ERAMERICANO DE CIENCIAS AGRICOLAS - OEA

ión del Plan de Acción del IICA en Uruguay

N DE TECNICAS ESTADISTICAS PARA
CION DE EMPRESAS AGROPECUARIAS

HUGO E. COHAN
ALFREDO ALONSO

Montevideo, Junio 1977

APLICACION DE TECNICAS ESTADISTICAS PARA TIPIFICACION DE EMPRESAS AGROPECUARIAS

Hugo E. Cohan (IICA - Uruguay)

Alfredo Alonso E. (DIEA-MAP-ROU)

Presentado al "I SEMINÁRIO DE MODERNIZAÇÃO DA EMPRESA RURAL"
Ministério da Agricultura - Fundação Getúlio Vargas
Rio de Janeiro, 25 al 27 de mayo de 1977

Revisado en Montevideo en junio de 1977

I N D I C E

* * * * * *

APLICACION DE TECNICAS ESTADISTICAS PARA TIPIFICACION DE EMPRESAS AGROPECUARIAS

I. INTRODUCCION Y OBJETIVOS DEL DOCUMENTO

En el presente documento se informa sobre lo actuado en Uruguay por un grupo de instituciones y técnicos interesados en el problema de tipificación de empresas agropecuarias. La presentación se ha elaborado como resumen de aportes escritos por el grupo de trabajo. Estos documentos constituyen por sí solos el conjunto de referencias citadas, en las que podrá encontrarse el material referencial más básico.

El grupo de trabajo que está en esta tarea incluye técnicos de varias reparticiones nacionales del Uruguay, nucleados en función de su interés profesional en el tema y sus vínculos operativos con el Instituto Interamericano de Ciencias Agrícolas. El apoyo metodológico fundamental ha provenido del Centro Interamericano de Enseñanza de la Estadística (CIENES)**.

El presente documento tiene por objetivos informar sobre:

a) un enfoque dado en Uruguay al problema de tipificación de empresas agropecuarias,

b) los avances logrados en el proceso iniciado en base a dicho enfoque, y

c) algunas conclusiones preliminares que pueden ser útiles para quienes estén en procesos similares.

II. LA TIPIFICACION DE EMPRESAS COMO PROBLEMA DE INTERES CONCEPTUAL

El hecho de que tipificación de empresas haya sido incluído en el temario de esta reunión, incluyendo ejemplos de aplicación en Brasil, sugiere que no es indispensable enfatizar su importancia en este documento.

Pese a la afirmación precedente, puede ser oportuno destacar el interés conceptual de esta cuestión. Si entráramos directamente a experiencias técnicas, correríamos el riesgo de perder de vista cosas que importan en cuanto al tema en sí y, muy especialmente, en cuanto al mismo uso razonable de las técnicas.

Comencemos aclarando que tipificar empresas, como parte de un conjunto posible de actividades clasificatorias de interés en economía agraria*, sólo puede hacerse con sentido pleno en el marco de un análisis más amplio. Los objetivos y requerimientos de ese análisis son indispensables para:

---

* que incluye "regionalizar", por ejemplo.

** afirmación que no responsabiliza a este Centro por los errores que puedan cometerse en el uso de las técnicas.

1) decidir la utilidad de tipificar, y

2) dar el marco teórico que permita elegir variables.

Tipificación de empresas es, entonces, una actividad que debe definirse en su utilidad y en su forma sólo a la luz del propósito que guía el trabajo que la motiva*.

Los usos más comunes de la tipificación en economía agraria suelen pertenecer a combinaciones entre las dos siguientes categorías de análisis más amplios:

1) respuesta diferenciada entre predios, con respecto a una política económica no diferenciada (reacción de oferta a precios, por ejemplo) y/o,

2) definición de la diferenciación de política necesaria para obtener una dada respuesta (proyectos de asistencia técnica, por ejemplo).

Para quien está en la práctica diaria de política económica (global o vía proyectos) éstas, o similares, aplicaciones definen los objetivos de la tipificación y la convierten en un claro procedimiento instrumental de relevancia táctica.

Hay posibilidades, sin embargo, para elevar el nivel de la tipificación a relevancia estratégica. Esta posibilidad surge al reconocerse que una correcta definición de predios tipos para propósitos de política económica, dudosamente se logre con independencia de nuestro progreso en varios frentes de comprensión de la realidad económico-social. Dentro de esa realidad, la unidad empresa es agente vital. Y partir de la empresa para llegar a una mejor comprensión de esa realidad puede ser una vía estratégica de investigación**.

III. EL POSIBLE ROL DE TECNICAS EN TIPIFICACION

Admitiendo que la tipificación tiene relevancia conceptual como problema táctico muy común y en función de sus posibilidades estratégicas, queda por discutir el rol de técnicas que puedan ayudar a mejor hacerla.

Dos puntos destacables llevaron en Uruguay a considerar el empleo de técnicas más sofisticadas que las usuales (uso de "informantes calificados" siendo la más usual).

Estos dos puntos, que se desarrollarán algo antes de indicarse las técnicas que se están usando, son:

---

* Véase una ampliación de este concepto en Cohan, H. (5)

** En sentido similar, véase el rol de la tipificación en Piñeiro, M. y Trigo, E. (12)

1) el status científico de las tipificaciones usuales, altamente subjetivas, y

2) los peligros de una mala tipificación.

Veamos algo, brevemente, sobre cada uno de estos dos temas.

Cabe aclarar ab-initio que ninguno de los problemas asociados a estos dos puntos se resuelven por la aplicación mecánica de técnicas más sofisticadas. Ellas, provenientes en esencia del arsenal estadístico, simplemente pueden ayudar, si bien usadas, a reducir el riesgo de caer en esos problemas.

El punto del status científico* puede resumirse en que, cualquiera sea la validez intrínseca y la utilidad de una tipificación muy subjetiva, es difícil hacerla aceptable por otros investigadores si no hay un mínimo de objetividad en proceso que condujo a generarla y a retenerla como útil. El recurrir a técnicas estadísticas y cuasi-estadísticas del tipo que reportaremos en este documento debe ayudar en esa búsqueda de objetividad**.

En cuanto al punto relacionado con los peligros de una mala tipificación, conviene plantear recaudos bien claros.

Cuando se decide tipificar, se implica que el universo tomado en conjunto no es lo suficientemente homogéneo para los propósitos de análisis. Esta conclusión trivial no siempre es deducible, sin embargo, de muchas tipificaciones que se encuentran en la práctica. Sin duda la hipótesis de necesaria diferenciación existió; pero la praxis suele resultar (con o sin uso de técnicas estadísticas) en empresas tipo poco o mal diferenciadas en cuanto a los propósitos de análisis. Y las tipificaciones parecen, por ende, el resultado de un síndrome tipificatorio inconsciente más que lo que debieran ser: una etapa en la aplicación del método científico.

Como lo indica Kaminsky***, la tipificación en economía agraria de-

---

* Véase la "Introducción del Editor" en (4)

** En la inevitable medida en que distintas concepciones de la realidad pueden todas pasar pruebas estadísticas, la afirmación del texto no implica nada más allá que un uso razonable de la estadística. Es decir: que al menos se rechace lo que los datos disponibles obligan a rechazar.

*** Véase (11). Esto podría ampliarse a una sugerencia metodológica generada por intervenciones verbales de M. Piñeiro y acordada como interesante criterio general por los participantes en la Reunión de Montevideo. La sugerencia sería: tipificar sólo en función de variables estructurales. Esto es claro cuando hay un objetivo y existe un modelo de comportamiento que indica cuál es el conjunto de variables observables que lo representa y cuál el ("estructural") que lo determina. Una vez formados los grupos, mediante variables de estructura, podría testarse la hipótesis de comportamiento diferenciado entre grupos, usándose las observaciones teóricamente "resultantes".

biera ser un medio para la identificación de estructuras diferenciadas a corregirse. Y una mala tipificación puede resultar en un mal remedio diferencial, peor incluso que la enfermedad, identificada como falta de política dirigida a corregir problemas generados por estructuras diferenciales.

Que el usar el conjunto de técnicas útiles para generar tipologías no basta para prevenir este error, lo precisa claramente Ferreira, cuando indica*: "... el investigador debe diferenciar aquellas situaciones en que un método revela o confirma una interesante estructura inherente a ciertos datos, o a cierto problema, y aquellas en que la metodología impone una estructura inexistente". De ahí que técnicas útiles para validar lo actuado tengan al menos tanta importancia como las disponibles para tipificar. Esto fue enfatizado por varios participantes en la Reunión de Tipificación realizada en Montevideo del 11 al 13 de mayo de 1977.

En este contexto deben interpretarse las técnicas que se están empleando en Uruguay y que se sumarizan en el siguiente capítulo.

IV. ALGUNAS TECNICAS EMPLEABLES EN TIPIFICACION

Aceptándose que algunas técnicas pueden ayudar a mejor tipificar, conviene indicar cuáles son y cómo operan ellas.

El material que se resume en este capítulo forma parte del conjunto de lo potencialmente útil para tipificar recomendado por técnicos del CIENES en un Seminario realizado en Montevideo a fines de 1975**.

Las técnicas corresponden al arsenal estadístico-matemático y tienen complejidades de interpretación y uso, previsibles, para el lego. Pero no sólo parecieron promisorias cuando se sugirió su aplicabilidad al tema sino que en un año de trabajar con ellas estamos logrando una convivencia razonable: nos están sirviendo cada vez más a medida que, empleándolas, mejor comprendemos sus alcances y limitaciones.

Difundimos su potencial y nuestras experiencias de uso porque pensamos que esa misma conveniencia fructífera puede lograrse en otros equipos de trabajo y porque, cuando más equipos intenten el proceso, más aprenderemos todos.

Nuestro resumen en esta sección cubrirá tres puntos: A saber:

1. Exploratorias de atributos.
2. Conformadoras de grupos, e
3. Identificadoras de calidad de la conformación elegida.

---

* Véase Ferreira, P. (8). También citado por Kaminsky, M. en (1)

** Véase (4)

Esta exposición seguirá el orden que sugiere nuestra poca experiencia acumulada, y no impone el uso de cada técnica ni en la forma ni el orden en que se presentan. Por cierto que tampoco agotará las posibilidades de uso de instrumental estadístico-matemático.

1. Técnicas exploratorias de atributos

Dado un universo de observaciones y variables (digamos n empresas con m atributos recogidos sobre cada una de ellas) se conforma una nube de n x m datos.

Hay varias cosas que podríamos saber sobre esta nube.

Tal vez, por ejemplo, hemos retenido sólo las m variables que, por hipótesis, mejor separan el total de n predios. "Mejor", recordemos, sólo puede interpretarse a la luz de nuestros objetivos. Y queremos ver si el conjunto de m x n datos se separa entre algunos vectores que podamos interpretar como "tipos" distintos y otros que no hagan sentido sustantivo.

O quizás, de manera más exploratoria que la precedente, nos preguntamos si el total de m x n no se descompondrá entre un subconjunto con mensaje claro, interpretable para nuestros propósitos, y otro que, si bien contribuya a explicar el 100% de la varianza de esta muestra, nos introduzca un ruido indeseable en términos de nuestra aptitud interpretativa. Es decir: si no será que podemos retener para siguientes etapas de análisis sólo n x g datos, donde g es un número inferior al m original.

A este efecto puede usarse la técnica de "Componentes Principales", para la que existen programas standard de IBM que se han incorporado a la biblioteca del IICA - Uruguay.

Componentes Principales resume el conjunto de datos por m vectores, uno por cada variable original, e indica qué correlación hay entre las m variable y dichos vectores. Estos vectores representan una estructura latente en el conjunto original de datos. Su sentido estadístico, que no trataré de explicar*, puede o no conformarse con una interpretación sustantiva de ellos. La interpretación corre por cuenta del usuario.

El usuario puede pedir sólo parte del total de m componentes, explicando ellos tan sólo un porcentaje, inferior al 100%, de la varianza total. El programa de cómputo produce, además, una rotación de estos vectores, la que suele facilitar la interpretación.

El uso de la técnica consiste en retener sólo aquellos componentes con interpretación clara y las variables (o algunas de ellas) más corre

* La técnica opera extrayendo vectores ortogonales entre sí que se van ubicando en la dirección de mayor variabilidad de datos. Véase Artigue, G. (3)

lacionadas con éstos.

En ocasiones, la estructura de correlaciones entre variables originales es lo suficientemente definida como para que con sólo tres o cuatro componentes se explique un buen porcentaje de la varianza y, simultáneamente, estas componentes puedan interpretarse con claridad.

En ocasiones puede también darse que, proyectando las n observaciones sobre las dos primeras componentes, se conforman grupos definibles ya como tipos sin necesidad de recurrir a técnicas adicionales.

De estos casos posibles mencionados pueden verse ejemplos en el trabajo de M. Kaminsky, en el volumen 2 de (4).

Nuestra experiencia en Uruguay nos ha sido útil especialmente en precisar hipótesis sobre variables para conformar grupos, o en sugerir variables representativas.

Claramente, la técnica ha dado más información que la contenida en los datos originales sólo en el muy preciso sentido de aprovechar correlaciones objetivas para mejor comprender.

Muy relacionada con esta técnica existe la denominada "Análisis Factorial". Este puede emprenderse con otra mecánica de cómputo o a partir de Componentes Principales.

Su sentido es algo diferente, en cuanto descompone la masa de datos en factores comunes (tantos como pida el usuario) y específicos. Los "comunes" retienen la influencia de elementos latentes en la estructura y los específicos conforman un receptáculo de la varianza que no puede atribuirse al juego de los factores comunes y, por ende, se asignan como propios o "específicos" de cada variable. Para mayor detalle sobre esto, véase el trabajo de G. Artigue (3).

2. Técnicas Conformadoras de Grupos*

Cuando se tiene una idea clara de qué variables usar para separar grupos de predios, existen técnicas denominadas "de conglomeración" que operan como ayuda casi mecánica para tipificar.

En principio, estos algoritmos podrían operar con el total de variables disponibles (generados por Censo o encuesta independiente, por ejemplo). Nuestra experiencia hasta el momento indica que no resulta

---

* Véase Alonso, A. (1)

usarlas en un sentido tan exploratorio. Más bien conviene partir de una noción de atributos relevantes, tal vez confirmada mediante uso de Componentes Principales.

Estas técnicas, de las que existen muchas alternativas, han sido programadas para el IICA en tres versiones básicas. A continuación se indicará cómo operan ellas.

2a) Van Rijsbergen. Hay que elegir un criterio de "distancia" entre dos empresas. Los criterios, para los cuales hay alternativas incluídas, se basan en comparar los valores que toman los atributos entre cualquier par de empresas*. Y hay que indicar qué distancia es la suficientemente estrecha como para que todo par de empresas que esté dentro de ese límite de tolerancia integre un mismo grupo. Con esta información, permitiendo parametrizaciones del nivel de tolerancia, el algoritmo indica cuántos grupos (clusters) se forman y qué empresas integran cada uno.

El programa adoptado para IICA incluye la muy recomendable opción de estandardizar variables, para que no pesen las unidades de medida. Y permite la opción de ponderar, si el usuario desea que pese más la diferencia entre ciertas variables que la diferencia entre otras.

2b) Sparks. Esta forma de constituir grupos sólo permite la opción de ponderar y estandardizar variables. Tomadas las correspondientes decisiones, el usuario debe indicar cuántos grupos quiere y proponer una empresa centro de cada grupo. El algoritmo va comparando las n observaciones en función de un concepto de distancia entre ellas, asigna a los grupos con respecto a cuyo centro hay menor distancia y, en sucesivas iteraciones, va recalculando centros. Al terminar propone cuál es la mejor conformación (menor distancia total acumulada) para el número de grupos propuesto y cuál es el centro final de cada grupo (la conocida "empresa representativa", si así se desea interpretarlo).

2c) Wishart. Esta mecánica constructiva parte de considerar que cada observación, cada empresa, es un cluster. A partir de allí se pregunta qué unión de clusters introduce el menor error** y va agrupando y corrigiendo distancias (para lo cual hay seis opciones de

---

* Por ejemplo, para las empresas 1 y 2 con sólo dos variables atributo (X^A y X^B):

$$D(1, 2) = \sqrt{(X_1^A - X_2^A)^2 + (X_1^B - X_2^B)^2}$$

** Error definido en términos de varianza dentro de clusters. Al comenzar, cuando cada empresa es un cluster, esta varianza es cero. Al terminar: toda la varianza de la muestra está en un cluster, si se optó por una de las seis formas disponibles para ir corrigiendo distancias cada vez que se unen grupos.

"distancia"). Termina con un solo cluster, que incluye al total de observaciones. El usuario debe elegir qué etapa del proceso, que es informado en detalle, presenta conformaciones útiles para sus objetivos.

Esta riqueza de alternativas, por otra parte sólo un porcentual de lo disponible en clustering, trae sus problemas. Salvo que el universo tenga una conformación estructural muy definida, es dudoso que todos los algoritmos generen una misma agrupación.

Cualquiera de estos mecanismos disponibles permite agrupar predios rápidamente, de manera precisa y sin necesidad de prefijar los usuales cortes transversales del tipo arbitrario: "hasta tantas hectáreas" con "tanto y cuanto porciento de marcación".

No tenemos una experiencia rotunda que aportar, salvo la de insistir en que toda técnica que se use sin un conocimiento razonable y en un proceso con objetivos imprecisos, trae problemas.

Siendo conveniente elegir algoritmos y variables (y tal vez ponderar de manera diferente diferencias entre éstas), conglomerar tiene mucho de arbitrario. No obstante: la arbitrariedad en el uso de estas técnicas desaparece si la tipificación se considera como un proceso científico. En efecto: debe haber un modelo de comportamiento y objetivos que expliquen el porqué de las variables retenidas y el por qué de las propiedades estadísticas del algoritmo elegido. Y debe preverse con anticipación una batería de tests estadísticos que, probablemente verificando hipótesis centrales, validen ese proceso en su conjunto.

3. Técnicas identificadoras de calidad de la conformación elegida

Como pertenecientes a este grupo de métodos, sólo haremos referencia directa al Análisis Discriminante. Más información puede encontrarse en (1), (4) y (8).

El Análisis Discriminante responde claramente al siguiente problema: tenemos K grupos conocidos, K particiones de una muestra de n empresas. Sabemos por qué armamos estos grupos en base a información sobre g variables. Y ahora, dada una nueva empresa cualquiera para la cual tenemos información en términos de las g variables, queremos saber a cuál de los K grupos asignarla.

El correspondiente programa de cómputo, desarrollado por IBM y disponible en IICA-Uruguay, requiere información sobre cuántas variables se han reunido sobre las n observaciones, cuántos grupos se han formado y qué observaciones (empresas, en nuestro caso) integran cada grupo. A partir de ello calcula e imprime una serie de información que también resulta útil para estimar la validez de la partición propuesta.

Una información útil que se genera en este proceso es el valor D^2

de Mahalanobis. Este valor indica la "calidad" de las agrupaciones propuestas, en el sentido de que se incrementa al aumentar la diferenciación entre las medias de los grupos. El D^2 se distribuye, bajo supuesto de normalidad y como propiedad asintótica, como ji-cuadrado con g por K-1 grados de libertad (número de variables por número de grupos menos uno) y permite una verificación estadística objetiva, aunque limitada, del grado de diferencia entre tipos.

Tenemos así un claro primer uso posible de esto en el sentido de mayor preocupación indicado en la Introducción a este trabajo: calcular grado de objetividad. Cabe reconocer, no obstante, que las muestras pequeñas con las que estamos trabajando (menos de 70 predios) al partirse en tres o más grupos y usarse muchas variables tienden a resultar estadísticamente aceptables, aún cuando se generen al azar. Esto no impide que se acepte la validez de una tipificación generada no al azar sino con sentido válido para los propósitos que la guía, pero deja insatisfacción en términos de la objetividad buscada, que esperamos se elimine al ir trabajando con mayor número de casos que, dado el número de variables, permitan mejor operar a la estadística D^2*.

El programa continúa con lo que es propósito sustancial del método: calcular funciones lineales, con las g variables como argumentos, que puedan aplicarse a nuevas empresas. Estas "funciones discriminantes", una para cada grupo propuesto, registran un término constante y g parámetros. Aplicándolas a las g variables de cada nueva empresa, "nueva" en el sentido de que no pertenezca al conjunto de n casos original, conceden a ésta "puntajes". La nueva empresa debería asignarse al grupo al que pertenece la función que genera el mayor puntaje.

En base a estas funciones discriminantes se ha propuesto, cosa que aún no hemos probado, definir tipos en base a una muestra del Censo, por ejemplo, y ubicar los restantes casos en uno u otro tipo conforme a estas funciones discriminantes**.

Para terminar, debe destacarse que el programa disponible ayuda a definir validez de la agrupación propuesta al aplicar las K funciones calculadas a las n empresas. Puede indicar así cuál de las funciones asigna mayor puntaje a cada predio. Por ende, si un predio propuesto en el grupo 1, al aplicársele la función 1, recibe el mayor puntaje de los K posibles, está bien clasificado. Si, en cambio, el mayor puntaje se lo otorga otra de las funciones, surge indicio de mala clasificación.

---

* La prueba D^2 es limitada, de todas maneras, como se explicará al comentarse el primer ejemplo de aplicación en el próximo capítulo. Y se están previendo pruebas estadísticas de más poder.

** Véase (4)

Es interesante destacar que hemos efectuado aplicaciones de esta técnica a grupos conformados por métodos no formales, a ejemplos de tipificaciones convencionales, y se han obtenido algunos resultados estadísticamente promisorios*. Para estos casos se puede contestar una duda planteada en la Introducción a la publicación del Seminario de 1975 (4). Aparentemente, no se puede rechazar universalmente la hipótesis de que la experiencia acumulada que hoy se vuelca en tipificar es válida. Esto era de esperar como resultado de trabajos bien hechos. Es que las técnicas, las expuestas u otras, no vendrán a sustituir el buen juicio. A lo sumo lo ayudarán y permitirán, si bien usadas, hacer más compartibles los resultados que él genera.

En las referencias (1), (4) y (8) se podrá observar el posible uso de dócimas paramétricas y no paramétricas, sobre cuya aplicabilidad a este ptobiema de testeo de conclusiones estamos experimentando.

V. ALGUNOS EJEMPLOS DE APLICACION

No disponemos aún en Uruguay de un caso real en el que se hayan empleado las técnicas en su conjunto. Pero sí pueden exponerse algunos casos prácticos que pueden completar la percepción de lo lograble con el uso de este bagaje**.

A estos efectos se han seleccionado los siguientes tres casos:

1. Prueba con Análisis Discriminante de una tipificación subjetiva, tomada de (7).
2. Análisis exploratorio de una encuesta con Análisis Factorial mediante Componentes Principales, también tomado de (7).
3. Aplicación de clustering a un proyecto de regionalización, tomada de (2).

1. Prueba de una clasificación***

a) El caso

El ejemplo de análisis discriminante es un caso en el que se disponía de una tipología que agrupó 28 predios ovejeros en tres estratos de tamaño (superficie), como se indica en el Cuadro No. 1.

---

* Aunque sólo se han usado tests de bajo poder.

** En torno a lo que se presentará como experiencias en análisis multivariante (especialmente: componentes), debe reconocerse que existen dudas metodológicas en cuanto al tratamiento de variables cualitativas que no respondan a una escala ordinal, en particular si se salen del caso mínimo 0, 1.

*** Véase Dabezies, M. y Sarroca, O. en (7)

Cuadro No. 1

Tipos usados en la aplicación de Análisis Discriminante

Tipos	Superficie(Hás.)	Número de casos
1	100 a 499	8
2	500 a 2.499	15
3	2.500 y más	5

La hipótesis era que con este criterio de agrupamiento bastaría para diferenciar tipos en cuanto a tecnología. Y ella se testeó calculándose las funciones discriminantes con cinco variables:

- porcentaje de mejoramiento del campo
- dotación total en invierno (UA/há)
- borregas segunda encarneradas sobre el total de ovejas de primera encarnerada (porcentaje).
- época de encarnerada (mes $\overline{X}$)
- uso de lombricida (0 = no, 1 = si)

b) Resultados y Conclusiones

En este caso el Discriminante dio resultado negativo para los objetivos perseguidos. En efecto, el D^2 de Mahalanobis dio menor que el X^2 de tablas al 5% de confianza. Los valores fueron los siguientes:

$$D^2 = 13.56$$

$$X^2_{10,.05} = 18.30$$

Estudiando el Cuadro No. 2, se observa que los porcentajes de predios mal clasificados fueron los siguientes:

	Porcentaje de observaciones mal clasificadas
Grupo 1	50
Grupo 2	46
Grupo 3	20
Sobre Total	43

Estos resultados implican exclusivamente que entre los grupos que se estructuraron, no existen diferencias de medias en términos de las variables tecnológicas consideradas.

Cuadro No. 2

Evaluación de las Funciones de Clasificación para cada Observación

Group 1

Observation	Probability associated with Largest Discriminant Function	Largest Function No.
1	0.62711	1
2	0.48295	2
3	0.44749	1
4	0.70079	1
5	0.63092	1
6	0.58099	2
7	0.75198	3
8	0.47293	2

Group 2

Observation	Probability associated with Largest Discriminant Function	Largest Function No.
1	0.46694	2
2	0.59538	1
3	0.88265	2
4	0.50944	2
5	0.63399	1
6	0.78844	2
7	0.77040	3
8	0.83268	2
9	0.72971	2
10	0.44870	1
11	0.38967	1
12	0.65743	2
13	0.52411	2
14	0.76311	3
15	0.68169	1

Group 3

Observation	Probability associated with Largest Discriminant Function	Largest Function No.
1	0.58061	2
2	0.78790	3
3	0.74701	3
4	0.75882	3
5	0.75156	3

Nota: Tomado de la salida de computadora

En este caso se reporta como ejemplo de una situación, de las pocas que hemos obtenido, en la que la hipótesis nula (igualdad de medias) no puede rechazarse. Obsérvese que, aparte de problemas estadísticos inherentes a este test, diferencias significativas de medias no serían suficientes para aceptar una clasificación subjetiva. En efecto, podría haber dentro de cada grupo varianzas demasiado grandes para los propósitos que guiaron el agrupamiento. La situación mejoraría en este aspecto si los grupos sometidos a análisis discriminante se hubieran constituído con, por ejemplo, un algoritmo de conglomeración que minimizara esa varianza interna del cluster*. Esta observación, recaudo operativo serio al uso del D^2, fue hecha por Pedro Ferreira en el reciente Seminario de Montevideo.

2. Análisis Exploratorio de una Encuesta**

a) El caso

También con el propósito de analizar el problema de tecnología en relación con otros factores, se aplicó análisis factorial mediante componentes a 56 observaciones (predios criadores) sobre los que se disponía de 24 variables.

Las variables se presentan a continuación, ordenadas según una hipótesis que sugería la eventual aparición de cuatro factores (Tecnología, Tamaño, Orientación de la empresa y Tenencia).

Tecnología:

- Porcentaje de Mejoramientos
- Dotación total en invierno (UA/há)
- Edad de entore (meses)
- Duración del entore (días)
- Diagnóstico de preñez (0 = no, 1 = si)
- Edad de destete (meses)
- Suplementación mineral (0 = no, 1 = si)
- Uso de lombricida (0 = no, 1 = si)
- % de Parición
- No. de tractores/hás. x 1.000
- Uso de crédito (0 = no, 1 = si)

Tamaño:

- Superficie total (hás)
- No. de potreros
- Superficie promedio por potrero (hás)

---

* Existen técnicas de conglomeración que minimizan varianza dentro de clusters y la maximizan entre ellos. Esto equivale a usar el D^2 como función objetivo de la técnica de conformación de grupos. Obviamente, un análisis de los grupos con Análisis Discriminante a posteriori, sólo tendría sentido si, usando otro conjunto de variables, no fuera una mera réplica del proceso de conglomeración.

** Véase Dabezies , M. y Sarroca, O. (7)

. Mano de obra total (E.H./há x 1.000)
. Mano de obra familiar (% sobre el total)

Orientación:

. Ovinos/vacunos
. Venta de Terneros de destete (0 = no, 1 = si)
. Recría de sus propios terneros (0 = no, 1 = si)
. Compra para recriar (0 = no, 1 = si)
. Hace inverne (0 = no, 1 = si)

Tenencia:

. Propietario (0 = no, 1 = si)
. Propietario - Arrendatario (0 = no, 1 = si)
. Arrendatario (0 = no, 1 = si)

A la variable tenencia, se le dio el tratamiento que se indica anteriormente transformándola en tres variables binarias (var. 22, 23 y 24).

b) Estudio de las correlaciones entre factores y variables

En el Cuadro No. 3 se presentan 10 de los 24 factores de la matriz, factores-variables, sin rotar, y los porcentajes acumulados de la varianza total que explica cada uno de ellos.

A los efectos de una mejor interpretación de los valores de las correlaciones, se estableció como límite inferior de significación de los mismos, el valor de tablas correspondiente al 1% con 54 grados de libertad (número de observaciones, menos 2).

Deben hacerse tres comentarios iniciales. A saber:

. Primero, el uso de variables cualitativas no ordinales no es fácil con estas técnicas. Probablemente convenga buscar forma de cuantitivizarlas o analizar por separado cada grupo.

. Segundo, deben advertirse los bajos valores de la varianza explicada por cada factor.

. Tercero, la interpretación de los factores no resulta muy evidente en todos los casos.

Tomando los primeros diez, el único factor que puede definirse claramente, es el primero, Se puede hacer por las variables más fuertemente correlacionadas con él, que son variables de "tamaño". Junto a éstas, aparecen asociadas en el mismo sentido, algunos indicadores de uso de prácticas Tecnológicas más avanzadas.

Cuadro No. 3

Matriz de Factores (sin rotar, 24 factores)*

Porcentaje de Varianza Acumulada para los 1ros. Diez Factores:		18,4 F_1	29,5 F_2	39,0 F_3	47,0 F_4	54,5 F_5	60,1 F_6	65,5 F_7	70,1 F_8	74,5 F_9	78,1 F_{10}
% Mejoramientos	1		-.39			.44	-.51				
Dotación	2			.61							
Edad entore	3		.39		-.39						
Duración entore	4	-.53						-.36			.39
Diagnóstico Preñez	5	.55									
Edad destete	6	-.47			-.54	.35					
Suplementación	7				-.64						
Lombricida	8	.64									
% Parición	9	.42							.61	.37	
Tractor/há.	10		-.49					-.69			
Crédito	11	.36	-.42								
Superficie	12	.76	.38								
No. Potreros	13	.75								-.34	
Sup.Hás/potrero	14	.59	.46	-.34							
Equiv. h/há.	15	-.63		.37							
%M. de O. Fam.	16	-.80		.34							
Ovino/Vacuno	17						.67				
Venta terneros	18				-.37	-.66					
Recría propia	19		.50			.57					
Recría compra	20		.45					-.46			
Inverne	21					.43				.48	.36
Propietario	22		-.50	-.74							
Prop.-Arrend.	23		.46	.35					.37		
Arrendatario	24			.60	.41						

* Para ayuda visual, se excluyen correlaciones no significativas al 1%. En este caso, ello resulta en retener aportes explicatorios de varianza del orden del 12%.

En el resto de los factores aparecen asociaciones no muy claras entre determinadas orientaciones, algunas variables de tecnología y otras de tamaño.

En el intento de clarificar la significación de estos factores, se hicieron dos nuevas corridas del programa, la primera rotando los primeros 9 factores que explican el 74.5% de la varianza total. La segunda, rotando los 7 primeros, los cuales explican el 66%.

Con la primera rotación se logró clarificar en parte el significado de algunos factores. Véase el Cuadro No. 4.

El factor 1, queda más identificado con "tamaño", ya que aumentan las correlaciones correspondientes a esas variables y desaparecen o disminuyen algunas de "tecnología". Este factor explica el 18.4% de la varianza total.

El factor 2, aparece ahora relacionado con variables indicadoras de un uso de tecnología avanzada (por ej.: diagnóstico de preñez, % de mejoramientos, suplementación mineral, uso de lombricida). Este factor explica por sí sólo el 11.1% de la varianza y, junto con el primero, el 29.5%.

Con algún esfuerzo interpretativo se pudo juzgar algo sobre los 7 factores restantes:

Los factores 3 y 4 aparecen fuertemente marcados por variables de tenencia. Hasta este último, se está con un 47% de la varianza explicada. El factor 5, se vincula claramente a orientación dentro de la ganadería: ciclo completo o cría. La varianza explicada hasta aquí es del 55%. El factor 6, se asocia a orientación ovina o vacuna. A mayor cantidad de ovinos, mayor dotación total en invierno. El factor 7, pone de manifiesto la posible relación entre el porciento de mejoramientos y la existencia de tractor. El factor 8, expresa que la dotación está directamente correlacionada con el porcentaje de parición. Por último, el factor 9, se puede definir como "orientación invernada", la cual aparece asociada directamente a compra de animales para recriar, mayor tamaño (sup. total) y mayor superficie promedio por potrero.

La segunda rotación, de la cual se buscaba que aportara una mayor claridad, resultó útil a este propósito sólo para destacar los dos primeros factores (Tamaño y Tecnología). Ellos resultaron más claros, pero en cambio, los restantes perdieron identidad.

c) Conclusiones

Atendiendo a la hipótesis de la cual se partió, al encarar el análisis Factorial, se puede decir que en líneas generales la misma se

Cuadro No. 4

Matriz de Factores Rotada (9 Factores)*

		F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9
% Mejoramientos	1		-.60					-.41		
Dotación	2						.36		.65	
Edad entore	3						-.56	.43		
Duración entore	4	-.39					-.42	-.45		
Diagnóstico Preñez	5	.40	-.51							
Edad destete	6	-.47		-.38	-.44					
Suplementación	7		-.79							
Lombricida	8	.42	-.39							
% Parición	9								.88	
Tractor/há.	10							-.83		
Crédito	11		-.60							
Superficie	12	.76								.40
No. Potreros	13	.79								
Sup.Hás/potrero	14	.59								.67
Equiv. h/há.	15	-.70			-.34					
% M. de O. Fam.	16	-.88								
Ovino/Vacuno	17						.86			
Venta terneros	18					-.83				
Recría propia	19					.82				
Recría compra	20				-.37					.36
Inverne	21					.34				.71
Propietario	22			-.80	.51					
Prop.-Arrend.	23				-.93					
Arrendatario	24			.93						

* Al igual que en el cuadro anterior, para claridad visual se excluyen correlaciones no significativas al 1%.

confirma. En efecto, basándose fundamentalmente en la rotación (Cuadro No. 4) se vio que aparecen más claramente los factores: tamaño y tecnología. Orientación y tenencia, con menor claridad, podrían asociarse a los factores 3 y 4.

El factor tamaño aparece como más importante que el tecnológico ya que aparece más claro y además explica un 18.4% de la varianza total, frente a un 11.1% del segundo.

El análisis más detallado de los factores lleva a la conclusión de que la conformación que a priori se pensó de cada uno de los factores, en algunos casos se confirma (tamaño) pero en otros no.

Esto último está ejemplificado por el factor tecnología. En efecto, las variables indicadoras de nivel tecnológico no aparecen agrupadas en un sólo factor, sino que tienden a distribuirse en más de uno. Resulta difícil pensar, entonces, en un factor "tecnología" único compuesto invariablemente por un determinado número de indicadores. Esto resulta así para las observaciones y las variables que se estudiaron.

En este mismo sentido, se concluye que, determinadas asociaciones encontradas entre variables tecnológicas y otros factores (por ej.: orientación), no necesariamente deben de poseer una explicación lógica desde el punto de vista físico. Las mismas pueden resultar del azar, para el universo estudiado (muestra reducida).

En cuanto al factor tenencia, su interpretación definitiva se hace difícil. La causa de esta circunstancia se cree que está en el tratamiento metodológico que se le dio a la variable.

En cuanto al aporte que el Análisis Factorial por Componentes hace a un replanteo clasificatorio, la más importante conclusión a la que se arriba, es que las características particulares que ya se comentaron respecto al factor tecnología, hacen pensar en una reformulación del papel que juega el mismo en el grupo de empresas estudiado. La tecnología por sí sóla o incluso junto con el factor tamaño, no serían suficientes a los fines de una futura clasificación. Esto se fundamenta en los porcentajes de la varianza total explicada por los factores que las representan.

3. Aplicación de Clustering a un Proyecto de Regionalización*

a) El caso

En este trabajo se presentó una agrupación de unidades administrativas como aporte a un Proyecto de Desarrollo Regional en Uruguay (convenio MAP-IICA).

La aplicación está orientada a la definición de subregiones dentro del departamento de Cerro Largo, que forma parte, junto con Tacuarembó y Rivera, de la región Plan del Proyecto.

* Véase Alonso, A. (2)

Este trabajo presenta una aplicación de dos métodos de clasificación disponibles en el IICA - Uruguay. El método de "Ward", que en forma más bien intuitiva aparece como muy indicado para clasificar empresas en un proceso de tipificación, se compara con el método de "Vecino más cercano", que opera en forma diferente y que no aparece como recomendable para los casos más generales de tipificación*.

b) Variables utilizadas

La información utilizada proviene del Censo General Agropecuario de 1970 tabulada a nivel de sector censal. El departamento de Cerro Largo se encuentra dividido en 47 sectores censales. Estos sectores constituyen los elementos a clasificar, como paso previo a la definición de subregiones. Dado que estos sectores difieren sustancialmente en tamaño, no se han considerado variables en términos absolutos, prefiriéndose la utilización de relaciones entre atributos para tratar de caracterizarlos.

Los atributos considerados se pueden agrupar en dos clases:

i. la primera, constituída por 6 variables, trata de identificar algunas características generales de los sectores, como ser: tamaño promedio de las explotaciones, intensidad de mano de obra, mano de obra familiar, grado de mecanización, tamaño promedio de los potreros, y mejoramiento de pasturas naturales.

ii. la segunda, constituída por 3 variables, trata de identificar los rubros predominantes: agricultura o ganadería y (dentro de esta última) producción ovina, bovina o de leche.

Las variables empleadas fueron:

- Tamaño promedio de las explotaciones
- Hectáreas por trabajador
- Mano de obra familiar
- Hectáreas por tractor
- Tamaño de los potreros
- Relación Agrícola-Ganadera
- Superficie mejorada
- Relación Ovino-Bovino
- Ganado lechero

---

* Los métodos son dos de las seis opciones incluídas en el programa de Wishart. Para un mayor detalle sobre estos métodos, ver Alonso, A. (1)

c) Métodos de conglomeración utilizados y resultados

Para agrupar los sectores censales se utilizó el algoritmo de Wishart aplicando el método de "Ward" que permite obtener una estimación de la varianza dentro de los grupos resultantes y el método de "Vecino más Cercano".

En ambas aplicaciones se estandarizaron las variables para obviar el problema de las unidades de medida y no se encontró razón para ponderar los atributos.

i. Método de Ward

En el Cuadro No. 5 se presentan los agrupamientos resultantes para la partición en 4 clusters.

Cuadro No. 5

Conglomerados resultantes de la aplicación del método de Ward

Cluster	Número de Sectores	Sectores Censales
I	12	3-4, 3-5, 5-5, 5-6, 5-7, 5-8, 11-1, 11-2, 14-1, 14-2, 14-3, 14-4
II	9	3-1, 3-2, 3-3, 10-1, 12-2, 13-1, 13-2, 13-5, 13-6
III	19	4-1, 4-2, 4-3, 4-4, 5-4, 6-2, 7-1, 7-2, 8-2, 8-3, 9-2, 9-3, 10-2, 10-3, 12-1, 12-3, 13-3, 13-4, 14-5
IV	7	5-1, 5-2, 5-3, 6-1, 8-1, 9-1, 10-4

En el Cuadro No. 6 aparece el valor de la función objetivo correspondiente a cada partición y el incremento que se produce cada vez que se unen dos conglomerados.

Al disminuir el número de clusters, estos se van haciendo cada vez menos homogéneos, aumenta la intravarianza y disminuye la varianza entre conglomerados.

Se puede apreciar que los incrementos que se van produciendo en la función objetivos crecen lentamente hasta llegar a la parti-

ción en 3 clusters y que a partir de ésta el valor de la intra varianza sube rápidamente. Esto permite afirmar que, dada la naturaleza de los datos, no se pueden elegir menos de 3 clusters, si se pretende trabajar con conglomerados razonablemente homogéneos.

La varianza dentro de los conglomerados para la partición en 4 clusters representa el 42% de la varianza total, de modo que la varianza entre los clusters es el 58% de la total. Con 3 clusters se tiene la varianza total dividida en partes iguales entre los conglomerados y dentro de los conglomerados.

De acuerdo con esto se puede considerar adecuada la partición en 4 clusters y a partir de ella se tratarán de definir las subregiones del departamento considerado.

Para facilitar la interpretación de los resultados, se presentan en el Cuadro No. 7 los valores medios de las variables consideradas para los clusters que surgen en la partición en 4 conglomerados junto con los valores medios correspondientes al departamento.

Cuadro No. 6

Función Objetivo, Valor Absoluto e Incremento para cada Partición

Número de Clusters	Valor de la Función Objetivo (Varianza dentro de cluster)	Incremento de la Función Objetivo (%)
7	116,21	---
6	130,84	12,59
5	151,96	16,14
4	178,51	17,47
3	212,27	18,91
2	301,64	42,10
1	422,99	40,23

La subregión I está formada por "sectores representados por explotaciones chicas, tecnificadas, de producción agrícola-lechera". La dotación de mano de obra por hectárea es alta, están muy mecanizadas y el tamaño de los potreros es reducido.

La subregión II está formada por "sectores representados por explotaciones chicas, familiares, de producción ovina". El nivel de mecanización es medio y la dotación de mano de obra es alta y de tipo familiar.

La subregión III está constituída por "sectores representados por explotaciones medianas, poco tecnificadas de producción

Cuadro No. 7

Valores Promedio Correspondientes a los Conglomerados

ıster	Tamaño Promedio	Hectáreas por Trabajador	Mano de Obra Familiar	Hectáreas Por Tractor	Tamaño Potreros	Relación Agrícola-Ganadera	Superficie Mejorada	Relación Ovino-Bovino	Ganado Lechero
I	154.17	69.39	0.78	966.42	45.36	8.27	8.19	1.24	7.42
:I	181.87	74.32	0.88	1.911.28	64.87	2.56	3.97	2.79	1.25
:I	420.94	189.34	0.69	3.464.57	105.00	1.20	5.33	1.92	0.70
:V	1.736.07	342.15	0.32	2.871.63	201.26	0.88	3.96	1.45	0.68
›arta ıto de ·ro ·go	315.65	131.66	0.72	1.824.56	87.43	2.43	5.29	1.73	1.84

ganadera". La mano de obra es baja, muy poco mecanizadas y con una proporción media de pasturas mejoradas.

La subregión IV está constituída por "sectores representados por explotaciones muy grandes, extensivas de producción bovina". Usan muy poca mano de obra, que es asalariada, y el tamaño de los potreros es elevado.

ii. Método de Vecino más Cercano

Al aplicar este método, los sectores tienden a unirse a clusters ya formados en lugar de convertirse en centros de nuevos clusters. Se forma así un gran conglomerado que contiene a la mayoría de los sectores, quedando sectores aislados que por alguna razón se diferencian netamente del resto.

Si se analiza, por ejemplo, la partición en 4 conglomerados que aparece en el Cuadro No. 8, se tienen 44 sectores en un conglomerado restando 3 clusters singulares.

Cuadro No. 8

Conglomerados Resultantes de la Aplicación del Método de Vecino más Cercano

Cluster	Número de Sectores	Sectores Censales
I	1	3-4
II	1	13-4
III	1	14-3
IV	44	El resto

Analizando la matriz de observaciones, se puede ver que los sectores que forman conglomerados singulares presentan valores extremadamente altos, con respecto al promedio general, para alguna de las variables consideradas.

Los sectores 3-4 y 13-4 presentan un porcentaje de superficie mejorada de 19.81 y 15.39 respectivamente, frente al promedio departamental que es de 5.29. El sector 14-3 presenta un porcentaje de ganado lechero de 27.38 frente a un promedio departamental de 1.84.

Dada su conformación, se puede concluir que los conglomerados obtenidos no sirven para definir sub-regiones en el sentido usual.

De acuerdo con los objetivos más generales de regionalización o

tipificación, este método no aparece como recomendable. Sin embargo, enfrentados a otro tipo de problemas, puede resultar de mucha utilidad. Aplicando este método, se podrían aislar las regiones que se encuentran en una situación más crítica de acuerdo con un problema planteado y que deben ser asistidas en forma prioritaria (erosión, por ejemplo).

d) Concordancia de la Clasificación Obtenida

La utilidad de las clasificaciones se debe plantear en la posibilidad que ofrezcan de ejemplificar algún principio de agrupamiento que sea coherente con los objetivos del trabajo planteado. Para ello puede estudiarse la concordancia que presenten los agrupamientos obtenidos con respecto a algún criterio o variable que resulte relevante en el contexto general.

En este ejemplo se plantean tablas de contingencia para docimar la existencia de asociación entre el hecho de que los sectores pertenecen a un conglomerado determinado y el valor que presentan para la variable: tamaño promedio de las explotaciones.

En una segunda etapa se trata de medir el grado de dependencia entre los criterios planteados mediante el estadígrafo de contingencia "C" de Pearson*.

Al analizar los resultados se deben tener presente las reservas que surgen del hecho de trabajar con un número reducido de observaciones, lo que determina que aparezcan muchas celdillas vacías.

En el Cuadro No. 9 se plantea una tábla de contingencia comparando los agrupamientos con el valor que presentan para la variable tamaño promedio de las explotaciones en cuatro tramos: menos de 200 hectáreas, de 200 a 500, de 500 a 1.000 y más de 1.000 hectáreas.

Al comparar el valor de T calculado con el valor de X^2 (con 9 grados de libertad a nivel 0.01), se rechaza la hipótesis nula de independencia entre las características consideradas.

Al medir el grado de asociación mediante el estadígrafo "C" de Pearson se observa que el valor calculado es alto comparado con el máximo que puede tomar para una tabla de contingencia de estas características.

Véase Cuadro No. 9

De acuerdo con los resultados obtenidos se podría concluir que la clasificación obtenida es concordante con la variable analizada; sin considerar otras pruebas de validación que deben ser realizadas, y

---

* Véase Kaminsky, M. (10)

aceptando que esta variable es relevante dentro del contexto general del problema planteado, la clasificación obtenida puede ser usada con un cierto grado de confianza para definir subregiones homogéneas dentro del universo considerado.

Cuadro No. 9

Tabla de Contingencia

Superficie \ Clusters	I	II	III	IV	f.m.
- 200	10 (4.60)	7 (3.45)	1 (7.28)	0 (2.68)	18
200 - 500	2 (4.34)	2 (3.26)	13 (6.87)	0 (2.53)	17
500 -1000	0 (1.53)	0 (1.15)	5 (2.43)	1 (0.89)	6
+ 1000	0 (1.53)	0 (1.15)	0 (2.43)	6 (0.89)	6
f.m.	12	9	19	7	47

$$T = \Sigma \frac{(F_o - F_t)^2}{F_t}$$

$T = 67,699$

$X^2_{(9,0.01)} = 21.666$

$$C = \sqrt{\frac{T}{T + N}}$$

$C = 0.768$

máx. $C = 0.866$

e) Conclusiones Generales de la Aplicación de Conglomeración

Resulta difícil, en el estado actual del conocimiento que se posee sobre aplicaciones de técnicas del Análisis de Conglomeración a la

tipificación o regionalización, el aconsejar el uso de determinados métodos.

De acuerdo con los objetivos específicos de los trabajos que se planteen y la naturaleza de los datos, los distintos métodos pueden presentar ventajas relativas que podrán evaluarse al realizar las aplicaciones concretas.

En forma general se puede decir que el método de "Ward" puede ser utilizado con éxito en procesos de regionalización, debido a que los agrupamientos obtenidos parecen adecuados para servir de base a la definición de regiones homogéneas.

El método de "Vecino más Cercano" no permite obtener conglomerados que puedan ser utilizados para definir regiones o empresas tipo en el sentido más general. Sin embargo, si se deseara seleccionar regiones que sean marcadamente diferentes del resto, de acuerdo con algún atributo o grupo de atributos dado, este método sería el más aconsejable.

De cualquier manera, al plantear un trabajo de regionalización en varias etapas partiendo de una regionalización a nivel global, para luego ir afinando a nivel de subregiones, se puede plantear la utilización de distintos métodos.

En cada etapa tanto las variables como los métodos de clasificación pueden variar sustancialmente en la medida en que cambia la naturaleza de los elementos a agrupar y posiblemente los objetivos del análisis.

VI. CONCLUSIONES

Las conclusiones que pueden presentarse en cuanto al uso de técnicas estadísticas en tipificación están en esencia contenidas a lo largo de este documento que reproduce las opiniones del grupo informal que trabaja en el tema en Uruguay.

Si acaso, tal vez convenga enfatizar que:

- tipificar necesita objetivos claros y se beneficia si opera en base a un modelo de comportamiento,
- al tipificarse pueden emplearse técnicas para conformar los grupos en forma deseable para los propósitos del trabajo,
- al igual que todo proceso científico, tipificar requiere la definición de pruebas objetivas de la validez de los grupos que se conformen,

- existen técnicas estadísticas y cuasi-estadísticas muy útiles para tipificar, que pueden ayudar en varias etapas del proceso, y que

- el uso mecanizado de dichas técnicas no garantiza la calidad de los agrupamientos que se construyan, si no está claro el proceso de con formación y validación al que las técnicas deben aportar objetividad.

* * * * * *

Mim.No.69
HEC/AA:mmc
13.06.77

REFERENCIAS CITADAS

1. ALONSO, A. "Algunas técnicas de conglomeración. Su naturaleza y sus posibilidades en tipificación de empresas". Mimeo 38 p. (1)

2. _____ "Subregionalización mediante análisis de conglomerados". Mimeo. 12 p. (1)

3. ARTIGUE, G. "Componentes principales y análisis factorial. Su naturaleza y sus posibilidades en tipificación". Mimeo. 41 p. (1)

4. COHAN, H. (ed.) Seminario sobre métodos y problemas en tipificación de empresas agropecuarias (2)

5. _____ "Tipificación de empresas y análisis de sistemas". En (2)

6. _____ "Ventajas y posibilidades del uso de técnicas estadísticas de tipificación aplicadas al análisis de empresas ganaderas". Mimeo. 10 p. IICA. Montevideo, 1976.

7. DABEZIES, M. y SARROCA, O. "Aporte para la tipificación de establecimientos ganaderos en la zona de Areniscas". Mimeo. 15 p. (1)

8. FERREIRA, P. "Algunos comentarios sobre evaluación de clusterings". Mimeo. 9 p. y anexo (1)

9. _____ "Técnicas disponibles para tipificación de empresas agropecuarias" En (2)

10. KAMINSKY, M. "Aplicaciones e ilustraciones de técnicas disponibles para tipificación de empresas agropecuarias". En (2)

11. _____ "Comentarios sobre procesos de tipificación y su validación". Mimeo. 24 p. y anexos. (1)

12. PIÑEIRO, M. y TRIGO, E. "Un marco general para el análisis del progreso tecnológico agropecuario. Las situaciones de cambio tecnológico". IICA Bogotá. Misceláneo No. 149, abril de 1977.

---

(1) Documento presentado a la Reunión Técnica sobre Tipificación de Empresas Agropecuarias. IICA-DIEA. Montevideo, mayo de 1977.

(2) "Seminario sobre métodos y problemas en Tipificación de empresas agropecuarias" IICA. Serie de Informes de Conferencias, Cursos y Reuniones No. 92. Montevideo, diciembre de 1975.