

EL ANALISIS DE LOS DATOS EN LOS ESTUDIOS DE INVESTIGACION EN ECONOMIA AGRARIA

Por
EMILIO GOMEZ MANZANARES
Doctor Ingeniero Agrónomo

1. INTRODUCCIÓN

EN un artículo anterior (1) destacábamos las siguientes fases comunes a todo trabajo de investigación en economía agraria:

1. Selección del estudio.
2. Planteamiento del problema.
3. Formulación de hipótesis en relación con el problema.
4. Determinación de los datos (información) necesarios para probar las hipótesis formuladas.
5. Recogida de la información.
6. Elaboración y análisis de los datos.
7. Presentación de los resultados.

En el citado artículo desarrollábamos con alguna extensión los puntos 4 y 5. Un ejemplo concreto de aplicación de este tipo de estudios a una situación real, abarcando los siete puntos mencionados, ha sido publicado en esta revista (2 y 3).

En el presente artículo nos ocupamos con más detalle del análisis de los datos del estudio, que corresponde a una parte del punto 6, tratando de cubrir prácticamente todos los diferentes casos que puedan presentarse.

2. DISTINTAS CLASES DE ESTUDIOS SEGÚN EL TIPO DE ANÁLISIS EMPLEADO

El análisis de los datos y la deducción de conclusiones para su presentación posterior son los objetivos finales a los que se encamina el estudio de investigación. La intensidad y el tipo de aná-

lisis que exigen los distintos estudios de investigación varían notablemente, pero un análisis concienzudo y completo de los datos es esencial en todo trabajo de investigación que se precie de serlo.

A los fines del análisis, los estudios con los que el economista agrario se enfrenta la mayoría de las veces pueden clasificarse en seis tipos principales: 1) Descriptivos, 2) De series cronológicas, 3) De relaciones funcionales, 4) De sustitución, 5) Matemáticos, 6) Diversos.

En el primer caso, el problema principal es el de describir estadísticamente el valor medio y el campo de variación de las variables en estudio en situaciones determinadas, en un lugar fijo y en un momento dado. En el caso de estudios basados en series cronológicas, se estudian las relaciones entre el año, mes, día u hora, y una o más variables dependientes. Se llaman estudios basados en relaciones funcionales aquellos en los que se hace un análisis de las relaciones existentes entre uno o más factores causales y otros factores dependientes o resultantes. Los estudios de sustitución ("budgeting", para los anglosajones) comprenden la estimación del resultado de sustituir factores o métodos. Los estudios matemáticos son principalmente los econométricos y los que resultan de aplicar a los datos económicos técnicas de investigación operativa. Además de los cinco tipos de análisis más corrientes mencionados, existen otros varios empleados en estudios de investigación especiales. A veces en un trabajo de investigación se utilizan algunos de los tipos más comunes al mismo tiempo que algunos de los que hemos llamado diversos. Presentamos a continuación algunos de los problemas y de los procedimientos de análisis más corrientes dentro de los distintos tipos considerados.

3. ESTUDIOS DESCRIPTIVOS

A veces el economista agrario debe recoger los datos necesarios para describir ciertos aspectos de la agricultura, precios, comercialización de productos agrícolas y problemas similares. Generalmente, la información disponible o recogida para el estudio se refiere solamente a una parte del total, de modo que resulta de la mayor importancia el disponer de una muestra verda-

deramente representativa si se pretende generalizar las conclusiones del estudio a la población completa (1).

Medidas de tendencia central

Sea o no representativo el grupo del que existen datos disponibles, se presenta de todas formas el problema de reducir la información existente en cada caso a una o unas pocas cifras. Es decir, es preciso calcular alguna medida de tendencia central. De entre las cinco clases de promedios: media aritmética, mediana, moda, media geométrica y media armónica, sólo las tres primeras son de uso corriente en la investigación económico-agraria. La media aritmética es igual a la suma de todos los datos dividida por el número de ellos; la mediana es el valor que divide en dos partes iguales al total de datos, una vez ordenados por orden de magnitud; la moda es el valor más frecuente. Con demasiada frecuencia el investigador tiende a servirse de la media aritmética antes de considerar la naturaleza de la distribución y el fin a que sirve la media. Deberá cerciorarse de si la distribución es simétrica o asimétrica, y en el último caso hasta qué grado lo es. Deberá también decidir si le interesa que los valores extremos influyan en el valor medio utilizado. En otras palabras, deberá tener presentes las características de los distintos promedios en función del objetivo deseado.

Medidas de dispersión

Aunque resulta necesario servirse de valores medios sobre grupos de explotaciones, viviendas, empresas, etc., el condensar todos los datos de un estudio en una sola cifra, oscurece totalmente la variación que existe entre los datos. De este modo, el investigador se enfrenta con dos problemas: de un lado, reducir, por ejemplo, los tamaños de 100 explotaciones diferentes a una sola cifra (el valor medio) para su comprensión rápida y fácil; de otro lado, mostrar de alguna forma la variabilidad de los tamaños de las 100 explotaciones en cuestión. Con frecuencia se olvida este segundo problema y el lector se encuentra sólo con cifras medias, sin indicación alguna sobre la variación existente entre los distintos casos individuales.

Existen cuatro métodos corrientes de medir tal variabilidad. El más sencillo es el recorrido o campo de variación, que es sim-

plemente la diferencia entre el dato mayor y el menor. Esta medida es simple y fácilmente comprensible. Sin embargo, es relativamente inexacta, pues un caso extremo en cualquiera de los datos de la serie puede ocasionar un campo de variación grande, cuando tal vez la mayoría de los datos se encuentran concentrados relativamente cerca del valor medio. Los intervalos llamados cuartiles proporcionan otro método de medir la dispersión de los datos, dispuestos por orden de magnitud. La serie se divide en cuatro partes iguales. El valor del primer cuartil es el del dato de la serie situado a un cuarto de distancia del dato más pequeño al mayor. El segundo cuartil, a la mitad de la distancia, y el tercero a tres cuartos de la misma. La diferencia entre el primero y el tercer cuartil, llamada intervalo inter-cuartil, se utiliza con frecuencia como medida de la dispersión, pues el 50 por 100 central de todos los datos está comprendido entre estos dos límites. A veces este intervalo inter-cuartil se divide en dos partes iguales llamadas desviación inter-cuartila y la dispersión se expresa como desviación con respecto a la mediana. En otras palabras, la desviación inter-cuartila a ambos lados de la mediana abarca la mitad central de los datos, suponiendo una distribución normal.

La desviación media mide la variabilidad de cada dato de la serie a partir de la media aritmética o de la mediana. En la desviación media se tienen en cuenta las desviaciones de todos los datos respecto al promedio, aunque no se tiene en cuenta su signo. La media aritmética de las desviaciones de todos los datos, formados con signo positivo, es la desviación media. Una desviación media a cada lado de la media aritmética incluye el 57 por 100 de los casos, suponiendo una distribución normal (*). Como medida de dispersión, la desviación media (DM) es fácilmente comprensible y relativamente fácil de calcular, pero suele utilizarse con menos frecuencia que la desviación típica, con la que se puede operar algebricamente.

La desviación típica, al igual que la desviación media, mide la variabilidad de los datos de una serie a partir de una medida de tendencia central, en este caso siempre la media aritmética. La desviación típica es la raíz cuadrada de la media aritmética de los cuadrados de las desviaciones de los datos. Por este procedimiento de obtención se aumenta la influencia de los valores extremos. Por eso, la desviación típica de una serie es siempre mayor que la desviación media (en el caso de la distribución nor-

mal, resulta de (*) que $DM = 0.79788\sigma$). Una desviación típica a cada lado de la media incluye el 68,27 por 100 de los casos en una distribución normal. La desviación típica es la más utilizada de las medidas de dispersión, principalmente porque se presta a análisis más profundos en relación con la curva normal.

Distribución de frecuencias

Si un estudio, o una parte del mismo, se realiza fundamentalmente con propósitos descriptivos, tal vez los promedios y medidas de dispersión discutidos anteriormente dejen algo que desear en cuanto simple descripción de la variación entre los casos individuales. Puede resultar conveniente en tales casos construir una distribución de frecuencias disponiendo los datos con arreglo a alguna medida cuantitativa, dividiéndolos en grupos según intervalos o clases de la misma dimensión y contando el número de casos en cada intervalo. Este procedimiento reduce los numerosos casos a tres a diez números, según el número de intervalos considerados y el campo de variación de los datos, y proporciona una descripción clara y sencilla de la variación de la serie en cuanto a la característica estudiada. La distribución de frecuencias puede representarse en forma de cuadro o estado, y también gráficamente. De hecho, aumenta la comprensión de la variación existente entre las explotaciones o empresas objeto del estudio.

4. ESTUDIOS BASADOS EN SERIES CRONOLÓGICAS

La sociedad es dinámica. Las circunstancias cambian continuamente. Algunas variables aumentan mientras otras disminuyen con el paso del tiempo. Las superficies de cultivo de distintos productos agrícolas varían de un año a otro; los rendimientos varían; el número de cabezas de ganado varía; los precios varían no sólo anualmente, sino mensualmente, semanalmente y aun dia-

$$\begin{aligned} (*) \text{ En efecto, D. M.} &= 2 \int_0^{\infty} xf(x) dx = 2 \int_0^{\infty} \frac{x}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \sqrt{\frac{2}{\pi}} \int_0^{\infty} (-x \\ &e^{-\frac{1}{2}x^2} dx) = \sqrt{\frac{2}{\pi}} \left[e^{-\frac{1}{2}x^2} \right]_0^{\infty} = \sqrt{\frac{2}{\pi}}, \text{ es decir, D. M.} = \sqrt{\frac{2}{\pi}} \text{ desviaciones típicas} \\ &= \sqrt{\frac{2}{\pi}} \sigma = 0,79788 \sigma. \end{aligned}$$

Para $D. M. = 0.79788\sigma$ encontramos en la tabla de áreas de la distribución normal el valor 0.2875, es decir, que el área comprendida entre una desviación media a cada lado de la media aritmética sería 2×28.75 por 100 = 57.5 por 100 c. q. d.

riamente. Otras series cronológicas de interés para el economista agrario presentan variaciones semejantes. Estas variaciones en el curso del tiempo pueden ser de naturaleza secular o a largo plazo; de naturaleza cíclica, abarcando al menos tres años; estacional, sucediendo cada año, o accidentales. Con frecuencia, la producción o el precio de un producto agrícola en un momento dado es el resultado de los cuatro tipos de variaciones mencionados. Para efectuar predicciones más realistas el economista agrario tratará de aislar y determinar el grado de variación en cada tipo de series cronológicas. En algunos análisis, el investigador deseará eliminar la tendencia secular, el efecto cíclico o la variación estacional en los datos que estudia.

Tendencia secular

La variación en el tiempo, llamada tendencia secular, es el ritmo o tasa de incremento o disminución registrado a lo largo de una serie de años. Una tendencia secular puede ser una línea recta o una curva. Con el paso del tiempo, el ritmo de crecimiento o decrecimiento puede acelerarse o retardarse, y la tendencia puede hasta cambiar de sentido. La tendencia secular puede referirse al cambio en el transcurso del tiempo de una explotación individual, al de una parte o la totalidad de las explotaciones de una comarca, de una provincia o de la nación, al de un producto agrícola determinado o al de algún aspecto o característica de la economía entera.

Existen varios métodos para calcular la tendencia secular. El investigador interesado únicamente en una rápida aproximación puede trazar a ojo una línea, que, a su parecer, se ajuste lo mejor posible al movimiento secular. Esta línea puede ser recta o curva, según indiquen los datos. Tal línea podría, sin embargo, no ser expresada en forma matemática. La curva, una vez trazada, no sería mejor que el juicio de la persona que la haya trazado.

Otro modo de determinar la tendencia secular es por el método de los promedios. En el caso del método de semipromedios, el período considerado se divide en dos partes, se halla la media de cada una y se representan los valores obtenidos en el punto medio de cada semiperíodo; finalmente se traza una recta que una ambos puntos. Naturalmente, la tendencia calculada de esta

forma resulta una línea recta. Los datos, por otra parte, pueden ser promediados por períodos de cinco años, por ejemplo, y cada uno de estos promedios pueden representarse gráficamente en los respectivos puntos medios de cada subperíodo. La línea de tendencia así terminada no tiene porqué ser una línea recta. Este último procedimiento se acerca ya al método llamado de los promedios móviles. En este método se utilizan períodos de tres, cinco, siete o diez años y se obtienen los promedios correspondientes, que se representan en los respectivos puntos medios. Se empieza con el primer grupo de años; para obtener el segundo grupo de años se suprime el primer año y se añade el siguiente al último; en cada paso sucesivo se elimina el año más antiguo del grupo anterior y se añade un año más moderno al último considerado, hallándose cada vez el promedio correspondiente. Aunque este método es sencillo y flexible, la línea de tendencia va siempre varios años atrasada y se produce un sesgo.

El método de mínimos cuadrados se emplea con frecuencia en la investigación económico-agraria para determinar la línea de tendencia secular. En este método, el ritmo medio anual de variación y el valor calculado de la tendencia al principio del período se determinan matemáticamente a partir de los valores correspondientes a cada año. En el caso de una recta que se ajusta a los datos anuales la ecuación es de la forma $Y = a + bX$, en la que Y representa el valor calculado para un año dado, a es el valor al principio del período, b el ritmo anual de variación y X el número de años a contar a partir del primer año del período. Este procedimiento da precisamente una tendencia en forma de línea recta, cuyo grado de ajuste puede determinarse (*).

(*) Los coeficientes a y b de la recta se obtienen fácilmente. En efecto, sea δ_i la distancia de un punto cualquiera de los datos a la recta en cuestión, medida paralelamente al eje de ordenadas. Se tendrá:

$$\delta_i = y(x_i) - h_i = a + bx_i - y_i; \delta_i^2 = (a + bx_i - y_i)^2; \Sigma \delta_i^2 = \Sigma (a + bx_i - y_i)^2$$

y por definición la recta buscada es la que hace $\Sigma \delta_i^2 = \text{mínimo}$. Para ello:

$$\frac{\delta \Sigma \delta_i^2}{\delta a} = 2 \Sigma (a + bx_i - y_i) = 0, \quad \frac{\delta \Sigma \delta_i^2}{\delta b} = 2 \Sigma x_i (a + bx_i - y_i) = 0,$$

o sea: $na + b \Sigma x_i = \Sigma y_i$, $\Sigma x_i a + \Sigma x_i^2 b = \Sigma x_i y_i$, de donde se obtiene:

$$a = \frac{\begin{vmatrix} \Sigma y_i & \Sigma x_i \\ \Sigma x_i y_i & \Sigma x_i^2 \end{vmatrix}}{\begin{vmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{vmatrix}} \quad b = \frac{\begin{vmatrix} n & \Sigma y_i \\ \Sigma x_i & \Sigma x_i y_i \end{vmatrix}}{\begin{vmatrix} n & \Sigma x_i \\ \Sigma x_i & \Sigma x_i^2 \end{vmatrix}}$$

Muchas tendencias seculares no son líneas rectas. En tales casos el método descrito resulta inapropiado. A veces una parábola de segundo grado describe mejor la tendencia. Entonces basta hallar la solución de un sistema de tres ecuaciones normales. Existen otras fórmulas matemáticas y curvas que se utilizan a veces para la descripción precisa de tendencias seculares curvilíneas (4).

Con frecuencia, el economista agrario desea obtener una medida de la tendencia secular con el solo objeto de eliminar su influencia en los datos, y así poder estudiar las variaciones cíclicas o estacionales.

Ciclos

Numerosas ocasiones tiene el economista agrario de estudiar las relativamente regulares y recurrentes fluctuaciones denominadas ciclos. La duración de tales fluctuaciones varía desde tres hasta cincuenta años, dependiendo del producto y de las condiciones que le afectan. La duración del ciclo de algunos productos es más bien regular, mientras para otros productos varía considerablemente. Existen ciclos en el número de cabezas de ganado, en las producciones y los precios, en la actividad de los negocios, en el tiempo (meteorología) y en otros muchos datos económicos.

El problema de la determinación del ciclo consiste en medir su duración y regularidad, y la intensidad de la fluctuación durante el ciclo. Frecuentemente, el producto que presenta una variación cíclica, por ejemplo en las cifras de producción, presenta, al mismo tiempo, una tendencia secular que debe eliminarse de los datos anuales para poder determinar el ciclo. Existen diversos procedimientos para ello, tales como calcular el porcentaje de la línea de tendencia, el porcentaje del año precedente o el porcentaje del promedio móvil (5).

Tras obtener los datos sobre la duración, la regularidad y la intensidad de un ciclo, tal información puede resultar útil para la predicción de la producción futura probable. El investigador encuentra, de este modo, de una gran utilidad el análisis de la tendencia secular y de los ciclos al proporcionarle una información básica para sus predicciones sobre la producción ganadera y los precios, así como sobre otros muchos aspectos semejantes.

Variaciones estacionales

La producción de la mayor parte de los productos agrícolas, los precios de muchos de ellos, la actividad de los negocios y muchas otras actividades económicas presentan una variación más o menos regular dentro del año, que se llama variación estacional. Por ejemplo, los productos son recolectados en un momento determinado del año, pero generalmente son puestos en el mercado por los agricultores a lo largo de los doce meses del año a un ritmo relativamente regular. Los precios percibidos por los agricultores cada mes vienen influidos por la producción, comercialización y situación de la demanda. Así surgen las variaciones estacionales en los precios.

Existen diversos métodos para determinar la intensidad de la variación estacional. Si la variación se refiere a cantidades físicas de un producto, un método muy sencillo consiste en promediar la producción de cada mes durante un cierto número de años. Después se calcula un índice para cada mes indicando el porcentaje de la media anual producido en el mes. Convendrá asegurarse de que no ha cambiado sensiblemente la estructura de las variaciones estacionales en el período considerado. Cuando existe tendencia secular en la producción de un determinado producto, el procedimiento anterior da un resultado sesgado. En tal caso, es preciso determinar el efecto de la tendencia secular en cada mes y corregirlo por el método llamado de corrección de la tendencia o por otro método. Si se realiza un estudio de la variación estacional del precio de un producto, resulta conveniente corregir los datos del efecto de un cambio en el nivel general de precios. Esto puede hacerse dividiendo los precios por el índice de precios al por mayor o por cualquier otro índice de precios.

Después de obtener el índice mensual de la producción o del precio de un producto, el investigador puede utilizar tal índice para eliminar la variación estacional, si lo desea, con objeto de tener una mejor base de comparación.

5. ESTUDIOS BASADOS EN RELACIONES FUNCIONALES

Todo el mundo observa la relación existente entre condiciones y resultados y formula conclusiones, basadas en una o más ob-

servaciones, sobre la manera de hacer máxima su renta o su satisfacción. En la agricultura, por ejemplo, el agricultor se enfrenta con el problema de la relación entre la escala de operación y la renta máxima, la combinación de especulaciones y el máximo beneficio, y otras muchas. E. O. Heady menciona ocho tipos de decisiones principales a tomar en el campo de la economía de la producción agrícola (6):

- I. El volumen de producción que debe alcanzarse a partir de determinados recursos en cantidad limitada.
- II. La combinación de recursos necesaria para producir un cierto volumen de producción y unos determinados productos.
- III. La combinación de especulaciones en un período de tiempo determinado.
- IV. La elección de la época de producción (o de venta de los productos) en función de las variaciones estacionales de precios.
- V. El nivel de conservación del suelo y de los capitales.
- VI. La escala óptima de operación.
- VII. El control de los recursos a utilizar en la producción y la óptima combinación de los mismos.
- VIII. La adaptación a los cambios e incertidumbre del mercado y del proceso productivo, teniendo en cuenta el crecimiento de la empresa agrícola con el tiempo.

El estudio de la relación existente entre los factores causales y los resultantes constituye un intento de determinación de la naturaleza y grado de bondad de tal relación, y la prueba, no prueba o modificación de las hipótesis formuladas con respecto a la relación en cuestión. Es esta una de las fases más importantes en el trabajo del investigador en economía agraria. El estudio de estas relaciones funcionales se realiza generalmente según uno de los siguientes métodos de análisis: *a)* Gráfico, *b)* Tabular, *c)* Por parejas, *d)* Correlación, y *e)* Econométrico.

a) Análisis gráfico

Constituye uno de los procedimientos más sencillos de estudiar las posibles relaciones entre dos factores o variables. Se realiza sobre papel de gráficos, situando a determinada escala el

valor del factor causal sobre el eje de abscisas y el factor resultante sobre el eje de ordenadas. Cada observación o caso es representado por un punto, de acuerdo con sus coordenadas sobre ambos ejes. El resultado es un diagrama de dispersión o nube de puntos en el cual la dispersión de los casos individuales en el interior del diagrama indica si existe alguna relación entre los factores. La dispersión de los puntos en relación con una línea trazada a mano o calculada indica el grado de bondad o ajuste de la relación. Si los puntos en el gráfico forman un modelo compacto en diagonal, entonces puede decirse que existe una relación bastante estrecha. Si todos los puntos forman una línea recta, se dice que existe una relación perfecta. La línea trazada puede ser igualmente curva.

Si bien el método de representar gráficamente las observaciones por medio de puntos es sencillo, no se presta a expresar matemáticamente la relación existente. Aún más importante, no se tiene en cuenta la influencia de otros factores que pueden figurar asociados con el factor resultante y pueden, de este hecho, deducirse conclusiones erróneas sobre el efecto que produce un solo factor causal representado. Con todo, la preparación de diagramas de dispersión suele constituir una técnica previa al empleo de otros métodos de análisis, y resulta muy útil para presentar algunas relaciones.

Cuando se desea conocer la influencia por separado de dos o más factores causales sobre un factor resultante, es preferible servirse de la correlación gráfica. Este procedimiento requiere un mayor conocimiento estadístico, pero es utilísimo cuando se pretende aislar los efectos de un solo factor.

b) *Análisis tabular*

La agrupación de fichas con arreglo a un factor causal o resultante, y la tabulación de las medias aritméticas para los distintos grupos es uno de los métodos más simples de estudiar relaciones. Sin embargo, escasean mucho las publicaciones en el campo de la estadística aplicada que traten de esta técnica. En breve, el procedimiento seguido consiste en: 1.º Agrupar las fichas con arreglo a algunos factores; 2.º Tabular los valores to-

tales y medios para los factores causales y resultantes en cada grupo, y 3.º Determinar relaciones entre los valores medios de los diferentes grupos.

Agrupación de las fichas.—Al hacer la clasificación en grupos, las fichas son agrupadas con arreglo a un factor, y entonces, divididas en un número discreto de grupos, dependiendo del número de fichas disponibles. Es importante que haya un número suficiente de fichas en cada grupo, de manera que proporcionen valores medios que sean significativos.

Al tratar de medir las relaciones existentes entre dos factores, de los cuales uno es la causa y el otro el resultado, es preferible realizar la agrupación con arreglo al factor causal que con arreglo al factor resultante. Por ejemplo, si se desea conocer la relación existente entre la dimensión de la explotación y la renta atribuible al trabajo, la agrupación deberá hacerse con arreglo al tamaño y se calculará la renta del trabajo correspondiente a cada grupo de tamaño, en lugar de clasificar según la renta del trabajo para determinar la dimensión media de las explotaciones en cada uno de los grupos de renta. Generalmente, es deseable tabular en cada grupo no sólo los factores causales y resultantes, sino varios otros factores, que se piensa pudieran tener influencia sobre el factor resultante.

Clasificar en grupos con arreglo al factor resultante constituye, a veces, un buen método para explorar los datos, con el fin de buscar inspiración sobre aquellos factores que valdría la pena utilizar para agrupar con arreglo al factor causal (con el fin de formular hipótesis). Por ejemplo, en estudios sobre gestión de explotaciones en los que se desea encontrar los factores que explican las grandes rentas atribuibles al trabajo (*), es práctica corriente clasificar primero con arreglo a la renta del trabajo y tabular todos los factores que podrían relacionarse con ella. Cualquier factor relacionado con la renta serviría entonces como factor de agrupación para una tabulación separada. En el trabajo corriente de extensión agraria sobre contabilidades de explotación, el procedimiento frecuentemente utilizado consiste en agrupar con arreglo a la renta y establecer comparaciones entre las

(*) La renta atribuible al trabajo se obtiene deduciendo del valor de la producción final todos los gastos anuales (reales y calculados) de la explotación, a excepción de los gastos de trabajo y gestión.

explotaciones más rentables y las menos rentables. Este procedimiento está justificado en el trabajo de extensión, siempre que se tomen las debidas precauciones al seleccionar los elementos a comparar, provechando que la validez de muchas de las comparaciones ha sido establecida previamente.

En algunos casos se obtienen relaciones entre dos factores que no presentan una relación de causa a efecto. Ambos pueden ser resultados (variables dependientes) de otro factor que es la causa (variable independiente). Por ejemplo, la edad de los explotantes y la edad de sus mujeres están relacionadas, pero no puede decirse que la una sea causa de la otra.

Factores asociados.—A veces, el análisis de los grupos establecidos muestra la existencia de una relación aparente entre dos factores, dándose el caso de que el factor que se supone ser la causa no es sino resultado de un tercer factor, que es el verdadero causante de las variaciones de los dos primeros. Por ejemplo, es probable que las explotaciones donde existen llaves inglesas en gran número presenten rentas superiores a las encontradas en explotaciones que poseen llaves inglesas en menor cantidad. Pero el número de llaves inglesas en la explotación no es la causa de la diferencia de rentas. Las grandes explotaciones tienden a poseer más llaves inglesas que las pequeñas explotaciones, y, por otra parte, las grandes explotaciones suelen tener rentas superiores a las pequeñas. La razón real por la que las explotaciones con mayor número de llaves inglesas son más remuneradoras es que son también las más grandes.

A veces un factor puede ser realmente causa de variaciones en otro factor y estar, a su vez, asociado a un tercer factor causal. La clasificación en grupos con arreglo al factor primero o tercero puede ocultar una parte de la influencia que ejercen ambos factores en el valor medio resultante para el factor segundo, en los distintos grupos. Dos factores que sirven para ilustrar este caso, en el análisis de las explotaciones agrícolas, son la dimensión de la explotación y la eficiencia del trabajo. Si se dividen las fichas de explotación en tres grupos, con arreglo al número de U. T. H./hombre (*), las explotaciones con los valores más al-

(*) U. T. H./hombre = unidades de trabajo humano por persona activa en la explotación. El número de U. T. H. representa en este caso las exigencias medias de trabajo de la explotación, teniendo en cuenta las especulaciones practicadas.

tos de este coeficiente probablemente mostrarán también los valores más altos de renta del trabajo. La razón es, en parte (tal vez, sobre todo), que la eficiencia del trabajo, medida por el número de U. T. H./hombre, es función del tamaño de la explotación. Las explotaciones con mayor número de U. T. H./hombre son probablemente las grandes explotaciones, y éstas son más remuneradoras que las pequeñas. Pero la eficiencia del trabajo ejerce, sin embargo, cierta influencia sobre la renta en explotaciones de un tamaño dado. Variaciones en el número de U. T. H./hombre causan variaciones en la renta del trabajo, pero no tan grandes como la tabulación directa mostraría.

Subagrupación

Cuando se dispone de un número suficiente de casos a estudiar, resulta a veces conveniente clasificarlos en grupos con arreglo a un factor y establecer entonces subgrupos dentro de cada grupo, con arreglo a otro factor. A veces los subgrupos son susceptibles de dividirse en nuevas agrupaciones con arreglo a un tercer factor. Tendremos entonces las llamadas tabulaciones de doble entrada o de triple entrada, respectivamente.

Este procedimiento de establecer subgrupos es una manera de eliminar parte de la influencia de los factores primeros en el resultado. En otras palabras, la tabla de doble entrada muestra el efecto de un factor con el otro mantenido constante a dos o más niveles. Supóngase, por ejemplo, que se desea medir la influencia de la intensidad ganadera y de la eficiencia del ganado sobre la renta del trabajo. Pueden dividirse las fichas de explotación en dos grupos con arreglo al número de U. G. (*) por hectárea, y cada uno de estos dos grupos puede subdividirse, a su vez, en dos subgrupos con arreglo a un índice de eficiencia del ganado (**). Los cuatro subgrupos representan, por consiguiente: 1) Explotaciones con mucho ganado y gran eficiencia; 2) Explotaciones con mucho ganado y poca eficiencia; 3) Explotaciones con poco ganado y gran eficiencia, y 4) Explotaciones

(*) U. G. = unidades ganaderas; constituyen un intento de reducir los efectivos de las distintas clases de ganado en la explotación a un común denominador. Las equivalencias se realizan teniendo en cuenta principalmente las necesidades energéticas de alimentación.

(**) Cualquier índice ponderado de los rendimientos físicos del ganado en la explotación, por ejemplo, en relación a los gastos de alimentación del mismo.

con poco ganado y poca eficiencia. Los valores medios obtenidos para la renta del trabajo en cada uno de los cuatro grupos pueden indicar que las explotaciones más rentables son las que disponen de mucho ganado, con gran eficiencia en la producción ganadera y que las menos rentables son las que tienen mucho ganado, pero con un bajo nivel de eficiencia. Esta conclusión —en el caso de que se desprendiese de los datos analizados— difiere considerablemente de la que se hubiera obtenido de estudiar únicamente la relación entre la intensidad ganadera y la rentabilidad de las explotaciones, por medio de un análisis tabular de una sola entrada.

Medias ponderadas y sin ponderar.—Si se comparan dos o más grupos de explotaciones con arreglo al número de hectáreas de S. A. U. (superficie agrícola útil) por U. T. H. (o cualquier otro coeficiente que exprese la relación de un factor a otro), ¿deberá considerarse como cifra media del grupo la media de los valores S. A. U./U. T. H. de cada explotación del grupo, o la que resulta de dividir el total de S. A. U. por el total de U. T. H.?

La primera de ambas cifras daría a cada U. T. H. el mismo peso en la determinación del número medio de S. A. U./U. T. H., independientemente del tamaño de las explotaciones (medido por el número de hectáreas de S. A. U.). La segunda cifra vendría ponderada con arreglo al tamaño de las explotaciones.

Hay ocasiones en las que la media ponderada resulta preferible. En muchos casos, sin embargo, la media sin ponderar es más apropiada. Supóngase que se compara la producción de huevos procedentes de gallinas alimentadas con suplemento proteínico y sin él. Los productores con aves alimentadas con suplemento proteínico, con 500 aves cada uno, obtienen 60 huevos por ave, salvo un solo propietario de 1.000 aves, que obtiene 180 huevos por ave. Todos han seguido el mismo sistema de alimentación; el propietario de las 1.000 aves obtuvo un resultado al cual no se deberá dar más peso que al obtenido por cada uno de los demás propietarios. La utilización de medias sin ponderar facilita, por otra parte, la determinación estadística de la significación de las diferencias entre valores medios.

Naturaleza de la relación.—La relación entre un factor causal y un factor resultante puede ser lineal o curvilínea. Es lineal

cuando, a cada variación unitaria del factor causal (variable independiente) corresponde una variación constante, del mismo o de distinto signo, del factor resultante (variable dependiente) en todo el intervalo de los datos. La relación es curvilínea cuando a variaciones unitarias de la variable independiente correspondan variaciones desiguales de la variable dependiente. En el análisis tabular no suele hacerse ningún supuesto anticipado acerca de la naturaleza de la relación: ésta suele descubrirse más adelante, en el análisis de los datos; podrá ser lineal o curvilínea, pero ello no afecta en nada al procedimiento de análisis empleado. En la mayoría de los otros métodos de análisis es preciso formular previamente algunas hipótesis antes de decidir qué técnica estadística de análisis se empleará.

Es relativamente sencillo en el análisis tabular agrupar los datos en tablas de doble y aun de triple entrada para estudiar problemas que comprenden dos o más variables independientes. No se precisa desarrollar complicadas fórmulas estadísticas, pero sí se requiere estar en posesión de un número considerable de fichas con objeto de disponer de un número adecuado en cada grupo.

La tabulación puede mostrar la existencia de una relación aparente entre dos factores, que puede ser simplemente accidental. Cuanto menor sea el número de casos promediados en cada grupo y mayor sea la variación entre los valores de cada grupo, tanto menor será la probabilidad de que la relación encontrada en la muestra (datos de que nos hemos servido) sea la verdadera relación, válida para toda la población o universo a estudiar. A veces el criterio de medida usado para diferenciar los grupos no es el más apropiado, aun cuando parezca resistir la prueba de significación estadística. Por ejemplo, con frecuencia se obtienen conclusiones diferentes cuando se utiliza la mediana, un coeficiente de correlación, o un coeficiente de regresión al establecer comparaciones en lugar de la media aritmética. La validez y consistencia de las relaciones obtenidas por el análisis tabular o por otros métodos deberán ser cuidadosamente comprobadas antes de generalizar las conclusiones.

Resumiendo sobre la utilidad del análisis tabular: es un método sencillo, que no requiere el empleo de fórmulas complejas, no exige el anticipar supuestos acerca de la naturaleza lineal o curvilínea de la relación existente, uno o dos factores indepen-

dientes además del factor que se analiza pueden ser tenidos en cuenta y mantenidos relativamente constantes a diferentes niveles, y resulta fácil la presentación de los datos en simples cuadros. Por otra parte, el análisis tabular exige un número considerable de casos para hacer cualquier subagrupación; el mantener factores independientes suplementarios a un nivel constante en los distintos grupos no resulta fácil; las tablas de doble y triple entrada son difíciles de analizar y presentar en forma simple; y el comprobar la consistencia de los resultados del análisis presenta también algunas dificultades.

c) *Análisis por parejas*

Es este un método relativamente poco utilizado en la investigación agro-económica. Es una especie de tabulación llevada al límite. Suele ser de utilidad especialmente en aquellos casos en que la clasificación de las fichas en grupos y subgrupos no permite aislar el efecto de un solo factor.

Cuando dos factores están relacionados y se desea averiguar si cada uno de ellos se relaciona independientemente con un tercer factor, un procedimiento sencillo para descubrirlo consiste en disponer los casos en parejas con arreglo a un factor y luego tabular la influencia del otro.

El tamaño de la explotación y la eficiencia del trabajo son ejemplos de factores que se prestan bien al empleo del método. Supóngase que se desea saber si la eficiencia del trabajo se relaciona con la renta del trabajo con independencia de la relación que presenta con el tamaño de la explotación. Las fichas de explotación pueden disponerse en orden, de mayor a menor, con arreglo a la superficie agrícola, en hectáreas. Cada pareja de fichas contiguas puede considerarse como representando un par de explotaciones prácticamente de la misma superficie. La eficiencia del trabajo en una de ellas será, con toda probabilidad, mayor que la de la otra, con la que constituye el par. Pues bien, se obtendrán los valores medios de la eficiencia del trabajo y de la renta del trabajo que corresponden a todas las fichas de "alta" eficiencia del trabajo (es decir, las constituidas por cada una de las fichas de cada pareja formada, que presenta una cifra de eficiencia del trabajo superior a la de la otra) y los mismos va-

lores medios relativos a las fichas de "baja" eficiencia del trabajo. El tamaño medio de las explotaciones en ambos grupos será sensiblemente el mismo, como resultado del procedimiento seguido para establecer los pares; por tanto, cualquier diferencia que aparezca entre la renta media del trabajo del grupo de "alta" eficiencia y la del grupo de "baja" eficiencia deberá atribuirse exclusivamente a diferencias en la eficiencia del trabajo.

El análisis por parejas constituye un procedimiento más eficaz para eliminar la influencia de los factores asociados que el análisis tabular. Según este último, las explotaciones se habrían clasificado primeramente en pequeñas, medias y grandes, y luego cada uno de estos grupos se habría subdividido con arreglo a la eficiencia del trabajo. Pero, aun dentro de cada grupo establecido, según el tamaño, existe cierta variación en este factor. Esta variación puede influir en el valor medio de la renta del trabajo de los tres subgrupos en que cada grupo está dividido. Por el contrario, en el análisis por parejas el factor tamaño queda prácticamente eliminado.

d) *Análisis de correlación*

Al igual que los otros métodos de análisis considerados, el análisis de correlación tiene como objeto principal el aislar e identificar la verdadera relación entre causas y efectos. El problema puede descomponerse en tres partes: 1.º La determinación de la forma que adopta la relación, es decir, la ecuación de la línea de regresión; 2.º La medida de la variación de los datos con respecto a la línea de regresión, es decir, la desviación típica (S_y), y 3.º La conversión de la desviación típica (S_y) en un valor *standard*, el coeficiente de correlación (r).

La determinación de la línea de regresión se realiza por el método de "mínimos cuadrados" que hemos descrito anteriormente, al tratar de la tendencia secular en las series cronológicas. La misma fórmula sirve para el caso: $Y = a + bx$. En este caso Y sigue siendo la variable dependiente; a es el valor calculado de Y cuando $x = 0$; b es la tasa calculada de variación de Y cuando se produce una variación unitaria de x , o, de otro modo, la pendiente o coeficiente angular de la línea de regresión; y x es la variable independiente. Por ejemplo, Y puede designar

la renta de la explotación, expresada en miles de pesetas, y x la superficie agrícola útil de la explotación, en hectáreas. Si se representan los datos en un diagrama de dispersión, figurando la superficie en el eje de abscisas y la renta en el de ordenadas, y se dibuja la línea de regresión calculada, no sólo se tendrá una medida gráfica (y analítica, por supuesto) de la relación existente entre las dos variables en cuestión, sino que podrá observarse al mismo tiempo la dispersión de los casos individuales alrededor de la línea de regresión. La desviación típica con respecto a la recta de regresión da la medida matemática de tal dispersión. La fórmula para el cálculo de ésta es

$$S_y = \sqrt{\frac{\sum (d^2)}{N}}$$

donde d es la desviación de los valores reales con respecto a la línea de regresión y N el número de casos. Es evidente que cuanto menor sea S_y más estrecha será la relación entre los factores considerados. El valor (S_y) puede utilizarse de la misma manera que la desviación típica en una distribución normal para expresar la dispersión. Así, por ejemplo, una desviación típica (S_y) a cada lado de la línea de regresión incluye el 68 por 100 de los casos, suponiendo que se trate de una distribución normal.

La desviación típica con respecto a la recta de regresión se expresa siempre en función de la variable Y (en el ejemplo considerado, en miles de pesetas de renta). Ahora bien, como con frecuencia suele resultar conveniente el poder comparar los valores de S_y obtenidos en dos series con distintas unidades para la variable dependiente Y , es necesario hacer tales valores comparables, calculando un coeficiente *standard* que no venga influido por el sistema de unidades de que se trate en cada caso. Este es el llamado coeficiente de correlación (r), en el cual la desviación típica con respecto a la línea de regresión viene dividida por la desviación típica de Y . La fórmula es:

$$r = \sqrt{1 - \frac{S_y^2}{\sigma_y^2}}$$

El valor de r varía entre $+1$ y -1 , y cuanto más próximo a $1,00$ más estrecha será la relación entre los factores dependiente e independiente. Si todos los casos individuales se encontrasen

sobre la recta de regresión, r sería igual a 1. Si, por otra parte, $r = 0,70$ en el análisis, solamente el 49 por 100 de la variación de Y está relacionado con variaciones en X (en efecto, el coeficiente de determinación r^2 resulta ser 0,49).

Esta breve presentación del análisis de correlación simple ha servido para mostrar el procedimiento básico que se sigue, pero muchos problemas de economía agraria son bastante más complejos que los que pueden tratarse adecuadamente por esta técnica. En la correlación simple, la naturaleza de la relación y el grado de bondad de la misma vienen ambos expresados por una sola cifra, pero únicamente se considera un factor despreciando la influencia de los otros, y se supone además que la relación es lineal, lo cual no es con frecuencia el caso.

El análisis de correlación múltiple ofrece la ventaja, sobre la correlación simple, de poder determinar la relación entre dos o más variables independientes y el factor dependiente. El procedimiento básico es el mismo. De nuevo se supone que las relaciones son lineales. También se hace el supuesto de que las relaciones son aditivas (las mismas con independencia del valor de las otras variables independientes), aun cuando es más probable que no lo sean. Tampoco suministra información alguna sobre la importancia relativa de los distintos factores que entran en juego.

El análisis de correlación parcial suprime algunos de los inconvenientes de la correlación múltiple al permitir la determinación de la importancia relativa de los varios factores independientes en su efecto sobre la variable dependiente. Este se realiza calculando la relación existente entre un solo factor independiente y la variable dependiente, cuando se mantienen constantes todos los demás factores independientes comprendidos en el estudio. Es, en particular, deseable cuando los factores independientes se relacionan entre sí. Las relaciones obtenidas por medio de la correlación parcial, sin embargo, se suponen también lineales y aditivas.

El análisis de correlación curvilínea se emplea con frecuencia en estudios de investigación agro-económica, cuando las relaciones en cuestión no son lineales, como era el caso en los tres tipos de correlación que hemos descrito. Si la relación entre un factor independiente y la variable dependiente es curvilínea, una curva de ajuste puede calcularse matemáticamente o simplemente dibujarse a ojo con objeto de mostrar la relación. Hay muchos tipos

de curvas con sus correspondientes fórmulas matemáticas, que resultan apropiadas para expresar tales relaciones. El problema de seleccionar la ecuación de la curva que mejor describa la relación entre los dos factores es la fase más importante del análisis. Una vez que la curva que mejor refleja la relación buscada ha sido hallada, se podrá determinar la bondad del ajuste a los datos reales por medio del cálculo del índice de correlación. El procedimiento es semejante al utilizado para el cálculo del coeficiente de correlación. La única diferencia es que el coeficiente de correlación r media la dispersión con arreglo a una recta, mientras que el índice de correlación indica la dispersión con arreglo a la curva utilizada para describir la relación.

Si se desea estudiar la influencia de más de un factor independiente en un análisis de relaciones curvilíneas, ello es posible sin más que realizar ciertos cambios de variables, antes de aplicar los métodos de mínimos cuadrados. Los números naturales que expresan los valores de los factores independientes pueden sustituirse por sus logaritmos, cuadrados, recíprocos, raíces cuadradas, etc., convirtiéndose la ecuación de la curva en cuestión, después del cambio de variables, en la de una recta de regresión en las nuevas variables. En un artículo anterior (7) hemos aplicado esta técnica a la determinación de funciones de producción del tipo Cobb-Douglas, que responden a la fórmula general:

$$Y = r X_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n}$$

Aplicando logaritmos se obtiene:

$$\log Y = \log r + \alpha_1 \log X_1 + \alpha_2 \log X_2 + \dots + \alpha_n \log X_n$$

y haciendo los cambios de variable:

$$\log Y = y; \log r = \alpha_0; \log X_1 = x_1; \log X_2 = x_2; \dots; \log X_n = x_n$$

y sustituyendo, queda:

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$$

que es la ecuación de una recta de regresión, fácilmente determinable por el método de mínimos cuadrados.

Es posible también estudiar relaciones no lineales (o curvilíneas) múltiples por los llamados métodos de aproximación; así, por ejemplo, los debidos a M. A. Ezekiel (5).

La decisión a adoptar sobre el método de análisis que resulte más adecuado para aislar y medir con precisión la relación funcional de las variables en estudio depende de varios factores; entre los más importantes se encuentran: el tipo de problema de que se trate, la cantidad y la calidad de los datos disponibles para el estudio, el número de variables independientes y la naturaleza de sus relaciones con el factor resultante, el grado de flexibilidad deseado en el análisis, el equipo de tabulación existente, así como el personal auxiliar disponible, el refinamiento de análisis deseado, la formación estadística del investigador y la utilización que se hará de los resultados.

6. ESTUDIOS DE SUSTITUCIÓN

Son los que se sirven principalmente de los métodos llamados de sustitución ("budgeting", en la terminología anglosajona). El análisis de sustitución hace uso generalmente de datos secundarios y, por consiguiente, hay en los estudios de investigación de este tipo una parte relativamente mayor de análisis propiamente dicho, a diferencia de otros estudios que requieren para empezar la recogida de datos primarios.

En el análisis de sustitución el investigador parte de algunos hechos que conoce sobre factores de producción ("inputs") y producciones ("outputs"), los dispone en una situación determinada o bien los sustrae del medio en que se encuentran, y crea una nueva situación sobre lo cual no dispone de datos empíricos. El puede, no obstante, determinar las características de la nueva situación, basándose en su conocimiento de las características y efectos de los hechos que conoce. La teoría económica constituye una ayuda valiosa para anticipar los resultados probables. La inclusión o exclusión de tales hechos y el cálculo de los resultados constituyen, en esencia, el análisis de sustitución.

Un ejemplo relativamente simple de análisis de sustitución en la investigación agro-económica es la determinación del efecto que tendrá sobre la renta de una explotación agrícola la sustitución de una parte creciente de la superficie destinada a cereales por una superficie equivalente de cultivos forrajeros, con vistas a mejorar la conservación del suelo. La unidad de análisis es, en este caso, el conjunto de la explotación. El investigador nece-

sitará conocer, para empezar, los efectos de los cambios del plan actual de producción sobre los rendimientos y la producción de cada cultivo. Probablemente un cambio en el plan de producción comportará, a su vez, cambios en las prácticas de cultivo y se necesitará conocer los efectos que éstos ocasionan. Con una mayor cantidad de forraje disponible para la alimentación del ganado el investigador precisará ahora conocer las tasas de sustitución forraje-grano para varios niveles de sustitución. Necesitará además conocer el efecto que ocasionan las modificaciones introducidas en los medios de producción, tales como la maquinaria y la mano de obra. En cuanto tales efectos sean diferentes en diferentes períodos de tiempo, el investigador precisará también conocer las relaciones temporales existentes. Con un conocimiento razonable de todos estos puntos, estará en posición de establecer un cuadro de "antes y después", que indicará el probable efecto combinado de todos los cambios apuntados. Probablemente hará una serie de estimaciones para comparar los efectos de diversos grados de sustitución.

Al preparar un cuadro de sustitución, con frecuencia resulta innecesario incluir todos los factores ("inputs") y productos ("outputs"), sino sólo aquellos que sufren algún cambio al operarse la sustitución. Por ejemplo, en el caso que estamos considerando, si no existe ningún motivo para esperar un cambio en el impuesto territorial o en el seguro de la vivienda y construcciones, tales conceptos pueden suprimirse tanto en la fase de "antes" como en la de "después". Si la mano de obra permanente y gran parte de la maquinaria van a permanecer constantes, tales costes "fijos" podrán omitirse y las comparaciones de factores limitarse a los costes "variables" que, naturalmente, cambiarán con los nuevos sistemas a analizar.

El análisis de sustitución requiere probablemente un mayor conocimiento técnico en su campo de aplicación y una mayor imaginación que otros tipos de análisis que se limitan a medir resultados del pasado o relaciones existentes. Resulta, por otra parte, menos costoso que medir los resultados de cambios reales. Permite el estudio y la verificación de cambios en la organización y gestión de la explotación cuando resulta imposible medir los resultados, y hace posible la predicción de resultados probables debidos a cambios introducidos en el sistema.

7. ESTUDIOS MATEMÁTICOS

Entran dentro de esta categoría amplia de estudios de investigación todos aquellos que requieren un tratamiento matemático especial de los datos y que exigen, por consiguiente, una formación matemática avanzada por parte del investigador.

Pertenecen a esta categoría los estudios econométricos, propiamente dichos, que implican, con frecuencia, el uso de ecuaciones simultáneas. Tales estudios se imponen, generalmente, cuando el problema a resolver es suficientemente concreto. También se incluyen en este grupo los estudios más en boga hoy en día: los que resultan de aplicar técnicas de la investigación operativa a problemas de economía agrícola. Así, por ejemplo, son ya numerosos los estudios de investigación agro-económica basados en la aplicación de conocidas técnicas de la investigación operativa, tales como la programación lineal, cuadrática y dinámica; la teoría de líneas de espera; la teoría de juegos de estrategia; los problemas de afectación; las cadenas de Markov y procesos estocásticos; los problemas de transporte; el control de inventarios; el análisis secuencial; los métodos de simulación (tipo Monte-Carlo); la teoría de la información, etc.

En un trabajo anterior (8) hemos expuesto los fundamentos de la programación lineal y hemos aplicado esta técnica a la resolución de problemas que se presentan al nivel de la empresa agrícola. En posteriores trabajos aplicaremos esta técnica a la resolución de problemas de porte macroeconómico y haremos uso también de otras técnicas de investigación operativa.

Conviene, tal vez, hacer hincapié en el bagaje matemático de que ha de disponer el investigador para realizar este tipo de estudios, bagaje que no ha de basarse tanto en las antiguas disciplinas de la ciencia matemática, tales como el cálculo infinitesimal e integral, trigonometría, geometría clásica y proyectiva, etcétera, cuanto en ramas más modernas, tales como el análisis vectorial y el cálculo matricial, la lógica simbólica, el cálculo de probabilidades, la teoría de conjuntos, y, naturalmente, la estadística matemática y las técnicas específicas ya mencionadas de la investigación operativa.

En la mayoría de los casos, la disponibilidad de calculadoras electrónicas (computadores digitales y analógicos) facilita notablemente la realización de este tipo de estudios.

8. ESTUDIOS DIVERSOS

En algunos estudios de investigación, la utilización de los datos no se ajusta muy bien a ninguno de los procedimientos de análisis que hemos considerado. Aunque resulta normal el considerar cualquier proceso de interpretación y utilización de los datos como parte de la fase analítica del estudio de investigación, a veces el uso de los datos implica una síntesis más que un verdadero análisis.

El método de sustitución ("budgeting") constituye un claro ejemplo. Cuando el investigador reemplaza ciertas cifras por otras para estudiar su efecto, en realidad está sintetizando una situación. Tal vez, sin embargo, la sustitución ha sido precedida de un análisis, realizado por él mismo o por algún otro investigador, con objeto de obtener las cifras que él emplea en el proceso de sustitución. Y aunque el proceso mecánico real de sustitución tiene más de síntesis que de análisis, el mismo proceso puede capacitarle para obtener un conocimiento sobre los efectos separados de factores individuales. De este modo, en realidad está analizando relaciones.

Otro ejemplo del uso de la síntesis es la búsqueda de la utilización más eficiente de la mano de obra en una fábrica de transformación de productos agrícolas. Un trabajo de investigación con este objetivo debería comprender el estudio de un cierto número de fábricas representativas, incluyendo un análisis de tiempos y movimientos en los diferentes procesos; como consecuencia de este análisis, se intentaría construir un procedimiento mejorado combinando los métodos más eficientes para cada uno de los diversos procesos. Esta síntesis constituye la aplicación funcional del análisis que la precedió. Puede, a su vez, continuarse con otros análisis para verificar las ventajas prácticas de la solución hallada.

Muchos estudios de investigación agro-económica se sirven de otros métodos diversos de interpretar y utilizar los datos que los que hemos descrito especialmente.

LITERATURA CITADA

- (1) GÓMEZ MANZANARES, E.: "Los estudios de investigación en economía agraria", REVISTA DE ESTUDIOS AGRO-SOCIALES, número 37, octubre-diciembre, 1961.
- (2) BOTELLA Y FUSTER, E., y GÓMEZ MANZANARES, E.: "Planteamiento y organización de un estudio económico de las explotaciones agrícolas de la zona de la Violada", REVISTA DE ESTUDIOS AGRO-SOCIALES, número 40, julio-septiembre, 1962.
- (3) BOTELLA Y FUSTER, E.: "Estudio económico de un grupo de explotaciones agrícolas en la zona de la Violada", REVISTA DE ESTUDIOS AGRO-SOCIALES, número 42, enero-marzo, 1963.
- (4) NEISWANGER, W. A.: "Elementary Statistical Methods", *The Mac Millan Co.*, 1943.
- (5) PEARSON and BENNET: "Statistical Methods", *John Wiley and Sons*, 1942.
- (6) HEADY, E. O., "Elementary Models in Farm Production Economic Research", *Journal of Farm Economics*, vol. XXX, número 2, 1948.
- (7) GÓMEZ MANZANARES, E., "Funciones de producción en la agricultura", REVISTA DE ESTUDIOS AGRO-SOCIALES, número 48, julio-septiembre 1964.
- (8) GÓMEZ MANZANARES, E.: "Programación de explotaciones agrícolas", REVISTA DE ESTUDIOS AGRO-SOCIALES, número 49, octubre-diciembre 1964.

RESUMEN

Como continuación de un artículo ya aparecido en esta revista, y formando parte de una serie sobre la investigación en economía agraria, el autor estudia los principales métodos de análisis de la información utilizados en la investigación económico-agraria. Los métodos estudiados son: los descriptivos, los basados en series cronológicas, los de relaciones funcionales, los de sustitución, los matemáticos y otros varios. De todos ellos hace una breve descripción, haciendo hincapié en la utilidad y oportunidad de su aplicación para estudiar determinados aspectos de la investigación en economía agraria, sin pretender profundizar excesivamente en la teoría. Los métodos de análisis de relaciones funcionales son tratados con algún detalle, especialmente en lo que se refiere al análisis gráfico, análisis tabular, análisis por parejas y análisis de correlación.

R E S U M E

Comme suite à un article, déjà publié dans cette revue, et faisant partie d'une série d'articles sur la recherche sur l'économie agraire, l'auteur étudie les principales méthodes d'analyse de l'information utilisées dans la recherche économique-agraire. Les méthodes étudiées sont: les méthodes descriptives, les méthodes basées sur des séries chronologiques, celles des relations fonctionnelles, celles de substitution, les méthodes mathématiques et diverses autres. Il fait de chacune une brève description en insistant sur l'utilité et l'intérêt de leur application pour étudier des aspects déterminés de la recherche en économie agricole, sans chercher à approfondir trop la théorie. Les méthodes d'analyse des relations fonctionnelles sont traitées un peu en détail, particulièrement en ce qui concerne l'analyse graphique, l'analyse tabulaire, l'analyse par couples et l'analyse de corrélation.

S U M M A R Y

In continuation of an article that has already appeared in this magazine and forming part of a series of articles on research into agrarian economy, the author studies the principal methods of analysis used in economic-agrarian research. The methods studied are: the descriptive ones, those based on time series, those based on functional relationships, those based on substitution, the mathematical ones and several others. He gives a brief description of them all and dwells upon the utility and opportunity of their application for the study of particular aspects of research into agrarian economy, without attempting to go too deeply into the theory. The methods of analysis by functional relationships are treated in some detail, especially with regard to graphic analysis, tabular analysis, analysis by pairs and analysis of correlation.
