

## Propuesta para evaluar el proceso de adopción de las innovaciones tecnológicas\*

### Proposal to evaluate the process of adoption of technological innovations

Blanca Isabel Sánchez Toledano<sup>1§</sup>, Jorge A. Zegbe Domínguez<sup>1</sup> y Agustín F. Rumayor Rodríguez<sup>1</sup>

<sup>1</sup>Campo Experimental Zacatecas. INIFAP. Carretera Zacatecas-Fresnillo, km 24.5. A. P. 18 Calera de V. R., Zacatecas. 98500. Tel. (478) 9-85-01-98 y 9-85-01-99. (jzegbe@zacatecas.inifap.gob.mx). <sup>§</sup>Autora para correspondencia: bsanchez@zacatecas.inifap.gob.mx.

#### Resumen

Es de importancia para las instituciones involucradas en el desarrollo del sector agropecuario conocer el grado de adopción y la velocidad con que las innovaciones tecnológicas son aceptadas por los usuarios. Es decir, cómo una novedad o innovación tecnológica deja de ser experimental y se transforma en una práctica de uso común en el sector social. Las variables de respuesta y el número de observaciones que integran las muestras bajo estudio proporcionan información valiosa por la diversidad de temas que pueden ser estudiados. Sin embargo, las variables de respuesta investigadas normalmente son expresadas en diferentes escalas cualitativas. En consecuencia, la información es analizada inapropiadamente o ésta es presentada descriptivamente. Por lo tanto, el objetivo de esta propuesta es presentar el uso de técnicas estadísticas para analizar y evaluar el proceso de adopción de innovaciones tecnológicas. La metodología que se propone incluye escalas cualitativas para analizarse con técnicas multivariadas como prueba de normalidad multivariada, análisis por componentes principales, aglomerados y correlación canónica. El uso de estas herramientas estadísticas se ejemplificó con un estudio de adopción tecnológica realizada con productores de cebada.

**Palabras clave:** análisis multivariado, diseño de encuesta productores de cebada.

#### Abstract

It is of importance to the institutions involved in the development of the livestock sector to know the extent of adoption and speed with which technological innovations are accepted by users. I.e., how a novelty or technological innovation is no longer experimental and is transformed into a practice commonly used in the social sector. Response variables and the number of observations that integrate the samples under study provide valuable information for the diversity of topics that can be studied. However, the investigated response variables are normally expressed at different qualitative scales. Therefore the information is analyzed inappropriately or it is presented descriptively. Therefore, the aim of this proposal is to present the use of statistical techniques to analyze and evaluate the process of adoption of technological innovations. The proposed methodology includes qualitative scales to be analyzed with multivariate techniques such as the multivariate normality test, principal component analysis, clusters and canonical correlation. The use of these statistical tools exemplified by a study of technology adoption made with barley producers.

**Key words:** multivariate analysis, survey design to barley producers.

\* Recibido: diciembre de 2012  
Aceptado: junio de 2013

## Introducción

Es de importancia para las instituciones involucradas en el desarrollo del sector agropecuario conocer el grado de adopción y la velocidad con que las innovaciones tecnológicas son adoptadas por los usuarios. Es decir, cómo una innovación tecnológica deja de ser experimental y se transforma en una práctica de uso común en el sector social.

La adopción de una innovación tecnológica se define como el proceso por el cual el productor agropecuario sustituye una práctica de uso común por otra novedosa; esto implica un proceso de aprendizaje y cambio del sistema de producción (Seré *et al.*, 1990). Por ejemplo, una práctica nueva o innovación tecnológica puede ser el desarrollo de un herbicida más económico y eficaz en el control de malas hierbas; otras podrían ser la generación de semilla mejoradas, una nueva raza de ganado, prácticas de prevención de incendios para el bosque, un ajuste en la fecha y densidad de siembra, maquinaria agrícola nueva, el cambio en la forma de siembra o de cosecha de un producto, entre otros (Seré *et al.*, 1990). También, una innovación tecnológica puede estar integrada por diferentes componentes tecnológicos, los cuales pueden ser fácilmente distinguidos por los usuarios (Hernández y Porras, 2006).

No obstante que el desarrollo de una tecnología es largo, una vez terminada, ésta enfrenta siempre la posibilidad de ser o no aceptada por el usuario potencial (i. e. por el agricultor, en este caso). Por lo tanto, conocer las razones que provocan uno u otro evento es útil para las instituciones dedicadas a la investigación agropecuaria y para otros organismos encargados de la difusión de la tecnología. Esto último facilita la búsqueda de estrategias que permitan incrementar la probabilidad de que las innovaciones tecnológicas sean adoptadas en el menor tiempo posible; pero además conocer el impacto de éstas en un sistema de producción en particular.

Al respecto, la FAO (1988) sugiere que la generación y adopción de las nuevas tecnologías deben realizarse paralelamente con el productor, tomando en consideración la propia idiosincrasia del productor, su cultura, sus intereses y las condiciones agroecológicas y económicas en la que éste se desarrolla. Estos aspectos representan, en la mayoría de los casos, una seria condición que limita la adopción de una tecnología. El análisis de encuestas a través

## Introduction

It is of importance to the institutions involved in the development of the agricultural sector to know the extent of adoption and speed with which technological innovations are adopted by users. I.e., how a technological innovation is no longer experimental and is transformed into a practice commonly used in the social sector.

The adoption of a technological innovation is defined as the process by which the producer replaces a practice commonly used by another novelty; this involves a process of learning and change in the production system (Sere *et al.*, 1990). For example, a new practice or technological innovation may be the development of a more economical and effective herbicide to control weeds; others could be improvement of seed generation, a new breed of cattle, forest fire prevention practices, an adjustment in planting date and plant density, new agricultural machinery, the change in the form of planting or harvesting of a product, among others (Sere *et al.*, 1990). Also, a technological innovation can be integrated by different technological components, which can be easily distinguished by the users (Hernández and Porras, 2006).

Although the development of a technology is long, once completed, it always faces the possibility of being or not accepted by the potential user (i.e., by the farmer, in this case). Therefore, to know the reasons that cause either event is useful for institutions dedicated to agricultural research and for other agencies responsible for technology diffusion. This latter facilitates the search of strategies to increase the probability that technological innovations are adopted as quickly as possible, but also to know their impact on a particular production system.

In this regard, FAO (1988) suggests that the generation and adoption of new technologies should occur in parallel with the producer, taking into account the idiosyncrasies of the producer, their culture, their interests and agro ecological and economic conditions in which develops. These aspects represent, in most cases, a serious condition that limits the adoption of a technology. The survey analysis by multivariate methods provides a clear and logical explanation of how much a particular factor contributes to the adoption decision (CIMMYT, 1993). The objective of this proposal is to present the use of statistical techniques to analyze and evaluate the process of adoption

de métodos multivariados proporciona una explicación clara y lógica de cuánto contribuye un determinado factor a la decisión sobre adopción (CIMMYT, 1993). El objetivo de esta propuesta es presentar el uso de técnicas estadísticas para analizar y evaluar el proceso de adopción de innovaciones tecnológicas. Esta metodología incluye escalas cualitativas para analizarse a través de técnicas multivariadas como análisis multivariadamente normal, aglomerados, componentes principales y correlación canónica. Estas técnicas, se ejemplifican con encuestas a productores de cebada como un caso de estudio.

### **Caso de estudio**

Con el fin de ilustrar la propuesta metodológica se usó un trabajo en donde se estudió la adopción del sistema de siembra en surcos doble hilera con productores de cebada de temporal en el estado de Zacatecas. El objetivo de esa investigación fue evaluar el grado de adopción de la tecnología siembra en surcos a doble hilera y pileteo, caracterizar a los usuarios de la tecnología e identificar los factores asociados a la adopción de esta tecnología en el estado de Zacatecas. La encuesta se aplicó a una muestra de 135 productores a finales de 2009 y principios de 2010 a productores de cebada en los municipios de Sombrerete, Miguel Auza, Sain Alto, Morelos, Calera, Ojocaliente, Pánuco, Pinos y Fresnillo en el estado de Zacatecas. El método de muestreo fue aleatorio simple.

### **Diseño de encuestas**

Existen varios tipos de estudios que pueden realizarse para evaluar la adopción de tecnologías, sin embargo, en esta investigación se describe el diseño y análisis de una encuesta formal. Las encuestas formales generan información cualitativa discreta útil para quienes toman las decisiones, y con ellas se pueden explotar mejor aspectos complejos que son necesarios para comprender la variabilidad de la adopción entre los productores. Sin embargo, aún cuando los errores del muestreo aleatorio son minimizados a través de encuestas formales, los errores no muestrales son frecuentes. Éstos son debidos al empleo de términos inadecuados al hacer las preguntas, falta de secuencia lógica en la presentación de éstas, preguntas fuera de contexto o innecesarias, ineficacia del encuestador al hacer la pregunta y la selección de un momento inoportuno para la aplicación de la encuesta (INEC, 2001). Por lo tanto, la aplicación de un encuesta piloto, minimiza esta fuente de errores (Malhotra, 2008).

of technological innovations. This methodology includes qualitative scales to be analyzed through multivariate techniques as regular multivariate analysis, clusters, principal components and canonical correlation. These techniques are exemplified by barley producer's surveys as a case study.

### **Case study**

In order to illustrate the proposed methodology was used a work where we studied the adoption of twin row planting system, with barley producers under rainfed conditions in the state of Zacatecas. The objective of this research was to evaluate the degree of adoption twin row and row diking, characterize technology users and identify factors associated with the adoption of this technology in the state of Zacatecas. The survey was applied to a sample of 135 producers in late 2009 and early 2010 to barley producers in the municipalities of Sombrerete, Miguel Auza, Sain Alto, Morelos, Calera, Ojocaliente, Panuco, Pinos and Fresnillo in the state of Zacatecas. The sampling method was simple random.

### **Survey design**

There are several types of studies that can be performed to assess the adoption of technology; however, this research describes the design and analysis of a formal survey. Formal surveys generate discrete qualitative information, useful for decision makers and with them can best exploit complex aspects that are necessary to understand the variability of adoption among farmers. However, even when random sampling errors are minimized through formal surveys, error samplings are frequent. These are due to the employment of inappropriate terms when asking the questions, lack of logical sequence in the presentation of these, questions out of context or unnecessary, inefficiency of pollster when asking the questions and the selection of an inopportune time for the implementation of the survey (INEC, 2001). Therefore, the application of a pilot survey minimizes this source of errors (Malhotra, 2008).

### **Qualitative indicators of the survey**

On the cover of the survey was collected information referring to the folio, producer's name, municipality and community, date, geographical coordinates or other variables that allows identification of the producer. Later were included five indicators that gave response to the

### Indicadores cualitativos de la encuesta

En la carátula de la encuesta se recolectó información referente al folio, nombre del productor, municipio y comunidad, fecha, coordenadas geográficas u otra variable que permita identificar al productor informante. Posteriormente se incluyeron cinco indicadores que dieron respuesta a los objetivos e hipótesis enfocados a la adopción de tecnologías. Los indicadores fueron: a) el proceso de adopción de la innovación; b) factores que afectan el proceso de adopción; c) impactos (económicos, sociales o ambientales) percibidos por el productor con la adopción; d) otros factores en la adopción de la innovación; y e) restricciones para la adopción de la innovación. La división de una encuesta en varias partes es una práctica conveniente.

### Escala y medición

Se usó una escala discontinua ordinal finita (1-5) como se indica: 1= total desacuerdo, nula, muy bajo, muy poca; 2= desacuerdo, bajo, poco; 3= indistinto, regular, medio; 4= acuerdo, bueno, alto, mucho, y 5= total acuerdo, muy bueno, muy alto. También, el uso de escalas diferentes, en la encuesta, no limita el proceso de análisis, debido a que éstas se pueden estandarizar antes de analizarse (Fernández, 2004). Sin embargo, los enunciados de las diferentes preguntas se plantean como aseveraciones, en vez de cuestionamientos (Sánchez *et al.*, 2012). En todos los casos, las respuestas se formulan de forma tal que, los informantes respondan ágilmente para clasificar la situación en relación al grado de adopción de una práctica o insumo de una manera suficientemente sencilla. Esto reduce el margen de error en las respuestas.

### Análisis estadísticos

Las investigaciones sociales proporcionan cuantiosa información por la diversidad de temas enfocados en ellas así como por el gran número de observaciones que integran las muestras (Poza, 2008). La técnica matemática que permite el análisis simultáneo de dos o más variables, la reducción de los datos, la descomposición en factores del fenómeno social, la clasificación y el ordenamiento de las unidades investigadas, es el análisis multivariado (INEI, 2002). Cada técnica multivariada tiene distintos fines y su aplicación depende de los objetivos e hipótesis a probar.

objectives and hypotheses focused on technology adoption. The indicators were: a) process of innovation adoption; b) factors affecting the adoption process; c) impacts (economic, social or environmental) raised by the producer with the adoption; d) other factors in the adoption of innovation; and e) restrictions for the adoption of innovation. The division of a survey in various parts is a convenient practice.

### Scale and measurement

It was used a discontinuous finite ordinal scale (1-5) as follows: 1= complete disagreement, null, very low, 2= disagree, low, little, 3= indistinct, regular, average, 4= agree, well, high, a lot, and 5 = totally agree, very good, very high. Also, the use of different scales, in the survey does not limit the analysis process, since these can be standardized before analysis (Fernández, 2004). However, the statements of the different questions are posed as statements rather than questions (Sánchez *et al.*, 2012). In all cases, the answers are formulated such way that the informants respond nimbly to classify the situation in relation to the degree of adoption of a practice or input in a simple enough manner. This reduces the margin of error in the answers.

### Statistical analyzes

Social research provided substantial information for the diversity of topics focused on them as well as the large number of observations that integrate the samples (Poza, 2008). The mathematical technique that allows the simultaneous analysis of two or more variables, data reduction, factorization of social phenomena, classification and the ordering of investigated units, is the multivariate analysis (INEI, 2002). Each multivariate technique has different purposes and its application depends on the objectives and hypotheses to be tested.

### Analysis of multivariate normality

In the statistical analysis, one of the fundamental assumptions is that a random variable is normally distributed (Park, 2008). When violates the assumption of normality, the interpretation and inference on this variable may not be valid or reliable (Shapiro and Wilk, 1965). However, to test the normality of a variable (univariate analysis) is relatively easy and there are several ways to correct the

### Análisis de normalidad multivariada

En el análisis de datos estadísticos, una de las asunciones fundamentales es que una variable aleatoria se distribuye normalmente (Park, 2008). Cuando se viola el supuesto de normalidad, la interpretación e inferencia sobre esta variable puede no ser válida ni fiable (Shapiro y Wilk, 1965). Sin embargo, probar la normalidad de una variable (análisis univariado) es relativamente fácil y existen varias formas de corregir la falta de normalidad a través de transformaciones apropiadas a una variable en particular (Steel *et al.*, 1997). En contraste, la normalidad multivariada implica en el sentido univariado, que un grupo de variables individuales son normales, si la combinación de éstas es también normal. Es decir, si un grupo de variables sigue una distribución multivariadamente normal, entonces cada una de las variables tiene una distribución univariadamente normal (Hair *et al.*, 1998). Por lo tanto, probar una normalidad multivariada resulta más difícil de ejecutar. Para ejemplificar esta primera parte, se utilizó una encuesta a productores de cebada de donde, al azar, se seleccionaron las siguientes variables: tiempo de conocer el sistema siembra en surcos doble hilera (V1), conocimiento de la sembradora para el sistema en surcos doble hilera (V2), tiempo de conocer la sembradora (V3), conocimiento de la cantidad de semilla recomendada para el sistema (V4), tiempo de conocer la cantidad de semilla recomendada (V5) y conocimiento de la pileteadora (V6), respectivamente. Para ese grupo de variables, la hipótesis nula se planteó que las seis variables consideradas multivariadamente no son significativamente diferentes de una distribución normal.

De acuerdo con Park (2008), la prueba de normalidad univariada para cada una de las seis variables incluidas en este ejemplo, fue rechazada según el estadístico de Kolmogorov-Smirnov (D) con  $p = 0.01$ . Este último resultado fue consistente cuando la normalidad de estas variables fue considerada colectivamente. Consecuentemente, los datos de estas variables fueron transformados (normalizados) a la raíz cuadrada del valor original más 0.5 (Steel *et al.*, 1997). Con esta transformación, no se rechazó la hipótesis nula con  $p = 0.089$ . Es decir, que las seis variables ahora transformadas y consideradas colectivamente no son significativamente diferentes a una distribución normal, y por lo tanto, los datos transformados pueden ser analizados colectivamente (Hair *et al.*, 1998).

lack of normality through appropriate transformations to a particular variable (Steel *et al.*, 1997). In contrast, multivariate normality implies univariate in the sense that a group of individual variables are normal, if the combination of these is also normal. That is, if a group of variables follows a normal multivariate distribution, then each of the variables has a regular univariate distribution (Hair *et al.*, 1998). Therefore, to test multivariate normality turns more difficult to execute. To illustrate this first part, was applied a survey to barley producers where, randomly selected the following variables: time of knowing the planting system of twin row (V1), knowledge of the planter for the twin row system (V2), time to know the planter (V3), knowledge of the amount of seed recommended for the system (V4), time of knowing the recommended amount of seed (V5) and knowledge of the dammer diker (V6), respectively. For this group of variables, the null hypothesis was proposed that the six variables considered multivariate are not significantly different from a normal distribution.

According to Park (2008), univariate normality test for each of the six variables included in this example, was rejected according to the statistic Kolmogorov-Smirnov (D) with  $p = 0.01$ . This last result was consistent when the normality of these variables was considered collectively. Consequently, the data for these variables were transformed (normalized) to the square root of the original value plus 0.5 (Steel *et al.*, 1997). With this transformation, was not rejected the null hypothesis with  $p = 0.089$ . That is, the six variables, now transformed and considered collectively are not significantly different to a normal distribution, and therefore, the transformed data can be analyzed collectively (Hair *et al.*, 1998).

### Principal component analysis

Principal component analysis (PCA) is a multivariate technique of synthesis information or reduction of the number of original variables involved in a study, preferably two or three explaining the total variation in the information without concern for a specific cluster of observations (Cruz *et al.*, 1994). PCA maximizes the variation between the original variables, identifies trends in a set of data and eliminates redundancy in univariate analysis when multicollinearity is involved in the information (Iezzoni and Pritts, 1991). The new variables generated, uncorrelated with each other are called principal components. These latter will explain most of the total variance in the data (Broschat, 1979).

## Análisis por componentes principales

El análisis de componentes principales (ACP) es una técnica multivariada de síntesis de información o reducción del número de variables originales involucradas en un estudio, preferiblemente dos o tres que expliquen el total de la variación en la información sin preocuparse por un agrupamiento específico de las observaciones (Cruz *et al.*, 1994). El ACP maximiza la variación entre las variables originales, identifica tendencias en un grupo de datos y elimina redundancia en un análisis univariado cuando se involucra multicolinealidad en la información (Iezzoni y Pritts, 1991). Las nuevas variables generadas, no correlacionadas entre sí se denominan componentes principales. Estos últimos explicarán la mayor cantidad de la de la varianza total en los datos (Broschat, 1979).

## Aplicación e interpretación

**a) Matriz de correlación.** El primer paso del ACP es ejecutar un análisis de correlación lineal simple entre todos los pares de variables de respuesta originales estandarizadas. Se seleccionaron las variables de la encuesta a productores de cebada que nos interesaban para el análisis, las cuales incluyeron el conocer, probar, adaptar, recomendar y adoptar. Para ejemplificarlo, se presenta la salida de las mismas seis variables cuya normalidad multivariada se probó en la sección anterior. En este caso se encontró una correlación de regular a buena entre las variables V1 y V3, V2 y V4, y V4 y V6. Las tres asociaciones indican una tendencia positiva (Cuadro 1). Por tanto, el análisis individual de estas variables se considera redundante, ya que las conclusiones serían similares para las variables. Mientras que las variables 5 y 6 presentaron una asociación regular pero negativa, es decir que mientras una de la variable incrementa su demisión, la otra la disminuye. Al efectuar el ACP, se espera que las variables con alta correlación sean incluidas en el mismo componente principal (CP).

**b) Definición de componentes principales relevantes.** Con base en la matriz de correlación, las variables originales son transformadas a factores o componentes principales (CPs) las cuales son combinaciones lineales de las variables estandarizadas no correlacionadas (ortogonales). Preferiblemente en los dos o tres CPs se busca que todas las variables estandarizadas expliquen la mayor cantidad de la varianza del grupo de datos. Generalmente se incluyen aquellas CPs que tienen un valor raíz igual o mayor a 1. En el ejemplo, se indica que el primer CP aporta 44% de la

## Application and interpretation

**a) Correlation Matrix.** The first step of PCA is to execute a simple linear correlation analysis between all pairs of standardized original response variables. Variables were selected from the survey to barley producers that interested us for analysis, which included knowledge, testing, adaptation, recommend and adopt. To illustrate, is presented the output of the same six variables whose multivariate normality was tested in the previous section. In this case was found a good correlation of regular to good between V1 and V3, V2 and V4, V4 and V6 variables. The three associations indicate a positive trend (Table 1). Therefore, the individual analysis of these variables is considered redundant, since the findings would be similar for the variables. While variables 5 and 6 showed regular association but negative, i.e. while one variable increases resignation the other decreases. By applying the PCA is expected that variables highly correlated are included in the same principal component (PC).

### Cuadro 1. Matriz de correlación entre variables de respuesta para medir la adopción de una innovación tecnológica.

**Table 1. Correlation matrix between response variables to measure the adoption of a technological innovation.**

	Variables de respuesta					
	V1	V2	V3	V4	V5	V6
V1	1.00	-0.10	0.65	-0.09	0.14	-0.12
V2		1.00	-0.22	0.87	-0.32	0.51
V3			1.00	-0.18	-0.03	0.22
V4				1.00	-0.29	0.60
V5					1.00	-0.57
V6						1.00

Tiempo de conocer el sistema siembra en surcos doble hilera= (V1); conocimiento de la sembradora para el sistema en surcos doble hilera= (V2); tiempo de conocer la sembradora= (V3); conocimiento de la cantidad de semilla recomendada para el sistema= (V4); tiempo de conocer la cantidad de semilla recomendada= (V5); y conocimiento de la pileteadora= (V6).

**b) Definition of relevant principal components.** Based on the correlation matrix, the original variables are to factors or to principal components (PCs), which are linear combinations of uncorrelated standardized variables (orthogonal). Preferably the two or three CPs is intended that all standardized variables explain the largest amount of variance of the data set. CPs generally includes those that have a root value equal or greater than 1. In the example, indicates that the first CP contributes 44% of the variance (proportion or percentage of the variance) and includes almost three CPs with a total variance of 87% (Table 2).

varianza (proporción o porcentaje de la varianza) e incluye casi tres CPs con una varianza total acumulada de 87% (Cuadro 2).

**c) Definición de las variables relevantes dentro de cada CP.** Una vez que se ha definido que los tres primeros CPs son los más importantes, es necesario ubicar cada una de las variables originales estandarizadas dentro del CP. Para esto se analiza los vectores raíz dentro de cada una de las combinaciones lineales de variables estandarizadas y determinar aquellas de mayor peso (Cuadro 2).

Así, el componente principal 1 (CP 1) se puede definir con la siguiente ecuación:

$$CP1 = -0.18(V1) + 0.53(V2) - 0.13(V3) + 0.54(V4) - 0.38(V5) + 0.48(V6)$$

En el primer CP, los vectores raíz con mayor peso y positivos estuvieron asociados con las variables 2, 4 y 6; por lo tanto, esta multivariable fue designada como "conocimiento y evaluación de la tecnología de siembra en surcos doble hilera", para abreviarla conocimiento y evaluación de la tecnología, e incluyó las variables que tuvieron que ver con el conocimiento y evaluación del sistema de siembra y los componentes tecnológicos individuales, excepto el pileteo. Este primer CP explicó 44% de la variación total.

En el segundo CP, los mayores pesos positivos fueron obtenidos por las variables 1 y 3, se le asignó el nombre de "conocimiento, experimentación, adopción y resultado económico de la pileteadora", abreviada como adopción y uso del pileteo, e incluyó las variables que implicaron tener información sobre este componente tecnológico y el proceso de evaluación por el productor. Este segundo CP explicó 27% de la variación total. Los dos CPs explicaron 71% variación total; mientras que el CP-3 sólo contribuyó con 16% de la variación total y su valor raíz fue menor que 1. Por lo tanto, se usaron los dos primeros CPs para la interpretación gráfica con la cual se pudo diferenciar cuatro grupos de productores (Figura 1).

### Análisis por aglomerados

El análisis por aglomerados (AA) es una técnica de análisis exploratorio de datos para resolver problemas de clasificación. El objetivo de esta herramienta consiste en ordenar objetos (personas, cosas, animales, plantas,

**c) Definition of relevant variables within each CP.** Once defined that the first three CPs are the most important, it is necessary to locate each one of the original standardized variables within the CP. For this, root vectors are analyzed within each of the linear combinations of standardized variables and determining those of greater weight (Table 2).

### Cuadro 2. Vectores raíz o coeficientes para la combinación de variables que integran los componentes principales del 1 al 3.

**Table 2. Root vectors or coefficients for the combination of variables that integrate the principal components 1 to 3.**

Variables	Vectores raíz		
	CP 1	CP 2	CP 3
V1*	-0.18	0.62	0.39
V2	0.53	-0.006	0.41
V3	-0.13	0.71	-0.09
V4	0.54	0.02	0.41
V5	-0.38	-0.15	0.65
V6	0.48	0.28	-0.28
Valores raíz	2.66	1.66	0.96
Varianza explicada (%)	44	27	16

\*V1= tiempo de conocer el sistema siembra en surcos doble hilera; V2= conocimiento de la sembradora para el sistema en surcos doble hilera; V3= tiempo de conocer la sembradora; V4= conocimiento de la cantidad de semilla recomendada para el sistema; V5= tiempo de conocer la cantidad de semilla recomendada; V6= conocimiento de la pileteadora.

Thus, the principal component 1 (PC 1) can be defined by the following equation:

$$CP1 = -0.18(V1) + 0.53(V2) - 0.13(V3) + 0.54(V4) - 0.38(V5) + 0.48(V6)$$

In the first PC, the root vectors with higher value and positive, were associated with variables 2, 4 and 6; therefore, this multivariate was designated as "knowledge and technology assessment of twin row system", to abbreviate it as knowledge and technology assessment, and included the variables that had to do with knowledge and planting system assessment and individual technology components except row diking. The first PC explained 44% of the total variation.

In the second PC, the higher positive values were obtained by variables 1 and 3, were given the name of "knowledge, experimentation, adoption and economic result of the dammer diker", abbreviated as adoption and use of row diking, and included the variables that involved having information about this technology component and the

variables, etc.) en grupos (aglomerados) de forma que el grado de asociación (similitud) entre miembros del mismo aglomerado sea mayor que el grado de similitud entre miembros de otros aglomerados (Hair *et al.*, 1998; Rencher, 2002). Es decir, cada aglomerado describe un conjunto de miembros con características similares, sin indicar en qué característica (s) son similares.

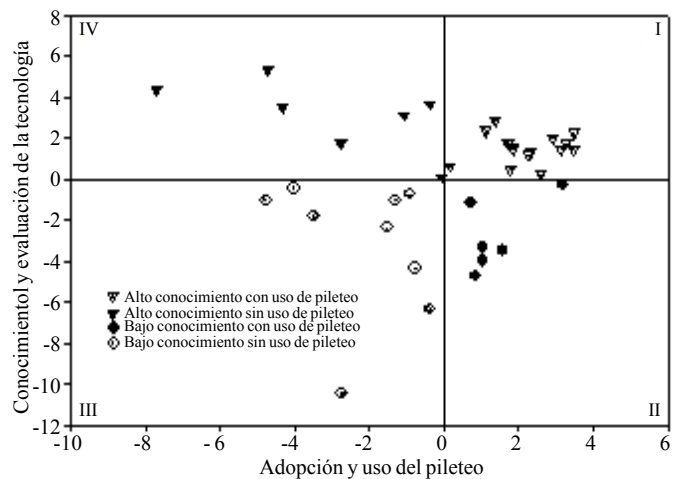
El AA es un método que permite descubrir asociaciones y estructuras en los datos que no son evidentes *a priori* pero que pueden ser útiles una vez que se han encontrado. Los resultados de un AA pueden contribuir a la definición formal de un esquema de clasificación tal como una taxonomía para un conjunto de objetos, sugerir modelos estadísticos para describir poblaciones, asignar nuevos individuos a las clases para diagnóstico e identificación, entre otros.

Se pueden encontrar dos tipos fundamentales de métodos de clasificación: jerárquicos y no jerárquicos (Rencher, 2002). En los primeros, la clasificación resultante tiene un número creciente de clases anidadas; mientras que en el segundo las clases no son anidadas. Los métodos pueden dividirse en aglomerativos y divisivos. En los primeros se parte de tantas clases como objetos tengan que clasificarse y en pasos sucesivos se obtiene clases de objetos similares. En el segundo se parte de una clase única formada por todos los objetos que se van dividiendo en clases sucesivamente. En este documento se describe fundamentalmente el método jerárquico aglomerativo.

En forma general, se anotan los pasos requeridos para una clasificación jerárquica de la siguiente manera:

a) decidir qué datos se colectarán en el aa para cada uno de los casos. Generalmente, se toma un número grande de variables todas del mismo tipo (continuas, categóricas, etc.) o dos variables como mínimo, ya que es difícil procesar distintas escalas; b) elegir una medida de la distancia entre los objetos a clasificar, que serán los aglomerados o clases iniciales. Para ello, existen multitud de métodos como son: ward, average, vecino más próximo, vinculación inter-grupos, entre otros; c) identificar qué aglomerados u objetos son más similares; d) unir aglomerados en un nuevo aglomerado que tenga al menos 2 objetos, de forma que el número de aglomerados irá decreciendo; e) calcular la distancia entre un nuevo aglomerado y el resto; y f) repetir desde el paso tres hasta que todos los objetos formen un solo aglomerado.

evaluation process for the producer. This second PC explained 27% of the total variation. The two CPs I explained 71% of total variation; while the PC-3 only contributed 16% of the total variation and its root value was less than 1. Therefore, the first two CPs were used for graph interpretation, from which were able to differentiate four groups of producers (Figure 1).



**Figura 1. Agrupación de productores adoptantes de la tecnología por las variables conocimiento y evaluación de la tecnología, y adopción y uso del pileteo.**

**Figure 1. Producers group adopters of technology by variables of knowledge and technology, assessment and adoption and use of row diking.**

### Cluster analysis

Cluster analysis (CA) is a technique for exploratory data analysis to solve classification problems. The purpose of this tool consists to arrange objects (persons, objects, animals, plants, variables, etc.) into groups (clusters), so that the association degree (similarity) between members of the same cluster is higher than the degree of similarity between members of other clusters (Hair *et al.*, 1998; Rencher, 2002). That is, each cluster describes a set of members with similar characteristics, without indicating what characteristics are similar.

CA is a method that allows discovering associations and structures in data that are not obvious *a priori*, but may be useful once found. The results of CA can contribute to the formal definition of a classification scheme such as taxonomy for a set of objects, suggesting statistical models to describe populations, assign new individuals to diagnostic classes and identification, among others.



Para medir la similitud entre dos objetos, se usan diferentes métodos, uno de los más usados es la distancia euclidiana. La distancia euclidiana es el intervalo entre dos puntos en un espacio euclidiano, la cual se deduce a partir del teorema de Pitágoras.

**Aplicación e interpretación**

El resultado del Sistema de Análisis Estadístico (SAS) incluye dos aspectos: la historia del aglomerado y el número de aglomerados que se formaron con una base de datos. Estos dos aspectos se ejemplifican con el estudio de adopción para el sistema de siembra en surcos doble hilera con productores de cebada de temporal en el estado de Zacatecas.

**a) Historia del conglomerado.** En el programa SAS usando la declaración PROC CLUSTER, al inicio de la salida del resultado se obtiene el historial del aglomerado. A través de éste se puede llevar a cabo el seguimiento en la formación de aglomerados de productores de cebada encuestados en el proceso de adopción de la tecnología del sistema de siembra en surcos doble hilera.

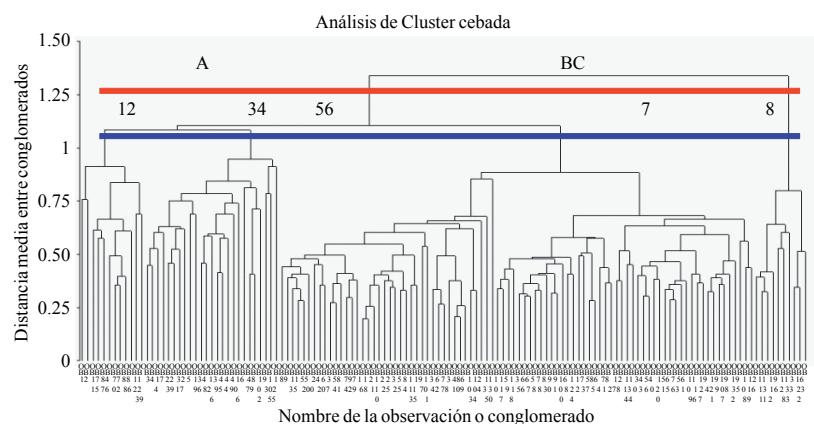
**b) Definición de aglomerados relevantes.** La representación gráfica del agrupamiento se hace a través de un dendrograma (Figura 2). En el eje de la X se anotan todos los sujetos (productores); mientras que en el eje de Y, se anota la distancia promedio entre productores. Si en el dendrograma se coloca una línea horizontal paralelo al eje de X, será posible definir el número de aglomerados, los productores incluidos en cada aglomerado e identificar las variables que hicieron diferentes a los grupos. El primer criterio integra tres aglomerados y el segundo integra ocho grupos de productores, de acuerdo con la altura a la que se dibuja la línea horizontal (Figura 2).

There can be found two fundamental methods of classification: hierarchical and non-hierarchical (Rencher, 2002). In the hierarchical, the resulting classification has an increasing number of nested classes; whereas in the non-hierarchical class are not nested. The methods can be divided into cluster and divisive. In the cluster, it starts from many classes as objects have to be classified and in successive steps are obtained similar object classes. The divisive starts from a unique class formed of all objects to be divided into classes on. This paper mainly describes the hierarchical clustering method.

In general, the steps required are entered for a hierarchical classification as follows:

- a) Decide what data will be collected in the CA for each of the cases. Generally, a large number of variables, all of the same type are taken (continuous, categorical, etc.) or two variables at a minimum, as it is difficult to process different scales;
- b) choose a measure of the distance between the objects to be classified, that will be the clusters or initial class. For this, there are many methods, such as: ward, average, nearest neighbor, link between-groups, among others;
- c) identifying which clusters or objects are alike;
- d) cluster attach into a new cluster having at least 2 items so that the number of clusters will decrease;
- e) to calculate the distance between a new cluster and the rest; and
- f) repeating from step three until all objects form a single cluster.

To measure similarity between two objects, different methods are used; one of the most used is the Euclidean distance. The Euclidean distance is the interval between two points in a Euclidean space, which is deduced from the Pythagorean Theorem.



**Figura 2. Dendrograma para productores usuarios y no usuarios de la tecnología de siembra en surcos doble hilera con piletteo para cebada maltera en el estado de Zacatecas, México.**

**Figure 2. Dendrogram for users and non-users producers of the twin bed technology with row diking for malting barley in the state of Zacatecas, Mexico.**

Con este análisis se pueden definir dos grandes grupos de productores, los adoptantes (aglomerado A) y los no adoptantes (aglomerado B) de la innovación tecnológica, además de un grupo atípico (aglomerado C) (Figura 2). Otra alternativa podría ser declarando ocho grupos de productores, cinco dentro de los adoptantes y otros tres grupos en los no adoptantes. La decisión sobre el número óptimo de aglomerados es subjetiva, especialmente cuando se incrementa el número de objetos. En este sentido, la experiencia del investigador es clave para definir los aglomerados o grupos de objetos relevantes.

### Correlación canónica

El análisis de correlación canónica (ACC) es un método de análisis multivariante desarrollado por Hotelling (1936). El objetivo de ACC es buscar las asociaciones y relaciones que puede haber entre dos grupos de variables y la validez de las mismas (Rencher, 1992). A diferencia de la correlación múltiple, el ACC estandariza las variables originales y luego correlaciona variables de dos grupos de variables (Rencher, 1988; Manly, 1986). Es decir, el ACC examina la relación lineal entre un grupo de variables, X, y un grupo de variables Y. La técnica consiste en encontrar una combinación lineal de las variables X, conocida ahora como variable V ( $V_1 = b_1X_1 + b_2X_2 + \dots + b_pX_p$ ) y otra combinación lineal de las variables Y, conocida como variable U ( $U_1 = a_1Y_1 + a_2Y_2 + \dots + a_qY_q$ ), de tal manera que la correlación entre U y V sea máxima (Hardoon *et al.*, 2003).

Después, podría ser necesario encontrar otras dos combinaciones lineales para cada grupo de variable  $V_2$  y  $U_2$ , que tenga correlación máxima, pero menor a la primera, y así sucesivamente se encuentran un conjunto de combinaciones lineales para cada grupo de variables que presenten correlación entre las diferentes Us y Vs (Vicario *et al.*, 1989). Estas combinaciones lineales se denominan variables canónicas, y las correlaciones entre los correspondientes pares de variables canónicas se denominan correlaciones canónicas (Hair *et al.*, 1998).

Después se interpretan las cargas canónicas para determinar la importancia de cada variable en cada función canónica. Las cargas canónicas reflejan la varianza que la variable observada comparte con el valor teórico canónico. Este análisis resulta similar al de componentes principales, en donde de acuerdo con el autovalor (valor raíz), se seleccionan aquellas variables con mayor peso dentro de cada variable canónica y se le asigna un nombre genérico.

### Application and interpretation

The result of the Statistical Analysis System (SAS) includes two aspects: the history of the cluster and the number of clusters that were formed with a database. These two aspects are exemplified by the study of adoption for twin bed planting system with barley producers under rainfed conditions in the state of Zacatecas.

**a) Cluster history.** In SAS using the PROC CLUSTER statement at the beginning of the output result is obtained the history of the cluster. Through this, it can be carry out the monitoring in the formation of clusters of surveyed barley producers in the process of technology adoption for the twin bed planting system.

**b) Defining relevant clusters.** The graphical representation of the cluster is done through a dendrogram (Figure 2). On the X axis are noted all subjects (producers), while in the Y axis, note the average distance between producers. If in the dendrogram is placed a horizontal line parallel to the X axis, will be possible to define the number of clusters, the producers included in each cluster and identify the variables that made different the groups. The first criterion integrates three clusters and the second integrates eight groups of producers, according to the height at which the horizontal line is drawn (Figure 2).

With this analysis can be defined two groups of producers, adopters (Cluster A) and non-adopters (cluster B) of technological innovation, plus an atypical group (cluster C) (Figure 2). Another alternative could be declaring eight groups of producers, five in the adopters and other three in the non-adopters groups. The decision on the optimal number of clusters is subjective, especially when the number of objects increases. In this sense, the experience of the researcher is key defining the clusters or groups of relevant objects.

### Canonical correlation

The canonical correlation analysis (CCA) is a multivariate analysis method developed by Hotelling (1936). The objective of a CCA is to find the associations and relations that can exist between two groups of variables and their validity (Rencher, 1992). Unlike the multiple correlations, CCA standardizes the original variables and then correlates variables of two group variables (Rencher, 1988; Manly, 1986). That is, CCA examines the linear relationship between a set of X variables, and a group of Y variables. The

## Aplicación e interpretación

El procedimiento e interpretación del ACC se ejemplifica con datos recabados en un estudio de adopción de innovaciones tecnológicas en cebada con productores de Zacatecas (Sánchez y Rumayor, 2011). El primer grupo incluyó diez variables (V13, V14, ..., V22) que describieron los componentes de adopción tecnológica. Éstas fueron designadas como variables "X". En el segundo grupo de datos se consideraron seis variables que definieron los beneficios económicos percibidos por los productores al adoptar la innovación (V7, ..., V12). Éstas se designaron como variables "Y". A través del ACC, las variables "X" y "Y" se convertirán en variables "V" y "U", respectivamente (Hardoon *et al.*, 2003).

**a) Correlación dentro de variables X y Y.** Las variables originales (no estandarizadas) de adopción tecnológica (X) mostraron una asociación positiva de regular a buena ( $r=0.51 - r=0.87$ ) entre ellas, excepto por el bajo coeficiente de correlación entre las variables V7 con V14 ( $r=0.21$ ), V9 con V14 ( $r=0.35$ ), V10 con V14 ( $r=0.27$ ), V11 con V12 ( $r=0.37$ ), V11 con V14 ( $r=0.15$ ), V14 con V15 ( $r=0.32$ ) y V14 con V16 ( $r=0.42$ ). En contraste, la asociación entre variables originales de beneficios económicos (Y), en términos generales, fue positiva y buena ( $r=0.63 - r=0.91$ ).

**b) Correlación entre variables X y Y.** La asociación entre variables de adopción tecnológica y beneficios económicos mostraron una correlación de regular a moderadamente alta y positiva, excepto por la correlación entre las variables V14 y V18, V14 con V20 y V22, y V12 y V21, cuyos coeficientes fueron moderadamente bajos y positivos (Cuadro 3).

technique consists in finding a linear combination of the X variables, now known as the variable V ( $V_1 = b_1X_1 + b_2X_2 + \dots + b_pX_p$ ) and other linear combination of the Y variables, known as variable U ( $U_1 = a_1Y_1 + a_2Y_2 + \dots + a_qY_q$ ), so that the correlation between U and V is a maximum (Hardoon *et al.*, 2003).

Then, it may be necessary to find two other linear combinations for each group of  $V_2$  and  $U_2$ , having maximum correlation, but less than the first, and so on are found a set of linear combinations for each set of variables presenting correlation between the different Us and Vs (Vicario *et al.*, 1989). These linear combinations are called canonical variables, and correlations between corresponding pairs of canonical variables are called canonical correlations (Hair *et al.*, 1998).

Afterwards the canonical loadings are interpreted to determine the importance of each variable in each canonical function. Canonical loading reflect the variance that the observed variable shares with the canonical theoretical value. This analysis is similar to principal components, where according to the eigenvalue (root value), are selected those variables with higher weight within each canonical variable and assigned a generic name.

## Application and interpretation

The procedure and interpretation of CCA is exemplified with data collected in a study of adoption of technological innovations in barley producers from Zacatecas (Sanchez and Rumayor, 2011). The first group included ten variables (V13, V14, ..., V22) which described the components of technological adoption. These variables were designated as

**Cuadro 3. Coeficientes de correlación entre variables originales de adopción tecnológica y beneficios económicos.**  
**Table 3. Correlation coefficients between original variables of technological adoption and economic benefits.**

Adopción Tecnológica (X)	Beneficios económicos (Y)					
	V17	V18	V19	V20	V21	V22
V7	0.67	0.64	0.66	0.62	0.53	0.59
V8	0.51	0.64	0.71	0.68	0.57	0.74
V9	0.78	0.85	0.87	0.87	0.72	0.90
V10	0.60	0.74	0.67	0.78	0.71	0.75
V11	0.74	0.66	0.71	0.61	0.56	0.73
V12	0.48	0.60	0.66	0.60	0.43	0.63
V13	0.78	0.76	0.77	0.73	0.50	0.78
V14	0.16	0.34	0.47	0.38	0.40	0.43
V15	0.84	0.70	0.82	0.71	0.57	0.69
V16	0.63	0.68	0.74	0.68	0.48	0.75

**c) Correlación canónica y prueba de hipótesis.** Los dos primeros pares de correlaciones canónicas generaron coeficientes de correlación altos y positivos. Aún cuando ambos fueron significativos ( $p=0.0001$ ), el primero explicó, en términos de regresión, 94% de la variación de las variables de beneficios económicos con respecto a las variables de adopción tecnológica (Cuadro 4). Esto último fue apoyado por el valor propio que para ambos pares de correlaciones canónicas fueron 79% y 13%, respectivamente.

**Cuadro 4. Correlación canónica (r), regresión canónica ( $r^2$ ), grados de libertad (gl) y F calculada (Fc) y significación asociada a los dos primeros pares de variables canónicas.**

**Table 4. Canonical correlation (r), canonical regression ( $r^2$ ), degrees of freedom (df) and F calculated (Fc) and significance associated with the first two pairs of canonical variables.**

Variabes canónica	r	$r^2$	gl	Fc	Significación
1	0.97	0.94	60	19.41	0.001
2	0.86	0.74	45	9.47	0001

**Correlación canónica.** Con base en los coeficientes de correlación canónica estandarizados (CCCE), la primera variable canónica de adopción tecnológica (adopción 1) estuvo positivamente dominada por las variables V7 (CCCE=0.50) y V9 (CCCE=0.82), lo cual se confirmó con sus correspondientes coeficientes de correlación (r) entre las variables originales y su variable canónica estandarizada ( $r=0.68$  y  $r=0.96$  para V7 y V9, respectivamente). En contraste, la primera variable canónica de beneficios económicos (beneficios económicos 1) fue dominada por la variable V22 (CCCE=0.50), lo cual fue confirmando con un alto y positivo valor de correlación ( $r=0.96$ ). Finalmente, la Figura 3 muestra que la adopción tecnológica influye positivamente en beneficios económicos. Los puntos que se observan en el lado izquierdo de la gráfica son productores con baja adopción de tecnología, y por lo tanto, con bajos beneficios económicos; mientras que lo opuesto se observa en el lado derecho de la Figura 3.

## Conclusiones

El protocolo propuesto puede seguirlo cualquier investigador con interés en llevar a cabo estudios de adopción. La aplicación de este tipo de análisis depende de los objetivos

"X". In the second set of data were considered six variables that defined the economic benefits received by producers to adopt the innovation (V7, ..., V12). These variables were designated as "Y". Through CCA, the "X" and "Y" variables will become "V" and "U" respectively (Hardoon *et al.*, 2003).

**a) Correlation within variables X and Y.** The original variables (not standardized) of technology adoption (X) showed a positive association of fair to good ( $r=0.51$  -  $r=0.87$ ) between them, except for the low coefficient of correlation between the variables V7 with V14 ( $r=0.21$ ), V9 with V14 ( $r=0.35$ ), V10 with V14 ( $r=0.27$ ), V11 with V12 ( $r=0.37$ ), V11 with V14 ( $r=0.15$ ), V14 with V15 ( $r=0.32$ ) and V14 with V16 ( $r=.42$ ). In contrast, the association between the original variables and economic benefits (Y), overall, were positive and good ( $r=0.63$  -  $r=0.91$ ).

**b) Correlation between variables X and Y.** The association between variables of technological adoption and economic benefits showed a correlation of regular to moderately high and positive, except for the correlation between variables V14 and V18, V14 with V20 and V22, and V12 and V21, whose coefficients were moderately low and positive (Table 3).

**c) Canonical correlation and hypothesis testing.** The first two pairs of canonical correlations generated high and positive correlation coefficients. Although both were significant ( $p=0.0001$ ), the first explained, in terms of regression, 94% of the variation in economic benefits variables regarding technology adoption variables (Table 4). The latter was supported by the eigenvalue that for both pairs of canonical correlations were 79% and 13%, respectively.

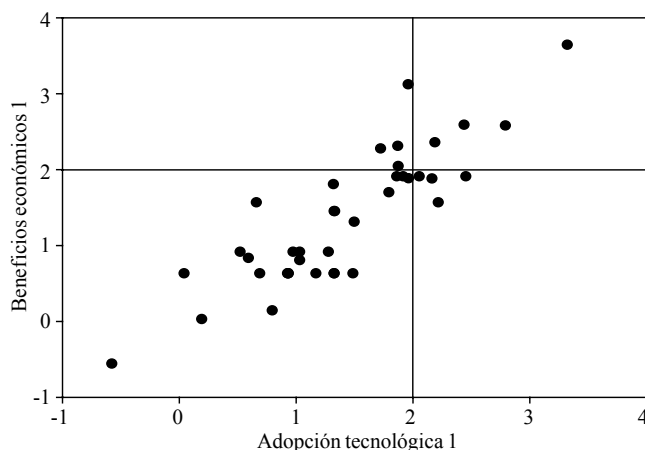
**a) Canonical correlation.** Based on the standardized canonical correlation coefficients (SCCC), the first canonical variable of technological adoption (adoption 1) was positively dominated by variables V7 (SCCC=0.50) and V9 (SCCC=0.82), which was confirmed with its corresponding correlation coefficients (r) between the original variables and standardized canonical variable ( $r=0.68$  and  $r=0.96$  for V7 and V9, respectively). In contrast, the first canonical variable of economic benefits (economic benefits 1) was dominated by the variable V22 (SCCC=0.50), which was confirmed with a high and positive correlation ( $r=.96$ ). Finally, Figure 3 shows that the technology adoption positively influences economic benefits. The dots observed on the left side of the graph are producers with low technology adoption and therefore low economic benefit; whereas the opposite is observed on the right side of Figure 3.

e hipótesis planteados en una investigación en particular. También, el uso de las técnicas multivariadas dependerá de la innovación tecnológica, escala utilizada, de la diversidad de clases entre variables y el juicio del investigador.

Se espera que con la metodología propuesta, los investigadores dedicados a esta área de investigación, dispongan de una estrategia para el diseño de encuestas, definición de una escala y técnicas de análisis estadísticas adicionales a las univariadas.

## Literatura citada

- Broschat, T. K. 1979. Principal component analysis in horticultural research. *HortScience* 14:114-117.
- Centro Internacional de Maíz y Trigo (CIMMYT). 1993. La adopción de tecnologías: guía para el diseño de encuestas. Programa de economía. México, D. F. CIMMYT. 92 p.
- Cruz, J.; Ganeshanandam, S.; Mackay, R.; Lawes, S.; Lawoko, O. and Woolley, J. 1994. Applications of canonical discriminant analysis in horticultural research. *HortScience* 29:1115-1119.
- Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO). 1988. Extensión rural: partiendo de lo posible para llegar a lo deseable. 2ª edición. Oficina Regional de la FAO para América Latina y el Caribe. Serie Desarrollo Rural Núm. 2. 50 p.
- Fernández, A. 2004. Investigación y encuestas de mercado. ESIC editorial. 2ª. Edición. España. 156 p.
- Hair, J.; Anderson, F. Tatham, R. and Black, C. 1998. Multivariate data analysis. Prentice-Hall Inc., New Jersey, USA. 730 p.
- Hardoon, D.; Szedmak, S. and Shawe, J. 2003. Canonical correlation analysis; an overview with application to learning methods. Department of Computer Science Royal Holloway, University of London Technical Report CSD-TR-03-02. 39 p.
- Hernández, C. y Porras, F. 2006. Estudio sobre la adopción de variedades mejoradas de frijol en las principales zonas productoras de frijol de la región Brunca de Costa Rica. *Agronomía Mesoamericana* 17(3):357-367.
- Hotelling, H. 1936. Relations between two sets of variables. *Biometrika* 28(32):1-377.
- Iezzoni, A. and Pritts, M. 1991. Applications of principal component to analysis in horticultural research. *HortScience* 26:334-338.
- Instituto Nacional de Estadísticas y Censos (INEC). 2001. Encuesta Nicaragüense de Demografía y Salud. 431 p.
- Instituto Nacional de Estadística e Informática (INEI). 2002. Guía para la aplicación del análisis multivariado a las encuestas de hogares. Lima, Perú. 59 p.
- Manly, B. 1986. Multivariate statistical, methods: a primer. Chapman and Hall Ltd., London, UK. 159 p.
- Malhotra, N. 2008. Investigación de mercados. 5ª edición. Prentice Hall. México. 811 p.
- Poza, C. 2008. Técnicas estadísticas multivariantes para la generación de variables latentes. *Rev. Escuela de Administración de Negocios*. 64(3):89-99.



**Figura 3. Variables canónicas adopción tecnológica 1 y beneficios económicos 1 generadas de un estudio de adopción de innovaciones tecnológicas para cebada de productores de Zacatecas (Sánchez y Rumayor, 2011)**

**Figure 3. Canonical variables for technological adoption 1 and economic benefits generated from a study of adoption of technological innovations for barley producers from Zacatecas (Sánchez and Rumayor, 2011).**

## Conclusions

The proposed protocol can be followed by any researcher interested in conducting studies of adoption. The application of this type of analysis depends on the objectives and hypotheses on a particular research. Also, the use of multivariate techniques will depend on technological innovation, scale used, class diversity among variables and the judgment of the researcher.

It is expected that the proposed methodology, the researchers dedicated to this area of research, will have a strategy for the survey design, definition of a scale and statistical analysis techniques in addition to the univariate.

*End of the English version*



- Rencher, A. C. 1988. Miscellanea on the use of correlations to interpret canonical functions. *Biometrika* 75:363-365.
- Rencher, A. C. 1992. Interpretation of canonical discriminant functions, canonical variates, and principal components. *Amer. Statist.* 46:217-225.