Anuario Facultad de Ciencias Económicas y Empresariales

Recibido: agosto 2013/Aceptado: sept. 2013

Sistema informático para el pronóstico de indicadores físicos y económicos para la producción y los servicios

Computer System for the Prognosis of Physical and Economic Indicators of Production and Services

Lic. Elio David Zaldívar-Linares; MSc. Alejandro Álvarez-Navarro

elio.zaldivar@eco.uo.edu.cu; alejandroa@eco.uo.edu.cu

Centro de Estudios de Investigaciones Económicas Aplicadas (CEIA); Facultad de Ciencias Económicas y Empresariales, Universidad de Oriente, Cuba

Resumen

El temprano pronóstico de los indicadores en una empresa reviste una gran importancia. Al aplicarse funciones de respuesta, a través del Análisis de Regresión Lineal Múltiple, se contribuye a su mejor obtención y al perfeccionamiento de la planificación, constituyendo esto un considerable beneficio para el país. El presente trabajo consiste en la confección de un sistema informático que supere los inconvenientes que presentan los programas profesionales permitiendo una mayor automatización del cálculo de los estimados en el entorno empresarial. El Sistema Informático para el Cálculo de Estimados de Producción y Servicios (SICEP) permite construir, resolver, analizar y aplicar modelos de regresión y ha sido validado con el sistema profesional Stathistical Package for Social Sciences (SPSS). Ha sido usado en varias entidades productoras seleccionadas de la provincia Santiago de Cuba, a partir de la zafra 2010-2011 y ha contribuido a un perfeccionamiento en la planificación que derivó en una disminución de los costos de producción de un 11 %.

Palabras Clave: regresión, estimación, estadística, modelación matemática, software.

Abstract

The early prognostic indicators in a enterprise is very important. The application of response functions through the Multiple Linear Regression Analysis, contributes to its best determination and the refinement of planning. This constitutes a considerable benefit to the country. This work involves the preparation of a computer system that overcomes the disadvantages of professional programs allowing greater automation of the calculation of the estimates in the business environment. The SICEP[1] lets you build, solve, analyze and apply regression models and has been validated with Stathistical Package for Social Sciences (SPSS) professional system. The system has been used in various production entities selected Santiago de Cuba province, from the 2010-2011 harvest impacting favorably in the precision of planning that determined the reduction of the production costs of value 11 %.

Keywords: regression, estimation, statistical, mathematical modeling, software.

[1] **Computer System for Calculating Production Estimates and Services , initials in Spanish.**

Introducción

En la producción y los servicios existen indicadores físicos y económicos de vital importancia en el desenvolvimiento de dichas actividades. Por tal motivo, en muchas ocasiones, en las entidades donde se llevan a cabo estos procesos es de interés poder pronosticar el comportamiento de estos parámetros en relación con otras variables de las cuales dependen, pues esto redundaría en una adecuada planificación y evaluación de sus metas, potenciando la eficiencia en la gestión económica.

Una de las herramientas estadísticas más efectivas y potentes para dicha estimación es el análisis de regresión, pues su flexibilidad y confiabilidad le permite al experto adaptarlo a las más disímiles problemáticas de la vida real tanto de índole económica como puramente científicas.

Sin embargo, la aplicación práctica del análisis de regresión tropieza con varios obstáculos. A pesar de existir varios paquetes estadísticos profesionales que abordan la regresión, como el Sistema Informático para el Cálculo de Estimados de Producción y Servicios (SPSS) en sus distintas versiones, estos tienen varias limitaciones para el caso específico de interés. Algunas de estas limitaciones son:

1. Rigidez en el conjunto de modelos de regresión soportados, lo que provoca la necesidad de alternar con hojas de cálculo para procesar las observaciones haciendo engorrosa la labor del especialista, a la vez que favorece la ocurrencia del error humano.

2. No automatización de las pruebas de hipótesis complementarias, lo que provoca que sea el especialista el que tenga que preparar los datos y ejecutar las pruebas que desee paso a paso.

Tomando en cuenta todo esto, el Centro de Estudios de Investigaciones Económicas Aplicadas (CEIA) se dio a la tarea de diseñar e implementar el SICEP enfocado, mediante el análisis de regresión, en resolver modelos de regresión lineal múltiple con el objetivo de viabilizar a los expertos la elección de la mejor función de respuesta para cada indicador y la aplicación de la regresión en general.

Fundamentación teórica

Enfoque matricial para el modelo de regresión lineal múltiple

La regresión lineal con *k*+1 variables

Cuando una función de regresión tiene más de una variable independiente se está en presencia de la regresión lineal múltiple. La regresión lineal múltiple, al igual que la simple, se basa en los criterios de varianza mínima, por tanto, todos los supuestos de la regresión simple tienen lugar aquí. La función de regresión múltiple da origen a un sistema de ecuaciones normales que deben ser resueltas utilizando los sistemas de ecuaciones lineales, expresadas a través de matrices; este es el enfoque que será utilizado.

La función de regresión poblacional (FRP) con k+1 variables que tiene la variable dependiente Y y k variables independientes[2] se puede escribir como:

$$\text{FRP: } Y_i = B_0 + B_1 X_{1i} + B_2 X_{2i} + \ldots + B_k X_{ki} + \mathring{a}_i \; i = 1, 2, , N$$

la cual se puede escribir matricialmente en forma desarrollada como:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_N \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & . & . & . & X_{k1} \\ 1 & X_{12} & X_{22} & . & . & . & X_{k2} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 1 & X_{1N} & X_{2N} & . & . & . & X_{kN} \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ . \\ . \\ . \\ B_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_N \end{bmatrix}$$

$$\underset{N \times 1}{\mathbf{y}} = \underset{N \times k+1}{\mathbf{X}} \; \underset{k+1 \times 1}{\mathbf{b}} + \underset{N \times 1}{\boldsymbol{\varepsilon}}$$

donde[3]

y = un vector de columna $N \times 1$ que contiene las observaciones de la variable dependiente Y.

[2] **En general la(s) variable(s) dependiente (independientes) que aparece(n) en la FRP son expresiones que dependen de la(s) variable dependiente (independientes) observada(s) del fenómeno en estudio.**

[3] **Los vectores son representados por letras minúsculas en negrillas y las matrices serán representadas por letras mayúsculas en negrillas.**

X = matriz de $N \times k+1$ que muestra las N observaciones de las k variables explicativas X_1 a X_k, con la primera columna de unos para representar el coeficiente independiente. (Esta matriz también se conoce como *matriz de datos*).

b = vector columna de $k+1 \times 1$ de los parámetros desconocidos de a .

å = vector columna de $N \times 1$ de las perturbaciones.

Estimaciones utilizando el método de los mínimos cuadrados ordinarios (MCO)

En el contexto del análisis de regresión se puede demostrar que los estimadores de MCO son mejores estimadores lineales insesgados (MELI). Según plantea el teorema de Gauss-Markov.[4]

Dicho esto, conociendo la representación matricial de la FRP y contando con un conjunto de muestras representativas podemos obtener el estimador de MCO para **b.** Antes se escribe la función de regresión muestral (FRM) para $k+1$ variables:

$$Y_i = \hat{B}_0 + \hat{B}_1 X_{1i} + \hat{B}_2 X_{2i} + \dots + \hat{B}_k X_{ki} + e,\ i = 1, 2, \dots, n$$

que se puede escribir más concretamente en notación matricial como:

y = **X** + **e**

donde $\hat{\mathbf{b}}$ es un vector columna de $k + 1$ elementos para los estimadores de MCO para los coeficientes de regresión (E($\hat{\mathbf{b}}$) = **b**) y **e** es un vector columna de N elementos para los residuos.

Los estimadores de MCO se obtienen minimizando la suma de los errores al cuadrado representada como $\sum e_i^2$. Cumpliendo este proceso obtenemos:

$$\hat{\mathbf{b}} = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{y}$$

A partir de los coeficientes estimados se realizan las pruebas estadísticas que se expondrán a continuación.

Inferencia estadística para los supuestos del modelo clásico de regresión lineal

Pruebas estadísticas

Para validar los coeficientes de regresión obtenidos se realizan la prueba de significancia global (basada en el estadístico F de Fisher) que determina la validez de la regresión en general y la prueba de coeficientes individuales (basada en el estadístico t de Student) que cuantifica la significación e influencia de cada coeficiente individualmente. Dichas pruebas por si solas no evidencian la necesidad de construir una alternativa a los paquetes estadísticos tradicionales para un entorno específico.

Como es conocido, los supuestos del modelo clásico de regresión lineal son los de aleatoriedad, normalidad y homocedasticidad. Cada uno de estos supuestos puede ser validado por un conjunto de procedimientos estadísticos quedando la selección a juicio del experto quien toma en consideración la naturaleza del proceso estudiado. Después de un profundo análisis basado en la experiencia de los especialistas del CEIA sobre la dinámica de los escenarios donde debía operar el SICEP se diseñó el sistema de pruebas que sería implementado. Dicho sistema se confeccionó a partir de las siguientes decisiones:

- para la prueba de aleatoriedad se debía usar la prueba de rachas [15].
- para la prueba de normalidad se debía usar el test de Kolmogorov-Smirnov [5]
- para la prueba de homocedasticidad se debía usar el test de White. El contraste de White para comprobar si existe homocedasticidad consiste en verificar si existe relación entre las variables independientes y la dependiente en una función de regresión que tiene como variable dependiente los errores cuadráticos en función de las variables exógenas, como también se le conoce a las variables independientes, sus cuadrados y sus productos cruzados. Si existe relación no hay homocedasticidad, en caso contrario, se dice que hay heterocedasticidad (distintas varianzas).

Esta última prueba, relativamente reciente, ha aparecido con poca o nula frecuencia en softwares

[4] El teorema de Gauss-Markov establece que en tre todos los estimadores de coeficientes que son lineales en las Y_i los estimadores de mínimos cuadrados tienen la variancia más pequeña.

profesionales, al punto de que aún no ha sido incorporada por el SPSS en su versión 20.0. El hecho de consumar la automatización del contraste de White revestiría en sí mismo un grado de novedad apreciable.

Proceso de selección de la mejor función respuesta [5]

El proceso de selección considera, primeramente, el coeficiente R^2 para determinar la bondad del ajuste; después, si la prueba F es significativa, queda demostrado que hay relación entre las variables independientes y la dependiente. Después se analizan las pruebas t, para conocer los coeficientes no significativos; luego se realizan las pruebas de normalidad, independencia de los errores y homocedasticidad que si son significativas demuestran que los estimadores son de máxima verosimilitud. Por último, se analiza la multicolinealidad, mediante el coeficiente R^2 y la significación de los coeficientes sabiendo que una alta "bondad de ajuste" y pocos coeficientes significativos, es síntoma de un grado considerable de multicolinealidad. En caso de que más de una curva cumpla con estos requisitos, se tiene en cuenta R^2 para escoger la que tenga un mayor valor para este coeficiente.

Desarrollar el proceso de selección completo para una sola persona sería en extremo engorroso pues los procedimientos estadísticos ofrecidas por los softwares profesionales:

• Solo pueden ser aplicados de manera individual previa preparación de los datos por parte del experto y se limitan a determinar un nivel de confianza (o error) del dictamen que emiten sobre la hipótesis que comprueban sin posibilidad de recibir un margen de error y en consecuencia decidir, quedando bajo la responsabilidad del usuario tomar nota del dato y emitir un criterio, proceso que no por sencillo anula la posibilidad del error humano

• En el mejor de los casos solo proveen un reducido grupo de patrones de regresión siendo la generalidad aquellos que solo proveen el modelo lineal de regresión, quedando para el usuario la misión de transformar los valores de las variables en las observaciones, presumiblemente en una hoja electrónica de cálculo según el modelo que desee comprobar.

En cuanto a este último punto, se puede abundar diciendo que proveer la crucial posibilidad de explotar el procedimiento de selección de la mejor función respuesta necesariamente presupone la existencia de un mecanismo cómodo, ágil y eficaz para gestionar el procesamiento automático de los valores recogidos para la obtención de la matriz de datos y el vector **y**. A esto se une la necesidad de incorporar disímiles modelos de regresión que, aunque algunos son conocidos con antelación, por lo general van surgiendo a medida que se avance en el análisis del proceso específico que se examina.

Una vez expuestos estos argumentos queda demostrada la utilidad del SICEP expuesta en el segmento introductorio.

Métodos utilizados

Características del sistema SICEP

Concepción de los modelos de regresión

Mediante la automatización de los procedimientos expuestos en la sección "Fundamentación teórica", tanto en lo concerniente a la solución como a su validación durante el procesamiento de una matriz de datos para una función de regresión determinada, se solucionaron las limitaciones de los paquetes estadísticos tradicionales referidas en el segundo punto del segmento introductorio. Como resultado de la implementación de la solución que se escogió para superar las limitaciones inherentes al primer punto de dicho segmento, el sistema permite crear modelos de regresión con variables genéricas, que serán asociadas a los factores objetivos que intervengan en el fenómeno objeto de estudio que puede estar ubicado dentro de una variada gama de posibles escenarios.

Habiendo mencionado los modelos de regresión, se puede decir que la búsqueda de la mejor función respuesta entre varias funciones de regresión es el procedimiento mediante el cual un fenómeno, cuyos

[5]Por función respuesta entendemos cualquier función de regresión cuyos coeficientes hayan sido cuantificados para una determinada matriz de datos

posibles factores causales y efecto resultante hayan sido sistemáticamente cuantificados, es probado con varios modelos de regresión. Luego, podemos concluir que un modelo no es más que una expresión que sirve de patrón para poder generar distintas funciones de regresión ante distintas situaciones. En este caso ante distintas cantidades de variables independientes.

Para fundamentar la notación que le permitirá al usuario crear sus propios modelos de regresión se comienza por proponer la forma general de una función de regresión. Dicha forma se obtiene de sintetizar la forma general de las FRP eliminando los convencionalismos inherentes a las perturbaciones y se expone a continuación:

$$Y = B_0 + B_1X_1 + B_2X_2 + \ldots + B_kX_k \quad (1)$$

donde

- B_i para $0 \le i \le k$ son los parámetros o coeficientes de la función.
- Se llama **término** a cada sumando de (1) que aparezca después de B_0 y contenga un parámetro (coeficiente).
- Y representa una función que depende de la variable dependiente observada que se denota como *y* (en lo adelante variable dependiente)
- X_i para $1 \le i \le k$ representan funciones que dependen de las variables independientes observadas $x_1, \ldots, x_{n-1}$ y x_n (en lo adelante variables independientes) donde, a la hora de evaluar, el subíndice de cada variable que aparece en la expresión de la función es la posición del argumento del cual toma su valor.

Sin embargo, en consecuencia con lo expresado al comienzo de esta sección, por lo general, un modelo en lugar de usar, como hemos visto hasta aquí, varios **términos estáticos** (que permanecen inalterables en todo momento), expresa mediante un **término dinámico** (que genera e introduce nuevos sumandos para la función de regresión), un conjunto de términos que siguen un comportamiento similar para distintos conjuntos de variables independientes. Esta concepción ofrece varias ventajas, entre las cuales se pueden citar:

- Ahorro de espacio en memoria y tiempo de máquina durante el proceso de edición y almacenamiento de modelos, ya que un modelo ocupa el mismo espacio y menos tiempo que una función de regresión estática y equivale a una cantidad de estas igual a la cantidad máxima de variables independientes con que se trabaje.

- Mayor adecuación a las pruebas estadísticas utilizadas que determinan la significación de cada término estático, lo cual elimina la incertidumbre del usuario en caso de que se generen algunos términos no previstos.

Para la especificación de la notación utilizada se adoptaron patrones para los **parámetros genéricos** (parámetros presentes en los términos dinámicos) y **metavariables** (variables dinámicas), los cuales se verán en el desarrollo de los siguientes ejemplos:

Suponiendo que se cuente con dos variables independientes

- Modelo log: y=ln(B0)+**..**B1j*(ln(z1))**..** genera y=ln(B0)+B1_1*(ln(x1))+B1_2*(ln(x2))

- Log-log: ln(y)=ln(B0)+**..**B1j*(ln(z1))**..** genera ln(y)=ln(B0)+B1_1*(ln(x1))+B1_2*(ln(x2))

Suponiendo que se cuente con tres variables independientes

- Lineal: y=B0+**..**B1j*(z1)**..** genera y=B0+B1_1*(x1)+B1_2*(x2)+B1_3*(x3)

- Cuadrático: y=B0+**..**B1j*(z1*z2)**..** genera: y=B0+B1_1*(x1*x1)+B1_2*(x1*x2)+ B1_3*(x1*x3)+B1_4*(x2*x2)+ B1_5*(x2*x3) +B1_6*(x3*x3)

- mientras y=B0+**...**B1j*(z1*z2)**...** +**..**B2j*(z1^2)**..** genera: y=B0+B1_1*(x1*x2) +B1_2*(x1*x3)+B1_3*(x2*x3)+B2_1*(x1^2) +B2_2*(x2^2)+ B2_3*(x3^2) que no es más que otra representación para el modelo cuadrático.

Como se puede apreciar, las metavariables están identificadas por la letra z y los parámetros genéricos por dos subíndices donde el segundo alude a la cantidad de términos estáticos que serán generados a partir de cada uno de los dinámicos.

Mediante esta notación, para crear funciones de regresión, a partir de un modelo, se sustituyen convenientemente las metavariables por variables estáticas, obteniéndose expresiones estáticas de manera tal que no se produzcan repeticiones.

El uso de dos combinaciones de puntos (".." y "…") permite más flexibilidad y poder expresivo de los modelos. La diferencia en los funcionamientos entre ambas combinaciones es que si la última combinación que aparece antes de una expresión dinámica es ". ." se permite la asignación a distintas metavariables de la misma variable independiente en cada una de las expresiones estáticas que se generen. Si el indicador es "…" esto último no se permite.

Funcionamiento del SICEP

Alcance y funcionalidades

Inicialmente se definieron tres ramas de aplicación para el SICEP. Las mismas son:

- Caña.
- Cultivos varios.
- Producción y servicios.

Para cada una de dichas ramas el sistema SICEP incluye las siguientes tareas:

- Gestiona series históricas y estimados.
- Realiza reportes estadísticos.
- Edita, comprueba y guarda modelos de regresión lineal múltiple.
- Genera funciones de regresión lineal múltiple a partir de los modelos definidos.
- Resuelve funciones de regresión lineal múltiple.
- Calcula estimados.

Al ejecutar la aplicación y después de autenticarse el usuario en el sistema aparece la ventana principal que se muestra en la figura 1. A partir de esta ventana se pueden acceder a todas las opciones del sistema.

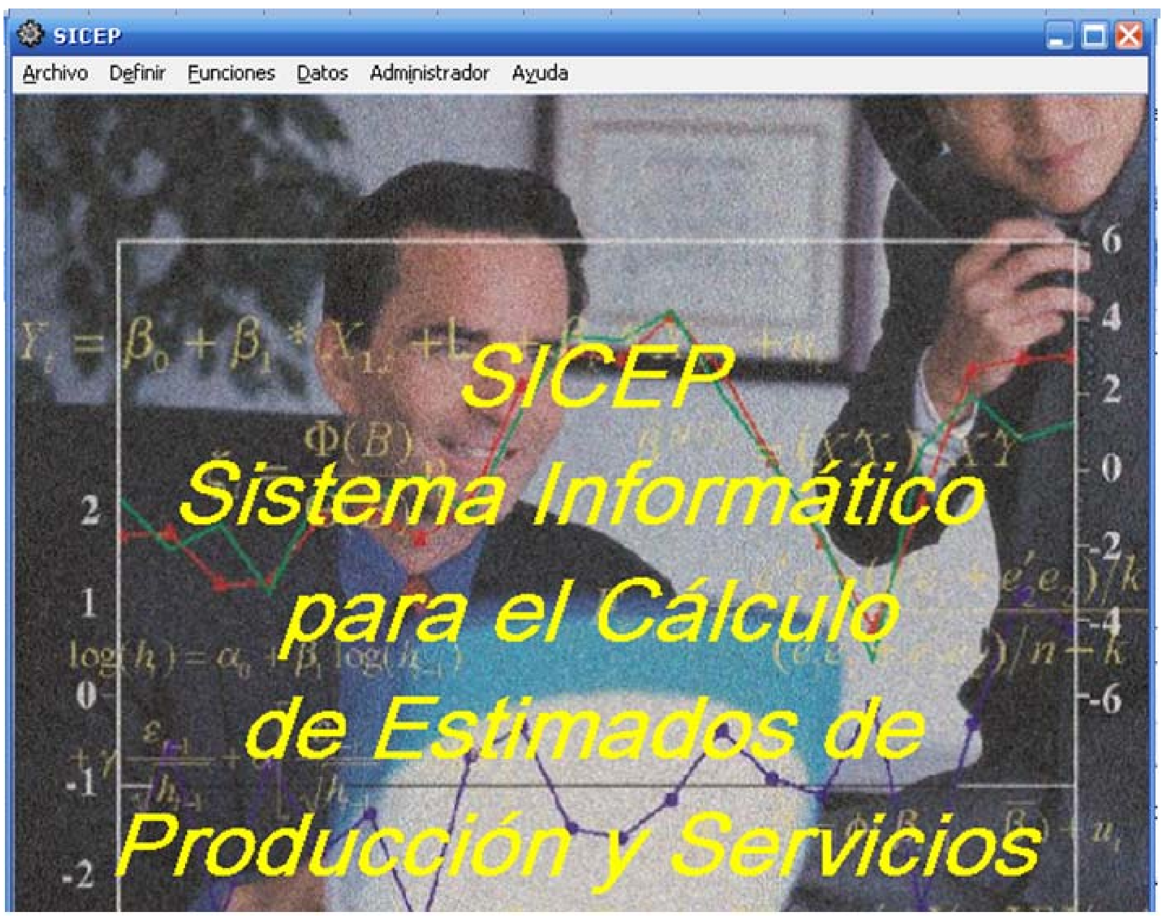

Figura 1: Ventana principal.

Edición de modelos

Para editar, comprobar y guardar modelos hay que presionar el submenú "Archivo" en la ventana principal y luego escoger la opción "Nuevo Modelo". Una vez hecho esto, aparece la ventana de edición de modelos, la cual se muestra en la figura 2. Esta ventana es un pequeño editor de modelos. En la figura se muestra en el cuadro de edición de modelos la metavariable z1. Como se recordará, esta se diferencia de las variables (las que empiezan con x) en que al expandirse su término se convierte en todas las variables independientes existentes en el proceso objeto de estudio, radicando en esta particularidad una de las principales potencialidades del sistema.

Trabajo con las muestras

La figura 3 muestra la ventana que se usa para el trabajo con las muestras. Para acceder a la misma, por ejemplo, para trabajar con los datos de la caña, es necesario ir al submenú "Datos" del menú principal y escoger la opción "Muestras Caña".

Esta ventana brinda al usuario las opciones:

• Cargar los datos desde un origen especificado por el usuario.

• Adicionar y configurar las variables objetos de estudio, así como asociarlas a las variables teóricas. (y, x1,..., etcétera).

• Editar manualmente los datos.

• Exportar a Word los datos e importarlos desde varios sistemas (Word, Excel, SPSS, etcétera)

• General datos aleatorios por filas y columnas dentro del rango y en la posición de la tabla que el usuario desee.

• Seleccionar los modelos para establecer las funciones de respuesta que se analizarán.

• Gestionar las distintas opciones estadísticas con el nivel de significación deseado.

Figura 2: Ventana de edición de modelos.

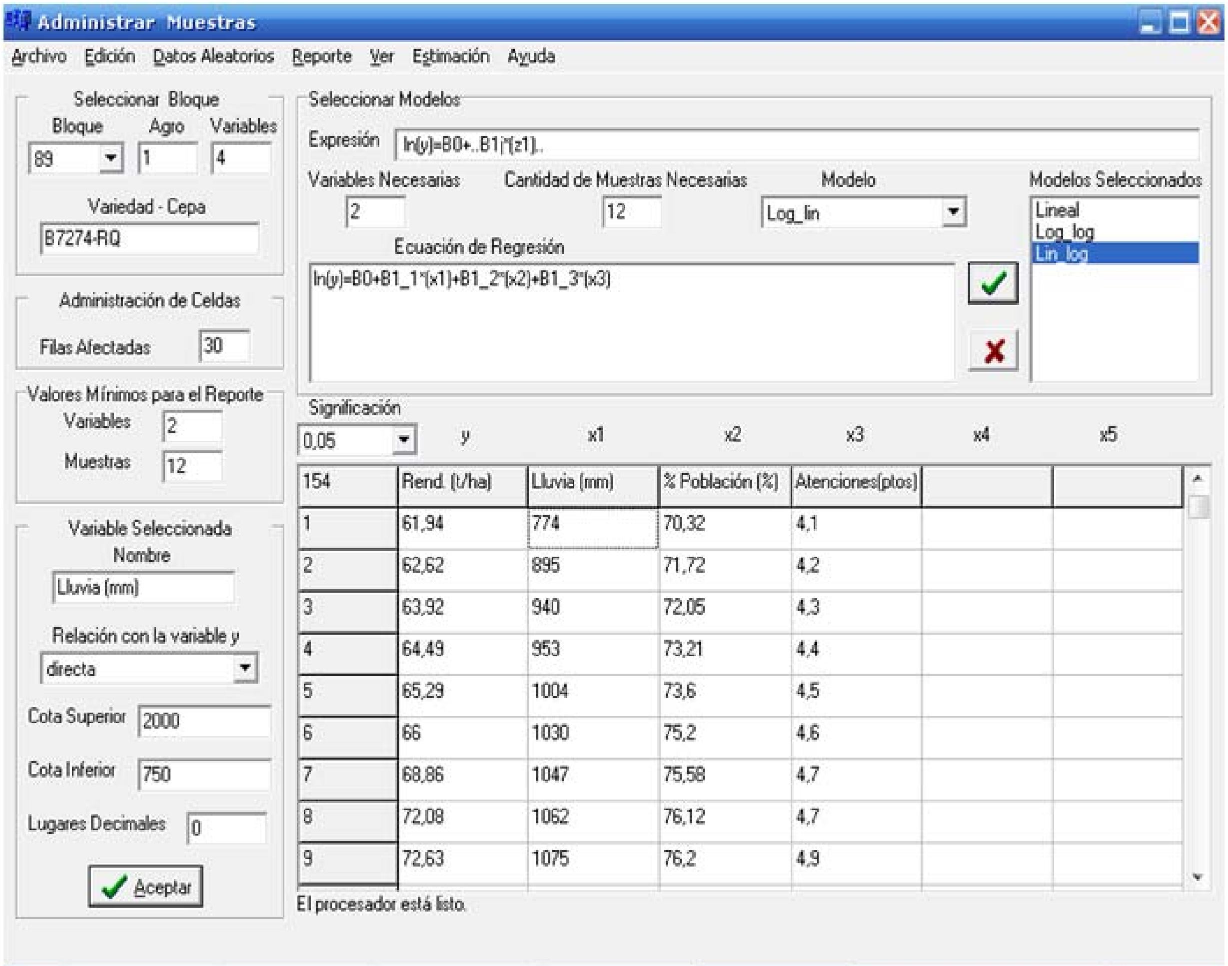

Figura 3: Ventana de administración de muestras.

Opciones estadísticas

Las opciones estadísticas aparecen distribuidas entre los submenús "Reporte", "Ver" y "Estimación" del menú principal de la ventana de gestión de muestras. Las distintas opciones estadísticas que brinda el sistema son:

• Cálculo y almacenamiento de estimados a partir de la mejor función de respuesta determinada por el sistema, considerando el nivel de confianza especificado manualmente, o de otra seleccionada por el usuario.

• Resumen estadístico general.

• Inferencia estadística que permite examinar la aceptación o no de cada una de las distintas pruebas estadísticas ante diferentes valores del nivel de confianza que puede ser 0,1; 0,5 o 0,01.

• Gráfica en 3D de las funciones de respuesta obtenidas.

Para concluir, y a modo de ejemplo, se muestra el resumen estadístico general que elabora el sistema en Word mediante la opción homónima del submenú "Reporte" del menú principal. Como se aprecia en la figura 4, el reporte refleja los resultados de las pruebas de hipótesis inherentes al procedimiento de selección de la mejor función respuesta ya explicado, relacionado con la significación de la función de regresión y de los coeficientes individuales y la comprobación de los supuestos de homocedasticidad, aleatoriedad y normalidad de las perturbaciones para el nivel de confianza especificado.

Reporte: Resumen Estadístico General
Bloque: 154
Variedad-Cepa: B7274-PQ
Agroecosistema: 1
Variables:
y: Rend. (t/ha)
x1: Lluvia (mm)
x2: % Población (%)
x3: Atenciones(ptos)
Significación: 0,05
Curva: Lineal
Ecuación de Regresión:
y=B0+B1_1*(x1)+B1_2*(x2)+B1_3*(x3)

Hipótesis:		Significación	Homos.	Aleatoriedad	Normalidad	
Prueba:		--Fisher--	--White--	--Rachas--	----Test KS---	
R^2	StdError	Pvalor	Pvalor	Pvalor	Calc	Tabla
0,9941	1,523354	0,0000	0,5436	0,1934	0,0745	0,1592
Decisión:		Aceptar	Aceptar	Aceptar	Aceptar	

Análisis de Coeficientes

Coeficiente	Valor	Pvalor	Significación
B0	5,935912	0,6197	No Significativo
B1_1	0,017420	0,0014	Significativo
B1_2	0,134876	0,5635	No Significativo
B1_3	7,529237	0,0001	Significativo

Ecuación de Regresión para la Homocedasticidad:

E^2=b0+b1_1*(x1)+b1_2*(x2)+b1_3*(x3)+b2_1*(x1*x1)+b2_2*(x1*x2)+b2_3*(x1*x3)+b2_4*(x2*x2)+b2_5*(x2*x3)+b2_6*(x3*x3)

Análisis de Coeficientes para Homocedasticidad

Coeficiente	Valor	Pvalor	Significación
b0	26,26755	0,9749	No Significativo
b1_1	-0,254600	0,5319	No Significativo
b1_2	-1,925514	0,9520	No Significativo
b1_3	65,47862	0,6339	No Significativo
b2_1	-0,000005	0,9721	No Significativo
b2_2	0,005141	0,5254	No Significativo
b2_3	-0,027589	0,6921	No Significativo
b2_4	0,036786	0,9050	No Significativo
b2_5	-1,642184	0,5467	No Significativo
b2_6	8,493553	0,4772	No Significativo

Figura 4: Reporte general del SICEP.

Resultados y discusión

Resultados

Este sistema fue introducido en la Empresa Azucarera "Paquito Rosales" con el fin de estimar el rendimiento cañero. Al no existir en las entidades cañeras información confiable de las series históricas de más de l0 años sobre el comportamiento de los rendimientos agrícolas se tomó como información de partida los valores mínimos y máximos de cada una de las variables (rendimiento, lluvia, atenciones culturales y % de población). Para el agroecosistema Cabaña los datos se presentan a continuación:

Tabla 1: Agroecosistema Cabaña. Año 2009

Cepas	Rendimiento (t/ha)		Lluvia (mm)		Atenciones culturales (puntos 1-10)		% de población (%)	
	Mín.	Máx.	Mín.	Máx.	Mín.	Máx.	Mín.	Máx.
PQ	60	120	750	2000	4	9	70	97
RQ	60	95	800	1100	4	9	80	92
F	70	110	750	1500	4	9	85	95
P	35	70	520	1800	4	9	85	95

Una muestra del cálculo del costo planificado utilizando el estimado del programador y el de la función de respuesta se presenta en la tabla 2. Para esto se parte del costo total de cultivo, calculado por el Departamento de Costo del extinto MINAZ Nacional, actualmente AZCUBA, para la Empresa Azucarera "Los Reynaldos" para una hectárea en secano y para el suelo sialitizado cálcico que corresponde al agroecosistema considerado. Este costo es de $434,57 y no considera el alza, corte y tiro de la caña, y el costo en divisas.

Se partió de considerar cuatro modelos fundamentales a saber:

- Lineal.
- Log-lin.
- Lin-log.
- Log-log.

Se obtuvo como función respuesta el modelo lineal y a partir de aquí se proyectaron los rendimientos en las demarcaciones identificadas por Bloque-Campo para las cepas que aparecían sembradas en dichas áreas.

Tabla 2: Agroecosistema Cabaña. Zafra 10-11

Cepas	Bloque-Campo	Rendimiento (t/ha)		Real (t/ha)	Costo plan de cultivo 434,57 ($/ha)*	
		Programador	Función		Programador ($/t)	Función ($/t)
PQ	154-508	54	63,68	61	8,04	6,92
RQ	89-114	80	89,73	86	5,43	4,84
F	82- 9	76,8	81	81	5,65	5,36
R	86-3	38,6	38,98	41	11,25	11,14

* Datos oficiales de la Empresa Azucarera Santiago de Cuba.

Como se muestra en el cuadro, los costos planificados a partir de los estimados hallados por la función son menores para todas las cepas que los que fueron previstos por el programador de la entidad y cuentan con el aval de estar fundamentados en rendimientos más próximos a los reales que fueron obtenidos empíricamente. Esto implica que los costos también serán menores a nivel de empresa azucarera.

Discusión

Un análisis de los resultados obtenidos en la tabla 2 arrojó que los mismos podían lograr la disminución de los costos en la producción cañera de un 11 %. El análisis de regresión también podría derivar en otros aportes tales como la determinación de la influencia de los factores analizados en las distintas variedades y cepas, así como proveer información precisa que sirva como punto de partida para decidir cambios y adiciones en las estructuras de siembra.

Conclusiones

Como parte de este trabajo se elaboró el sistema SICEP, validado en cuanto a su precisión y confiabilidad por el SPSS 11,5; el cual ha sido introducido con éxito en algunas entidades vinculadas a la caña, los cultivos varios y la producción y los servicios. En el caso de la caña el sistema ha sido validado en Unidades Empresariales de Base (UEBs) atención a productores y UEB centrales azucareros seleccionados de la provincia Santiago de Cuba, a partir de la zafra 2010-2011 con un impacto de disminución de los costos de producción de caña un 11% por concepto de una mejor estimación de los rendimientos cañeros. Como ha quedado evidenciado en este trabajo el sistema SICEP es más recomendable que los paquetes estadísticos tradicionales en situaciones donde el análisis de regresión demande:

- ***Flexibilidad de los modelos soportados.***
- ***Facilidades para la utilización de nuevos modelos por el usuario.***
- ***Automatización de todos los pasos que el análisis de regresión y los supuestos del modelo clásico implican.***

Recomendaciones

Los autores recomiendan a los directivos de otras entidades económicas y productivas la utilización del SICEP para el pronóstico de indicadores físicos y económicos.

Bibliografía

1. ALMAZÁN DEL OLMO, Oscar. "Viabilidad y alternativa económica". *Revista Bohemia*. 2002, Nº 23, p. 15-19.
2. ÁLVAREZ, S. *Adoosi Visual: Metodología de análisis y diseño orientada a objetos de sistemas informativos en medios ambientes visuales.* Departamento de Informática CEIS. ISPJAE. Abril de 2002.
3. BETANCOURT LOYOLA, Manuel. "Sistemas automatizados para el tiro de la caña y la siembra de variedades y cepas". Tesis de Maestría, Facultad de Matemática y Computación. Universidad de Oriente, Cuba, 2000.
4. CRAIG, Larman. *UML y Patrones. Introducción al análisis y diseño orientado a objetos.* La Habana: Editorial Pueblo y Educación, 2005.
5. CUÉ MUÑIZ, Juan L., y otros.*Estadística*. La Habana, (La Habana: [s/n]. 1987.
6. DANÍLINA, N.I., y otros. *Matemática de cálculo*. Mocú: Editorial Mir, 1990.
7. EILON, Samuel. Industrial Engineering Tables. La Habana: Editorial Pueblo y Educación, 1979.
8. GUERRA BUSTILLO, Caridad W., y otros. *Estadística*. La Habana: Editorial Pueblo y Educación, 1991.
9. GUJARATI, D. N. *Econometría.* Quinta edición. México: Editorial Grijalba, 2005.
10. HANSEN, GARY; HANSEN, James . *Database Management and Design*. New Jersey, 2003.
11. HERRERA MACHADO, Juan. "La eficiencia agroindustrial". Material del MINAZ, noviembre, 1990, p. 13-15.
12. JACOBSON, Ivar; BOOCH, Grady; Rumbaugh, James : *El lenguaje de modelado unificado.*Addison Wesley, Reading, Massachusetts, EE.UU. 1999.
13. _____: *El proceso unificado de desarrollo de software.* Addison Wesley, Reading, Massachusetts, EE.UU. 1999
14. KATRIB MORA, Miguel. *Lenguajes de programación y técnicas de compilación*. La Habana: Editorial Pueblo y Educación, 1988.
15. MILLER R., Irwin, y otros. Probabilidad y Estadística para ingenieros. Editorial Félix Varela, La Habana, Cuba, 2004.
16. RUMBAUGH, y otros. *Object Modelling Techniquesn and Design.* Prewntice Hall, 2002.
17. ULLMAN, Jefrey D. *Principles of Database Systems.* Computer Science Press. Second Edition, New York, EE.UU., 2002.
18. YOURDON, E; Coad, P. ***Object-oriented analysis.*** Second Edition, Yourdon Press, Englewood Cliffs, New Jersey, EE.UU, 1991.