza tomando cada valor y dividiéndolo por el valor más grande de la variable, obteniéndose valores estandarizados con un máximo de 1 y un mínimo variable pero nunca menor de 0.

³¹ La transformación denominada "Media 1", divide cada valor por la media de la variable. Los resultados obtenidos poseen una media igual a 1, y una escala cuya unidad de medida es la media de la distribución de la variable.

³² La transformación denominada como "Desviación típica 1" divide cada valor por la desviación típica de la variable. Los valores estandarizados poseen una desviación típica igual a 1 y una escala en donde la unidad de medida es la desviación típica media de la variable.

matriz inicial de las $n \times n$ explotaciones agropecuarias (o $p \times p$ variables) cambia, dado que dos de sus filas –y dos de sus columnas– han sido fundidas en una, transformándose en una matriz $(n - 1) \times (n - 1)$.

Luego de recalculer las distancias, en la siguiente etapa del análisis se vuelven a seleccionar los dos elementos de la matriz más próximos entre sí, los cuales son fundidos en un nuevo conglomerado. Los dos elementos fundidos en esta segunda etapa pueden ser dos casos individuales o un caso individual y el conglomerado ya formado en la primera etapa. En este momento, la matriz de distancias de dimensiones $(n-1) \times (n-1)$ se transforma en una matriz de distancias de dimensiones $(n-2) \times (n-2)$, lo que exige volver a calcular las distancias del nuevo conglomerado respecto al resto de elementos de la matriz.

El proceso continúa etapa tras etapa hasta que, finalmente, se consigue fundir en un único conglomerado a todos los elementos de la matriz de distancias (de dimensiones finales 2×2). De modo que los métodos de conglomeración son los procedimientos mediante los cuales es posible volver a calcular las distancias entre los nuevos elementos en cada etapa del proceso de fusión.

El método de conglomeración utilizado para recalculer las distancias en cada etapa del proceso de fusión puede determinar de manera sustantiva la calidad de la solución alcanzada. Ante esta situación, es recomendable explorar diferentes métodos de conglomeración con el propósito de analizar, comparar y evaluar sus resultados y alcances en relación a los objetivos que se persiguen.

La exploración y el análisis de las tablas referidas al historial de conglomeración y los gráficos dendogramas que los software generan, permiten observar la presencia de situaciones diferenciales en cuanto a los resultados obtenidos y los posibles alcances que pueden establecerse en función de los propósitos establecidos.

Frecuentemente, los métodos de aglomeración suelen ofrecer dos tipos de soluciones en relación a la clasificación de unidades o explotaciones. En un primer esquema, hallamos un grupo de métodos que arrojan como resultado central la posibilidad de distinguir casos con características muy particulares, agrupando por un lado a la gran mayoría de las explotaciones y diferenciando, por el otro, grupos con unos pocos casos. Dentro de este grupo se encuentran los métodos: "agrupación de medianas",³³ "entre-grupos",³⁴ "vecino más próximo",³⁵ y "agrupación de centroides".³⁶

En el caso del segundo esquema de clasificación, encontramos un grupo de métodos que se caracterizan por brindar soluciones en donde los conglomerados se encuentran integrados por cantidades similares de explotaciones, o al menos no presentan distribuciones fuertemente sesgadas hacia un conglomerado en particular. Los métodos que integran este segundo grupo de

³³ En el método de aglomeración *agrupación de medianas* los dos conglomerados (o elementos) que se combinan reciben idéntica ponderación en el cálculo del nuevo centroide combinado, independientemente del tamaño de cada uno de los conglomerados (o elementos). Esto permite que, al momento de caracterizar los conglomerados resultantes, los conglomerados pequeños tengan la misma importancia que los conglomerados grandes. Dado un conglomerado AB y un elemento C , la nueva distancia del conglomerado al elemento se calcula como: $d_{(AB)C} = (d_{AC} + d_{BC} / 2) - (d_{AB} / 4)$

³⁴ El método *entre grupos* o de vinculación promedio aprovecha la información de todos los miembros de los dos conglomerados que se comparan. La distancia entre dos conglomerados se calcula como la distancia promedio existente entre todos los pares de elementos de ambos conglomerados: $d_{AB} = (1 / n_A n_B) \sum_{i \in A} \sum_{j \in B} d_{ij}$

³⁵ El método de conglomerados definido como *vecino más próximo* o también denominado como de vinculación simple o enlace simple, comienza seleccionando y fundiendo los dos elementos de la matriz de distancias que se encuentran más próximos. La distancia de este nuevo conglomerado respecto a los restantes elementos de la matriz se calculan como la menor de las distancias entre cada elemento del conglomerado y el resto de los elementos de la matriz. En los pasos sucesivos la distancia entre dos conglomerados se calcula como la distancia entre sus dos elementos más próximos. Así la distancia d_{AB} entre los conglomerados A y B se calcula mediante: $d_{AB} = \min (d_{ij})$

³⁶ El método *agrupación de centroides* calcula la distancia entre dos conglomerados como la distancia entre sus vectores de medias. Con este método, la matriz de distancias original sólo se utiliza en la primera etapa. En las etapas sucesivas se utiliza la matriz de distancias actualizada en la etapa previa. En cada etapa el algoritmo utiliza la información de los dos conglomerados (o elementos) fundidos en la etapa previa y el conglomerado (o elemento) que se intentará fundir en esa etapa. La distancia entre el conglomerado AB y el elemento C se calcula como: $d_{(AB)C} = (n_A / n_A + n_B) d_{AC} + (n_B / n_A + n_B) d_{BC} - [n_A n_B / (n_A + n_B)^2] d_{AB}$

soluciones son el denominado "método del vecino más lejano",³⁷ el "método ward"³⁸ y "método Intra-grupos".

En el caso que hemos analizado, el primer grupo de métodos tiende a mantener y/o sobrerrepresentar las diferencias, mientras que el segundo grupo de métodos presenta características opuestas, es decir, tiende a subrepresentar las diferencias generando agrupamientos más diversos y heterogéneos en su interior.

IV. Aportes y limitaciones de las técnicas consideradas en la construcción de tipologías de sujetos agrarios.

En este punto del trabajo creemos importante retomar algunas cuestiones planteadas al inicio, cuando sosteníamos el carácter conceptual que asume toda construcción de tipologías. En este sentido, las técnicas que hemos presentado anteriormente no deben pensarse como estrategias independientes y neutrales, sino que por el contrario, es preciso señalar el carácter teórico que debe guiar su aplicación y evaluar el alcance que poseen en la determinación de los tipos agrarios.

Las técnicas consideradas permiten retrotraer la reflexión al punto de partida con el que iniciamos este artículo y que implica la propia construcción de las categorías sociales. Dado que estas categorías generalmente se encuentran inmovilizadas bajo conceptualizaciones sedimentadas, las estrategias metodológicas que hemos presentado permiten poner en cuestión la construcción de los tipos sociales agrarios y en consecuencia, desagregar las dimensiones y los indicadores fundidos en dichas categorías.

La posibilidad de captar las particularidades de un determinado proceso histórico también es un aspecto relevante a resaltar y que no suele considerarse habitualmente. La particular combinación de dimensiones y el peso que poseen en la determinación de los grupos que se construyen, permite una mejor comprensión de los fenómenos socioeconómicos que influyen en la conformación de la estructura agraria. La contribución de las técnicas al proceso de visualización de las dimensiones que dan cuenta de los tipos agrarios, es un aporte central, que también permitiría una evaluación más precisa del alcance de las tipologías clásicas de la sociología rural en el presente.

Desde el clásico modelo tripartito de agro inglés definido por Marx, pasando por las elaboraciones de Kautsky y Lenin, hasta las clasificaciones más recientes de la sociología rural, la construcción de tipologías ha tenido como principal carencia la estimación de la magnitud de los tipos agrarios y su confrontación empírica. En este sentido, la posibilidad de avanzar en una estimación de la cantidad y la significación territorial de los tipos construidos podría aportar un mejor conocimiento de las características que definen el perfil social del agro. Pero además nos coloca en una posición más sólida para el análisis teórico de las dinámicas que asumen los diferentes tipos, evitando discusiones que se asientan en combinaciones teóricas, sin su correspondiente correlato empírico.

³⁷ El método de vinculación por el *vecino más lejano*, se comporta de manera opuesta al método de vinculación por el vecino más próximo. La distancia entre dos conglomerados se calcula como la distancia entre sus dos elementos más alejados. Es decir, la distancia entre dos conglomerados A y B se calcula como: $d_{AB} = \max (d_{ij})$

³⁸ Bajo la premisa de perder la menor cantidad de información posible como resultado de la fusión de los elementos, la cantidad de información se cuantifica como la suma de las distancias al cuadrado de cada elemento respecto al centroide del conglomerado al que pertenece (*SCE = Suma de Cuadrados Error*). Para ello, se comienza calculando, en cada conglomerado, el vector de medias de todas las variables, es decir, el centroide multivariante. A continuación, se calculan las distancias euclídeas al cuadrado entre cada elemento y los centroides (vector de medias) de todos los conglomerados. Por último, se suman las distancias correspondientes a todos los elementos. En cada paso se unen aquellos conglomerados (o elementos) que dan lugar a un menor incremento de la *SCE*, es decir, de la suma de cuadrados de las distancias intra-conglomerado. La *SCE* se define como: $SCE = \sum_{i=1}^n [\sum_{j=1}^n x_{ij}^2 - 1 / n_j (\sum_{i=1}^n x_{ij})^2]$

El uso de las técnicas presentadas permite entonces realizar una evaluación de la eficacia de las categorías construidas. El mantenimiento de los tradicionales esquemas ha implicado una suerte de despojo de su carácter histórico, que las técnicas consideradas podrían revertir.

V. Comentarios finales

Debemos señalar finalmente las implicaciones que presentan las diferentes estrategias de clasificación consideradas en los apartados anteriores con respecto a la construcción de los tipos sociales en el espacio agrario pampeano.

La complejidad que asumen los productores agrarios impone ciertas limitaciones a la aplicación de esquemas univariados para la clasificación de los mismos.

Asimismo, la opción de incluir de modo simultáneo muchas variables bajo el esquema de tablas de contingencia, si bien permite reconocer la diversidad de sujetos agrarios y la especificidad del agro pampeano, suele generar una gran cantidad de tipos sociales que le quitan practicidad.

Entonces, parece interesante avanzar en la posibilidad de encontrar una alternativa a estas estrategias en el uso de esquemas multivariados que, como señalábamos más arriba, no presentan restricciones en la cantidad de variables a incluir en el modelo de clasificación, y a su vez, ofrecen una solución acotada en términos del número de tipos agrarios que se construyen.

En particular, la utilización de la técnica de conglomerados jerárquicos se muestra como la alternativa más idónea para la construcción de los tipos sociales. En principio, la técnica de conglomerados implica una etapa previa de reflexión teórica y de definición empírica en relación a los indicadores que serán incluidos en el modelo de clasificación.

Pero también, la aplicación de este procedimiento requiere del análisis de las diferentes alternativas que surgen de la combinación de los distintos métodos y formas de calcular la distancia entre los elementos de interés. Sobre todo porque no existe consenso sobre cuál es el método y el tipo de distancia más adecuado para cada problema en particular. Por lo tanto, en esta etapa se debe realizar una evaluación de las opciones que resulten más adecuadas en términos de los objetivos teóricos y operativos que se plantea cada aplicación.

Los resultados obtenidos deben evaluarse en función de los planteos teóricos desarrollados en la primera etapa, como así también en relación a las características que definen la problemática abordada. En este sentido, la aplicación de esta técnica implica asumir una actitud de apertura conceptual hacia la posibilidad de obtener tipos sociales agrarios novedosos o poco convencionales en términos de las tradiciones teóricas con las que solemos trabajar.

Finalmente, cabe señalar como última cuestión el aporte que las diferentes técnicas brindan en la estimación de la significación numérica y territorial de los distintos tipos agrarios. Este aspecto es fundamental dado que ofrece información relevante para el diseño de políticas públicas diferenciales y específicas para cada tipo de sujeto agrario.

REFERENCIAS BIBLIOGRÁFICAS.

ANDER-EGG, Ezequiel (1995) *Técnicas de Investigación Social*. Buenos Aires: Humanitas.

BALSA, Juan Javier (2002) "La conceptualización de los productores rurales y su transformación a lo largo del siglo XX. Reflexiones a partir de la confrontación con los estudios sobre el agro norteamericano". Jornadas *Empresas y empresarios rurales en la Argentina del siglo XVIII al XX*, Universidad Argentina de la Empresa. Buenos Aires, 10 y 11 de junio de 2002.

_____ (2006) *El desvanecimiento del mundo chacarero*. Bernal, Buenos Aires: Universidad Nacional de Quilmes Editorial.

_____ (2007) "Fuentes y metodologías para el abordaje de algunos problemas en la Historiografía del agro pampeano del Siglo XX", en GRACIANO, O. y S. LÁZZARO (comp.) *La Argentina rural del siglo XX. Fuentes, problemas y métodos*. Buenos Aires: La Colmena.

BAR, Aníbal (2000) "Un aporte a la discusión sobre el status metodológico de las variables y escalas de medición". *Cinta de Moebio*, N° 7. Facultad de Ciencias Sociales, Universidad de Chile.

BARTON, Allen (1973) "Concepto de espacios de atributos en sociología", en BOUDON, Raymond y P. LAZARFELD. *Metodología de las ciencias sociales, I. Conceptos e índices*. Barcelona: Laia.

BECKER Howard (1956) "Sociología interpretativa y tipología constructiva", en: GURVITCH, George y Wilbert MOORE (ed.) *Sociología del siglo XX*. Buenos Aires: Ateneo.

BLALOCK, Hubert (1991) *Causal Inferences in Nonexperimental Research*. Chapel Hill: University of North Carolina Press.

CARACCILO DE BASCO, Mercedes y Guillermo NEIMAN (1979) *Evaluación socio-económica de las explotaciones agropecuarias constituidas en la provincia de Buenos Aires en el período abril 1976 a diciembre 1977*. Publicación ESR/119/79. Buenos Aires: SEAyG.

COMITÉ INTERAMERICANO DE DESARROLLO AGRÍCOLA (1965) *Tenencia de la tierra y desarrollo socioeconómico del sector agrícola*. Argentina: Washington, Ed. Unión Panamericana, OEA.

CORONEL DE RENOLFI, Marta y Sigfredo Francisco ORTUÑO PÉREZ (2005) "Tipificación de los sistemas productivos agropecuarios en el área de riego de Santiago del Estero, Argentina", en: *Problemas del Desarrollo. Revista Latinoamericana de Economía*, vol. 36, N° 140. Disponible en: <http://www.revistas.unam.mx/index.php/pde/article/view/7572>

DE MARTINELLI, Guillermo (2007) "Construcción de tipologías para explotaciones agropecuarias pampeanas. Algunas consideraciones sobre problemas, fuentes y métodos", en GRACIANO, O. y S. LÁZZARO (comp.) *La Argentina rural del siglo XX. Fuentes, problemas y métodos*. Buenos Aires: La Colmena.

DJENDEREDJIAN, Julio (2001) "Análisis estadístico aplicado a un caso de frontera ganadera: sociedad y economía en el Norte Entrerriano a fines de la época colonial", ponencia presentada en las *VIII Jornadas Interescuelas / Departamentos de Historia*. Salta.

FARIÑA GÓMEZ, Beatriz; GORDO GÓMEZ, Pablo; DE LOS RÍOS RODICIO, Ángel y Beatriz RODRÍGUEZ PRADO (2004) "Identificación y tipificación de los espacios rurales en la Comunidad Autónoma de Castilla y León. Una aproximación para el análisis de las dinámicas espaciales", ponencia presentada al *V Congreso de Economía Agraria*. Santiago de Compostela. Disponible en: http://www.emp.uva.es/~pablog/espacio_rural_cl.pdf

FIGUERAS, Manuel (2001) "Análisis de conglomerados o cluster", [en línea] *5campus.org, Estadística*. Disponible en <http://www.5campus.org/leccion/cluster>

KORN, Francis (1971) "El significado del término 'variable' en sociología", en: AAVV. *Conceptos y variables en la investigación social*. Buenos Aires: Cuadernos de Investigación social, Ediciones Nueva Visión.

LAGO MARTÍNEZ, S.; ÁLVAREZ, G. y M. MAURO (2000) "La conformación de subregiones al interior de cuatro Provincias Argentinas, según el impacto en el Desarrollo de las Condiciones de Vida". *Revista electrónica de Epistemología de las Ciencias Sociales Cinta de Moebio*, N° 9, Noviembre. Universidad de Chile, Facultad de Ciencias Sociales.

LAZARFELD, Paul (1973) "De los conceptos a los índices empíricos", en BOUDON, Raymond y P. LAZARFELD. *Metodología de las ciencias sociales, I. Conceptos e índices*. Barcelona: Laia.

MARTÍNEZ, Alfredo, Silvia CLOQUELL y otros (1982) "Diagnóstico de las limitantes al aumento de la productividad en el sur santafecino". Convenio INTA-MAG-UNR. Rosario. *Inédito*

MCKINNEY, John C. (1968) *Tipología constructiva y teoría social*. Buenos Aires: Amorrortu editores.

NATALI, Pamela Mariel (s/f). "Análisis Clusters: una aplicación a los departamentos de la provincia de Córdoba". Disponible en: <http://www.eco.unrc.edu.ar/wp-content/uploads/2010/04/TRABAJO-N%C2%BA-32.pdf>

NATERA RIVAS, Juan José (2005) "La población rural del noroeste argentino. Un ensayo reclasificación departamental a través de métodos de análisis multivariante", en: *Revista Geografía, Espacio y Sociedad*. Vol. 1, N° 1. Disponible en: http://www.mapasderecursos.org.ar/archivos_biblioteca/1256242123_Natera_Rivas_La_poblacion_rural_del_NOA.pdf

OLIVA, Miguel (1997) "Evolución del Mercado Laboral en los aglomerados urbanos de Argentina en el período 1994-1996", *Serie de Estudios Metodológicos* N° 6. Disponible en: http://www.catedras.fsoc.uba.ar/salvia/catedra/series/6_1.htm

PAZ, Raúl (1994) *Aproximación cuantitativa del sector campesino en la provincia de Santiago del Estero*. Santiago del Estero: Ed. Programa Social Agropecuario - Secretaría de Agricultura, Ganadería y Pesca de la Nación.

PEÓN, César (1992) *Sociología rural latinoamericana. Hacendados y campesinos*. Buenos Aires: CEAL.

SAMAJA, Juan (1993) *Epistemología y metodología*. Buenos Aires: EUDEBA.

SANCHEZ CARRIÓN, Juan Javier (1992) *Análisis de tablas de contingencia*. Madrid: CIS-Siglo XXI.

SOLÁ, Felipe (1985) "Empresas agrícolas, diferenciación, rentabilidad e impactos de políticas alternativas". *Documento n°12*. Buenos Aires: CISEA.

TORT, María Isabel (1983) "Tecnificación agraria y formas de producción. Su impacto en las condiciones de vida rural. Un estudio de caso". Tesis del curso tutorial de posgrado, FLACSO-Argentina. *Inédita*

TORT, M. I. y M. ROMÁN (2005) "Explotaciones familiares: diversidad de conceptos y criterios operativos", en: GONZÁLEZ, María del Carmen (coord.). *Productores familiares pampeanos: Hacia la comprensión de similitudes y diferenciaciones zonales*. Buenos Aires: Astralib.

TRUFFER, Isabel; TENTOR, Gonzalo; LUJÁN, Ana y CECCOTI, Laura (2003) "Las zonas agro ecológicas y los tipos sociales agrarios (Entre Ríos-Argentina)", en *Actas de las III Jornadas Interdisciplinarias de Estudios Agrarios y Agroindustriales*, Facultad de Ciencias Económicas, UBA, noviembre, Buenos Aires.

WEBER, Max ([1922] 1984) *Economía y Sociedad. Esbozo de sociología comprensiva*. México: FCE.

WILLER, David (2001) *La sociología científica. Teoría y método*. Buenos Aires: Amorrortu.

Autor.

Guillermo De Martinelli.

Universidad Nacional de Quilmes / Universidad Nacional de la Plata / Consejo Nacional de Investigaciones Científicas y tecnológicas (CONICET). Argentina.

Profesor en Historia (UNLP). Posgrado en Generación y Análisis de Información Estadística, Universidad Nacional de Tres de Febrero.

E-mail: gmartinelli@unq.edu.ar <mailto:jymena@terra.com.pe>

Citado.

DE MARTINELLI, Guillermo (2011). "De los conceptos a la construcción de los tipos sociales agrarios. Una mirada sobre distintos modelos y las estrategias metodológicas". *Revista Latinoamericana de Metodología de la Investigación Social - ReLMIS*. N° 2. Año 1. Oct. 2011 - Marzo 2012. Argentina. Estudios Sociológicos Editora. ISSN 1853-6190. Pp. 24 - 43. Disponible en:

<http://www.relmis.com.ar/ojs/index.php/relmis/article/view/22/19>

Plazos.

Recibido: 06 / 06 / 2011. Aceptado: 29 / 09 / 2011.