

# **ANÁLISIS ESTADÍSTICO DE LA PRODUCCIÓN Y EL CONSUMO DE TRIGO EN EL ECUADOR**

Christian Rodríguez Jacho<sup>1</sup>, Fernando Sandoya Sánchez<sup>2</sup>

<sup>1</sup>Ingeniero en Estadística Informática 2000

<sup>2</sup>Director de Tesis. Matemático, Escuela Politécnica Nacional, 1993, Especialización en Estadística e Investigación de Operaciones, Profesor de la ESPOL desde 1995

## **RESUMEN**

El presente trabajo consiste en el Análisis de la producción y el consumo de trigo en el Ecuador, con el fin de aplicar diferentes técnicas estadísticas como por ejemplo estadísticas descriptivas, análisis de regresión, análisis de series de tiempo y análisis multivariado.

En la primera parte se desarrolla el análisis estadístico de la producción y el consumo de trigo en el Ecuador mediante modelos de regresión de mínimos cuadrados para estas variables, así como también se obtienen proyecciones para el próximo año.

## **INTRODUCCIÓN**

El presente trabajo es un análisis de la producción interna y el consumo de trigo en el Ecuador, con el cual se estima la producción interna en toneladas métricas y la superficie cosechada en hectáreas en diferentes años, para esto se propone un modelo de regresión de mínimos cuadrados, independiente para cada una de las variables. El consumo de trigo lo estimamos a través del principal producto elaborado a partir de este, como es la harina de trigo, porque se lo utiliza para la fabricación de pan, pastelería, galletería, para la fabricación de fideos y otros; estos productos desempeñan una función primordial en la alimentación humana, siendo esta la razón para que el cultivo de trigo se encuentre difundido en todas las latitudes y tenga repercusiones sociales, políticas y económicas en el ámbito mundial.

Para el efecto del análisis, se ha utilizado estadísticas descriptivas para observar de forma general el comportamiento de las variables. Análisis de series de tiempo para elaborar un modelo que pronostique el comportamiento de una variable en un período determinado, y el análisis multivariado para hacer inferencias de una población con múltiples variables, basados en información que proporciona una muestra.

## **CONTENIDO**

Todo el análisis, que se muestra posteriormente, corresponde a la producción, costo, precio y ventas de una empresa representativa de la Industria molinera de nuestro país, que tiene sus operaciones en la ciudad de Guayaquil, con lo que tratamos de reflejar y analizar el comportamiento del sector en el Ecuador.

## Producción y consumo de harina de Trigo.

Hemos escogido para el análisis 14 variables que influyen y determinan la producción y consumo de harina de Trigo en el Ecuador en el período de Enero de 1994 hasta Noviembre de 1999.

Las primeras 4 variables se relacionan al número de quintales de harina producidos y vendidos, así como también a la cantidad de quintales que se necesita para la producción mensual. Las 8 siguientes es su mayoría son costos relacionados a la producción y venta de los quintales de harina producidos, y las 2 restantes se relacionan con precios promedios de venta y valor de recuperación de producir un quintal de harina. Gracias a las herramientas estadísticas del análisis multivariante determinaremos las variables que son más representativas.

A continuación se detalla con más precisión las variables para el análisis y su notación:

$X_1$ : Cantidad de quintales de harina producidos

Corresponde a la cantidad mensual de quintales de harina producidos.

$X_2$ : Cantidad de quintales de harina vendidos

Corresponde a la cantidad mensual de quintales de harina vendidos.

$X_3$ : Cantidad de quintales de trigo usados

Corresponde a la cantidad mensual de quintales de trigo usados para la producción.

$X_4$ : Cantidad de quintales de afrechillo producidos

Corresponde a la cantidad mensual de quintales de afrechillo resultado de la pilación de trigo.

$X_5$ : Costo de trigo usado para producir un quintal de harina

Corresponde al valor promedio mensual en sucres del insumo trigo para producir un quintal de harina.

$X_6$ : Costo de aditivos por cada quintal de harina

Corresponde al valor promedio mensual en sucres de los aditivos para producir un quintal de harina.

$X_7$ : Costo de sacos y piolas por cada quintal de harina

Corresponde al valor promedio mensual en sucres de los insumos saco y piola para producir un quintal de harina.

$X_8$ : Costo de mano de obra por cada quintal de harina

Corresponde al valor promedio mensual en sucres de la mano de obra utilizada para producir un quintal de harina.

$X_9$ : Costo de servicios básicos

Corresponde al valor promedio mensual en sucres de los servicios básicos, como agua, luz, teléfono, para producir un quintal de harina.

$X_{10}$ : Costo de servicio técnico

Corresponde al valor promedio mensual en sucres del servicio técnico utilizado para producir un quintal de harina.

$X_{11}$ : Costo total de fabricación de un quintal de harina

Corresponde al valor promedio total mensual en sucres para producir un quintal de harina.

$X_{12}$ : Precio promedio de venta por cada quintal de harina

Corresponde al valor promedio mensual en sucres de comercialización de un quintal de harina.

$X_{13}$ : Costo promedio de venta de un quintal de harina

Corresponde al valor promedio mensual en sucres de comercializar un quintal de harina.

$X_{14}$ : Recuperación de Subproducto

Corresponde al valor promedio mensual en sucres de recuperación como resultado de la pilación del trigo.

## Resultados de las variables de la producción y consumo de harina de trigo.

### Variable:

#### $X_1$ : Cantidad de quintales de harina producidos

El 98.12% de la variabilidad de la cantidad de quintales de harina producidos ( $X_1$ ), es explicada por el modelo de regresión lineal de mínimos cuadrados.

La cantidad de quintales de harina producidos se puede estimar, mediante el modelo:

$$X_1 = -67.494 + 0.761X_3$$

donde  $X_3$  es la cantidad de quintales de trigo usados.

### Series de Tiempo

El modelo ARMA(1,2) se ajusta a la serie de la variable  $X_1$ , los resultados se muestran en la tabla a continuación:

TABLA I  
COEFICIENTES PARA EL MODELO ARMA(1,2) DE LA VARIABLE  $X_1$

Tipo	Coeficientes	Inferior 95%	Superior 95%
k	136634.884	135003.763	138266.004
AR	-0.881	-1.121	-0.642
MA	-1.337	-1.674	-1
MA	-0.489	-0.357	0.429
Valor MCE=1.73115E+08			
Prueba de Box-Ljung		Q=26.063 p=0.112	

En la tabla I tenemos los intervalos de predicción para el modelo ARMA(1,2) para la variable ( $X_1$ ) cantidad de qq de harina de trigo producidos, en la misma se detalla, para cada período el valor inferior del intervalo, el valor de la predicción y el valor superior correspondiente, luego se muestra el gráfico de estas predicciones (figura 1), junto con la serie anterior.

TABLA II  
INTERVALO DE PREDICCIÓN (95% DE CONFIANZA) PARA  $X_1$

<i>Inferior</i>	<i>Predicción</i>	<i>Superior</i>
55,125.40	80,913.74	106,702.08
40,266.53	68,609.47	96,952.41
47,747.68	76,180.25	104,612.81
41,007.37	69,509.33	98,011.28
46,831.64	75,387.35	103,943.06
41,610.60	70,207.98	98,805.38
46,142.04	74,771.73	103,401.43
42,095.67	70,750.43	99,405.19
45,619.57	74,293.77	102,967.97
42,482.30	71,171.59	99,860.88
45,221.67	73,922.66	102,623.66
42,788.50	71,498.58	100,208.66

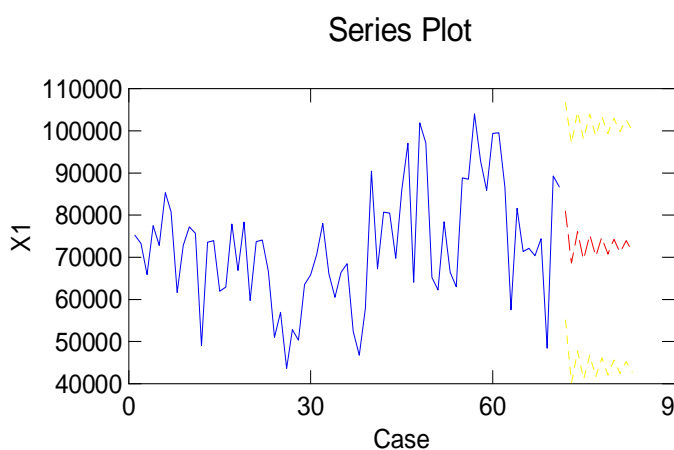


FIGURA 1. PREDICIONES DE LA VARIABLE  $X_1$

**Variable:**

**$X_2$ : Cantidad de quintales de harina vendidos**

El 85.75% de la variabilidad de la cantidad de quintales de harina vendidos ( $X_2$ ), es explicada por el modelo de regresión lineal de mínimos cuadrados.

La cantidad de quintales de harina vendidos se puede estimar, mediante la siguiente ecuación:

$$X_2 = 7,596.536 - 0.883X_1$$

donde  $X_1$  es la cantidad de quintales de harina producidos.

**Series de tiempo**

El modelo ARMA(1,2) se ajusta a la serie de la variable  $X_2$ , los resultados se muestran en la tabla a continuación:

TABLA III  
COEFICIENTES PARA EL MODELO ARMA(1,2) DE LA VARIABLE  $X_2$

Tipo	Coeficientes	Inferior 95%	Superior 95%
k	132947.212	132547.308	133347.115
AR	0.145	-1.141	-0.563
MA	0.168	-1.708	-1.038
MA	0.125	-0.763	-0.264
Valor MCE=1.46600E+08			
Prueba de Box-Ljung		Q=22.788 p=0.120	

En la tabla IV tenemos los intervalos de predicción para el modelo ARMA(1,2) para la variable ( $X_2$ ) cantidad de qq de harina de trigo vendidos, en la misma se detalla, para cada período el valor inferior del intervalo, el valor de la predicción y el valor superior correspondiente, luego se muestra el gráfico de estas predicciones, (figura 2) junto con la serie anterior.

TABLA IV  
INTERVALO DE PREDICCIÓN (95% DE CONFIANZA) PARA  $X_2$

Inferior	Predicción	Superior
52,071.57	75,802.98	99,534.38
43,822.93	70,577.16	97,331.40
45,986.11	72,791.55	99,596.98
44,061.58	70,904.14	97,746.70
45,643.34	72,512.85	99,382.36
44,252.63	71,141.69	98,030.75
45,407.12	72,310.38	99,213.65
44,400.69	71,314.26	98,227.83
45,242.23	72,163.30	99,084.36
44,513.13	71,439.63	98,336.13
45,125.99	72,056.44	98,986.89
44,597.38	71,530.70	98,464.02

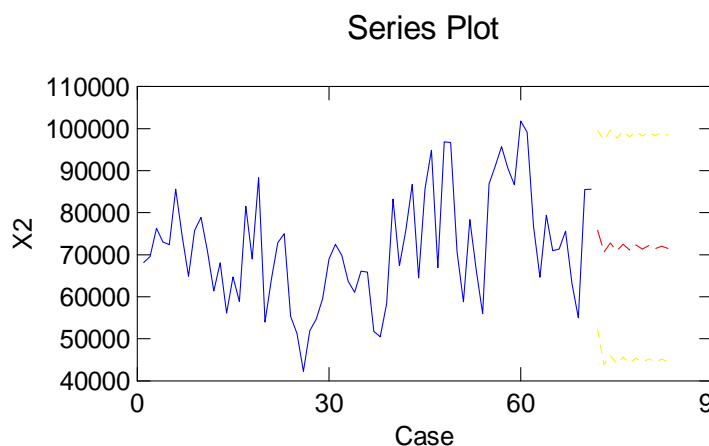


FIGURA 2. PREDICIONES DE LA VARIABLE  $X_2$

**Variable:**

**X<sub>3</sub>: Número de quintales de trigo usados**

TABLA V  
ESTADISTICA DESCRIPTIVA:  
NUMERO DE QUINTALES USADOS DE TRIGO  
Año: 1994-1999

Tamaño de la muestra n (meses)	71
Máximo valor: en Diciembre de 1997	13,6685.4
Mínimo valor: en Febrero de 1996	58,479
Rango	78,206.4
Mediana	94,081.2
Media	95,451.854
Intervalo de confianza para la media	$91,119.564 \leq \mu \leq 99,784.143$
Desviación estándar	18,303.170
Desviación estándar del error	2,172.187
Varianza	3.35006e+08
Sesgo	0.219
Kurtosis	-0.334

La tabla nos muestra toda la información válida para un análisis, por ejemplo tenemos a la media mensual que nos indica que la cantidad promedio de quintales usados de trigo es 95,451.854 qq mensualmente. La desviación estándar nos indica que la cantidad de quintales usados de trigo se desvía en 18,303.170 qq de la media mensual. También muestra un intervalo de 95% de confianza para la media, que nos indica que de 100 veces al menos 95 veces la media de la cantidad de quintales usados de trigo caerá dentro de [91,119.564;99,784.143] qq.

Variable dependiente:

X<sub>12</sub>: Precio promedio de venta por cada quintal de harina.

Variable independiente: (X<sub>5</sub>) Costo de trigo usado para producir un quintal de harina.

El 74% de la variabilidad del precio promedio de venta por cada qq de harina de trigo (X<sub>12</sub>), es explicada por el modelo de regresión cuadrático de mínimos cuadrados.

El precio promedio de venta por cada quintal de harina se puede estimar, mediante el modelo:

$$X_{12} = 26.471 - 2.653X_5 - 0.157X_5^2$$

donde X<sub>5</sub> es el Costo de trigo usado para producir un quintal de harina.

### **Análisis Multivariado**

El resultado obtenido después de haber utilizado esta técnica de reducción de variables, que consiste en:

Suponga que se tiene p variables que se desea analizar (en nuestro caso son 14 variables), la técnica se basa en el hecho de que existen k variables "latentes" (k < p), que pueden explicar de manera aceptable la variabilidad contenida en las p variables originales, reemplazándolas, con lo que se obtiene

una menor cantidad de variables a analizar, es decir, si se tenía  $p$  variables con  $n$  observaciones (en nuestro caso  $n=71$  observaciones), habiendo encontrado las  $k$  componentes o factores (en nuestro caso  $k=4$ ), se tiene solo  $k$  variables con  $n$  observaciones.

La tabla VI muestra las 4 nuevas variables encontradas o componentes principales, las mismas que explican el 84,67% de toda la información existente en el análisis. La posición  $ij$  denota el coeficiente de la variable  $j$  en la componente  $i$  para formar el vector propio.

TABLA VI  
COMPONENTES PRINCIPALES DEL ANÁLISIS

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14
C1	-0.333	-0.321	-0.320	-0.243	0.304	0.087	0.262	0.270	0.210	0.025	0.339	0.322	0.342	-0.056
C2	0.310	0.323	0.337	0.385	0.344	-0.072	0.073	0.005	-0.068	0.159	0.287	0.245	0.294	-0.388
C3	0.058	-0.025	0.087	0.153	0.185	-0.406	0.062	-0.358	0.478	-0.514	0.108	0.016	0.076	0.352
C4	0.098	0.074	0.131	0.238	-0.071	0.634	0.554	0.174	0.219	-0.285	-0.088	-0.144	-0.086	0.039

A cada componente se le da un nombre, al cual se le denominará factor, este reunirá a las todas las variables del análisis pero su nombre lo determinará las variables que mayor pesen en el componente.

A continuación se detallará los factores encontrados que determinan la producción y consumo de harina de trigo en una Industria molinera representativa del Ecuador.

#### Factor 1

El primer factor aporta con el mayor porcentaje de varianza, 41.04%, se explica en términos de las variables ( $X_5$ ) Costo de trigo usado para producir un quintal de harina, ( $X_{11}$ ) Costo total de fabricación de un quintal de harina, ( $X_{12}$ ) Precio promedio de venta por cada quintal de harina, ( $X_{13}$ ) Costo promedio de venta de un quintal de harina, todas las variables están asociadas con valores en dólares, en los que se incurre para realizar la distribución de harina de trigo, por lo que se ha considerado definir a este factor como "Comercialización de harina de trigo".

#### Factor 2

El segundo factor aporta con un 22.25% de la varianza total y esta explicada por las variables: ( $X_1$ ) Cantidad de quintales de harina producidos, ( $X_2$ ) Cantidad de quintales de harina vendidos, ( $X_3$ ) Cantidad de quintales de trigo usados, ( $X_4$ ) Cantidad de quintales de afrechillo producidos, ya que todas las variables representan unidades (en quintales) de producción y venta, llamaremos a esta componente: "Producción de harina de trigo"

#### Factor 3

El tercer factor aporta con un 11.26% de la varianza total y esta explicada por las variables: ( $X_6$ ) Costo de aditivos por cada quintal de harina, ( $X_7$ ) Costo de sacos y piolas por cada quintal de harina, ( $X_8$ ) Costo de mano de obra por cada quintal de harina. Las variables que más aportan están encerradas en un

círculo cercano a la abscisa para la componente 3. A este factor lo denominaremos “Costos directos de producción”.

#### Factor 4

El cuarto factor aporta con un 10.11 % de la varianza total y esta explicada por las variables: ( $X_9$ ) Costo de servicios básicos, ( $X_{10}$ ) Costo de servicio técnico, ( $X_{14}$ ) Recuperación de Subproducto. Las variables que más aportan están encerradas en un círculo cercano a la ordenada para la componente 4. A este factor lo denominaremos “Costos indirectos”.

## CONCLUSIONES

1. En el Ecuador aún no existe cultura estadística, ya que en el desarrollo de esta investigación se pudo constatar que tanto empresas privadas como gubernamentales no prestan las facilidades necesarias para obtener información que permita llevar a cabo una investigación de este tipo.
2. El volumen de importación de trigo en toneladas métricas desde 1991 hasta 1999 registradas por el Banco Central del Ecuador es de 3,456,090.81 Tm. Y el valor de esas importaciones de trigo fue de 663,688,703.36 Usd, es decir el dinero invertido por el país en este rubro representa una importante causa de la salida para el Ecuador.
3. El volumen de importación de trigo se puede estimar por medio de la siguiente ecuación  $y = 128,650,000 - 2,745,534.65x + 14,677.27x^2$   
Donde  $y$  es el volumen de importación de trigo y  $x$  es el año de importación.
4. El país desde donde se importa la mayor cantidad de trigo es Estados Unidos, desde 1991 hasta marzo de 1999 el Ecuador importó de esta región un total de 1,715,295,51 toneladas métricas, aproximadamente el 51% del total importado desde esa fecha hasta la actualidad. El segundo país abastecedor de trigo es Canadá con 923,519.36 toneladas métricas, aproximadamente el 27% del total importado.
5. El *origen* del trigo para consumo humano proviene en una gran mayoría de Estados Unidos de Norteamérica con un 48% en el período de 1991-1999. El Ecuador solo ha producido para consumo alrededor de 185,859.31 toneladas métricas aproximadamente el 5% del total de trigo para consumo que es de 3,641,950.12 toneladas métricas.
6. La producción de trigo está decreciendo considerablemente, debido a muchos factores, podríamos citar alguno de ellos como por ejemplo la falta de incentivos y consideraciones al productor agrícola que se dedica a la cosecha de este grano.
7. En 1983 la producción de trigo en el Ecuador fue de alrededor 20,334 toneladas métricas y para 1999 se proyectó en 17,312 toneladas métricas con tendencia la baja.



8. La curva de regresión para estimar la producción de trigo en el Ecuador es:

$$y = -18974500 + 618853.386x - 6698.41x^2 + 24.092x^3$$

Donde  $y$  es el la producción en Tm de trigo y  $x$  es el año de producción.

9. La superficie cosechada con trigo desde 1983 hasta 1999 tiende una forma decreciente, teniendo en 1983 35,700 ha cosechadas y en 1999 26,878 ha cosechadas, con lo cual se obtuvo un descenso del 24.71% en las ha cosechadas.

10. Una de las causas por la baja producción de trigo en el Ecuador se refleja en su rendimiento que desde 1983 hasta 1999 ha sido en promedio de 0.68.

11. La cantidad de quintales de harina vendidos se puede estimar, mediante el modelo de regresión lineal:

$$X_2 = 7,596.536 - 0.883X_1$$

Donde  $X_1$  es la cantidad de quintales de harina producidos.

12. El precio promedio de venta por cada quintal de harina se puede estimar, mediante el modelo de regresión cuadrático:

$$X_{12} = 26.471 - 2.653X_5 - 0.157X_5^2$$

Donde  $X_5$  es el Costo de trigo usado para producir un quintal de harina.

13. La cantidad de quintales de harina producidos se puede estimar, mediante el modelo de regresión lineal:

$$X_1 = -67.494 + 0.761X_3$$

Donde  $X_3$  es la cantidad de quintales de trigo usados.

14. La determinación de la producción y el consumo de harina de trigo para una empresa representativa de la Industria molinera puede ser agrupado en cuatro factores los cuales reúnen el 84.67 % de toda la información contenida en las 14 variables incluidas en el análisis.

Estos factores son:

#### Factor 1

El primer factor aporta con el mayor porcentaje de varianza, 41.04%, se explica en términos de las variables ( $X_5$ ) Costo de trigo usado para producir un quintal de harina, ( $X_{11}$ ) Costo total de fabricación de un quintal de harina, ( $X_{12}$ ) Precio promedio de venta por cada quintal de harina, ( $X_{13}$ ) Costo promedio de venta de un quintal de harina, todas las variables están asociadas con valores en dólares, en los que se incurre para realizar la distribución de harina de trigo, por lo que se ha considerado definir a este factor como "Comercialización de harina de trigo".

#### Factor 2

El segundo factor aporta con un 22.25% de la varianza total y esta explicada por las variables: ( $X_1$ ) Cantidad de quintales de harina producidos, ( $X_2$ ) Cantidad de quintales de harina vendidos, ( $X_3$ ) Cantidad de quintales de trigo usados, ( $X_4$ ) Cantidad de quintales de afrechillo producidos, ya que

todas las variables representan unidades (en quintales) de producción y venta, llamaremos a esta componente: “Producción de harina de trigo”

#### Factor 3

El tercer factor aporta con un 11.26% de la varianza total y esta explicada por las variables: ( $X_6$ ) Costo de aditivos por cada quintal de harina, ( $X_7$ ) Costo de sacos y piolas por cada quintal de harina, ( $X_8$ ) Costo de mano de obra por cada quintal de harina.

Las variables que más aportan están encerradas en un círculo cercano a la abscisa para la componente 3. A este factor lo denominaremos “Costos directos de producción”.

#### Factor 4

El cuarto factor aporta con un 10.11 % de la varianza total y esta explicada por las variables: ( $X_9$ ) Costo de servicios básicos, ( $X_{10}$ ) Costo de servicio técnico, ( $X_{14}$ ) Recuperación de Subproducto. Las variables que más aportan están encerradas en un círculo cercano a la ordenada para la componente 4. A este factor lo denominaremos “Costos indirectos”.

15. Las provincias en las que se cultiva son Bolívar, Chimborazo, Pichincha, Carchi e Imbabura, siendo la de mayor producción la provincia de Bolívar, en efecto, se encontró que en el total producido entre 1983 y 1997 esta región contribuyó a la producción nacional con un 24% del total.
16. En los primeros meses del año 2000 se dará un decrecimiento en el volumen de producción de qq de harina de trigo, tal como lo muestran las predicciones generadas utilizando el modelo ARMA(1,2)

<i>Mes</i>	<i>Inferior</i>	<i>Predicción</i>	<i>Superior</i>
Diciembre	55,125.40	80,913.74	106,702.08
Enero	40,266.53	68,609.47	96,952.41
Febrero	47,747.68	76,180.25	104,612.81

17. En los primeros meses del año 2000 se dará un decrecimiento en el volumen de ventas de qq de harina de trigo, tal como lo muestran las predicciones generadas utilizando el modelo ARMA(1,2).

<i>Mes</i>	<i>Inferior</i>	<i>Predicción</i>	<i>Superior</i>
Diciembre	52,071.57	75,802.98	99,534.38
Enero	43,822.93	70,577.16	97,331.40
Febrero	45,986.11	72,791.55	99,596.98

## REFERENCIAS

### a) Página Web

1. [www.sica.gov.ec/cadenas/trigo](http://www.sica.gov.ec/cadenas/trigo)

### b) Tesis

2. Ch. Rodríguez Jacho, “Análisis estadístico de la producción y el consumo de trigo en el Ecuador” (Tesis, Instituto de Ciencias Matemáticas, Escuela Superior Politécnica del Litoral, 2000)

**c) Artículo**

3. CAPA SANTOS HOLGER, Modelación de Series Temporales, Escuela Politécnica Nacional, Junio de 1998

**d) Libro**

4. HAIR, ANDERSON, TATAHM, BLACK, Multivariate Data Analysis, Quinta Edición, Editorial Prentice Hall

**e) Libro**

5. ABASCAL GRANDE, Métodos Multivariantes para la Investigación Comercial, Primera Edición, Octubre de 1985, Editorial Ariel