

T
519.33
ART
C2



ESCUELA SUPERIOR POLITECNICA DEL LITORAL
Instituto de Ciencias Matemáticas

**ANALISIS ESTADISTICO DE LA PRODUCCION
BANANERA EN EL ECUADOR**

TESIS DE GRADO

**Previa a la obtención del Título de:
INGENIERO EN ESTADISTICA INFORMATICA**



Presentada por:

María Arteaga García



D-20538

GUAYAQUIL - ECUADOR

AÑO

2000

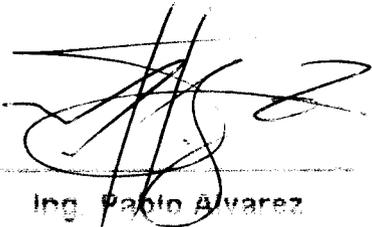
AGRADECIMIENTO

A todas las personas e instituciones que colaboraron en la realización de esta investigación, y en especial, a mis padres por su ayuda incondicional, y al Ing. Gaudencio Zurita, por ser mi guía durante mi carrera universitaria.

DEDICATORIA

A mi madre

TRIBUNAL DE GRADUACION



Ing. Pablo Alvarez
SUBDIRECTOR DEL ICM



Ing. Gaudencio Zurita
DIRECTOR DE TESIS



Ing. Marcelo Espinosa
VOCAL



Mat. Fernando Sandoya
VOCAL

DECLARACION EXPRESA

"La responsabilidad del contenido de esta Tesis de Grado, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITECNICA DEL LITORAL."

(Reglamento de Graduación de la ESPOL)

A handwritten signature in black ink, written in a cursive style, which reads "María Arteaga García". The signature is written over a horizontal line.

María Arteaga García

RESUMEN

En este trabajo se realiza el análisis estadístico de la producción bananera en el Ecuador. En el primer capítulo se detallan las generalidades del banano, su ecología y variedades. Además, se explica la evolución de las exportaciones bananeras desde 1927 hasta la actualidad.

En el segundo capítulo se presenta el marco teórico, es decir la descripción de las técnicas estadísticas utilizadas en la realización del análisis univariado y multivariado, y en la elaboración de los modelos de predicción.

En el tercer capítulo se aplican las técnicas univariadas y multivariadas, descritas en el capítulo anterior, a las variables de estudio. El análisis univariado incluye las estadísticas básicas: mínimo, máximo, media, mediana, desviación estándar, varianza, sesgo y kurtosis; histogramas de frecuencias, ojivas, diagramas de cajas y pruebas de bondad de ajuste (Kolmogorov-Smirnov). El análisis multivariado incluye la matriz de correlación y la aplicación de la técnica de componentes principales a los datos originales y estandarizados. Además se realiza el análisis de componentes principales de las dos clases de datos, utilizando el método de rotación: Varimax.

En el cuarto capítulo se establecen modelos de predicción para las variables de interés, utilizando dos técnicas: series temporales y modelos de regresión. La técnica de series temporales no dió buenos resultados, ya que no permitió modelar las principales variables de estudio: producción y exportación bananera. En cuanto a los modelos de regresión, se obtuvieron seis modelos satisfactorios, ya que superan la prueba de parámetros, cumplen con el supuesto de normalidad de los residuos y tienen una alta potencia de explicación. Además se realiza la prueba de independencia de residuos,, utilizando el estadístico de Durbin-Watson.

En el quinto capítulo constan las conclusiones y recomendaciones resultantes del análisis estadístico realizado en este trabajo.

INDICE GENERAL

	Pág.
RESUMEN	II
INDICE GENERAL	III
ABREVIATURAS	IV
INDICE DE FIGURAS	V
INDICE DE TABLAS	VI
INTRODUCCION	3
I. PRODUCCION BANANERA EN EL ECUADOR	4
1.1 Descripción y ecología del Banano	4
1.2 Origen y llegada del banano al Ecuador	16
1.3 Boom bananero en el Ecuador	24
1.4 La industria bananera del Ecuador	63
II. DESCRIPCION DE LAS TECNICAS ESTADISTICAS A UTILIZAR	68
2.1 Análisis univariado	68
2.2 Análisis multivariado	88
2.3 Series temporales	104
2.4 Modelos de regresión	120
III. ANALISIS DE DATOS UTILIZANDO TECNICAS MULTIVARIADAS	133
3.1 Análisis univariado	137
3.2 Análisis multivariado	182
3.2.1 Matriz de correlación de las variables de estudio	182
3.2.2 Análisis de componentes principales	183

IV. ANALISIS DE DATOS UTILIZANDO SERIES TEMPORALES Y MODELOS DE REGRESION	194
4.1 Series temporales	195
4.2 Modelos de regresión	216
V. CONCLUSIONES Y RECOMENDACIONES	221
APENDICES	
BIBLIOGRAFIA	

ABREVIATURAS

cm	Centímetro
FOB	Free on Board (franco a bordo)
frec	Frecuencia
h	hora
ha	Hectárea
IPC	Índice de precios al consumidor
Km.	Kilómetro
lb	Libra
LCI	Límite de confianza inferior
LCS	Límite de confianza superior
m	Metro
mm	milímetro
PIB	Producto Interno Bruto
R	Números reales
t	Tiempo
TM	Tonelada métrica
Z	Números enteros

INDICE DE FIGURAS

		Pág.
Figura 1.1	Planta de banano	5
Figura 1.2	Transportación de los racimos de banano	12
Figura 1.3	Manos de banano	13
Figura 1.4	Desmanada de los bananos	14
Figura 1.5	Manos seleccionadas para ser empacadas	15
Figura 1.6	Caja de banano	16
Figura 1.7	Ingreso por exportaciones de banano en el Ecuador (1927-1947)	21
Figura 1.8	Producto Interno Bruto del Ecuador (1927-1947)	23
Figura 1.9	Participación de las exportaciones bananeras en el PIB del Ecuador (1927-1947)	25
Figura 1.10	Producto Interno Bruto del Ecuador (1948-1965)	27
Figura 1.11	Exportaciones de banano en el Ecuador (1948-1964)	30
Figura 1.12	Participación de las exportaciones bananeras en el PIB del Ecuador (1948-1964)	32
Figura 1.13	Exportaciones de banano en el Ecuador (1965-1976)	38
Figura 1.14	Evolución de la extensión de las áreas sembradas con banano en el Ecuador (1965-1976)	42
Figura 1.15	Exportaciones de banano en el Ecuador (1977-1998)	46

Figura 1.16	Superficie cultivada con banano Cavendish en el Ecuador (1980-1997)	48
Figura 1.17	Porcentajes de las principales exportadoras: 1998	52
Figura 1.18	Nivel de tecnificación de los cultivos en 1999	60
Figura 1.19	Producción de banano en el Ecuador en 1999	62
Figura 1.20	Desperdicios de banano en el Ecuador	63
Figura 1.21	Exportaciones de elaborados de banano en el Ecuador (1983-1996)	66
Figura 2.1	Función empírica de una muestra	71
Figura 2.2	Histograma de frecuencias correspondiente a la producción bananera del Ecuador	72
Figura 2.3	Ojiva	73
Figura 2.4	Diagrama de cajas correspondiente a la superficie cultivada con banano	74
Figura 2.5	Gráfica correspondiente a la diferencia D	88
Figura 2.6	Valores propios (eigenvalores)	101
Figura 2.7	Proceso (simulado) media móvil de orden 1, $Z_t = U_t - 0.2 U_{t-1}$	110
Figura 2.8	Proceso (simulado) autoregresivo de orden 1, $Z_t = U_t - 0.8 Z_{t-1} + U_t$	111
Figura 3.1	Exportaciones de banano en el Ecuador en miles de TM. (1969-1998)	138
Figura 3.2	Ojiva y diagrama de cajas de las exportaciones bananeras en TM.	139
Figura 3.3	Histograma de frecuencias de las exportaciones bananeras en TM.	140
Figura 3.4	Exportaciones de banano en el Ecuador en miles de dólares FOB. (1969-1998)	141

Figura 3.5	Ojiva y diagrama de cajas de las exportaciones bananeras en miles de dólares FOB.	142
Figura 3.6	Histograma de frecuencias de las exportaciones bananeras en miles de dólares FOB	143
Figura 3.7	Producción de banano en el Ecuador en miles de TM. (1969-1998)	145
Figura 3.8	Ojiva y diagrama de cajas de la producción bananera en TM.	146
Figura 3.9	Histograma de frecuencias de la producción bananera en TM.	147
Figura 3.10	Participación de las exportaciones bananeras en el PIB. (1969-1998)	148
Figura 3.11	Ojiva y diagrama de cajas de la participación de las exportaciones bananeras en el PIB	149
Figura 3.12	Histograma de frecuencias de la participación de las exportaciones bananeras en el PIB	150
Figura 3.13	Participación de las exportaciones bananeras en las exportaciones totales del Ecuador(1969-1998)	151
Figura 3.14	Ojiva y diagrama de cajas de la participación de las exportaciones bananeras en las exportaciones totales	152
Figura 3.15	Histograma de frecuencias de la participación de las exportaciones bananeras en las exportaciones totales	153
Figura 3.16	Índice de precios al consumidor (alimentos) (1969-1998)	154
Figura 3.17	Ojiva y diagrama de cajas del índice de precios al consumidor (alimentos)	155

Figura 3.18	Histograma de frecuencias del índice de precios al consumidor (alimentos)	156
Figura 3.19	Índice de precios al consumidor (1969-1998)	157
Figura 3.20	Ojiva y diagrama de cajas del índice de precios al consumidor	158
Figura 3.21	Histograma de frecuencias del índice de precios al consumidor	159
Figura 3.22	Cotización del dólar en el Ecuador (1969-1998)	160
Figura 3.23	Ojiva y diagrama de cajas de la cotización del dólar	161
Figura 3.24	Histograma de frecuencias de la cotización del dólar	163
Figura 3.25	Superficie cultivada con banano en el Ecuador (1969-1998)	164
Figura 3.26	Ojiva y diagrama de cajas de la superficie cultivada con banano	165
Figura 3.27	Histograma de frecuencias de la superficie cultivada con banano	166
Figura 3.28	Población económicamente en el Ecuador (1969-1998)	167
Figura 3.29	Ojiva y diagrama de cajas de la población económicamente activa	168
Figura 3.30	Histograma de frecuencias de la población económicamente activa	169
Figura 3.31	Ojiva de la cotización del dólar	170
Figura 3.32	Ojiva de la superficie cultivada	171
Figura 3.33	Ojiva de la población económicamente activa	172
Figura 3.34	Producción bananera en el Ecuador	176
Figura 3.35	Participación de las exportaciones bananeras en el PIB	177

Figura 3.36	Participación de las exportaciones bananeras en las exportaciones totales	177
Figura 3.37	Índice de precios al consumidor	179
Figura 3.38	Superficie cultivada con banano	181
Figura 3.39	Población económicamente activa	181
Figura 3.40	Valores propios	184
Figura 3.41	Valores propios	188
Figura 3.42	Gráfica de las cargas de los factores	191
Figura 4.1	Exportaciones bananeras del Ecuador en miles de TM.	196
Figura 4.2	Autocorrelaciones de las exportaciones bananeras (TM.)	196
Figura 4.3	Autocorrelaciones parciales de las exportaciones bananeras (TM.)	198
Figura 4.4	Exportaciones bananeras del Ecuador en miles de dólares FOB.	198
Figura 4.5	Autocorrelaciones de las exportaciones bananeras (dólares FOB)	200
Figura 4.6	Autocorrelaciones parciales de las exportaciones bananeras (dólares FOB)	200
Figura 4.7	Producción bananera del Ecuador en miles de TM.	202
Figura 4.8	Autocorrelaciones de la producción bananera (TM.)	202
Figura 4.9	Autocorrelaciones parciales de la producción bananera (TM.)	204
Figura 4.10	Participación de las exportaciones bananeras en el PIB	204
Figura 4.11	Autocorrelaciones de la participación de las exportaciones bananeras en el PIB	206



Figura 4.12	Autocorrelaciones parciales de la participación de las exportaciones bananeras en el PIB	206
Figura 4.13	Predicciones del modelo ARMA(1,0) para la participación de las exportaciones bananeras en el PIB	208
Figura 4.14	Cotización del dólar en el Ecuador	208
Figura 4.15	Autocorrelaciones de la cotización del dólar	210
Figura 4.16	Autocorrelaciones parciales de la cotización del dólar	210
Figura 4.17	Índice de precios al consumidor (alimentos)	212
Figura 4.18	Autocorrelaciones del índice de precios al consumidor (alimentos)	212
Figura 4.19	Autocorrelaciones parciales del índice de precios al consumidor (alimentos)	214
Figura 4.20	Población económicamente activa del Ecuador	214
Figura 4.21	Autocorrelaciones de la población económicamente activa	215
Figura 4.22	Autocorrelaciones parciales de la población económicamente activa	215
Figura 4.23	Gráfica de dispersión de las exportaciones bananeras contra el índice de precios al consumidor (alimentos)	218
Figura 4.24	Gráfica de la dispersión de los residuos contra los valores estimados del modelo 1	219
Figura 4.25	Gráfica de probabilidad normal del modelo 1	219

Figura 4.26	Gráfica de dispersión de las exportaciones bananeras contra la cotización del dólar	222
Figura 4.27	Gráfica de la dispersión de los residuos contra los valores estimados del modelo 2	223
Figura 4.28	Gráfica de probabilidad normal del modelo 2	223
Figura 4.29	Gráfica de la dispersión de los residuos contra los valores estimados del modelo 3	226
Figura 4.30	Gráfica de probabilidad normal del modelo 3	227
Figura 4.31	Gráfica de la dispersión de los residuos contra los valores estimados del modelo 4	230
Figura 4.32	Gráfica de probabilidad normal del modelo 4	230
Figura 4.33	Gráfica de la dispersión de los residuos contra los valores estimados del modelo 5	234
Figura 4.34	Gráfica de probabilidad normal del modelo 5	234
Figura 4.35	Gráfica de la dispersión de los residuos contra los valores estimados del modelo 6	238
Figura 4.36	Gráfica de probabilidad normal del modelo 6	238
Figura 4.37	Gráfica de la dispersión de los residuos contra los valores estimados del modelo 7	242
Figura 4.38	Gráfica de probabilidad normal del modelo 7	242

Figura 4.39	Gráfica de la dispersión de los residuos contra los valores estimados del modelo 8	246
Figura 4.40	Gráfica de probabilidad normal del modelo 8	246

INDICE DE CUADROS

	Pág.
Cuadro I	Exportaciones de banano y plátano en Ecuador (1927-1947) 20
Cuadro II	Producto Interno Bruto del Ecuador (1927-1947) 22
Cuadro III	Participación de las exportaciones bananeras en el PIB del Ecuador (1927-1947) 24
Cuadro IV	Producto Interno Bruto del Ecuador (1948-1965) 26
Cuadro V	Exportaciones de banano en el Ecuador (1948-1964) 29
Cuadro VI	Participación de las exportaciones bananeras en el PIB del Ecuador (1948-1964) 31
Cuadro VII	Exportaciones de banano en el Ecuador (1965-1976) 37
Cuadro VIII	Superficie de bananos Gross Michel y Cavendish en el Ecuador (1965-1976) 39
Cuadro IX	Evolución de la extensión de las áreas sembradas con banano en el Ecuador (1965-1976) 41
Cuadro X	Exportaciones de banano en el Ecuador (1977-1998) 45
Cuadro XI	Superficie del Ecuador cultivada con banano Cavendish (1980-1997) 47
Cuadro XII	Áreas productivas de banano afectadas por el fenómeno de "El Niño" 50
Cuadro XIII	Principales países compradores del banano ecuatoriano (1993-1998) 51
Cuadro XIV	Exportaciones bananeras en el Ecuador por tipo de caja (1993-1997) 55

Cuadro XV	Exportaciones de banano de Ecuador	58
Cuadro XVI	Nivel de tecnificación de los cultivos en el Ecuador (May-99)	60
Cuadro XVII	Producción de banano en el Ecuador en 1999	61
Cuadro XVIII	Desperdicios de banano (1983-1995)	62
Cuadro XIX	Exportación de elaborados de banano en el Ecuador (1982-1996)	65
Cuadro XX	$Z_t = U_t - 0.2 U_{t-1}$	109
Cuadro XXI	$Z_t = 0.8 Z_{t-1} + U_t$	111

INTRODUCCION

El objetivo de esta investigación es analizar algunas características de uno de los principales productos de exportación de nuestro país, como es el banano. Para el efecto se determinarán las características básicas del banano; su llegada al Ecuador y la evolución en el tiempo, de la exportación de este producto, desde el año 1927. En este trabajo, se aplicaron técnicas de regresión y series temporales en la elaboración de modelos de predicción para las exportaciones bananeras. Además utilizando otros métodos estadísticos univariados y multivariados, se determinará si existe dependencia entre las variables de interés. Dependiendo de los resultados obtenidos se establecerán las correspondientes conclusiones y recomendaciones.

CAPITULO 1

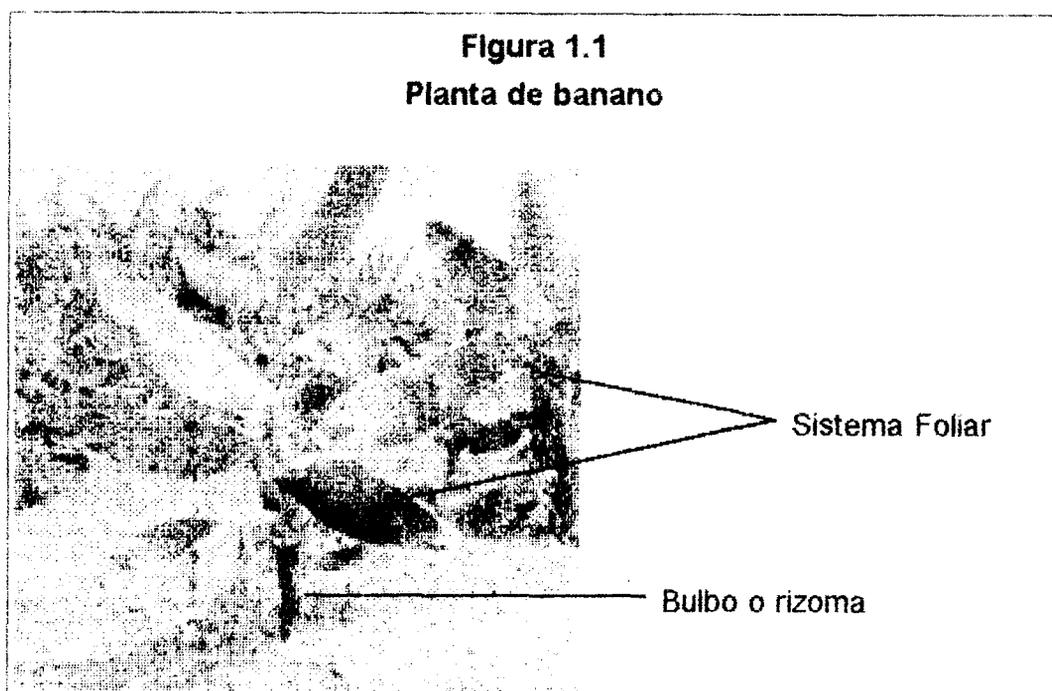
1. HISTORIA DE LA PRODUCCION BANANERA EN EL ECUADOR

1.1 Descripción y ecología del banano

Descripción

El banano es una planta musácea, que se originó a través de una serie de cambios genéticos a partir de especies silvestres no comestibles, de frutos pequeños y con numerosas semillas. Las transformaciones para llegar a la ausencia de semillas en el fruto (partenocarpia) y los cambios en el número de cromosomas, que determinan las características hereditarias, dieron origen, después de muchos siglos de evolución natural, al banano comestible y comercial. Las especies silvestres de banano son diploides (sus células poseen dos cromosomas) y las comerciales son triploides (sus células poseen tres cromosomas).

Para la producción comercial los triploides son más ventajosos, porque son más vigorosos, sus frutos son más grandes y carecen de semillas, razón por la que se propaga vegetativamente, es decir que no se reproduce sexualmente.



La planta de banano está conformada de diferentes partes a las cuales denominaremos órganos. Los órganos principales son: el rizoma o bulbo, que es el tallo subterráneo de crecimiento longitudinal, que se encuentra engrosado por las sustancias de reserva necesarias para la planta. En su interior se originan las raíces. El sistema radicular, que está conformado por las raíces del banano, las cuales se caracterizan por ser superficiales y distribuirse en el suelo,

en una capa de 30 a 40 cm. Se encuentran en mayor cantidad en la capa de 15 a 20 cm. El sistema foliar (hojas) se origina en la parte superior del bulbo o rizoma. Al principio la hoja emerge enrollada en forma de cigarro, una vez que ha salido las dos terceras partes de su longitud, comienza a desarrollarse primero del lado derecho. La presencia de la coloración verde se hace inmediatamente. La inflorescencia está constituida por el conjunto de ramificaciones florales de la planta. (Ver figura 1.1)

Variedades

Hay dos tipos de bananos comerciales : Gross Michel y Cavendish. La variedad Cavendish a su vez comprende las sub-variedades Lacatán o Filipino, Poyo o Valery, Gran Cavendish y Cavendish Enano.

- **Gross Michel:** Las plantas de esta variedad son las más grandes y robustas. El tallo es fuerte , tiene una longitud de (7 ± 1) m., y presenta una coloración verde claro con un tono rosa por alguna parte. La maduración es regular y homogénea. El sabor delicioso y la textura delicada por lo que se le conoce con el nombre de "guineo o plátano de seda". Es muy sensible a la enfermedad llamada "Mal de Panamá".

- **Cavendish:** La característica más importante y común es su resistencia a la enfermedad "Mal de Panamá".

El **Lacatán o Filipino** puede alcanzar alturas de (5 ± 1) m. Los racimos son alargados y los frutos son curvados. El fruto es muy sensible a los parásitos, después de cosechado. La maduración es más delicada y el aspecto final menos atractivo. El **Poyo (Robusta o Váleriy)**, es de mediana altura, (3.4 ± 0.6) m. Sus hojas son más cortas. Son de rápido desarrollo y se pueden sembrar en mayor densidad, lo que redundaría en mayor productividad. El racimo es más compacto y los frutos son largos y rectos. El **Gran Cavendish** conocido como Gran Enano, Grand Naine o Mons Mari, es más pequeño que el anterior, llega a una altura de (2.9 ± 0.1) m. Es una variedad que se caracteriza por su productividad. El **Cavendish Enano** es el más pequeño, su tallo llega a una altura de 2 m. Los frutos son curvos.

Ecología

El banano es una planta que se desarrolla en condiciones óptimas en las regiones tropicales, que son húmedas y cálidas. Presenta un crecimiento continuo, cuya inflorescencia aparece cuando se detiene la producción de hojas y raíces. Su velocidad de crecimiento es impresionante, y ese vigor vegetativo sólo puede darse bajo

condiciones ecológicas apropiadas. Los factores ecológicos que mayormente influyen en el rendimiento de la cosecha son:

Localización geográfica: Las condiciones climáticas óptimas para la producción de banano, se ubican entre una latitud de 15 grados norte y 15 grados sur del Ecuador. Entre los 15 y 23 grados norte y sur , el banano ocasionalmente sufre "quemaduras" por baja temperatura que afectan el crecimiento de la planta. Entre los 23 y 30 grados norte y sur, el banano se convierte en un cultivo estacional, es decir, que se cosecha únicamente en los meses calientes.

Altitud del suelo: Las variaciones en altitud modifican en forma muy notoria los hábitos de crecimiento en las plantas del banano. A mayor nivel de altitud, mayor prolongación del ciclo biológico. La zona de altitud entre 0 y 300 m. sobre el nivel del mar, donde la temperatura no es limitante en el desarrollo del banano, el régimen de lluvias es el que determina la producción, bajo condiciones de suelos y cultivo apropiados.

Lluvia y humedad: La planta de banano, por su estructura botánica, requiere de una gran disponibilidad de humedad permanente en los suelos. Para la obtención de cosechas económicamente rentables, se considera suficiente suministrar de 100 a 180 mm. de agua por mes

para cumplir con los requerimientos de la planta. Es necesario regar el suelo cuando las lluvias no son regulares.

Temperatura: La temperatura tiene un efecto preponderante en el desarrollo y crecimiento de la planta de banano. El clima ideal para el cultivo del banano es Tropical húmedo. La temperatura mínima es de 18 grados centígrados y su máxima es de 34.8. Exposiciones a temperaturas mayores o menores causan deterioro y lentitud en el desarrollo, además de daños en la fruta. Existe un aletargamiento en el crecimiento a medio día como consecuencia de un déficit hídrico pasajero que acompaña a las altas temperaturas; en estos casos el crecimiento nocturno podría ser mayor al diurno en el momento que la planta se hidrata.

Movimiento de aire, vientos: El sistema foliar de la planta de banano soporta vientos cuyas velocidades sean de (25 ± 5) Km./h . Si los vientos tienen velocidades de (47.5 ± 7.5) Km./h , los daños provocados en las plantas son menores, mientras que si las velocidades superan los 55 Km./h, la destrucción ocasionada es total.

Luminosidad: La fuente de energía que utilizan las plantas verdes es la radiación solar. La duración del día influye en el ciclo de desarrollo. Si la luminosidad es baja, el ciclo de desarrollo se alarga. En las

zonas de alta luminosidad, la duración del ciclo de desarrollo es de (85 ± 5) días, en las zonas de luminosidad media es de (95 ± 5) días y en las zonas de baja luminosidad es de (98.5 ± 13.5) días.

Suelos: Para escoger un suelo bananero, se requiere conocer la variedad a plantar, su potencial de productividad y el nivel de tecnología a emplear. Los suelos más aptos para el desarrollo comercial del cultivo del banano son los que ofrecen a las raíces el máximo de posibilidades en profundidad y, por lo tanto, de disponer de mayores reservas de agua y elementos minerales. En general, se consideran mejores aquellos que presenten textura franco, franco arenosa, franco arcilloso, franco arcillo limoso y franco limoso. No son apropiados los suelos que tienen más del 40% de arcilla o que tienen alto grado de compactación. La profundidad debe ser de (1.35 ± 0.15) m. como mínimo. Debe poseer un buen drenaje interno, alta fertilidad y óptimas propiedades de retención del agua; los suelos arcillosos y compactos, si bien retienen la humedad, en cambio, no facilitan el drenaje ni el desarrollo de las raíces y los arenosos se secan pronto y las plantas sufren por falta de agua. Los suelos bananeros deben ser ricos en potasio, calcio y magnesio y tener un pH de 6.5.

Riego: La planta de banano demanda elevados requisitos de agua. De acuerdo con la forma en que se proporcione el agua, los riegos en

banano pueden ser por: gravedad o por aspersión . El riego por gravedad es aplicable en plantaciones que se desarrollan en suelos de buena permeabilidad y cuando se cuenta con un caudal de agua de bajo costo, pero tiene el inconveniente de producir alta erosión del suelo. El riego por aspersión tiene algunas ventajas: menor erosión, ahorro de agua, alto rendimiento de producción, permite su aplicación conjunta con productos químicos y puede ser instalado en terrenos irregulares.

Drenaje: El drenaje permite eliminar el exceso de agua en el suelo, provocado por la lluvia, el riego y el desbordamiento de los ríos. La planta de banano es muy susceptible a los excesos y fluctuaciones de agua, por lo que el mantenimiento de un buen sistema de drenaje es absolutamente necesario para obtener un fruto de excelente calidad y por ende una cosecha de gran rendimiento.

Manejo del banano desde la cosecha hasta el empaque

Previo a la industrialización, el banano es sometido a un proceso post-cosecha. El proceso que se aplica en la cosecha, transporte y empaque del banano presenta muchas diferencias de una plantación a otra, en una hacienda tecnificada el procedimiento que se sigue es el siguiente:

Para el corte y manejo del racimo dentro de la plantación se conforma una cuadrilla de cosecha conformada por el cortador, el ayudante y los cargadores; el cortador estimará el número de racimos listos para cortar y quitará todas las hojas que puedan dificultar el corte del racimo, el cargador pondrá el soporte en forma de cuna contra el racimo; con la cuna en posición, el cortador pica el tallo de la planta en forma que caiga suavemente el racimo dentro de la cuna y esta sobre el hombro del cargador. De allí se procede a transportar los racimos hacia la empacadora, mediante funiculares. (Ver figura 1.2).

Figura 1.2

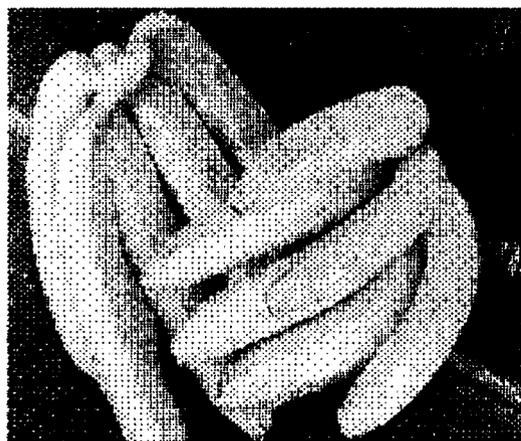
Transportación de los racimos de banano



Ya receptada la fruta en la empacadora se procederá a revisarla y se verificará el cumplimiento de los requisitos de aceptación.

rechazándose todos los racimos que estén pasados o faltos de maduración, bananos cortos, estropeados, dañados por insectos, hojas o animales en todas las manos (Ver figura 1.3).

Figura 1.3
Manos de banano



La fase siguiente es la desmanada (Separación o desprendimiento de los bananos del racimo), para la cual se requiere a tres personas: una que desprende las manos de banano del racimo, otra que los pasa al tanque y la tercera que ayuda a enganchar los racimos, quitar los tallos y piolas y mantener limpia el área (Ver figura 1.4). Para la limpieza y saneo de las manos de banano se emplea un tanque dividido en dos partes, con agua corriente y rebose en cada lado, de modo que las manos de banano dejados por los desmanadores queden en el lado junto a ellos y aquí será donde trabajarán los

lavadores y saneadores, los cuales tendrán que revisar mano por mano antes de ser depositados en la otra parte del tanque, de donde deberán ser tomados por los pesadores.

Figura 1.4
Desmanada de los bananos



Los pesadores son los responsables de seleccionar las manos que se empacarán, y de garantizar el peso exacto de la fruta que entra en las cajas, así como de la cantidad de manos que entran por caja (Ver figura 1.5). Los bananos que recibirán el sello serán los de la parte inferior de la mano, comenzando por el segundo dedo (banano) del lado izquierdo, y se continuará pasando uno hasta el final de la mano,

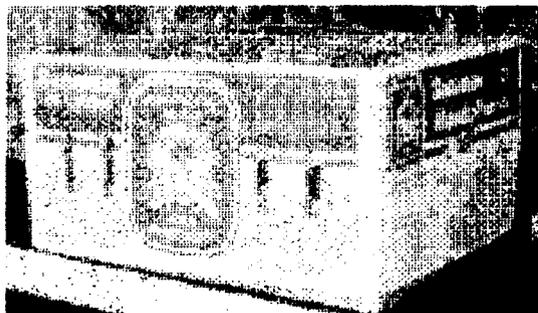
no importa si el último sello queda en el último dedo del lado derecho; la colocación debe ser restringida al centro de la mano.

Figura 1.5
Manos seleccionadas para ser empacadas



La fruta se empacará de acuerdo a la cantidad de manos que se tenga. Si al terminar de empacar una caja se nota que las manos sobrepasan el nivel de la caja, será necesario volverla a empacar porque quiere decir que no se han seguido las instrucciones correspondientes. La marcada será hecha sellando sobre el fondo y la tapa; el sello contendrá indicaciones de la fecha, la hacienda y la empacadora (Ver figura 1.6).

Figura 1.6
Caja de banano



1.2 Origen y llegada del banano al Ecuador

El banano es considerado como el fruto más antiguo del mundo, pues su planta data de tiempos prehistóricos. El Sureste Asiático se considera el lugar de origen de los bananos, su cultivo se desarrolló simultáneamente en Malasia y en las Islas Indonesias. Sin embargo, el origen exacto no es completamente claro. Según el antropólogo Herbert Spiden, el banano probablemente es oriundo de las húmedas regiones tropicales del sureste de Asia, incluyendo el noreste de la India, Birmania, Camboya y partes del Sur de China , así como las Islas Mayores de Sumatra, Java, Borneo, las Filipinas y Taiwán. En esos lugares, las variedades sin semilla del verdadero banano de consumo doméstico, se encuentran en estado silvestre.

Desde épocas remotas las hojas del banano se han utilizado como envoltura o como fuente de fibra, y la fruta como alimento.

Antiguamente, gran porcentaje de las variedades conocidas tenían una alta proporción de semillas. Con los años, al cultivarse el banano se obtuvieron formas sin semilla. El banano tan solo era conocido en Asia hasta que los mercaderes árabes lo introdujeron en Africa durante sus expediciones, en las cuales comerciaban y obtenían esclavos.

Los cultivos de Africa Occidental ya estaban establecidos en el siglo VI cuando llegaron los europeos. Desde el Africa Occidental fue llevado a las Islas Canarias por los portugueses, en la primera parte del siglo XIV, y desde allí a Santo Domingo en 1516 por Fray Tomás de Berlanga, de donde se propagó a otras islas y de allí al continente a través de México, Honduras y Panamá. Desde Panamá llegó el cultivo a Ecuador, pero no existen datos que ayuden a precisar la época de introducción del banano al país.

Para describir la evolución de las exportaciones bananeras ecuatorianas se determinarán tres etapas: antes (1927-1947), durante (1948-1964) y después del Boom Bananero (desde 1965). En esta sección se analizará la primera etapa, es decir, todo lo que aconteció antes de que nuestro país se convirtiera en uno de los principales países exportadores de banano en el mundo.

El Ecuador ha mantenido un ritmo creciente en sus exportaciones desde comienzos del presente siglo (1927), superando duras crisis. La primera crisis en la exportación bananera ecuatoriana se presenta durante la Primera Guerra Mundial que motivó un notable descenso en las exportaciones bananeras porque a diferencia de 1914 en que 77.047 racimos fueron exportados, en 1919 tan solo se exportaron 39.562 racimos. La drástica caída en las exportaciones bananeras se recuperó en 1920 con la exportación de 112.973 racimos. A comienzos de la década de 1930, la gigantesca United Fruit Company instaló una división bananera en la parroquia Tenguel, provincia del Guayas.

United Fruit: El 30 de marzo de 1899 en el estado de New Jersey, la Boston Fruit Company y la United Fruit Company (UFCO) se fusionaron para formar la compañía más grande de bananas en el mundo, con plantaciones en Colombia, Costa Rica, Cuba, Jamaica, Nicaragua, Panamá y Santo Domingo. Dicha compañía compraría fincas de banano para cultivarlas, siendo la responsable de convertir al banano en una fruta de consumo mundial. En 1901, ingresó en Guatemala, país que se transformó en su base de operaciones. United Fruit Company cambió más tarde su nombre a United Brands y entró en dificultades financieras durante el año 1970. Las tierras de UFCO fueron compradas por la Corporation Del Monte. La United Brands suspendió sus operaciones en Ecuador a partir de 1965,

desde entonces compra banano en cantidades limitadas a intermediarios como la Exportadora Bananera Noboa.

En el Cuadro I constan los ingresos en miles de dólares generados por las exportaciones bananeras realizadas entre los años 1927 y 1947. En la figura 1.7 se presenta la gráfica de los ingresos por exportaciones. Se pueden notar las fluctuaciones de los ingresos causadas por la depresión de los años 32, consecuencia de la crisis económica generada, entre otras razones, por la Primera Guerra Mundial, en que las exportaciones bananeras ecuatorianas registran sus más bajos volúmenes; durante 1932 se exportaron 324.387 racimos equivalentes a 40.000 dólares.

Superada la crisis (1931-1933) para 1934 las exportaciones bananeras se recuperan a 1'452.230 racimos que generaron ingresos de 258.000 dólares, fecha a partir de la cual se incrementan las exportaciones hasta su nueva caída generada por la Segunda Guerra Mundial entre los años 1941 y 1945.

Dentro de este período (1941-1945), es en 1944 que las exportaciones bananeras nacionales tuvieron su punto más bajo al registrar un volumen de ventas de 540.974 racimos que significó 135.000 dólares. A partir de 1946 las exportaciones bananeras se

recuperan registrando una exportación de 1'301.409, con ingresos de alrededor de 551.000 dólares. Antes de 1948 el banano no era considerado como uno de los principales productos de exportación, tales como el cacao y el café, ya que en 1947 sólo representaba el 4% del total de las exportaciones nacionales.

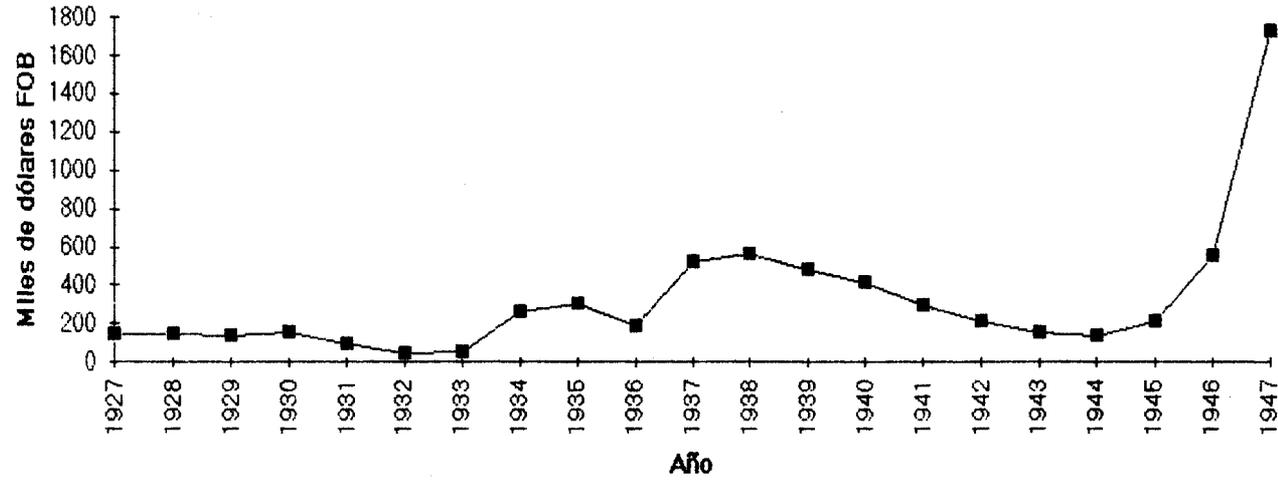
Cuadro I
EXPORTACIONES DE BANANO Y PLATANO EN ECUADOR
1927-1947

Año	Miles de dólares FOB	Año	Miles de dólares FOB	Año	Miles de dólares FOB
1927	144	1934	258	1941	291
1928	143	1935	303	1942	208
1929	133	1936	186	1943	154
1930	154	1937	517	1944	135
1931	91	1938	563	1945	211
1932	40	1939	483	1946	551
1933	55	1940	413	1947	1730

Fuente: El Banano en el Ecuador, Larrea M. Carlos, Corporación Editora Nacional, 1987

En el cuadro II consta los valores del Producto Interno Bruto registrados entre los años 1927 y 1947. Se puede distinguir en la figura 1.8 dos etapas de crecimiento: en la primera etapa comprendida entre 1927 y 1939, el PIB registró un incremento del 43% (79 a 113 millones de dólares), mientras que en la segunda

Figura 1.7
Ingreso por Exportaciones de Banano en el Ecuador (1927-1947)



Fuente: El Banano en el Ecuador, Larrea Carlos, Corporación Editora Nacional, 1987

etapa comprendida entre 1940 y 1947, el PIB registró un incremento del 143% (148 a 360 millones de dólares).

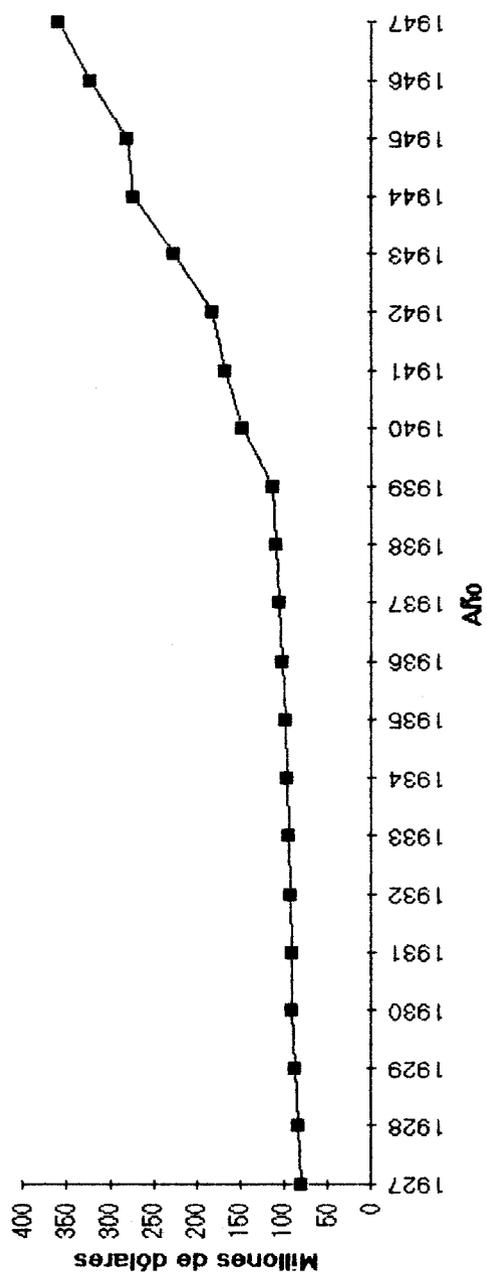
Cuadro II
PRODUCTO INTERNO BRUTO DEL ECUADOR
1927-1947

Año	Millones de dólares	Año	Millones de dólares	Año	Millones de dólares
1927	79	1934	96	1941	168
1928	83	1935	98	1942	183
1929	87	1936	102	1943	228
1930	91	1937	105	1944	273
1931	91	1938	109	1945	280
1932	92	1939	113	1946	323
1933	94	1940	148	1947	360

Fuente: Setenta años de Información Estadística 1927-1997, Banco Central del Ecuador, 1997

En el cuadro III constan los porcentajes de la participación de las exportaciones bananeras en el PIB entre los años 1927 y 1947. En la figura 1.9 se observa que el porcentaje de participación tuvo un importante descenso en el año de 1932 (depresión económica) registrando un porcentaje de 0.04. En 1938 se da el porcentaje de participación más elevado del período 1927-1947, el cual fue de 0.52. Otra etapa de decrecimiento del porcentaje de participación ocurrió entre los años 1943 y 1945 (Segunda Guerra Mundial), siendo 1944

Figura 1.8
Producto Interno Bruto del Ecuador (1927-1947)



Fuente: Setenta Años de Información Estadística, BCE, 1997

el año más crítico, con un porcentaje de 0.05. Ya en 1947 se incrementa el porcentaje de participación a 0.48.

Cuadro III
PARTICIPACION DE LAS EXPORTACIONES BANANERAS EN EL PIB
DEL ECUADOR
1927-1947

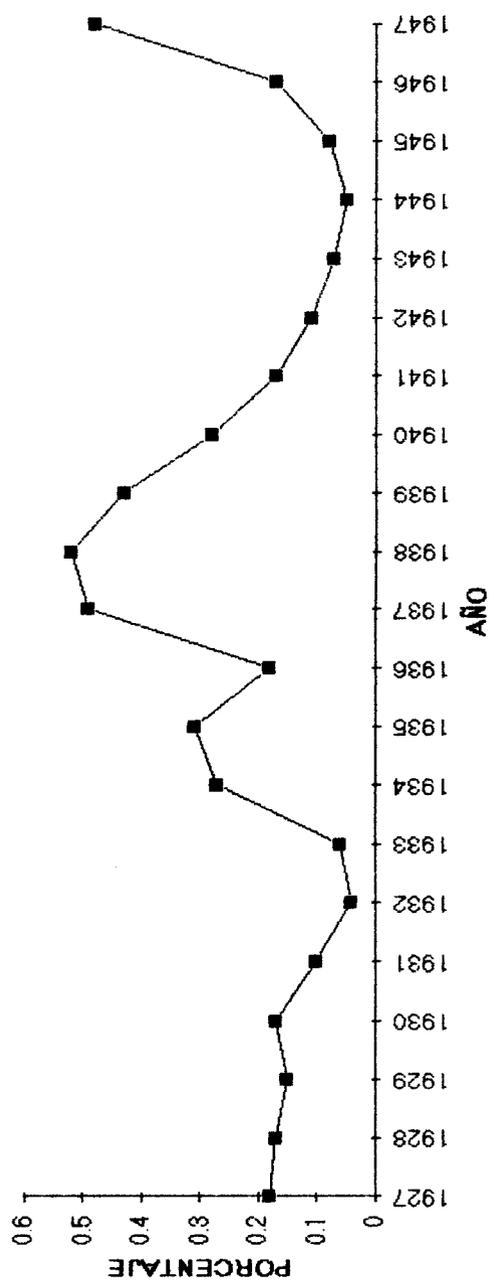
Año	Porcentaje	Año	Porcentaje	Año	Porcentaje
1927	0.18	1934	0.27	1941	0.17
1928	0.17	1935	0.31	1942	0.11
1929	0.15	1936	0.18	1943	0.07
1930	0.17	1937	0.49	1944	0.05
1931	0.1	1938	0.52	1945	0.08
1932	0.04	1939	0.43	1946	0.17
1933	0.06	1940	0.28	1947	0.48

Fuente. Setenta años de Información Estadística 1927-1997, Banco Central del Ecuador, 1997

1.3 Boom Bananero en el Ecuador

En esta sección analizaremos la evolución de las exportaciones bananeras ecuatorianas en la segunda y la tercera etapa, es decir, durante y después del Auge Bananero respectivamente. El Auge Bananero está comprendido entre los años 1948 y 1965. En este periodo se exportaba petróleo crudo, banano y plátano, café y elaborados, cacao y elaborados y camarón. El régimen de Galo Plaza (1948-1952) se preocupó de atraer la inversión extranjera para la agricultura, dando lugar a la "fiebre del banano". En 1948 la

Figura 1.9
Participación de las Exportaciones Bananeras en el PIB del Ecuador (1927-1947)



Fuente: Setenta Años de Información Estadística, BCE, 1997

cotización del dólar era de 13,5 sucres, mientras que en 1965 fue de 18,18 sucres. La población ecuatoriana aumentó de 3'270.403 a 5'194.239 habitantes.

En el cuadro IV consta el PIB correspondiente al período 1948-1965. En la figura 1.10 se puede observar que con el auge bananero, el Ecuador entró en un largo proceso de crecimiento sostenido: el PIB ascendió un 50.43% entre 1948 y 1954 y un 46.46% entre 1954 y 1965.

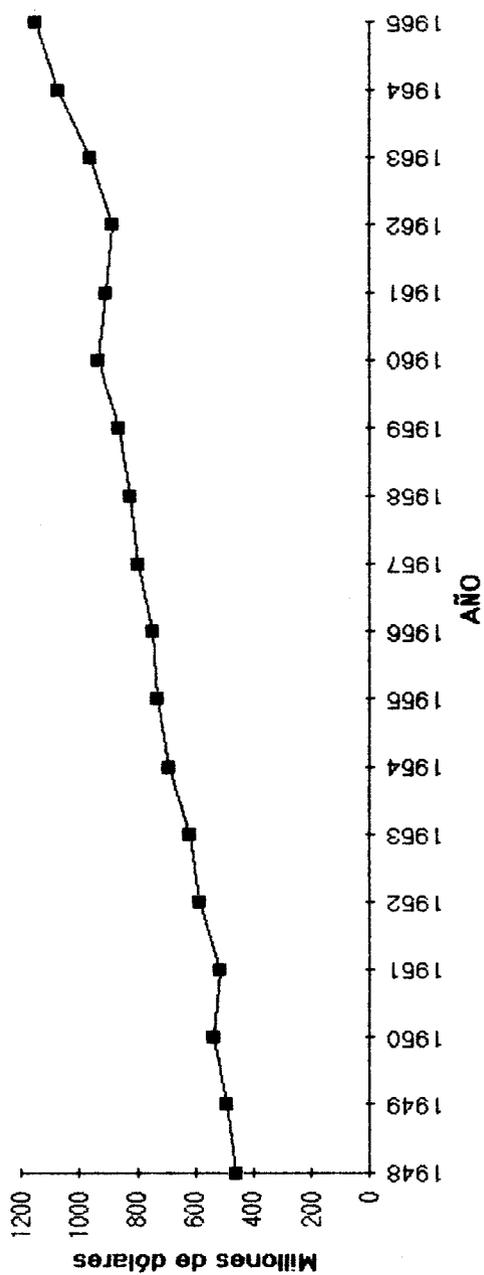
**Cuadro IV
PRODUCTO INTERNO BRUTO DEL ECUADOR
1948-1965**

Año	Millones de dólares	Año	Millones de dólares	Año	Millones de dólares
1948	460	1954	692	1960	938
1949	492	1955	734	1961	907
1950	535	1956	749	1962	888
1951	517	1957	799	1963	966
1952	588	1958	826	1964	1075
1953	623	1959	864	1965	1151

Fuente. Setenta años de Información Estadística 1927-1997, Banco Central del Ecuador, 1997

El ingreso del Ecuador como proveedor importante del mercado mundial (es el primer abastecedor mundial desde 1952), se produce

Figura 1.10
Producto Interno Bruto PIB del Ecuador (1948-1965)



Fuente: Setenta Años de Información Estadística, BCE, 1997

en un mercado establecido, con proveedores y consumidores definidos desde los inicios de las exportación de banano, a fines del siglo XIX. Antes de la Segunda Guerra Mundial, el mercado libre estuvo abastecido sobre todo por los grandes enclaves de la United Fruit y de la Standard Fruit, ubicados en Costa Rica, Guatemala, Honduras y Panamá, y un tercio del total provenía de un grupo de países en el cual las transnacionales no habían realizado grandes inversiones (principalmente México, Nicaragua, Cuba, República Dominicana, Haití y Belice). A partir de 1948, las exportaciones ecuatorianas reemplazaron a las de este último grupo.

En el cuadro V constan los datos sobre las exportaciones bananeras del período 1948-1964. En la figura 1.11 se muestra el rápido crecimiento de las exportaciones desde 1948 hasta el año 1952. A partir de 1953 el crecimiento es bastante lento.

En los primeros años del auge bananero, el principal mercado fue el norteamericano. Su participación superó, casi todos los años, el 50% del total y bordeó el 80% hasta 1954. Las ventas a Europa Occidental, especialmente a Alemania, cubrieron casi toda la diferencia. A partir de 1952 nuestro país toma el liderazgo de primer exportador de banano en el mundo. En 1955 se exportaron 23'874.310 racimos que significaron ingresos de 36'761.000 dólares. En 1964 el Ecuador abastecía el 25% de la oferta internacional,

superando levemente a los cuatro exportadores centroamericanos juntos.

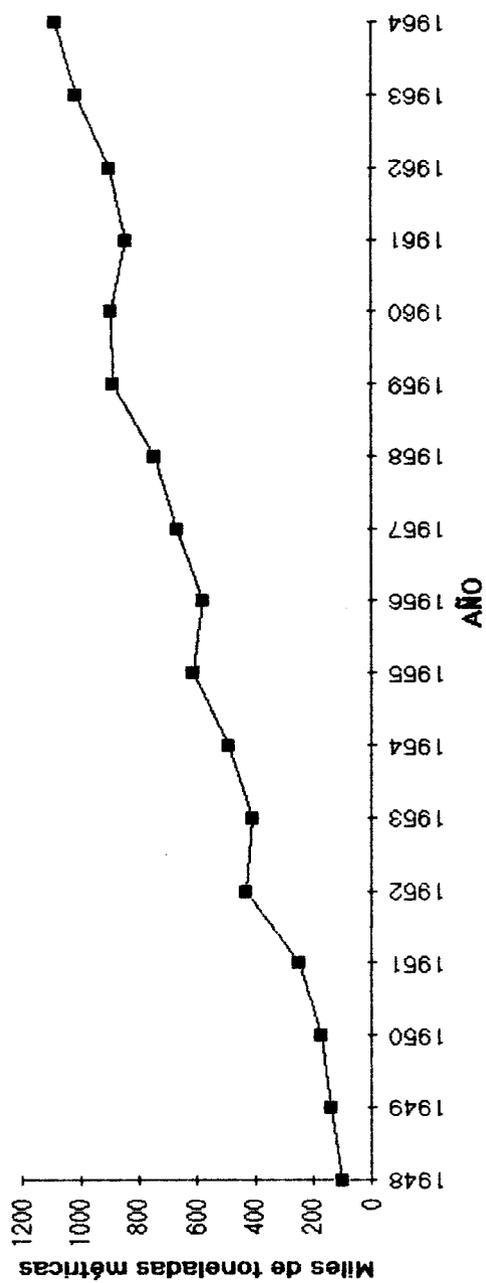
Cuadro V
EXPORTACIONES DE BANANO EN EL ECUADOR
1948-1964

Año	Miles de TM.
1948	99.6
1949	138
1950	169.6
1951	246.5
1952	429.8
1953	406.4
1954	492.2
1955	612.6
1956	578.9
1957	669.1
1958	742.7
1959	885.6
1960	895.1
1961	842.3
1962	897.8
1963	1014.3
1964	1086.8

Fuente: El Banano en el Ecuador, Larrea M. Carlos, Corporación Editora Nacional, 1967

En el cuadro VI constan los porcentajes de participación de las exportaciones bananeras en el PIB. En la figura 1.12 se nota el rápido crecimiento del porcentaje de participación entre los años 1948

Figura 1.11
Exportaciones de Banano en el Ecuador (1948-1964)



Fuente: El Banano en el Ecuador, Larrea Carlos, Corporación Editora Nacional, 1967

y 1964. A partir de 1956 este porcentaje comienza a decrecer lentamente, presentando una notable recuperación (de 4.10 a 7.02%) en 1963.

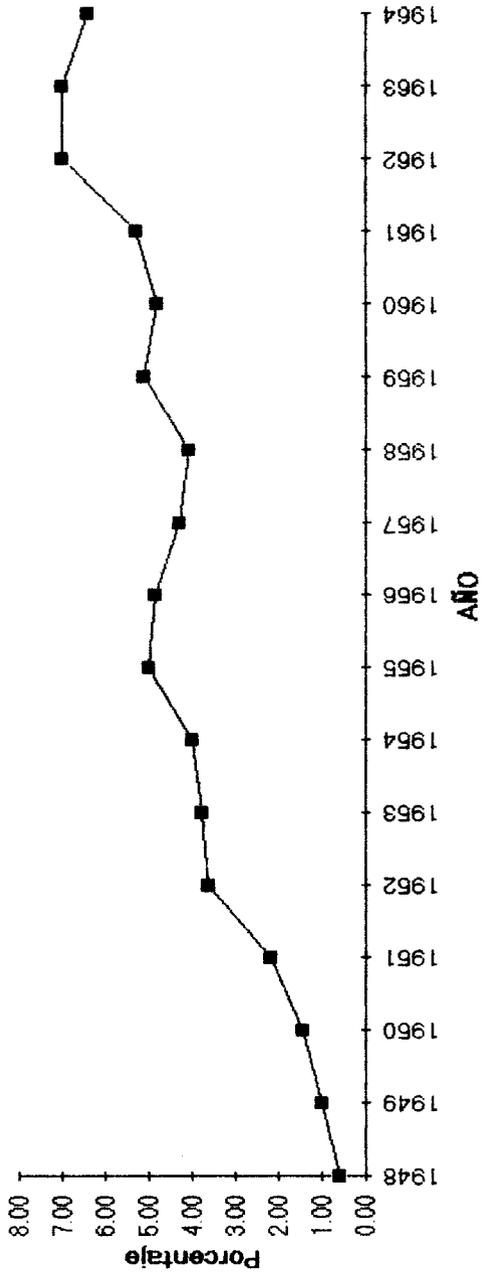
Cuadro VI
PARTICIPACION DE LAS EXPORTACIONES BANANERAS EN EL PIB
DEL ECUADOR
1948-1964

Año	Porcentaje
1948	0.6
1949	0.4
1950	1.47
1951	2.18
1952	3.64
1953	3.8
1954	4
1955	5.01
1956	4.87
1957	4.32
1958	4.1
1959	5.12
1960	4.83
1961	5.29
1962	7.02
1963	7.02
1964	6.41

Fuente: Setenta años de Información Estadística 1927-1997, Banco Central del Ecuador, 1997

Distintas empresas o grupos participaron en el mercado, sin que existiera un fuerte predominio de ninguno. La empresa más importante de esta etapa, United Fruit, no superó el 25% de las

Figura 1.12
Participación de las Exportaciones Bananeras en el PIB del Ecuador (1948-1964)



Fuente: Setenta Años de Información Estadística, BCE, 1997

exportaciones. La segunda empresa en importancia fue la Fruit Trading Corporation, que estableció un enclave en la provincia de Esmeraldas; alcanzó, en su mejor momento, entre 1952 y 1955, aproximadamente el 30% del total, el mismo que declinó persistentemente hasta la desaparición de la empresa en 1965. El grupo EFE-UBESA, conformado por empresas alemanas abriendo el mercado europeo, mantuvo una participación similar. Las restantes firmas tuvieron una participación menor.

En esta etapa, ciertas empresas extranjeras hicieron algunos ensayos para participar en la producción o constituir enclaves, pero sus intentos fracasaron. El caso más importante es el de la United Fruit, que en 1933 compró la hacienda de Tenguel y, años después, era propietaria aproximadamente de una docena de predios que totalizaron 77.700 ha.; el mayor de ellos tuvo una superficie de 29000 ha. La empresa exportaba su propia producción y la fruta comprada a productores independientes que representaban el mayor volumen de sus exportaciones.

Hacia fines de los años 50, la compañía enfrentó una serie de problemas financieros, laborales, políticos y técnicos: el rendimiento sobre el capital se redujo de un 13.1% en 1953 a solo un 0.8% en 1960. Tuvo conflictos serios con los campesinos que culminaron en la invasión de Tenguel en 1962; también surgieron dificultades con el

gobierno ecuatoriano y aparecieron plagas en las plantaciones (Mal de Panamá). Todo ello obligó a la empresa a deshacerse de sus propiedades por venta o entrega a los campesinos.

La Standard Fruit no intervino directamente en la producción y su participación, aunque pequeña fue incrementándose, presentando un comportamiento regular y estable. Se consolidó la Exportadora Bananera Noboa, que comenzó a operar como exportadora independiente en 1956. En el APENDICE A se puede observar que la Exportadora Bananera Noboa ha incrementado año tras año, su participación en las exportaciones bananeras, de 9.09 en 1959 a 15.83% en 1964. Entre los años 1959 y 1962, la United Fruit es la exportadora con mayor porcentaje de participación, mientras que en 1963 y 1964 el líder es el Grupo UBESA, debido a que la United Fruit había decrementado su porcentaje de participación a medida que pasaban los años; en 1964 la participación de la United Fruit tan solo fue de 13.98%.

Las exportaciones crecieron lentamente entre 1955 y 1960. Tanto el presidente Velasco Ibarra así como el presidente Carlos Julio Arosemena trataron de establecer planes para acrecentar las exportaciones pero fracasaron. En 1963 toma el poder una junta militar, pretendió establecer reformas agrarias, administrativas, financieras y educativas, pero fundamentalmente nunca llegaron y en

1966 termina su mandato. En 1972 con el derrocamiento del presidente Velasco Ibarra, se inicia la "era petrolera" (Boom Petrolero). La población ecuatoriana aumentó de 5'194.239 a 7'070.954 habitantes.

Durante el auge bananero (1948-1964), las exportaciones bananeras se incrementaron rápidamente, pero desde 1965 hasta 1976, ocurrieron algunas crisis causadas por: el retiro de las compañías extranjeras, la eliminación de las plagas en las plantaciones de la United Fruit en América Central y el lanzamiento mundial de la variedad "chiquita banana"(Cavendish), desplazando a la variedad ecuatoriana (Gross Michel).

A partir de 1965 las exportaciones ecuatorianas comenzaron a ser desplazadas tanto de los Estados Unidos de norteamérica como de Europa Occidental. Las ventas del Ecuador a Estados Unidos se redujeron a la mitad entre 1964 y 1973, mientras que en el caso europeo la declinación fue menor. Los nuevos mercados permitieron apenas compensar esta caída. Y empeoró la situación, ya que entre 1974 y 1975, hubo sobreoferta de banano y abandonos de áreas por la presencia de la sigatoka negra, que es un hongo microscópico que ataca únicamente las hojas de la planta del banano.

En el cuadro VII constan las exportaciones bananeras en toneladas métricas en el período 1965-1976. En la figura 1.13 se puede observar la influencia que tuvieron el Mal de Panamá (años 1966-1969) y la Sigatoka Negra (1974-1975). Hubo un gran descenso entre los años 1964 y 1965 de 1086.8 a 874.6 miles T.M., desde 1966 hasta 1968 hay una ligera recuperación, gracias a la introducción de la variedad Cavendish. En 1972 se inicia el Boom Petrolero, al año siguiente se produce un estrepitoso descenso de las exportaciones bananeras (948.5 miles de T.M), debido a la plaga de la Sigatoka Negra, que provocó el abandono de las plantaciones, hubo otro descenso en 1976 (937.3 miles T.M), ya que no era rentable la producción de banano, porque la enfermedad impedía un proceso de cosecha normal y por ende la obtención de frutos era irregular.

Las exportaciones provenientes de Costa Rica, Honduras, Guatemala y Panamá iniciaron una acelerada expansión que casi triplicó su volumen hasta 1971. A partir de ese mismo año, las transnacionales bananeras impulsaron también la exportación bananera de Filipinas hacia Japón desplazando al Ecuador. Todo esto se motivó por el cambio de la estructura de costos a nivel internacional, que se derivó de la introducción masiva de la variedad *Cavendish*. Los costos de la United Fruit en Honduras bajaron de un promedio de 2.54 dólares por caja para el intervalo 1960-1964 a 1.4 para el período 1965-1970.

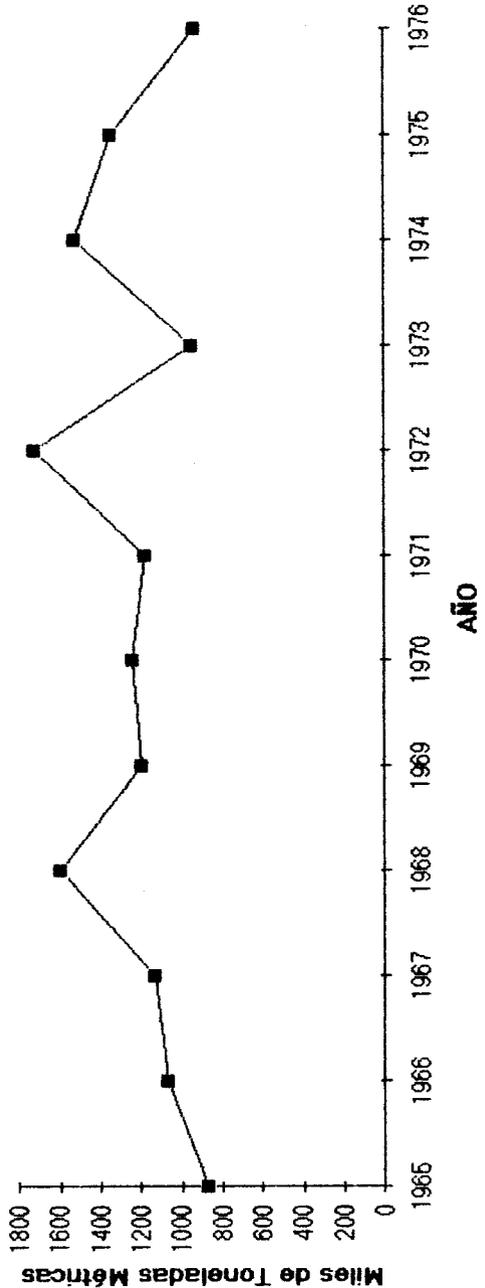
Cuadro VII
EXPORTACIONES DE BANANO EN EL ECUADOR
1965-1976

Año	Miles de TM.
1964	1086.8
1965	874.6
1966	1070.7
1967	1131.8
1968	1597.7
1969	1198.6
1970	1246.3
1971	1179.7
1972	1726.1
1973	948.5
1974	1525.1
1975	1348.5
1976	937.3

Fuente: El Banano en el Ecuador, Larrea M. Carlos, Corporación Editora Nacional, 1987

También las ventajas comparativas del Ecuador provenientes de sus condiciones naturales perdieron importancia: el menor desarrollo del **mal de Panamá** en el Ecuador desapareció como ventaja por la invulnerabilidad del *Cavendish*, y la ausencia de ciclones en Ecuador, perdió significación por la mayor resistencia de la nueva variedad. Desde 1967 los productores ecuatorianos sembraron y comenzaron a exportar banano *Cavendish* (ver cuadro VIII), al parecer introducido por su propia iniciativa; más tarde el Estado estimuló el cambio de variedad, que hacia 1976 se había operado casi por completo, pero ello no revirtió la relación de costos comparativos.

Figura 1.13
Exportaciones de Banano en el Ecuador (1965-1976)



Fuente: El Banano en el Ecuador, Larrea Carlos, Corporación Editora Nacional, 1987

Cuadro VIII
SUPERFICIE DE BANANOS GROSS MICHEL Y CAVENDISH
EN EL ECUADOR
1965-1976

Año	G. Michel (has.)	%	Cavendish (has.)	%
1965	144088	100	—	—
1966	163772	100	—	—
1967	146233	91.3	13995	9.7
1968	131898	84.1	24978	15.9
1969	108280	73.4	32249	26.6
1970	74045	59.6	50200	40.4
1971	59665	52.1	54921	47.9
1972	47067	46.6	53927	53.4
1973	42399	45.8	50107	54.2
1974	35545.5	39.3	54955.5	60.7
1975	33464.5	38.3	53938.5	61.7
1976	23432	27.3	62472.5	72.7

Fuente: Listados de Sistemas Mecanizados del P.N.B. 1977

La United Fruit suspendió, desde 1965, sus operaciones regulares en el Ecuador (aproximadamente un sexto de las exportaciones del país) y siguió comprando, aunque de manera irregular e inferior a los valores acostumbrados antes de la crisis. La Standard Fruit mantuvo e incrementó sus operaciones cubriendo parte del mercado dejado por United Fruit.

El tiempo de producción ecuatoriana coincide con el de las demandas del hemisferio norte cuando las centroamericanas no lo hacen, por esa razón Ecuador generalmente exporta altos volúmenes. En 1974, Ecuador recuperó el mercado de Estados Unidos debido a dos fenómenos: el conflicto político que tuvieron las transnacionales con los países que constituyeron la UPEB (Unión de Países Exportadores de Banano), y la destrucción provocada por el huracán Fifi en Honduras, actuando Ecuador como abastecedor de reserva.

El cambio de variedad produjo una transformación profunda en la distribución espacial de la producción bananera, así como en la superficie cultivada y en la estructura de la tenencia de la tierra. En el cuadro IX constan el número de hectáreas sembradas con banano en el periodo 1965-1976. Los rendimientos por hectárea del *Cavendish* duplican a los del *Gross Michell*; en consecuencia, al estancarse los volúmenes exportados, la superficie se redujo sustancialmente de 210351 has. registradas en 1965 a 120425 has. en 1976. Este decrecimiento continuo de las extensiones de las áreas bananeras se observa en la figura 1.14.

La United Fruit, que fue la principal compañía en la etapa anterior (Auge Bananero), suspendió sus operaciones en 1965 y, a partir del año 1974 mantuvo un volumen irregular y limitado de compras. La Exportadora Bananera Noboa se consolidó, logrando una clara

primacia respecto a las compañías restantes (Ver APENDICE B). Su participación ascendió del 20.12% en 1965 al 44.58% en 1976; desde 1970 ocupa el primer lugar en el mercado. Durante esta etapa el comportamiento de la Standard Fruit fue irregular, pero, a diferencia de la United, mantuvo al 15% del mercado como mínimo.

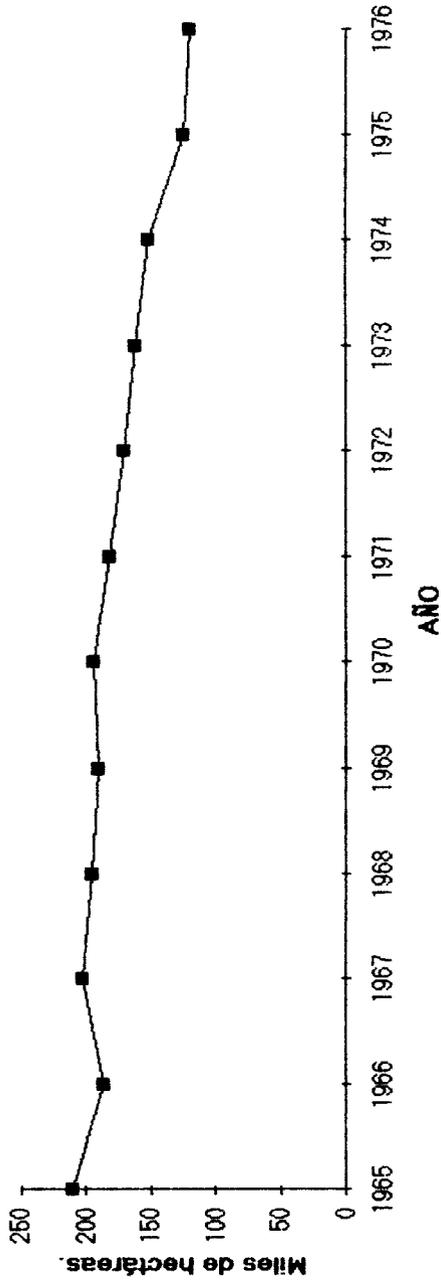
Cuadro IX
EVOLUCION DE LA EXTENSION DE LAS AREAS
SEBRADAS CON BANANO EN EL ECUADOR
1965-1976

Año	Hectáreas
1965	210351
1966	186688
1967	202716
1968	195095
1969	190170
1970	193560
1971	181050
1972	170740
1973	161907
1974	151779
1975	124635
1976	120425

Fuente: Primer Compendio Agropecuario del Ecuador 1965-1994, 1995

La crisis afectó gravemente a algunas compañías medianas y pequeñas, siendo el más significativo el caso del grupo alemán EFE-UBESA, que desde 1967 perdió posiciones en el mercado, a lo cual se sumó la desaparición de EFE en 1970. Es probable que la

Figura 1.14
Evolución de la Extensión de las Areas Sembradas con Banano en Ecuador
(1965-1976)



Fuente: Primer Compendio Agropecuario del Ecuador, 1994

creciente influencia de United Fruit y Standard Fruit en Europa Occidental haya ocasionado la pérdida del mercado germano occidental para el Ecuador. En el período 1977-1998 la población ecuatoriana se incrementó de 7'238.607 a 11'348.000 habitantes. La cotización del dólar se incrementó de 26.7 a 6480 sucres. En Ecuador, la "fiebre petrolera" comienza a disminuir en los 80's, ya que decrece el consumo de hidrocarburos. En 1986 ocurre el Boom Camaronero, llegando a convertirse en el primer producto de exportación (sin considerar al petróleo), desplazando así al banano.

La Standard Fruit aumentó significativamente su cuota entre 1977 y 1984 (del 18 al 37%). En 1982, Ecuador vendió a Estados Unidos 649000 toneladas, equivalente al 51.5% de las exportaciones nacionales de banano. Entre 1983 y 1984, el mercado norteamericano alcanzó el 60% de las exportaciones del país, duplicando el porcentaje de 1973 (29%).

La expansión de las exportaciones a Estados Unidos ha ido acompañada de tendencias declinantes en la mayor parte de los restantes mercados. Países como : Bélgica, Italia e Irlanda, han mantenido también una tendencia declinante. Mientras que en 1974 y 1975 superaron las 300.000 toneladas, desde 1981 sus exportaciones están por debajo de las 200.000 toneladas.

Desde 1977, el mercado de los países compradores no tradicionales como los países árabes y Nueva Zelandia, ha sufrido deterioros. A partir de los 80's surgen nuevos mercados: Argentina y Uruguay.

En el cuadro X constan las exportaciones de banano realizadas en el periodo 1977-1998. En la figura 1.15 se observan los periodos de crisis en las exportaciones bananeras ecuatorianas que se registra durante los años 83 y 84 a consecuencia de las fuertes adversidades invernales de 1983, en que las exportaciones descendieron a 800.15 miles de toneladas métricas que significaron ingresos de aproximadamente 152'926.000 dólares frente a los 213'297.000 generados durante 1982 por la exportación de 1254.38 miles de toneladas métricas. Desde 1983 la producción, en su mayoría de la variedad Cavendish, ha estado en alza desde el fenómeno de El Niño. El alza ha sido más pronunciada en la primera mitad de los 90.

En el cuadro XI constan los valores de la extensión sembrada con la variedad Cavendish y la correspondiente productividad del periodo 1980-1997. La sustitución total de la variedad Gross Michel por Cavendish, que permitía rendimientos más altos por hectárea, produjo una reducción sistemática del área cultivada, hasta llegar a 1980 con sólo 60.568 has. en producción de banano Cavendish. Esta cifra disminuyó aún más en los años subsiguientes, para registrar su punto más bajo en 1983 con 49.561,5 has. Se produce luego un

repunte moderado y se alcanzan las 65.028 has. A partir de 1989, y ante el incremento de las ventas en Europa, se inicia un proceso sostenido y casi explosivo de crecimiento del área de cultivo, para alcanzár el tope de 134.504,1 has. en 1992. En los últimos años ha disminuido un poco.

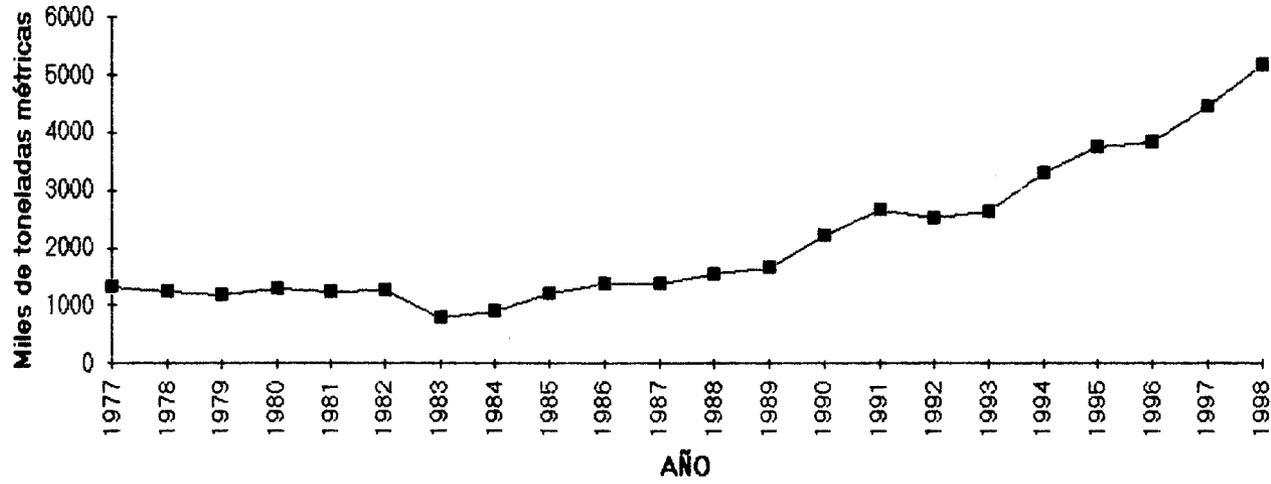
Cuadro X
EXPORTACIONES DE BANANO EN EL
ECUADOR
1977-1998

Año	Miles de TM.	Año	Miles de TM.
1977	1317.7	1988	1534.82
1978	1223.8	1989	1648.87
1979	1170.1	1990	2200.78
1980	1290.6	1991	2654.34
1981	1235.4	1992	2511.13
1982	1254.38	1993	2632.76
1983	800.15	1994	3300.5
1984	906.3	1995	3735.5
1985	1207.86	1996	3842.4
1986	1361.8	1997	4456.27
1987	1381.51	1998	5154.56

Fuente: Economía Ecuatoriana en Cifras 1970-1998, Serrano Alberto, 1999

La relación de producción bananera a área de cultivo nos da un rendimiento de 37.05 tm/ha en 1995 y presenta una caída a 33.8 tm/ha en 1996. En 1997 se incrementa a 34.48 tm/ha (Ver figura 1.16).

Figura 1.15
Exportaciones de Banano del Ecuador (1977-1998)



Fuente: El Banano en el Ecuador, Larrea Carlos, Corporación Editora Nacional, 1987

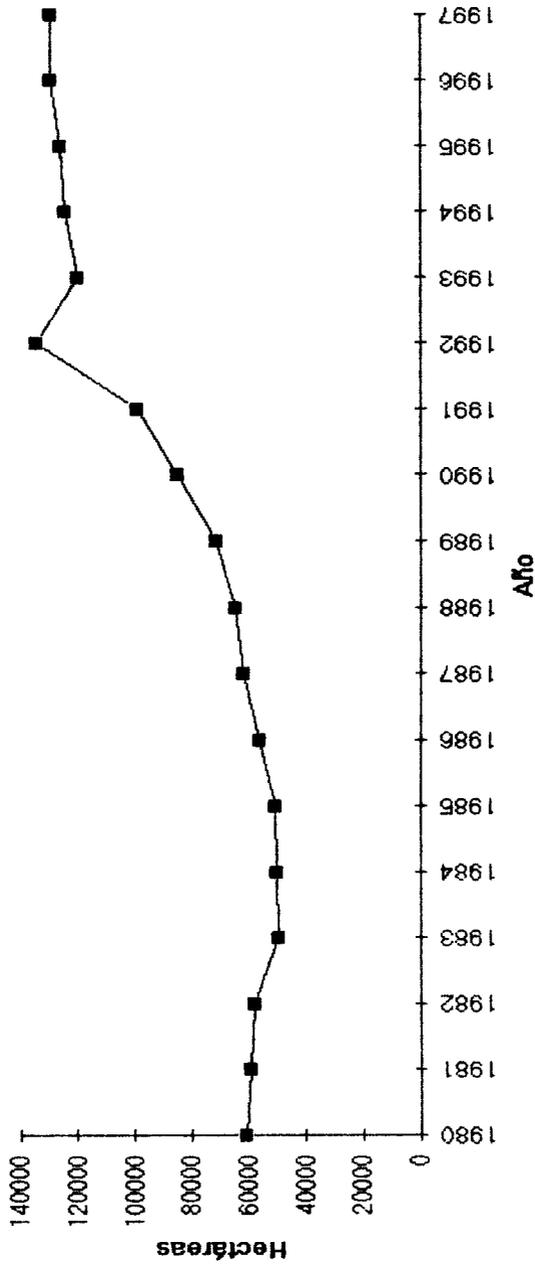
El periodo comprendido entre 1990 y 1993 registra volúmenes de producción superiores a dos pero inferiores a tres millones de toneladas métricas por año. Desde 1994 se registran volúmenes de producción superiores a los tres millones de toneladas métricas por año.

Cuadro XI
SUPERFICIE DEL ECUADOR
CULTIVADA CON BANANO
CAVENDISH
1980-1997

Año	Hectáreas sembradas	Productividad (TM/Ha)
1980	60568	31.56
1981	59077	33.01
1982	58196.5	32.98
1983	49561.5	26.61
1984	49993	30.34
1985	50636	34.71
1986	56175.4	33.36
1987	61717.4	33.3
1988	65028	32.72
1989	71601.2	30.14
1990	85187.9	29.47
1991	99040.1	28.74
1992	134504.1	25.79
1993	119425.2	33.92
1994	124400	31.31
1995	125603.6	37.05
1996	129125.1	33.8
1997	129230	34.48

Fuente: Economía Ecuatoriana en Cifras 1970-1998, Serrano Alberto, 1999

Figura 1.16
Superficie cultivada con banano Cavendish en el Ecuador (1980-1997)



Fuente: Economía Ecuatoriana en Cifras 1970-1998, Serrano Alberto, 1999

Entre 1992 y 1998, la actividad bananera registra un marcado crecimiento en sus exportaciones al registrar nuevos récords en sus ventas (año 1997). Pero, los productores se han visto afectados por la reducción de sus rentabilidades, ya que se incrementó notablemente el costo de mano de obra e insumos.

El año de 1997 fue un año record para el sector bananero ya que generó 1327 millones de dólares, cifra que será difícil de superar en el presente siglo debido a los bajos precios del banano en el mercado internacional. Esta cifra fue apenas 84 millones de dólares inferior a lo que el Estado recibió por concepto de exportaciones de petróleo crudo. Pero en 1998, el banano generó 1070 millones de dólares, es decir, 279 millones de dólares más que el petróleo, representando el 25.8% de las exportaciones totales, y de no haber sido por la disminución del volumen de ventas al exterior por los efectos de El Niño el porcentaje hubiera sido mayor.

En 1998 el mercado chino solo compró a Ecuador 65998 toneladas, mientras que en Rusia las ventas bajaron un 70%. También el fenómeno "El Niño" provocó estragos en las áreas productivas de banano, ocasionando gran cantidad de pérdidas (Ver Cuadro XII).

Cuadro XII
AREAS PRODUCTIVAS DE BANANO AFECTADAS POR EL FENOMENO DE "EL NIÑO"

		Pérdidas en cajas por no exportación a razón de 1581 cajas/ha/año
Hectáreas perdidas por huracanadas o inundaciones	3.867.60	6.114.675.60
Hectáreas sobresaturadas de humedad	16.026.90	17.736.970.00
Total de hectáreas afectadas	19.894.50	23.851.645.60
Valor de las pérdidas		101.369.494.50

Valor promedio de una hectárea en producción es 45'940000 sucres o 10000 dólares de 4594 sucres por dólar.

Las hectáreas perdidas representan el 15.3% del total inscritas. La pérdida estimada fue de 913.953'330.000 sucres

ZONAS AFECTADAS	HECTAREAS PERDIDAS	HECTAREAS SOBRESATURADAS
Norte	10	57
Central	240	560
Subcentral	839	2500
Milagro	149	493
Triunfo	118	2001.4
Naranjal	1550.6	8670.5
Sur-Guabo	5	345
Sur-Machala	956	1400
TOTAL	3867.6	16026.9

Las pérdidas en ventas por disminución de producción con relación al año 1997 alcanzan la suma de 190 millones de dólares.

Fuente: Revista El Agro, Mayo 1999 No. 38

El cuadro XIII presenta una tabla de posiciones de los países compradores de banano en el período 1993-1998.

Cuadro XIII
PRINCIPALES PAISES COMPRADORES DEL
BANANO ECUATORIANO
1993-1998

Lugar	1993	1994	1995	1996	1997	1998
1	U.E.	U.E.	U.E.	U.E.	U.E.	EEUU
2	EEUU	EEUU	EEUU	EEUU	EEUU	U.E.
3	Lat. Amér.	Lat. Amér.	Resto de Europa	Resto de Europa	Asia	Países del Este
4	Asia	Resto de Europa	Lat. Amér.	Asia	Oceanía	Asia

Fuente: Banco Central del Ecuador, Boletín Anuario #21, 1999

Unión Europea (U.E.): Conformada por Italia, Alemania, Bielorusia, Portugal, Grecia, Finlandia, España, Suecia, Gran Bretaña, Irlanda, Austria y Francia.

Resto de Europa: ex URSS, Polonia, ex Checoslovaquia, ex-Yugoslavia, Hungría y otros.

Extremo Oriente: China Popular, Japón y Corea.

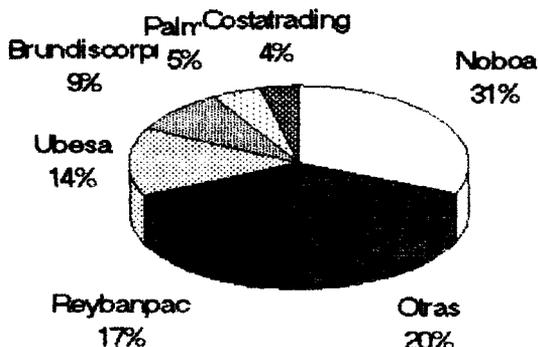
Latinoamérica: Argentina, Chile, Uruguay y otros.

Oceanía: Tiene como representante a Nueva Zelanda.

La Unión Europea ocupó el primer lugar desde 1993 hasta 1997. En 1998, Estados Unidos se convierte en nuestro principal comprador. En cuanto a los mercados de: Latinoamérica, resto de Europa, Asia y Oceanía, siempre se ubican entre el tercer y cuarto lugar.

Seis empresas manejaron en 1998 el 79.8% de las exportaciones de banano medidas en toneladas métricas. Bananera Noboa es el principal exportador con el 31.6% del total. A continuación está Reybanpac con el 17% . El tercer lugar lo tiene Ubesa, subsidiaria de Standard Fruit, con 14.4%. El cuarto lugar lo tiene Brundiscorpi con el 9.2%, el quinto Agrocomercio Palmar con el 4.6% y le sigue Costatrading con el 3.5%. Así lo podemos observar en la figura 1.17.

Figura 1.17
Principales exportadoras: 1998



Fuente: Suplemento "Líderes", Diario "El Comercio", Mayo 1999

Exportadora Bananera Noboa: Comenzó a operar como exportadora independiente en 1956, creció rápidamente hasta 1964 y durante los primeros años de la crisis (1965-1976), mantuvo una participación del 20% del mercado. Entre 1971 y 1977, la compañía creció velozmente hasta llegar a controlar casi la mitad de las exportaciones del Ecuador. Esta expansión ocurrió cuando las transnacionales bananeras como la Standard Fruit pasaban momentos de crisis. A nivel nacional, Exportadora Bananera Noboa ha sido considerada como la empresa no petrolera más grande del Ecuador. Durante la etapa de auge (hasta 1964) la empresa consiguió el mercado constituido por: Bélgica, Italia, Holanda, países nórdicos y naciones de Europa Occidental, sus ventas a Estados Unidos ocupaban el segundo lugar. Pero, la principal plaza que provocó la expansión de Exportadora Noboa fue Japón en 1972, hasta que este mercado fue abastecido por las transnacionales desde Filipinas.

Exportadora Noboa incluye empresas claves tales como: plantaciones, fumigadoras, una flota de camiones refrigerados, la principal empresa nacional productora de cajas de cartón (Industria Cartonera Ecuatoriana), que abastece a las bananeras y la demanda nacional.

Reybanpac (Rey Banano del Pacífico): A diferencia de otras empresas medianas y pequeñas tales como EFE-UBESA o Fruit

Trading, Reybanpac logró ingresar en 1978 , sobrevivir a la crisis y mantenerse en el mercado. Su participación como exportadora independiente creció notablemente entre 1978 y 1983. Se cree que su éxito se debe a su asociación con **Del Monte**, siendo su productor asociado e intermediario de exportación. Al igual que la Exportadora Noboa, también posee su propia cartonera: Andina.

Estas dos importantes exportadoras le deben su éxito a la estrategia de poseer sus propias productoras cartoneras. Las productoras cartoneras comenzaron a tomar gran importancia en los años sesenta, debido a que se dejó de exportar banano en racimos a partir de 1960, y ya en 1963 se introduce la exportación en cajas pero en cantidades pequeñas, continuando el predominio de la exportación en racimos; a partir de 1964 el proceso se invierte, aumentando rápidamente la exportación en cajas y reduciéndose en forma drástica las ventas en racimo; en 1972 sólo se exportaron 24.483 racimos y en 1976 esta modalidad desapareció por completo. Las cartoneras se han convertido en importantes fuentes de empleo y son un factor clave en las exportaciones de banano, ya que se encargan de elaborar los diferentes tipos y diseños de cajas que difieren según el peso y el destino y son los siguientes:

TIPO	PESO (Kg.)	DESTINOS
22-XU	19.52	Todos los países
115 K	15.44	Todos los países
208	12.71	Todos los países
25-27	11.81	Todos los países
22-XU-CS	22.7	Países del Cono Sur
115 K-DP	15.44	Autorizada para exportar barraganete

En el cuadro XIV constan las exportaciones bananeras realizadas en el período 1993-1997, clasificadas por tipo de caja. Es notorio que el tipo de caja con mayor demanda es la 22 XU o de 43 lbs. La caja 208 o de 28 lbs. en comparación con los tipos de caja restantes presenta una alta demanda.

Cuadro XIV
EXPORTACIONES BANANERAS EN EL ECUADOR POR TIPO
DE CAJA
1993-1997

Año	22 XU (43 LB)	115K (34LB)	208 (28LB)	25-27 (26LB)	22XUCS (50LB)
1993	116247960	-----	16006257	10411191	-----
1994	152886820	-----	14568749	7501658	-----
1995	152670438	405364	13607557	8199242	20604014
1996	146900465	202102	2165610	21156269	20390611
1997	167492946	-----	53325884	6568630	18323362

Fuente: El Agro, Noviembre 1999, Edición No. 30

La historia de los precios pagados por caja de banano por los exportadores al productor puede resumirse en que hasta 1992 los productores bananeros recibían y ganaban más dinero que lo que actualmente obtienen. En 1991 los precios para la caja 22 XU desde marzo hasta abril era de 4,6 dólares, valor que se incrementó de abril a mayo a 5 dólares y que en octubre ascendió a 6 dólares.

A partir de octubre de 1991 hasta mayo de 1992 los precios para la caja de banano se mantuvieron dentro de una fluctuación de $5,4 \pm 2$ dólares. A partir de septiembre de 1992, a pesar de los procesos inflacionarios, los precios para la caja de banano ecuatoriano se fijan hasta diciembre del mismo año en 4,5 dólares, valor que posteriormente se reduce aún más a 4,1 dólares por caja hasta enero de 1994, fecha a partir de la cual se mantuvo en el país el precio de 4.5 dólares para la caja de banano 22-XU.

El año de 1999 ha sido difícil para los productores debido a la falta de demanda y la caída de los precios en los mercados internacionales, así cuando se exporta a los países del Cono Sur, la caja debe contener 50 libras netas y se comercializa a 8000 sucres, en cambio la que se exporta a Estados Unidos y Europa, tiene un peso de 40 libras y un valor de 11000 sucres. Ninguno de estos precios se igualan al precio oficial de 3.05 dólares por caja. Todo esto influye en que los productores prefieran perder la fruta y se refleja en que el

30% de la producción se comercializa y el 70% restante se pierde en las plantaciones.

La caja de 43 libras en la Unión Europea costaba 19 dólares, ahora se cotiza a 13; en Estados Unidos bajó de 13 a 7 dólares y los mercados emergentes se cotiza entre 4.5 y 5 dólares.

El 9 de junio de 1999, el Gobierno bajó el precio de la caja de banano de 43 libras a 2.6 dólares, precio que deberá pagar el exportador al productor. También se estableció una excepción para los precios de las cajas de banano que se exporten a los mercados emergentes, considerados los países de Europa del Este, China, Rusia y Corea del Sur, cuyo precio mínimo será de 2.2 dólares la caja.

Actualmente el Ecuador se ve amenazado con una sobreproducción bananera que no encuentra nuevos mercados a nivel internacional y que provoca la baja de los precios de la fruta. Se afirma que la oferta mundial de exportaciones crecerá en el orden del 4% al 5% sin sembrar una hectárea más de banano. Pero la demanda crecerá en apenas 1,6% con lo que la situación del sector será difícil. Por eso, se redefinirá la meta de exportaciones a no más de 950 millones de dólares. El monto de ventas en 1998 fue de 1.071 millones de dólares.

Del crecimiento de exportaciones, el 70% va hacia los mercados emergentes, que hoy comprenden los países colapsados por la guerra de Kosovo. Debido a los errores de bombardeo de la OTAN hacia Yugoslavia, se dejó de exportar a través de los puertos ubicados en las costas del mar Adriático, que separa a Italia de Yugoslavia. A través de los puertos Rijeka, Trieste y Stil, se ingresa banano para su distribución hacia una parte de Rusia, Lituania, Suiza, Polonia y directamente a Eslovenia, Rumania y Yugoslavia (Ver cuadro XV). Semanalmente ingresaban 1.3 millones de cajas por estos puertos, pero debido a los incesantes bombardeos de la OTAN, hasta mayo de 1999 no habían podido recibir 7.8 millones de cajas.

Cuadro XV
EXPORTACIONES DE BANANO DE ECUADOR A

En TM.	1997	1998	*1999
Suiza	10220	-	-
Eslovenia	397	-	-
Lituania	117968	-	-
Polonia	55917	86514	18450
Rumania	43756	16194	-
Yugoslavia	63648	24410	2924
Rusia**	427224	407826	66550

* Enero a Marzo

** Sólo una parte del banano que se exporta a Rusia entra por la zona de conflicto

Fuente: Diano El Universo, Mayo 1999

Debido a la sobreoferta los bananeros están realizando la práctica del "vampirazo" que es una técnica que consiste en clavarle una estaquilla a una planta de banano para que no produzca el racimo. Así se alivia el exceso de fruta en temporada baja y se evitan los gastos excesivos que conlleva el mantener las plantaciones en tiempos de precios bajos.

La sobreoferta es interna, porque el Gobierno ha permitido la siembra indiscriminada de miles de hectáreas ilegales (hectáreas no consideradas en las superficies de cultivo de banano, establecidas oficialmente), lo que ha llevado a una caída de los precios de la fruta y a una reducción de la calidad de las siembras. Se estima que el área sembrada ilegalmente suma 7796 hectáreas, repartidas en medio centenar de haciendas, la mayoría propiedad de grandes exportadoras.

En el cuadro XVI se presentan las hectáreas cultivadas hasta mayo de 1999, clasificadas por su nivel de tecnificación. En la figura 1.18 se puede notar que 14448 (10,09%) de las hectáreas cultivadas les falta tecnología.

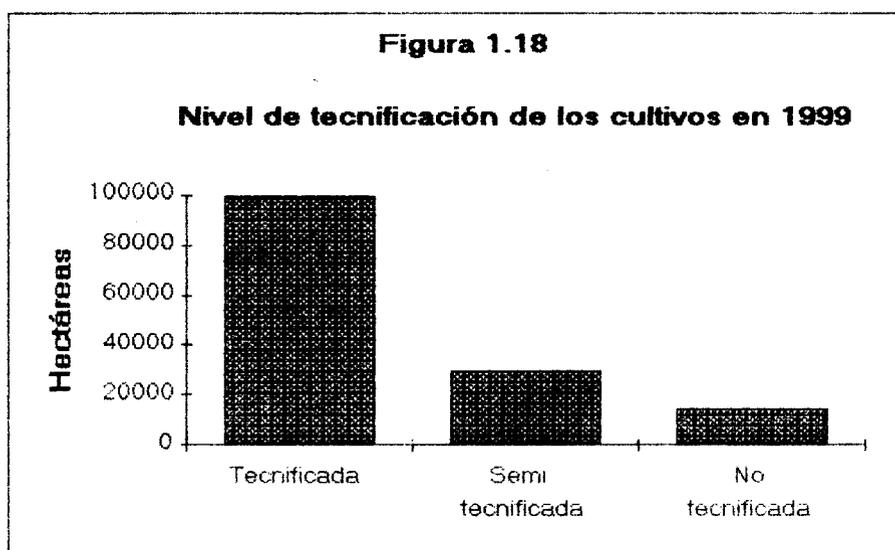
Cuadro XVI
NIVEL DE TECNIFICACION DE LOS CULTIVOS EN EL ECUADOR
(May-99)

Nivel	Hectáreas cultivadas	%
Tecnificadas	99442	69.5
Semitecnificadas	29201	20.41
No tecnificadas	14448	10.09

Fuente: Agroinversiones

Figura 1.18

Nivel de tecnificación de los cultivos en 1999



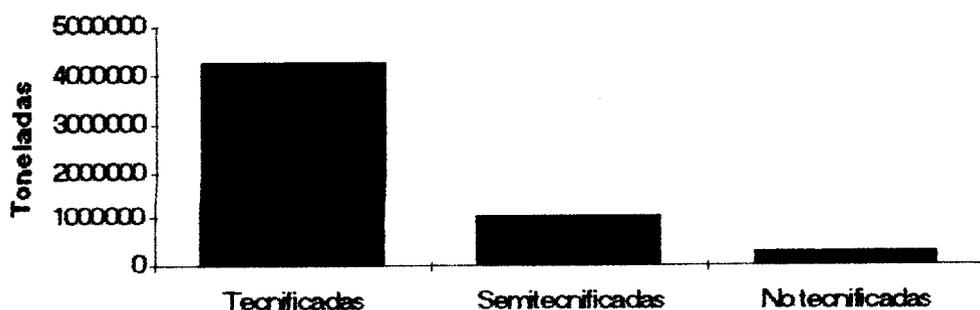
El cuadro XVII presenta las diferentes áreas de cultivo: tecnificadas, semitecnificadas y no tecnificadas, y su respectiva productividad, con lo que se puede calcular la producción esperada para el año 1999 que será de 5'560.885 toneladas métricas. En la figura 1.19 se presenta un histograma de estos datos.

**Cuadro XVIII
DESPERDICIOS DE BANANO
1983-1995**

Año	Toneladas Métricas
1983	276172
1984	308183
1985	321708
1986	287490
1987	419292
1988	323664
1989	172654
1990	150648
1991	85231
1992	581479
1993	695286
1994	150741
1995	283200

Fuente: Compendio Estadístico Agropecuario del Ecuador, 1994

**Figura 1.19
Producción de Banano en Ecuador :1999**



Fuente: Diario El Universo, Mayo 19 de 1999

Cuadro XVII
PRODUCCION DE BANANO EN EL ECUADOR : 1999

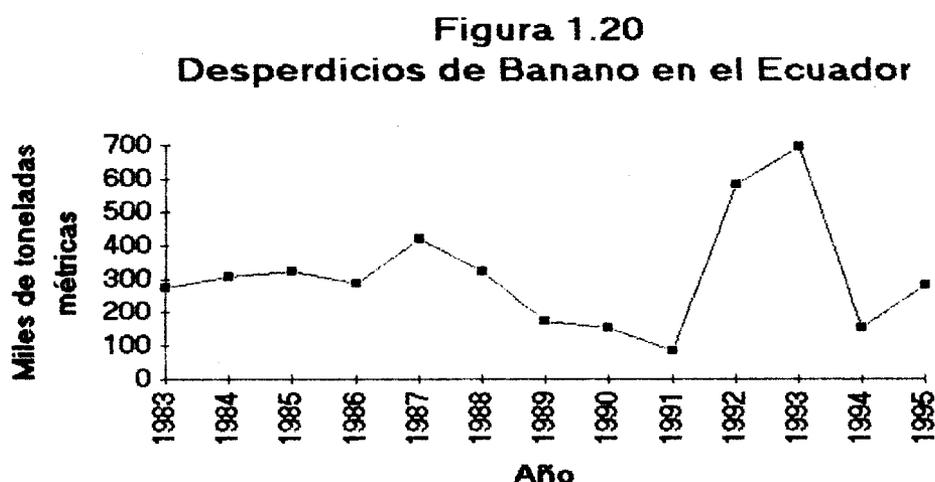
	Productividad (Tm/has/año)	Producción esperada (toneladas)
Tecnificadas	45	4249890
Semitecnificadas	35	1022035
No tecnificadas	20	288960
TOTAL		5560885

Fuente: Diario El Universo, Mayo 19 de 1999

Pérdidas post-cosechas

Las pérdidas post-cosechas del banano son ocasionadas por los daños de: manejo (estropeo, mal empaque), medio ambiente, insectos y animales, por enfermedades, genéticos y fisiológicos y por material extraño. En el cuadro XVIII se presentan las cifras en toneladas métricas de los desperdicios de banano generados en el período 1983-1995.

En la figura 1.20 se puede notar que entre 1983 y 1991 hubo un decremento continuo en la cantidad de desperdicios generados en la cosecha de banano. Pero, en 1992 se incrementa notablemente la cifra de desperdicios a 581479 toneladas métricas y es en 1993 que alcanza el mayor valor, siendo éste 695286 toneladas métricas.



Fuente: Compendio Estadístico Agropecuario del Ecuador, BCE, 1994

1.4 La industria bananera del Ecuador

El aumento que viene registrando la producción bananera no guarda relación con el ritmo de crecimiento que registran las exportaciones. Este comportamiento ha dado como resultado que una buena parte de la producción bananera, aproximadamente entre el 15 y 40% quede sin utilizarse, lo cual constituye un desperdicio de recursos. Estos excedentes son empleados en la industrialización.

En los años 1955 y 1956 el Banco Central del Ecuador registró 11000 y 13000 dólares, por concepto de banano transformado. La industrialización del banano ecuatoriano comenzó en 1958, año en que se instaló la primera planta de banano deshidratado en la provincia del Guayas. Actualmente la industrialización se enfoca preferentemente en los siguientes productos: banano en almíbar y en rodajas deshidratadas, congelado, deshidratado en hojuela, bebidas alcohólicas y etanol , harina y polvo , jaleas, néctares y bebidas, preparación de jarabe de glucosa y fructuosa , puré , rodajas fritas, sabor y aroma , vinagre , aprovechamiento de residuos de banano con fines textiles y almidón .

El proceso de deshidratado es muy sencillo: en estado maduro, comestible y desprovisto de corteza, dispuesto en charoles metálicos los bananos son introducidos en un horno de calefacción lenta que permite la eliminación progresiva de los altos contenidos de agua; esto produce concentraciones de azúcares y otras sustancias lo que origina un sabor agradable, obteniendo así el banano deshidratado.

Posteriormente se procede al envase, para lo cual se utiliza el papel celofán para protegerlo de cuerpos extraños y particularmente de la humedad del medio ambiente. En las exportaciones que datan de 1955 se utilizaban cajas de madera que actualmente se reemplazaron por cartón.

En el cuadro XIX constan la cantidad de elaborados de banano exportados (kilos) y los respectivos ingresos obtenidos (miles de dólares FOB) en el período 1982-1996 (Ver figura 1.21).

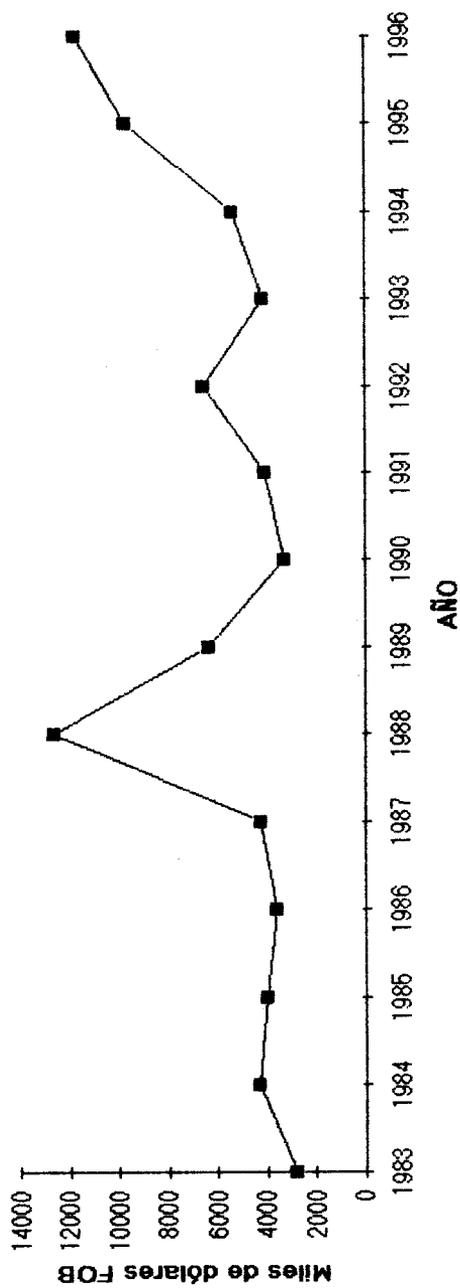
**Cuadro XIX
EXPORTACION DE ELABORADOS DE
BANANO EN EL ECUADOR
1982-1996**

Año	Miles de kilos	Miles de dólares FOB
1982	2112	5719
1983	1561	2792
1984	2160	4281
1985	2142	4010
1986	1996	3594
1987	3583.1	4210
1988	10413.4	12618
1989	8893.2	6355
1990	6938.6	3271
1991	8380.1	4066
1992	12998.7	6528
1993	2600	4133
1994	3467.3	5376
1995	9534.5	9667
1996	26785.6	11734

Fuente: Banco Central del Ecuador, Boletín Anuario #19, 1997

Las exportaciones de banano deshidratado se han dirigido de modo casi exclusivo al mercado europeo y japonés. En el mercado europeo

Figura 1.21
Exportaciones de Elaborados de Banano en el Ecuador (1983-1996)



Fuente: Banco Central del Ecuador. Boletín Anuario # 19, 1997

a los países: Holanda, Alemania, Inglaterra, Francia, Bélgica, Suiza, Italia, en Norteamérica, Estados Unidos y en Latinoamérica, Chile. Las exportaciones de puré, flakes y polvo de banano se distribuyen entre Estados Unidos, Japón, Europa, en orden de importancia.

CAPITULO 2

2. DESCRIPCION DE LAS TECNICAS ESTADISTICAS A UTILIZAR

Previo al análisis estadístico multivariado es necesario realizar el *análisis estadístico univariado* de las características, lo cual incluye la utilización de técnicas descriptivas e inferenciales.

2.1 Análisis univariado

A partir de un subconjunto de observaciones seleccionadas de una población denominado *muestra aleatoria*, se obtienen la *media aritmética* y la *varianza muestral*. Además se elaboran los siguientes gráficos: *función empírica*, *histograma de frecuencias*, *ojiva* y *diagrama de cajas*.

La media aritmética y la varianza muestral son fundamentales para describir el comportamiento de los datos de una muestra aleatoria dada, de allí que sea importante conocer sus respectivas definiciones.

Sean X_1, X_2, \dots, X_n , elementos de una muestra aleatoria de tamaño n , la *media aritmética*, denotada por \bar{X} , está definida por:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n} = \frac{\sum_{i=1}^n X_i}{n}$$

Sean X_1, X_2, \dots, X_n , elementos de una muestra aleatoria de tamaño n , y sea \bar{X} la media aritmética de la muestra, la *varianza muestral*, denotada por s^2 , se define de la siguiente forma:

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

Los gráficos facilitan el análisis de las características de estudio, en cuanto a su dispersión, tendencia central, frecuencias relativas y acumuladas. A continuación se definirán los gráficos más importantes:

función empírica, histograma de frecuencias, ojiva y diagrama de cajas.

Sean $X_{(1)}, X_{(2)}, \dots, X_{(k)}, \dots, X_{(n)}$, los estadísticos de orden correspondientes a una muestra aleatoria de tamaño n , esto es, $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$, la *función empírica*, denotada por $\hat{F}(x)$, se define de la siguiente forma:

$$X_{(1)} = \text{mín} (X_1, X_2, \dots, X_n)$$

$X_{(k)}$ es el k -ésimo estadístico de orden ; $2 \leq k \leq n-1$

$$X_{(n)} = \text{máx} (X_1, X_2, \dots, X_n)$$

$$\hat{F}(x) = \begin{cases} 0; & x < x_{(1)} \\ \frac{k}{n}; & x_{(k)} \leq x < x_{(k+1)} \\ 1; & x \geq x_{(n)} \end{cases}$$

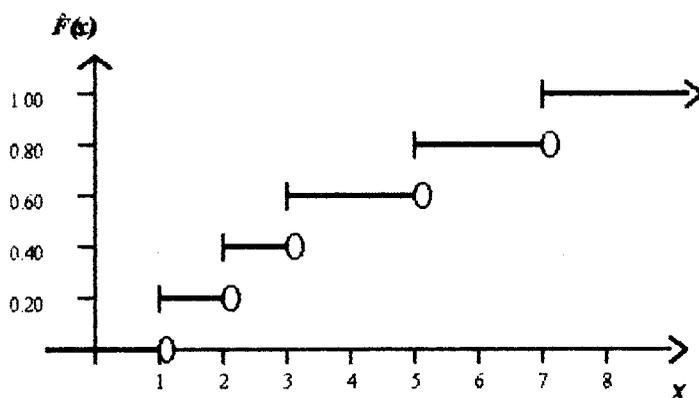
Por ejemplo, si tenemos la siguiente muestra de tamaño 5 que es : $x_1 = 2, x_2 = 7, x_3 = 3, x_4 = 5, x_5 = 1$, obtendremos los siguientes estadísticos de orden: $x_{(1)} = 1, x_{(2)} = 2, x_{(3)} = 3, x_{(4)} = 5, x_{(5)} = 7$, resultando la siguiente función empírica:

$$\hat{F}(x) = \begin{cases} 0; & x < 1 \\ \frac{1}{5}; & 1 \leq x < 2 \\ \frac{2}{5}; & 2 \leq x < 3 \\ \frac{3}{5}; & 3 \leq x < 5 \\ \frac{4}{5}; & 5 \leq x < 7 \\ 1; & x \geq 7 \end{cases}$$

cuya gráfica se presenta en la figura 2.1.

Figura 2.1

Función empírica de una muestra

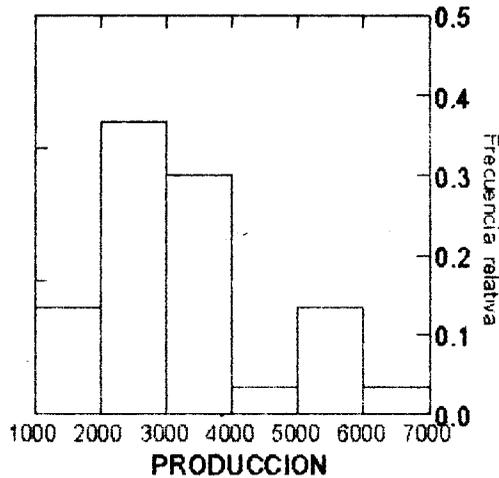


Histograma de frecuencias

Las clases son intervalos consecutivos y de igual longitud de tal forma que la primera incluya al mínimo valor $X_{(1)}$ y la última incluya al máximo valor $X_{(n)}$. Un *histograma de frecuencias* (Figura 2.2) está formado por una serie de barras que tienen sus bases sobre un eje horizontal (eje X) e iguales al ancho de su clase y su altura es igual a la frecuencia relativa de clase.

Figura 2.2

Histograma de frecuencias correspondiente a la producción bananera del Ecuador

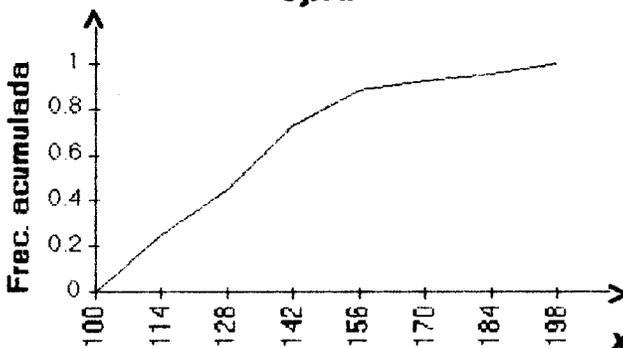


Ojiva

La *ojiva* (Figura 2.3) , llamada también *polígono de frecuencias acumuladas*, es el gráfico de líneas trazado sobre las clases definidas para el histograma de frecuencias. La ojiva tiene una frecuencia acumulada relativa igual a cero en el mínimo valor de la primera clase e igual a uno en el máximo valor de la última clase. En la gráfica de la ojiva se pueden determinar los percentiles P_1, P_2, \dots, P_{99} , donde el 1% de las observaciones son menores o iguales que P_1 , el 2% de las observaciones son menores o iguales que P_2 , y así sucesivamente; los deciles D_1, D_2, \dots, D_9 , donde el 10% de las observaciones son menores o iguales que D_1 , el 20% de las observaciones son menores o iguales que D_2 , y así sucesivamente; y los cuartiles Q_1, Q_2 y Q_3 , donde el 25% de las observaciones son menores o iguales que Q_1 , el 50% de las observaciones son menores o iguales que Q_2 y el 75% de las observaciones son menores o iguales que Q_3 .

Figura 2.3

Ojiva

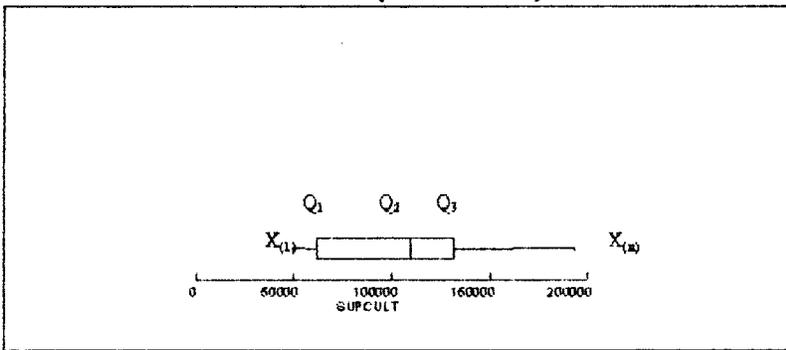


Diagramas de cajas

Un diagrama de cajas trata de mostrar al mismo tiempo la tendencia central y dispersión de los datos . El gráfico puede presentar valores aberrantes, los cuales se representan mediante asteriscos (*). En el diagrama se presentan los cuartiles Q_1 , Q_2 y Q_3 , y dos *segmentos de recta* , uno que es la línea que se traza desde el mínimo valor hasta el primer cuartil y el segundo que es la línea que se traza desde el tercer cuartil hasta el valor máximo. El valor de Q_2 representa la mediana muestral (ver figura 2.4).

Figura 2.4

Diagrama de cajas correspondientes a la superficie cultivada con banano (hectáreas)



Para realizar un análisis estadístico, es necesario conocer las definiciones básicas correspondientes a la teoría de la probabilidad: espacio muestral, conjunto potencia, función probabilidad, variables aleatorias y variables aleatorias conjuntas.

Probabilidad

El *espacio muestral* está dado por el par (Ω, \mathcal{F}) , donde Ω es el conjunto de todos los resultados posibles de un experimento y \mathcal{F} es el conjunto potencia de Ω . Si Ω es finito y tiene n elementos, el conjunto potencia \mathcal{F} tendrá 2^n elementos. Consideremos el siguiente experimento: Lanzar una moneda y ver que lado sale. El conjunto de resultados posibles es {sello, cara} = {s,c}. El conjunto potencia correspondiente es: $\mathcal{F} = \{\emptyset, \{s\}, \{c\}, \{s,c\}\}$. Podemos notar que el espacio muestral tiene dos elementos, de allí que el conjunto potencia esté conformado por 2^2 (cuatro) elementos.

La *función probabilidad* es una función $P: \mathcal{F} \rightarrow [0,1]$, que a cada elemento de \mathcal{F} le asigna un número, real entre 0 y 1, y sólo uno, bajo las siguientes condiciones:

1. $0 \leq P(E) \leq 1; \forall E \in \mathcal{F}$
2. $P(\Omega) = 1$
3. $P(E_1 \cup E_2) = P(E_1) + P(E_2)$ siempre que $E_1 \cap E_2 = \emptyset$

Dado un espacio muestral (Ω, \mathcal{F}) una *variable aleatoria* es una función $X: \Omega \rightarrow \mathbb{R}$, que a cada elemento de Ω le asigna uno y solamente un número real. Las variables aleatorias pueden ser *continuas o discretas*.

Una variable aleatoria X se dice *discreta* si solamente puede tomar un número finito o infinito contable de valores. Un conjunto de elementos es infinito contable si se puede establecer una relación uno-uno con el conjunto de los enteros positivos. Si X es discreta, la probabilidad de que X tome el valor x , $f(x) = P(X = x)$, se define como la suma de las probabilidades de todos los puntos muestrales de (Ω, \mathcal{F}) que tiene asignado el valor x .

Para una variable aleatoria discreta X , $P(X = x)$ debe cumplir las siguientes propiedades:

1. $0 \leq P(X = x) \leq 1$
2. $\sum_{\text{toda } x} P(X = x) = 1$

Por ejemplo, la distribución de probabilidades de la variable aleatoria X dada por:

$$P(X = x) = \begin{cases} x^2 / 30 & ; x = 1, 2, 3, 4 \\ 0, & \text{resto de } x \end{cases}$$

Una variable aleatoria X se dice *continua* si toma cualquier valor en un intervalo de números reales. Con cada variable aleatoria continua se asocia una *función de densidad de probabilidad* f . Su valor es $f(x)$.

Para una variable aleatoria continua X , $f(x)$ debe cumplir las siguientes propiedades:

$$1. f(x) \geq 0$$

$$2. \int_{-\infty}^{+\infty} f(x) dx = 1$$

$$3. P(a \leq X \leq b) = \int_a^b f(x) dx$$

Por ejemplo, la función densidad de la variable aleatoria X dada por:

$$f(x) = \begin{cases} (e^{-x/100})/100, & x > 0 \\ 0, & \text{resto de } x \end{cases}$$

Una vez detallado lo concerniente a las propiedades y distribuciones de probabilidad, procederemos a definir la media y la varianza poblacional para las variables aleatorias discretas y continuas.

Esperanza matemática de una función

La esperanza matemática de una función g , está dada por:

$$E[g(x)] = \sum_x g(x)P(X = x), \text{ si } X \text{ es discreta}$$

y,

$$E[g(x)] = \int_{-\infty}^{+\infty} g(x)f(x)dx, \text{ si } X \text{ es continua y la integral existe}$$

Media poblacional

Si X es una variable aleatoria discreta con función de probabilidad $f(x)$ y $g(x) = x$, entonces, la *esperanza matemática de X* se denota $E(X)$, entonces:

$$\mu = E(g(x)) = E(X) = \sum_x xf(x),$$

Si X es una variable aleatoria continua con función de densidad $f(x)$, entonces, el valor esperado de X se denota $E(X)$ y está definido por:

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

siempre que exista la integral.

Varianza

La *varianza* de una variable aleatoria X con media m , es la esperanza matemática de la función $g(x) = (x - \mu)^2$. Es decir,

$$\text{Var}(X) = \sigma^2 = E[(X - \mu)^2] = E(X^2) - \mu^2$$

La desviación estándar de X es la raíz cuadrada positiva de $\text{Var}(X)$, esto es

$$\sigma = \pm\sqrt{\sigma^2}$$

La distribución de un grupo de datos, puede presentar diferentes patrones de comportamiento, que pueden ser explicados mediante los coeficientes de sesgo y kurtosis.

Sesgo

Describe la *asimetría* de los grupos de datos con respecto a la media.

El *coeficiente de sesgo* es:

$$\gamma_1 = \frac{E[(X - \mu)^3]}{\{E[(X - \mu)^2]\}^{3/2}}$$

Si el coeficiente de sesgo es positivo, la distribución está *sesgada a la izquierda*, lo que implica que los datos se concentran al lado derecho.

Si el coeficiente de sesgo es cero, la distribución es *simétrica*. Si el coeficiente de sesgo es negativo, la distribución está *sesgada a la derecha*, lo que implica que los datos se encuentran cargados hacia el lado izquierdo.

Kurtosis

Es una medida de la *picudez* del grupo de datos. También se lo toma como una medida de la *concentración de los datos en las colas*. El *coeficiente de kurtosis* está dado por:

$$\gamma_2 = \frac{E[(X - \mu)^4]}{(E[(X - \mu)^2])^2}$$

El coeficiente de kurtosis es igual a 3 *para una variable aleatoria normal*, la cual recibe el nombre de *mesocúrtica*. Una distribución, cuyas colas presentan mayor concentración de datos que en la normal, se denomina *leptocúrtica* y tiene un coeficiente *mayor que 3*. Una distribución, cuyas colas presentan menor concentración de datos que la normal se denomina *platicúrtica* y tiene un coeficiente *menor que 3*.

La función generadora de momentos determina de manera única a una variable aleatoria, ya que pueden existir variables aleatorias diferentes, con medias y varianzas iguales.

Función generadora de momentos

Sea X una variable aleatoria, se define como *función generadora de momentos* al valor esperado de e^{tx} , esto es:

$$M(t) = E[e^{tx}] = \begin{cases} \sum_{\text{todas } x} e^{tx} f(x), & \text{si } x \text{ es discreta, donde } t \in (-h, h) \\ \int_{-\infty}^{+\infty} e^{tx} f(x) dx, & \text{si } x \text{ es continua, donde } t \in (-h, h) \end{cases}$$

además, se puede probar que:

$$\left. \frac{\partial M(t)}{\partial t} \right|_{t=0} = \mu \quad \text{y} \quad \sigma^2 = \left. \frac{\partial^2 M(t)}{\partial t^2} \right|_{t=0} - \left(\left. \frac{\partial M(t)}{\partial t} \right|_{t=0} \right)^2$$

Se ha tratado todo lo concerniente a las variables aleatorias, pero no se ha considerado el siguiente caso: que más de una variable aleatoria sea definida en un espacio muestral al mismo tiempo, lo cual da origen a las variables aleatorias conjuntas.

Variabes aleatorias conjuntas

En un experimento aleatorio pueden intervenir dos o más variables aleatorias, por lo que es necesario determinar la *distribución de*

probabilidad conjunta, si son discretas y la *función de densidad de probabilidad conjunta*, si son continuas.

La *distribución de probabilidad conjunta* de las variables aleatorias discretas X_1, X_2, \dots, X_p , es:

$f_{X_1, X_2, \dots, X_p}(x_1, x_2, \dots, x_p) = P(X_1 = x_1, X_2 = x_2, \dots, X_p = x_p)$
 para todos los puntos (x_1, x_2, \dots, x_p) en el rango de X_1, X_2, \dots, X_p .

donde:

$$\sum_{x_1} \sum_{x_2} \dots \sum_{x_p} f(x_1, \dots, x_p) = 1$$

La *función de densidad de probabilidad conjunta* de las variables aleatorias continuas X_1, X_2, \dots, X_p , se denota como

$f_{X_1, X_2, \dots, X_p}(x_1, x_2, \dots, x_p)$

donde:

$$\int_{-\infty}^{+\infty} f(x_1, \dots, x_p) dx_1, \dots, dx_p = 1$$

Es un hecho que los valores exactos de los parámetros poblacionales no son conocidos, por esa razón se emplean estimadores puntuales. Por ejemplo, los estimadores de la media y la varianza de una población dada, son la media aritmética y la varianza muestral, respectivamente.

Estimadores

Sea X_1, X_2, \dots, X_n , una muestra aleatoria de tamaño n con parámetro poblacional θ , entonces utilizando la información proporcionada por esta muestra aleatoria, se puede definir una función $\hat{\theta} : R^n \rightarrow R$ que no incluya al parámetro poblacional y que es denominada *estimador del parámetro poblacional* θ .

La importancia del teorema del límite central radica en que enuncia una aproximación para la distribución muestral de \bar{X} , que se puede utilizar para una población con una distribución cualquiera.

Teorema del límite central

Sean X_1, X_2, \dots, X_n variables aleatorias independientes e idénticamente distribuidas con media μ y varianza $\sigma^2 < \infty$, y sea

$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$, bajo estas condiciones la variable aleatoria:

$$U_n = \frac{\left(\frac{\sum X_i}{n} - \mu \right) \sqrt{n}}{\sigma}$$

converge en distribución a una variable aleatoria normal estándar.

Ahora que se ha enunciado el teorema del límite central, nos referiremos a los estimadores por intervalos, conocidos como intervalos de confianza, los cuales definiremos a continuación.

Intervalos de confianza

Una *estimación por intervalos* de un parámetro poblacional desconocido θ es un intervalo de la forma $i \leq \theta \leq s$, donde los puntos extremos i y s dependen del valor numérico del parámetro estimado $\hat{\theta}$ y de su distribución para una muestra dada. Dichos puntos reciben el nombre de *límites de confianza inferior y superior*, respectivamente.

La probabilidad de seleccionar una muestra que dé lugar a un intervalo que contenga el valor real del parámetro θ está dada por:

$$P(i \leq \theta \leq s) = 1 - \alpha \quad \text{donde } 0 < \alpha < 1$$

Si se cumplen las condiciones del teorema del límite central, el intervalo de $(1 - \alpha)100\%$ de confianza para un parámetro objetivo θ .

$$\hat{\theta} - Z_{\alpha/2}\sigma_{\hat{\theta}} \leq \theta \leq \hat{\theta} + Z_{\alpha/2}\sigma_{\hat{\theta}}$$

donde:

$Z_{\alpha/2}$ es el punto de la distribución normal estándar que corresponde al porcentaje $\alpha/2$, y

$\sigma_{\hat{\theta}}$ es la desviación estándar del estimador del parámetro objetivo θ .

Una vez que se han detallado los diferentes estimadores (puntuales y por intervalos), procedemos a definir lo que es una prueba de hipótesis.

Prueba de hipótesis

Una *hipótesis estadística* es una proposición sobre los parámetros o la distribución de una población. El procedimiento de toma de decisión sobre la hipótesis se conoce como *prueba de hipótesis*. Los elementos de una prueba estadística son: *hipótesis nula* H_0 , *hipótesis alterna* H_1 , *estadístico de la prueba* y la *región crítica o de rechazo*.

La *hipótesis nula* H_0 es aquella hipótesis que se desea probar, mientras que la *hipótesis alterna* H_1 es la hipótesis que debe

aceptarse en el caso de rechazar H_0 . El *estadístico de prueba* es una función de las mediciones muestrales y la *región crítica* es un subconjunto de R^n en el que se rechaza la hipótesis nula.

Hay dos tipos de errores que se puede cometer al contrastar hipótesis: el *error tipo I* se comete cuando se rechaza H_0 siendo esta verdadera. La probabilidad de error tipo I, llamada también *nivel de significancia* y se denota por α . El *error tipo II* se comete si se acepta H_0 cuando H_1 es verdadera. La probabilidad de error tipo II se denota por β .

Supóngase que se tiene el siguiente contraste de hipótesis:

$$H_0: \theta = \theta_0$$

vs.

$$H_1: \theta \neq \theta_0$$

donde θ es un parámetro cualesquiera. La potencia de la prueba estadística, denotada por $k(\theta)$, es la probabilidad de rechazar la hipótesis nula H_0 .

Para poder determinar la distribución más adecuada para las variables de estudio analizadas, se emplean las pruebas de bondad de ajuste. A continuación se explicará la prueba de Kolmogorov-Smirnov.

Prueba de bondad de ajuste: Kolmogorov-Smirnov

La *prueba de Kolmogorov-Smirnov* es una *prueba de bondad de ajuste*, esto es, permite determinar si la distribución propuesta es un modelo satisfactorio para la población estudiada, evaluando la concordancia entre los valores de la distribución empírica de la muestra y la función de distribución especificada. La prueba se basa en la *máxima diferencia absoluta* D entre los valores de la distribución acumulada empírica de una muestra aleatoria de tamaño n y una distribución acumulada teórica específica.

La prueba de hipótesis se establece así:

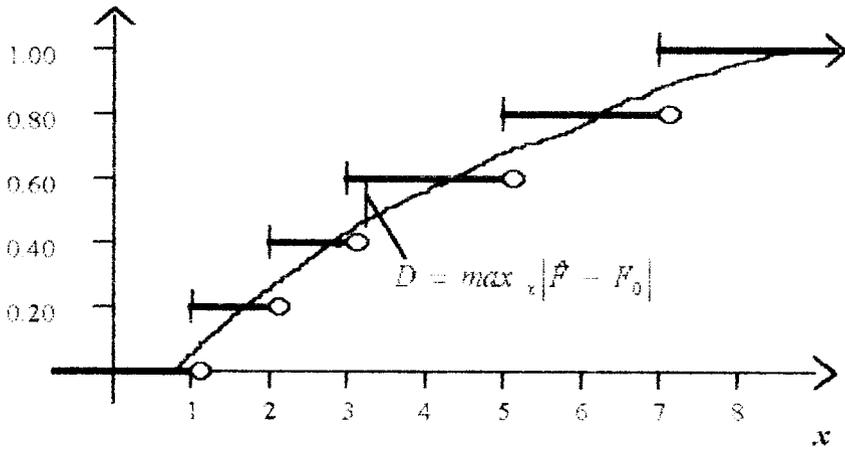
H_0 : La población muestreada tiene función de distribución $F_0(x)$

vs.

H_1 : No es verdad H_0 .

Figura 2.5

Gráfica correspondiente a la diferencia D



donde:

 $\hat{F}(x)$: Distribución empírica de la muestra $F_0(x)$: Función de distribución, postulada en la hipótesis nula H_0

2.2 Análisis multivariado

Vectores y matrices aleatorios

Sean X_1, X_2, \dots, X_p , p variables aleatorias, se define un vector aleatorio p -dimensional de la siguiente forma:

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} \in \mathbb{R}^p$$

Una vez definido vector aleatorio X , ahora trataremos lo que se conoce como matriz aleatoria. *Matriz aleatoria* es una matriz cuyos elementos son variables aleatorias, de allí que una matriz de datos sea una matriz aleatoria definida por p variables aleatorias o características ($p \geq 1$) y n observaciones registradas para cada variable. Por ejemplo, una encuesta conformada por p preguntas aplicadas a n personas. Cada fila representará las respuestas dadas por una persona y cada columna las respuestas de n personas a una misma pregunta. La *matriz de datos* resultante de orden $n \times p$ (n filas y p columnas) será de la siguiente forma:

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1k} & \cdots & X_{1p} \\ X_{21} & X_{22} & \cdots & X_{2k} & \cdots & X_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ X_{j1} & X_{j2} & \cdots & X_{jk} & \cdots & X_{jp} \\ \vdots & \vdots & & \vdots & & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nk} & \cdots & X_{np} \end{bmatrix}$$

X_{jk} : Representa el valor de la k -ésima variable en la j -ésima observación, ya que las filas representan las observaciones y las columnas a las variables.

Ya hemos definido los vectores y matrices aleatorios, ahora conoceremos los vectores de medias y las matrices de varianzas y covarianzas, elementos claves del análisis multivariado.

Vector de medias y matrices de varianzas y covarianzas

Sea $\mathbf{X}^T = [X_1, X_2, \dots, X_p]$ un vector aleatorio de orden $p \times 1$. Cada elemento X_i de \mathbf{X} es una variable aleatoria con su propia distribución de probabilidad. Entonces el *vector de medias* se define como:

$$\boldsymbol{\mu} = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{bmatrix}$$

La covarianza es una medida de la relación lineal que puede existir entre dos o más variables. Sean X_i y X_j variables aleatorias, la *covarianza* entre X_i y X_j está definida por:

$$\sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)]$$

Si X_i, X_j son variables aleatorias continuas con función de densidad conjunta $f_{ij}(x_i, x_j)$, la covarianza está dada por:

$$\sigma_{ij} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x_i - \mu_i)(x_j - \mu_j) f_{ij}(x_i, x_j) dx_i dx_j$$

Si X_i, X_j son variables aleatorias discretas con función de densidad conjunta $P(X_i = x_i, X_j = x_j) = f_{ij}(x_i, x_j)$, la covarianza está dada por:

$$\sigma_{ij} = \sum_i \sum_j (x_i - \mu_i)(x_j - \mu_j) f_{ij}(x_i, x_j)$$

donde :

$\mu_i = E[X_i]$ y $\mu_j = E[X_j]$; $i, j = 1, 2, \dots, p$; son respectivamente las medias de las variables aleatorias X_i y X_j . Cuando $i = j$, la covarianza de X_i, X_j $\text{cov}(X_i, X_j)$ se convierte en la varianza de X_i $\text{Var}(X_i)$.

Hemos detallado como se obtiene la covarianza entre dos variables aleatorias, sean estas discretas o continuas, ahora procederemos a conocer lo que es una matriz de varianzas y covarianzas.

La *matriz de varianzas y covarianzas* está dada por $\Sigma = E[(X - \mu)(X - \mu)^T]$

donde:

$$\mu = E[\mathbf{X}] = \begin{bmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_p] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$$

$$\Sigma = E \left[\begin{bmatrix} (X_1 - \mu_1) \\ (X_2 - \mu_2) \\ \vdots \\ (X_p - \mu_p) \end{bmatrix} E \left[(X_1 - \mu_1) \quad (X_2 - \mu_2) \quad \cdots \quad (X_p - \mu_p) \right] \right]$$

$$\Sigma = \begin{bmatrix} E(X_1 - \mu_1)^2 & E(X_1 - \mu_1)(X_2 - \mu_2) & \cdots & E(X_1 - \mu_1)(X_p - \mu_p) \\ E(X_2 - \mu_2)(X_1 - \mu_1) & E(X_2 - \mu_2)^2 & \cdots & E(X_2 - \mu_2)(X_p - \mu_p) \\ \vdots & \vdots & \ddots & \vdots \\ E(X_p - \mu_p)(X_1 - \mu_1) & E(X_p - \mu_p)(X_2 - \mu_2) & \cdots & E(X_p - \mu_p)^2 \end{bmatrix}$$

$$\Sigma = \text{Cov}(\mathbf{X}) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_{pp} \end{bmatrix}$$

donde $\sigma_{ii} = \sigma_i^2 = \text{Var}(X_i) \quad i = 1, 2, \dots, p$

La estandarización de las covarianzas da lugar a las correlaciones. El *coeficiente de correlación* ρ_{ij} entre X_i y X_j está definido de la siguiente forma:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}} \sqrt{\sigma_{jj}}}$$

Además se puede probar que:

$$-1 \leq \rho_{ij} \leq 1$$

A partir de la matriz de varianzas y covarianzas, se obtiene la *matriz de correlación* ρ , la cual se define de la siguiente forma:

$$\rho = \begin{bmatrix} \frac{\sigma_{11}}{\sqrt{\sigma_{11}} \sqrt{\sigma_{11}}} & \frac{\sigma_{12}}{\sqrt{\sigma_{11}} \sqrt{\sigma_{22}}} & \dots & \frac{\sigma_{1p}}{\sqrt{\sigma_{11}} \sqrt{\sigma_{pp}}} \\ \frac{\sigma_{12}}{\sqrt{\sigma_{11}} \sqrt{\sigma_{22}}} & \frac{\sigma_{22}}{\sqrt{\sigma_{22}} \sqrt{\sigma_{22}}} & \dots & \frac{\sigma_{2p}}{\sqrt{\sigma_{22}} \sqrt{\sigma_{pp}}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{1p}}{\sqrt{\sigma_{11}} \sqrt{\sigma_{pp}}} & \frac{\sigma_{2p}}{\sqrt{\sigma_{22}} \sqrt{\sigma_{pp}}} & \dots & \frac{\sigma_{pp}}{\sqrt{\sigma_{pp}} \sqrt{\sigma_{pp}}} \end{bmatrix}$$

$$\rho = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{12} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1p} & \rho_{2p} & \cdots & 1 \end{bmatrix}$$

Ya se ha descrito todo lo referente a las matrices de varianzas-covarianzas y de correlación, a continuación trataremos la independencia de variables aleatorias.

Independencia de variables aleatorias

Sean X_i y X_j variables aleatorias, las mismas son *independientes* cuando su *función de densidad conjunta* $f_{ij}(x_i, x_j)$ puede ser expresada como el producto de sus funciones de *densidad marginales* $f_i(x_i)$ y $f_j(x_j)$, esto es,

$$f_{ij}(x_i, x_j) = f_i(x_i) f_j(x_j) \text{ para todos los pares } (x_i, x_j) \in \mathbb{R}^2$$

Sean X_1, X_2, \dots, X_p , p variables aleatorias continuas, se dice que las mismas son *independientes* si su *función de densidad conjunta* puede ser factorizada como:

$$f_{12 \dots p}(x_1, x_2, \dots, x_p) = f_1(x_1)f_2(x_2) \dots f_p(x_p)$$

La independencia estadística establece una importante implicación para la covarianza. Si X_i y X_j son independientes entonces la $\text{cov}(X_i, X_j) = 0$, lo cual se puede demostrar de la siguiente forma:

$$\sigma_{ij} = E[(X_i - \mu_i)(X_j - \mu_j)] = E(X_i X_j) - \mu_i \mu_j$$

Si existe independencia de variables, entonces $E(X_i X_j) = E(X_i)E(X_j)$. Por lo tanto,

$$\sigma_{ij} = E(X_i)E(X_j) - \mu_i \mu_j = \mu_i \mu_j - \mu_i \mu_j = 0.$$

No es cierto que si la $\text{cov}(X_i, X_j) = 0$, entonces X_i y X_j son independientes.

Por ejemplo, si tenemos las variables X : número de hijos y Y : ingreso mensual, y su covarianza $\text{cov}(X, Y)$, resulta igual a cero, no se puede afirmar que el número de hijos y el ingreso anual son variables independientes.

En el análisis multivariado ocurre lo mismo que en el univariado, nunca se conocen los verdaderos parámetros sino que se emplean

estimadores. Por esa razón, vamos a tratar los estimadores del vector de medias y de la matriz de varianzas y covarianzas.

Estimador del vector de medias

Sea el vector aleatorio $X^T = [X_1, X_2, \dots, X_p]$, cuyo vector de medias poblacionales es μ y matriz de covarianzas es Σ . Entonces \bar{X} es un estimador insesgado de μ y su matriz de covarianzas es Σ/n , es decir que:

$$E[\bar{X}] = \mu$$

$$\text{Cov}[\bar{X}] = \Sigma/n$$

Sea $\bar{X} = (X_1 + X_2 + \dots + X_n) / n$ entonces :

$$\begin{aligned} E[\bar{X}] &= E(X_1/n + X_2/n + \dots + X_n/n) \\ &= E(X_1/n) + E(X_2/n) + \dots + E(X_n/n) \\ &= (1/n)E(X_1) + (1/n)E(X_2) + \dots + (1/n)E(X_n) \\ &= \mu/n + \mu/n + \dots + \mu/n \\ &= \mu \end{aligned}$$

por lo tanto la media muestral \bar{X} es un estimador insesgado de la media poblacional μ .

Varianza generalizada

Para medir la dispersión de una variable se emplea la *varianza muestral*. Cuando son p variables, la *matriz de varianzas y covarianzas de la población* es estimada por la *matriz de varianzas y covarianzas muestral* $\mathbf{S} = \hat{\Sigma}$, que se define como:

$$\mathbf{S} = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{12} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{1p} & S_{2p} & \cdots & S_{pp} \end{bmatrix} = (S_{ik}) \in M_{pp}$$

$$\text{siendo } S_{ik} = \frac{1}{n-1} \sum_{j=1}^n (X_{j,i} - \bar{X}_i)(X_{j,k} - \bar{X}_k)$$

$$i, k = 1, 2, \dots, p$$

La matriz $\hat{\Sigma}$ contiene p varianzas y $\frac{p(p-1)}{2}$ covarianzas. A veces es necesario determinar un solo valor numérico, al igual que la usual varianza muestral para una variable, que describa la variación expresada por \mathbf{S} . Este valor está dado por el *determinante de la matriz \mathbf{S}* y se lo denomina *varianza muestral generalizada*.

Una vez que se han definido los estimadores del vector de medias y de la matriz de varianzas y covarianzas, trataremos la técnica multivariada conocida como componentes principales.

Componentes principales

El análisis de componentes principales, es una variación del *análisis de factores*, permite explicar la matriz de varianzas y covarianzas de un conjunto de variables, a través de las combinaciones lineales de variables observables. Los objetivos son: reducción de datos e interpretación.

Aunque se requieren las p componentes para reproducir la variabilidad total del sistema, a menudo es suficiente considerar un número k de componentes principales tal que $k < p$, ya que contienen gran cantidad de información requerida, con lo cual se logra reducir los datos. Un análisis de componentes principales permite descubrir relaciones entre variables que muchas veces no son consideradas y por ende afecta a nuestras interpretaciones finales. Este análisis no supone que las poblaciones consideradas son normales p -variadas.

Algebraicamente, las componentes principales son combinaciones lineales de p variables aleatorias X_1, X_2, \dots, X_p , que se obtienen a partir de la matriz de covarianzas Σ o de la matriz de correlación ρ de X_1, X_2, \dots, X_p .

Sea el vector aleatorio $\mathbf{X}^T = [X_1, X_2, \dots, X_p]$ que tiene la matriz de covarianzas Σ con valores propios (eigenvalores) $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, obtenidos a partir de la ecuación:

$$\det(\Sigma - \lambda \mathbf{I}) = 0$$

Además definamos p variables no observables Y_1, Y_2, \dots, Y_p , como una combinación lineal de X_1, X_2, \dots, X_p , así:

$$\begin{aligned} Y_1 &= \mathbf{b}_1^T \mathbf{X} = b_{11}X_1 + b_{12}X_2 + \dots + b_{1p}X_p \\ Y_2 &= \mathbf{b}_2^T \mathbf{X} = b_{21}X_1 + b_{22}X_2 + \dots + b_{2p}X_p \\ &\vdots \\ Y_p &= \mathbf{b}_p^T \mathbf{X} = b_{p1}X_1 + b_{p2}X_2 + \dots + b_{pp}X_p \end{aligned}$$

donde la varianza y la covarianza de cada una de las combinaciones lineales está dada por:

$$\text{Var}(Y_i) = \text{Var}(\mathbf{b}_i^T \mathbf{X}) = \mathbf{b}_i^T \Sigma \mathbf{b}_i \quad i = 1, 2, \dots, p$$

$$\text{Cov}(Y_i, Y_k) = \mathbf{b}_i^T \Sigma \mathbf{b}_k \quad i, k = 1, 2, \dots, p$$

Los componentes principales son aquellas combinaciones lineales no correlacionadas Y_1, Y_2, \dots, Y_p , cuyas varianzas son lo más grandes

posible. La primera componente es la combinación lineal con máxima varianza, cumpliéndose que:

$$\text{Var}(Y_1) \geq \text{Var}(Y_2) \geq \dots \geq \text{Var}(Y_p)$$

Se puede probar que los vectores \mathbf{b} 's son los vectores propios de Σ , los cuales se caracterizan por ser ortogonales y unitarios, esto es:

$$\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0 \text{ para } i \neq j \text{ y } \|\mathbf{b}_i\| = 1$$

$$Y_1 = \mathbf{b}_1^T \mathbf{X} \text{ maximiza } \text{Var}(\mathbf{b}_1^T \mathbf{X}) \text{ sujeto a } \langle \mathbf{b}_1, \mathbf{b}_1 \rangle = 1$$

$$Y_2 = \mathbf{b}_2^T \mathbf{X} \text{ maximiza } \text{Var}(\mathbf{b}_2^T \mathbf{X}) \text{ sujeto a } \langle \mathbf{b}_2, \mathbf{b}_2 \rangle = 1 \text{ y } \text{Cov}(\mathbf{b}_1^T \mathbf{X}, \mathbf{b}_2^T \mathbf{X}) = 0$$

⋮

$$Y_k = \mathbf{b}_k^T \mathbf{X} \text{ maximiza } \text{Var}(\mathbf{b}_k^T \mathbf{X}) \text{ sujeto a } \langle \mathbf{b}_k, \mathbf{b}_k \rangle = 1 \text{ y } \text{Cov}(\mathbf{b}_i^T \mathbf{X}, \mathbf{b}_k^T \mathbf{X}) = 0 \text{ para } k < i$$

donde la varianza de cada componente es igual al correspondiente valor propio, es decir que se puede probar que:

$$\text{Var}(Y_i) = \lambda_i = \mathbf{b}_i^T \Sigma \mathbf{b}_i \quad i=1, 2, \dots, p$$

$$\text{Varianza total} = \lambda_1 + \lambda_2 + \dots + \lambda_p$$

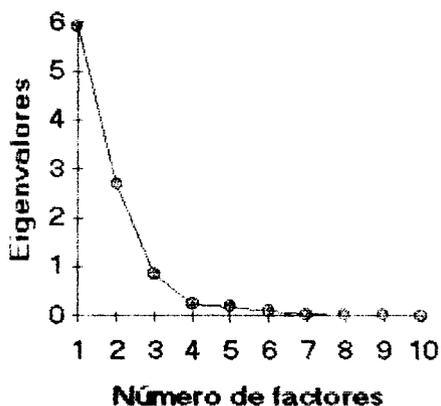
La proporción de la varianza total explicada por la i -ésima componente principal está dada por:

$$\frac{\lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \quad i=1,2,\dots,p$$

Para determinar el número de componentes principales a seleccionar, se debe considerar la cantidad de varianza total explicada por los componentes. Una ayuda visual para determinar el número de componentes adecuados, es la gráfica de los eigenvalores ordenados descendientemente (Ver figura 2.6).

Figura 2.6

Valores propios (eigenvalores)



Componentes principales obtenidas de variables estandarizadas

Dadas X_1, X_2, \dots, X_p , con medias $\mu_1, \mu_2, \dots, \mu_p$, las componentes principales pueden ser obtenidas también de variables estandarizadas,

$$\begin{aligned} Z_1 &= \frac{(X_1 - \mu_1)}{\sqrt{\sigma_{11}}} \\ Z_2 &= \frac{(X_2 - \mu_2)}{\sqrt{\sigma_{22}}} \\ &\vdots \\ Z_p &= \frac{(X_p - \mu_p)}{\sqrt{\sigma_{pp}}} \end{aligned}$$

Las componentes principales de \mathbf{Z} pueden ser obtenidas utilizando los vectores propios de la matriz de correlación ρ de \mathbf{X} . Los resultados son similares a los anteriores, empleando la matriz de covarianzas Σ , con la diferencia de que son más simplificados, debido a que la varianza de cada Z_i es equivalente a la unidad. La notación se mantiene, Y_i representa la i -ésima componente y $(\lambda_i, \mathbf{e}_i)$ para los valores y vectores propios de Σ o ρ , los cuales difieren entre sí. Las componentes de variables estandarizadas son de la forma:

$$\begin{aligned} Y_1 &= \mathbf{e}_1^T \mathbf{Z} = e_{11} Z_1 + e_{12} Z_2 + \dots + e_{1p} Z_p \\ Y_2 &= \mathbf{e}_2^T \mathbf{Z} = e_{21} Z_1 + e_{22} Z_2 + \dots + e_{2p} Z_p \\ &\vdots \\ Y_p &= \mathbf{e}_p^T \mathbf{Z} = e_{p1} Z_1 + e_{p2} Z_2 + \dots + e_{pp} Z_p \end{aligned}$$

Además, la varianza total de la población es simplemente p , la suma de los elementos de la diagonal de la matriz de correlación ρ , y la proporción de la varianza total explicada por la k -ésima componente principal de Z es:

$$\lambda_k / p \quad k = 1, 2, \dots, p$$

donde $\lambda_1, \lambda_2, \dots, \lambda_p$ representan los valores propios de ρ .

Rotación de factores

La *rotación de factores* o transformación ortogonal de los factores de carga, corresponde a la rotación rígida o reflexión de los ejes coordenados (noventa grados), lo cual incide en los factores resultantes. Se la emplea cuando las cargas originales muestran resultados difíciles de interpretar y es necesario rotarlas para poder mejorar el enfoque de dichas cargas, y así determinar los elementos constitutivos de los factores. También puede aplicarse la *rotación oblicua*, que corresponde a la rotación no rígida de los ejes coordenados, es decir diferente a noventa grados.



2.3 Series temporales

Para poder emplear un modelo de *análisis de series temporales* es necesario referirnos a qué son los *procesos estocásticos*. Un *proceso estocástico* es una familia de variables aleatorias $\{X_t\}$ donde t es un punto en un espacio T , llamado espacio parametral y donde para cada $t \in T$; X_t es un punto en un espacio S , llamado espacio de estados. Un ejemplo de un proceso estocástico es la *caminata aleatoria*, en la cual cada cambio sucesivo de la variable está ligado al valor previo y una distribución de probabilidad con media cero. La variable X_t entonces tomará valores de acuerdo a:

$$X_t = X_{t-1} + U_t$$

donde U_t es una variable aleatoria con media cero y que se registra independientemente cada período.

Como ya se ha presentado la definición de proceso estocástico, procederemos a tratar lo que se conoce como serie temporal. Las series temporales son procesos estocásticos $\{X_t\}$, donde t pertenece a Z que representa al conjunto de los números enteros. Están constituidas de observaciones registradas en intervalos regulares de tiempo.

Para comprender el análisis de series temporales es importante conocer el concepto de *estacionaridad*, la cual es una propiedad que requiere que la distribución conjunta sea invariante en el tiempo t , es decir,

$$\text{distribución } (X_1, \dots, X_{t+k}) = \text{distribución } (X_{t+m}, \dots, X_{t+k+m})$$

donde t es algún punto en el tiempo y k y m son enteros.

Un proceso estacionario $\{X_t\}$, donde t pertenece al conjunto de los números enteros Z , es *débilmente estacionario* si cumple las siguientes propiedades:

1. $E(X_t^2) < M < \infty$
2. $E(X_t) = \mu, \forall t \in Z$
3. $Cov(X_t, X_{t+h}) = \gamma_h, \forall h \in Z$

Los modelos para series de tiempo estacionarias se caracterizan por ser procesos estocásticos lineales. Un proceso estocástico se dice lineal si puede expresarse como una secuencia de ruidos aleatorios U_t , de la siguiente manera:

$$X_t = \sum_{j=0}^{\infty} a_j U_{t-j} \quad t \in \mathbb{Z}$$

$$\text{donde } a_0 = 1, \sum_{j=0}^{\infty} |a_j| < \infty$$

Estos ruidos aleatorios independientes e idénticamente distribuidos, se denominan *ruidos blancos*, los cuales cumplen las siguientes características:

1. $\text{Var}(X_t) = \sigma^2 < \infty$
2. $E(X_t) = 0 \quad \forall t \in \mathbb{Z}$
3. $\text{Cov}(X_t, X_{t+h}) = \begin{cases} \sigma^2; h \neq 0 \\ 0; h = 0 \end{cases}$

La función de autocorrelaciones permite analizar la relación existente entre las observaciones de la serie temporal y ayuda a determinar los modelos de predicción más adecuados. Las *autocorrelaciones*, son el resultado de la estandarización de las autocovarianzas, que indican la relación lineal existente entre las observaciones, al ser estas divididas para la varianza γ_0 .

La correlación entre Z_t y Z_{t+j} se denota por ρ_j , por lo tanto al conjunto de autocorrelaciones se lo denominará *función de autocorrelaciones*, la cual está dada por:

$$\rho_0 = \frac{\gamma_0}{\gamma_0} = 1$$

$$\rho_1 = \frac{\gamma_1}{\gamma_0}$$

$$\rho_2 = \frac{\gamma_2}{\gamma_0}$$

...

Las autocovarianzas y autocorrelaciones están relacionadas con las covarianzas y correlaciones (ver página 90). Las covarianzas y correlaciones establecen la relación lineal entre dos o más variables, sin considerar al tiempo, mientras que las autocovarianzas y autocorrelaciones indican la relación lineal existente entre diferentes observaciones en la misma serie temporal.

Un estimador de la función de autocorrelaciones ρ_j , denotado por c_j , está dado por:

$$c_j = \frac{1}{T} \sum_{t=1}^{T-j} [(Z_t - \bar{Z})(Z_{t+j} - \bar{Z})] \quad j = 1, 2, \dots$$

donde

$$T = \text{total de observaciones}, \quad \bar{Z} = \frac{1}{T} \sum_{t=1}^T Z_t$$

Ya se han tratado los conocimientos básicos sobre series temporales, a continuación vamos a definir los modelos para series de tiempo estacionarias, los cuales son: *media móvil MA*, *autoregresivo AR* y la *combinación autoregresivo - media móvil ARMA*. Estos modelos se caracterizan por ser *procesos estocásticos lineales discretos*.

Procesos media móvil

Un proceso estocástico $\{X_t\}$, donde t pertenece al conjunto de los números enteros Z , es un proceso media móvil de orden q , que se denota $MA(q)$, si:

$$X_t = \sum_{j=0}^q \theta_j U_{t-j} \quad t \in Z$$

donde $\theta_0 = 1$, $\theta_j \neq 0$, $\{U_t\}$ es un ruido blanco.

En el cuadro XX, se presentarán los valores de un proceso media móvil de orden 1 con parámetro $\theta_1 = 0.2$ y los ruidos blancos U_t son normales con media 0 y varianza 1. La gráfica resultante para catorce valores de t , se muestra en la figura 2.7.

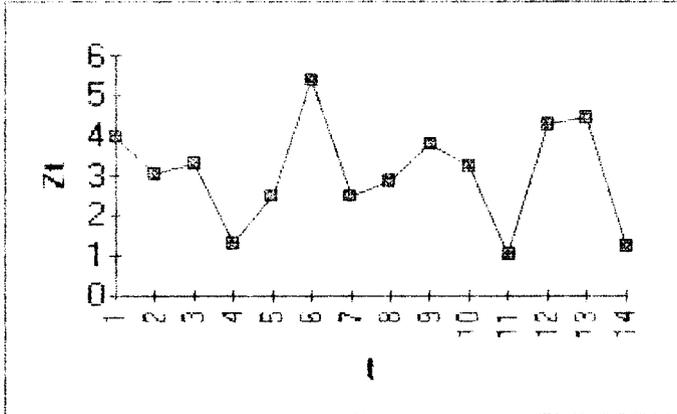
Cuadro XX

$$Z_t = U_t - 0.2U_{t-1}$$

t	Z_t	U_t
0	-	-0.723
1	3.951	0.8058
2	3.063	0.2246
3	3.32	0.3645
4	1.288	-1.639
5	2.506	-0.822
6	5.395	2.2309
7	2.47	-0.084
8	2.887	-0.129
9	3.784	0.7586
10	3.26	0.4113
11	1.042	-1.875
12	4.304	0.9294
13	4.462	1.6482
14	1.219	-1.452

Figura 2.7

Proceso (simulado) media-móvil de orden 1, $Z_t = U_t - 0.2U_{t-1}$



Procesos autoregresivos

Un proceso estocástico $\{X_t\}$, donde t pertenece al conjunto de los números enteros Z , es un proceso autoregresivo de orden p si:

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + U_t = \sum_{j=0}^p \phi_j X_{t-j} + U_j$$

donde $\phi_p \neq 0$, $\{U_t\}$ es un ruido blanco.

En el cuadro XXI, se presentarán los datos de un proceso autoregresivo de orden 1 con parámetro $\phi_1 = 0.8$ y los ruidos blancos U_t son normales con media 0 y varianza 1. La gráfica resultante para catorce valores de t , se presenta en la figura 2.8.

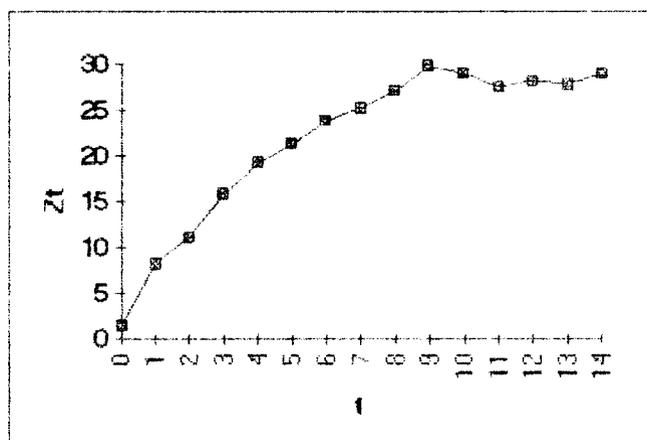
Cuadro XXI

$$Z_t = 0.8 Z_{t-1} + U_t$$

t	Z_t	U_t
0	1.526	-0.141
1	8.1906	0.9698
2	11.113	-1.439
3	15.713	0.8225
4	19.164	0.5936
5	21.319	-0.013
6	23.717	0.6618
7	25.269	0.2951
8	27.048	0.8331
9	29.706	2.0673
10	28.872	-0.892
11	27.411	-1.687
12	28.167	0.239
13	27.706	-0.828
14	28.969	0.8039

Figura 2.8

Proceso autoregresivo (simulado) de orden 1, $Z_t = 0.8 Z_{t-1} + U_t$



Proceso autoregresivo media móvil

Un proceso estocástico $\{X_t\}$, donde t pertenece al conjunto de números enteros Z , de la forma *autoregresiva y media móvil de órdenes p y q* conocidos como *ARMA(p, q)* presentan la siguiente forma:

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \delta + U_t - \theta_1 U_{t-1} - \dots - \theta_q U_{t-q}$$

La condición de estacionaridad se satisface si las raíces de la *ecuación característica* $(1 - \phi_1 B - \dots - \phi_p B^p) = 0$ se encuentran dentro del círculo unidad.

Una vez definidos los modelos para series de tiempo estacionarias, pasamos a tratar lo concerniente a la función de autocorrelaciones parciales. Para la estimación de las *autocorrelaciones parciales* se consideran las *ecuaciones de Yule-Walker* para el proceso autoregresivo de orden p , denotado $AR(p)$, que están dadas por:

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ &\dots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \end{aligned}$$

Si se conoce p y los valores de ρ_1, \dots, ρ_p , se podría resolver este sistema de p ecuaciones y p incógnitas ϕ_1, \dots, ϕ_p . Un estimador de ϕ_1

podría obtenerse solucionando la ecuación de Yule-Walker para $p = 1$, en la cual ρ_1 es reemplazado por el estimador r_1 , es decir:

$$r_1 = \hat{\phi}_1$$

donde $\hat{\phi}_1$ es el estimador resultante de ϕ_1 . Ahora si ϕ_1 es diferente de cero, podríamos concluir que estamos tratando con un proceso de al menos orden 1. Para determinar si el proceso es de orden 2 o más, podríamos resolver las ecuaciones de Yule-Walker para $p = 2$,

$$\begin{aligned} r_1 &= \hat{\phi}_1 + \hat{\phi}_2 r_1 \\ r_2 &= \hat{\phi}_1 r_1 + \hat{\phi}_2 \end{aligned}$$

Si el resultado estimado $\hat{\phi}_2$ difiere de cero, se puede concluir que el proceso es al menos de orden 2. Ahora se supondrá que el proceso se repite para valores grandes de p . Si denotamos el valor de $\hat{\phi}_j$ como $\hat{\phi}_{j|j}$, que forma parte de la solución del sistema para $p = j$, entonces los valores $\hat{\phi}_{j|j}$ son considerados como las autocorrelaciones parciales estimadas del proceso.

Se ha analizado todo lo concerniente a los modelos para series de tiempo estacionarias, a continuación se definirán los modelos para series de tiempo no estacionarias.

La *diferenciación* es una forma de transformar una serie no estacionaria a estacionaria. Consideremos el proceso autoregresivo AR(1) :

$$Z_t = \phi_1 Z_{t-1} + U_t$$

y los valores que podría tomar el *parámetro* ϕ_1 . Si el *valor absoluto* de ϕ_1 es menor que 1, entonces el proceso es *estacionario*, caso contrario las series crecen *exponencialmente*, es decir no son estacionarias. La *no estacionaridad es homogénea* porque las diferencias en el proceso es invariable, es decir que las series de tiempo de diferenciaciones son estacionarias porque las diferencias se denotan como:

$$X_t - X_{t-1} = U_t$$

y la distribución de U_t es fija. Además si definimos a W_t como la *sucesión de diferencias*:

$$W_t = X_t - X_{t-1}$$

entonces el *modelo general* puede ser escrito:

$$W_t = \phi_1 W_{t-1} + \dots + \phi_p W_{t-p} + U_t - \theta_1 U_{t-1} - \dots - \theta_q U_{t-q}$$

En algunos casos las primeras diferenciaciones no son estacionarias pero las segundas diferenciaciones si lo son. Las segundas

diferencias son las diferencias de las primeras diferenciaciones, es decir que si Y_t son las segundas diferencias de Z_t , entonces

$$\begin{aligned} Y_t &= W_t - W_{t-1} \\ &= (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) = X_t - 2X_{t-1} + X_{t-2} \end{aligned}$$

Se denota el *grado de diferenciación* por d , entonces un modelo $ARIMA(p, d, q)$ es un *proceso* que después de diferenciarlo d veces se transforma en un proceso $ARMA(p, q)$.

El parámetro p estará dado por el número de períodos, en que los primeros valores de las autocorrelaciones parciales sobrepasen el intervalo de confianza, mientras que el parámetro q estará dado por el número de períodos, en que los primeros valores de las autocorrelaciones sobrepasen el intervalo de confianza correspondiente.

Si la serie no es estacionaria, se producen autocorrelaciones muestrales con valores altos durante muchos períodos o pueden presentar proximidad entre sí, los primeros valores de la función. Es importante mantener el más bajo nivel posible de diferenciación para obtener un modelo estacionario, ya que al sobrediferenciar la serie se incrementaría innecesariamente la varianza del modelo.

Una vez que se han definido los modelos para series de tiempo no estacionarias, pasaremos a tratar los modelos para series de tiempo estacionales.

La estacionalidad es la tendencia a repetir un modelo de comportamiento cada cierto periodo. Las series estacionales se caracterizan por revelar una fuerte correlación en el período estacional (trimestres, semestres, años, etc). Para modelar las series con esta clase de comportamiento se emplean los modelos ARIMA estacionarios $SARIMA_S [(p,d,q),(P,D,Q)]$.

La identificación de los parámetros estacionales P,D,Q , se la realiza en forma similar a la de los parámetros estacionarios p,d,q , esto es, observando las gráficas de las autocorrelaciones y las autocorrelaciones parciales. Para determinar el parámetro P , se verificará el número de períodos en que las autocorrelaciones parciales sobrepasen el intervalo de confianza alrededor de los períodos múltiplos de s . En el caso de Q , se verificará el número de períodos en que las autocorrelaciones sobrepasen el intervalo de confianza alrededor de los períodos múltiplos de s . Los procesos estacionales se subdividen en procesos estacionales media móvil y autoregresivos.

Considérese un proceso media móvil cuyo orden es un entero múltiplo de s , y los coeficientes diferentes de cero son aquellos cuyos subíndices son enteros múltiplos de s . Este proceso tiene la siguiente apariencia:

$$Z_t = U_t - \theta_s U_{t-s} - \dots - \theta_{Qs} U_{t-Qs}$$

donde Q es el mayor múltiplo de s , el orden del proceso es Qs . Para distinguir el *proceso estacional* del *proceso MA*, se adoptará la notación:

$$\Delta_j = \theta_{js}$$

por lo que la ecuación del proceso estacional MA de orden Q se presenta así:

$$Z_t = U_t - \Delta_1 U_{t-s} - \dots - \Delta_Q U_{t-Qs}$$

En un proceso estacional MA la función de autocorrelación será diferente de cero solamente en los periodos $s, 2s, \dots, Qs$. Es importante comprender que la *función de autocorrelación* implica que una observación dada está correlacionada solamente con observaciones futuras y anteriores por 12 periodos, en el caso anual.

En general, para un proceso estacional MA de orden Q , las correlaciones se dan solamente para los períodos estacionales Q .

El *proceso estacional autoregresivo AR* es de la forma:

$$Z_t = \phi_s Z_{t-s} + \dots + \phi_{P_s} Z_{t-P_s} + U_t$$

donde P es el mayor múltiplo de s presente en el modelo. Es conveniente establecer una notación especial para el modelo estacional,

$$\phi_{js} = \Gamma_j$$

entonces el modelo estacional AR de orden P se transforma en:

$$Z_t = \Gamma_1 Z_{t-s} + \dots + \Gamma_P Z_{t-P_s} + U_t$$

Las autocovarianzas son diferentes de cero solamente en los períodos que son enteros múltiplos de s .

Se han explicado los diferentes modelos empleados para las series temporales estacionarias, no estacionarias y estacionales. Ahora se procederá a detallar la forma de seleccionar el modelo adecuado.

Una vez que se ha optado por un modelo a utilizar, se debe verificar que los parámetros propuestos sean adecuados, es decir que sus correspondientes intervalos de confianza no incluyan al cero, ya que alguno de los parámetros podrían tomar el valor de cero, lo cual afectaría a la efectividad del modelo. Si el modelo supera la prueba de parámetros, éste debe ser sometido al análisis de residuos (ruidos blancos), el cual indicará si debe modificarse el número de parámetros del modelo.

Si más de un modelo propuesto supera las pruebas de parámetros y residuos, se seleccionará el modelo final de acuerdo a dos criterios: varianza y predicción. Si se toma como criterio la varianza, se escogerá el modelo que presente el mínimo valor de varianza. Si se considera el criterio de predicción, se escogerá el modelo, cuya predicción, se ajuste mejor a la serie real.

2.4 Modelos de regresión

El *análisis de regresión* permite modelar el comportamiento de una variable cuantitativa que se relaciona o depende de otras variables, denominadas *variables de explicación*.

El modelo probabilístico de la forma:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \epsilon_i \sim N(0, \sigma^2), \text{cov}(\epsilon_i, \epsilon_j) = 0 \text{ para } i \neq j \\ i = 1, 2, \dots, n$$

donde ϵ_i representa el error aleatorio, recibe el nombre de *modelo de regresión lineal simple*, tiene una variable independiente o *de explicación* X y la variable dependiente o *respuesta* Y . Además, se supone que el error aleatorio ϵ_i tiene media cero y varianza σ^2 , y que los errores aleatorios correspondientes a diferentes observaciones, son variables aleatorias no correlacionadas, es decir $\text{cov}(\epsilon_i, \epsilon_j) = 0$; $i \neq j$.

Supóngase que la relación entre Y y X es una recta, lo cual es una regresión lineal simple. El valor esperado de Y para cada valor x de X es

$$E(Y / X = x) = \beta_0 + \beta_1 x$$

donde la ordenada al origen β_0 y la pendiente β_1 son los coeficientes desconocidos de la regresión.

El método que estima los coeficientes β_0 y β_1 , minimizando las sumas de los cuadrados de las diferencias entre los valores observados y los estimados, se lo denomina *método de mínimos cuadrados*.

El modelo ideal es $Y = \beta_0 + \beta_1 x$, pero como existe error, cada observación se la expresa mediante la ecuación:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, 2, \dots, n$$

y la suma de los cuadrados de las diferencias entre los valores observados y_i y los estimados por la recta de regresión es

$$L = \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2$$

Si minimizamos L , se obtienen las siguientes ecuaciones:

$$\left. \frac{\partial L}{\partial \beta_0} \right|_{\beta_0, \beta_1} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\left. \frac{\partial L}{\partial \beta_1} \right|_{\beta_0, \beta_1} = -2 \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i) x_i = 0$$

que al ser simplificadas, dan origen a las *ecuaciones normales de mínimos cuadrados*, cuyo desarrollo, da como resultado a los estimadores de mínimos cuadrados $\hat{\beta}_0$ y $\hat{\beta}_1$.

Las ecuaciones normales son las siguientes:

$$n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n x_i = \sum_{i=1}^n y_i$$

$$\hat{\beta}_0 \sum_{i=1}^n x_i + \hat{\beta}_1 \sum_{i=1}^n x_i^2 = \sum_{i=1}^n y_i x_i$$

Las *estimaciones de mínimos cuadrados* de la ordenada al origen y la pendiente del modelo de regresión lineal simple son:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - \left(\sum_{i=1}^n y_i \right) \left(\sum_{i=1}^n x_i \right) / n}{\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 / n}$$

donde $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$, $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$

Así, la línea de *regresión estimada o ajustada* es:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

donde las observaciones correspondientes a las variables Y y X cumplirán la relación:

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + e_i, \quad i = 1, 2, \dots, n$$

donde $e_i = y_i - \hat{y}_i$ recibe el nombre de *residuo*. El residuo representa el error en el ajuste del modelo en la i-ésima observación y_i .

Se puede probar que $\hat{\beta}_0$ y $\hat{\beta}_1$ son *estimadores insesgados* de la ordenada al origen β_0 y de la pendiente β_1 . Además, se puede demostrar que la covarianza de las variables aleatorias $\hat{\beta}_0$ y $\hat{\beta}_1$ es diferente de cero.

Una vez que se han explicado los estimadores de mínimos cuadrados de los coeficientes del modelo de regresión lineal simple, pasaremos a tratar lo concerniente al análisis de varianza.

El *análisis de varianza* determina la significancia de una regresión, ya que particiona la variabilidad total en la variable de respuesta Y: la *variabilidad considerada por la recta de regresión* y la *variación residual que queda sin explicar por la recta*, se denotan por SCReg y SError, respectivamente. La suma de SCReg y SError se denomina suma total de cuadrados de Y y se denota SCTotal.

$$SCTotal = SCReg + SCError$$

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Tabla de análisis de varianza para regresiones lineales simples

Fuente de variación	Grados de libertad	Sumas cuadráticas	Medias cuadráticas	F
Regresión	1	SCReg	MReg	MReg/MError
Error	n-2	SCError	MError	
Total	n-1	SCTotal		

La suma total de cuadrados SCTotal tiene (n-1) grados de libertad, la suma cuadrática de la regresión SCReg y la suma cuadrática del error SCError tienen 1 y (n-2) grados de libertad, respectivamente, para regresiones lineales simples. También se puede demostrar que:

$$\hat{\sigma}^2 = SCError / (n - 2)$$

Las *medias cuadráticas* son el resultado de dividir las sumas de cuadrados para sus respectivos grados de libertad, así tenemos la *media cuadrática del error MError* y la *media cuadrática de la regresión MReg*. Además si la hipótesis nula $H_0: \beta_1 = 0$ es verdadera, el estadístico F_0 sigue la distribución $F_{1, n-2}$, con lo que H_0 debe rechazarse si $f_0 > f_{\alpha, 1, n-2}$.

Para verificar si un modelo de regresión lineal simple es adecuado, es necesario elaborar, además de la tabla de análisis de varianza, los intervalos de confianza para los coeficientes β_0 y β_1 .

Bajo la hipótesis de que las observaciones están distribuidas de manera normal e independiente, el *intervalo de $(1-\alpha)100\%$ de confianza para la pendiente β_1* en una regresión lineal simple es

$$\hat{\beta}_1 - t_{\alpha/2, n-2} \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}} \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2, n-2} \sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}$$

y el *intervalo de $(1-\alpha)100\%$ de confianza para la ordenada al origen β_0*

$$\hat{\beta}_0 - t_{\alpha/2, n-2} \sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{x^2}{S_{xx}} \right]} \leq \beta_0 \leq \hat{\beta}_0 + t_{\alpha/2, n-2} \sqrt{\hat{\sigma}^2 \left[\frac{1}{n} + \frac{x^2}{S_{xx}} \right]}$$

donde:

$$S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2$$

Una vez que se han definido los modelos de regresión lineal simple, los coeficientes y sus correspondientes intervalos de confianza, pasaremos a tratar la predicción de nuevas observaciones.

Los modelos de regresión son importantes porque permiten predecir observaciones futuras de Y , correspondientes a un nivel dado de la variable de regresión X . La nueva observación es independiente de las observaciones utilizadas para desarrollar el modelo de regresión. Un intervalo de predicción para una observación futura Y_0 del $(1-\alpha)100\%$ de confianza en el valor X_0 , está dado por:

$$\hat{y}_0 - t_{\alpha/2, n-2} \sqrt{\hat{\sigma}^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]} \leq y_0 \leq \hat{y}_0 + t_{\alpha/2, n-2} \sqrt{\hat{\sigma}^2 \left[1 + \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}} \right]}$$

Antes de seleccionar un modelo de regresión para realizar predicciones, debemos emplear el análisis de residuos, el cual es útil para verificar si los errores cumplen el supuesto de normalidad con varianza constante, para ello se construye una gráfica de probabilidad normal con los residuos y se realiza la prueba de Kolmogorov-Smirnov. También son útiles las gráficas de los residuos contra los valores estimados y contra la variable independiente X . De acuerdo a los patrones mostrados, se determina si hay desigualdad en la varianza o si el modelo es inadecuado.

Cuando varios modelos cumplen con todos los requerimientos, se utiliza como criterio de selección al coeficiente de determinación R^2 . El coeficiente de determinación indica la adecuación del modelo de regresión y está dado por:

$$R^2 = \frac{SCReg}{SCTotal} = \frac{SCTotal - SCEError}{SCTotal} = 1 - \frac{SCEError}{SCTotal}$$

La potencia de explicación de un modelo está dado por:

$$R^2 (100\%)$$

Hay que tener en cuenta que un valor grande de R^2 , es decir mayor que 0.9, no necesariamente implica que el modelo de regresión proporcionará predicciones precisas de observaciones futuras.

Ya hemos explicado los modelos de regresión lineal simple, de igual forma, detallaremos todo lo concerniente a los modelos de regresión lineal múltiple.

Modelo de regresión lineal múltiple

Un modelo de regresión en que la *variable dependiente* o de *respuesta* Y está relacionada con más de una variable de explicación, recibe el nombre de modelo de regresión lineal múltiple. El modelo

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_{p-1} x_{i,p-1} + \epsilon_i \quad i = 1, 2, \dots, n$$

donde: $\epsilon_i \sim N(0, \sigma^2)$, $cov(\epsilon_i, \epsilon_j) = 0$ para $i \neq j$

recibe el nombre de modelo de regresión lineal múltiple con $(p-1)$ variables de explicación. Los parámetros β_j , $j = 0, 1, \dots, p-1$, se conocen como coeficientes de regresión. Los modelos polinomiales y

los que incluyen efectos de interacción también pueden analizarse con los métodos de regresión lineal múltiple. Al igual que el modelo de regresión lineal simple, también se emplea el método de mínimos cuadrados para estimar los coeficientes de regresión.

Al realizar el ajuste de un modelo de regresión lineal múltiple es conveniente expresar las operaciones matemáticas en forma matricial. Sean $(p-1)$ variables de explicación, n observaciones y el modelo que relaciona los regresores con la respuesta es:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_{p-1} x_{ip-1} + \epsilon_i, \quad i = 1, 2, \dots, n$$

Este modelo, llamado modelo lineal general, es un sistema de n ecuaciones que puede expresarse en notación matricial como

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

donde

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p-1} \\ 1 & x_{21} & x_{22} & \dots & x_{2p-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np-1} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

\mathbf{y} es un vector de observaciones de $(n \times 1)$, \mathbf{X} es una matriz $(n \times p)$ de los niveles de las variables independientes, $\boldsymbol{\beta}$ es un vector de $(p \times 1)$

formado por los coeficientes de regresión y ϵ es un vector de $(n \times 1)$ de errores aleatorios.

Para encontrar el vector de estimadores de mínimos cuadrados, $\hat{\beta}$ se emplean las ecuaciones normales, cuya forma matricial es:

$$X^T X \hat{\beta} = X^T y$$

y por consiguiente el estimador de mínimos cuadrados de β es:

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

La forma matricial del modelo ajustado es

$$\hat{y} = X \hat{\beta}$$

Los estimadores de mínimos cuadrados de los coeficientes de regresión $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$, son estimadores insesgados de los coeficientes de regresión $\beta_1, \beta_2, \dots, \beta_k$. La matriz de covarianza de los coeficientes de regresión $\hat{\beta}$ está dada por:

$$\text{cov}(\hat{\beta}) = \hat{\sigma}^2 (X^T X)^{-1}, \text{ por tanto}$$

$$\text{cov}(\hat{\beta}) = \text{MCError} (X^T X)^{-1}$$

Al igual que para el modelo de regresión lineal simple, la tabla de análisis de varianza es importante para el modelo de regresión lineal múltiple, ya que permite determinar si el modelo es adecuado.

El análisis de varianza es similar al de la regresión lineal simple, en el cálculo de las sumas de cuadrados y las medias cuadráticas. Para este caso la suma de cuadrados de la regresión y del error tienen $(p-1)$ y $(n-p)$ grados de libertad, respectivamente.

Tabla de análisis de varianza para regresiones lineales múltiples

Fuente de variación	Grados de libertad	Sumas cuadráticas	Medias cuadráticas	F
Regresión	$p-1$	SCReg	MReg	MReg/MError
Error	$n-p$	SError	MError	
Total	$n-1$	SCTotal		

p : número de parámetros

Es necesario determinar la significancia de la regresión a través de una prueba, la cual determina si existe una relación lineal entre la variable de respuesta Y y un subconjunto de las variables de regresión X_1, X_2, \dots, X_k .

Las hipótesis apropiadas son:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \beta_j \neq 0 \text{ al menos para una } j$$

y el estadístico de prueba para H_0 es:

$$F_0 = \frac{SCReg/(p-1)}{SCErr/(n-p)} = \frac{MCReg}{MCErr}$$

Si se rechaza H_0 implica que al menos uno de los estimadores de los Beta's es diferente de cero y por ende tiene una contribución significativa en el modelo.

Para los modelos de regresión simples y múltiples, se asume que los residuos son independientes, lo cual puede ser comprobado mediante el estadístico de Durbin-Watson.

Estadístico de Durbin-Watson para las autocorrelaciones

Cuando el análisis de regresión está basado en datos de series temporales, es necesario evaluar si los residuos en el modelo de regresión están correlacionados o no, para lo cual se emplea el estadístico de Durbin-Watson. Esta prueba consiste en determinar si el parámetro de autocorrelación ρ es cero, lo cual implicaría que los residuos son independientes. Las alternativas consideradas son:

$$C_1: \rho = 0$$

$$C_2: \rho > 0$$

Para evaluar estas alternativas, se utiliza el estadístico D , el cual está dado por:

$$D = \left(\sum_{t=2}^n (e_t - e_{t-1})^2 \right) \div \left(\sum_{t=1}^n e_t^2 \right)$$

donde n es el número de observaciones y e_t son los residuos.

Durbin y Watson han determinado límites de confianza superior e inferior, denotados d_s y d_i respectivamente, para el estadístico D . La selección de alternativas se realizará de acuerdo a las siguientes reglas:

Si $D > d_s$, se seleccionará a C_1

Si $D < d_i$, se seleccionará a C_2

Si $d_i \leq D \leq d_s$, la prueba es inconcluyente

CAPITULO 3

3. ANALISIS DE LOS DATOS UTILIZANDO TECNICAS MULTIVARIADAS

En este capítulo se realizará el análisis univariado y multivariado de las siguientes variables, relevantes para la producción bananera ecuatoriana: exportaciones de banano en miles de toneladas métricas, exportaciones de banano en miles de dólares FOB, producción de banano en miles de toneladas métricas, porcentaje de participación de las exportaciones bananeras en el PIB, porcentaje de participación de las exportaciones bananeras en las exportaciones totales, índice de precios al consumidor (total), índice de precios al consumidor (alimentos), cotización anual del dólar, superficie cultivada medida en hectáreas y población económicamente activa del Ecuador. Se presentarán 30 observaciones de cada variable, las cuales representan los valores anuales registrados entre los años 1969-1998, obteniéndose una matriz de datos de 30 filas (observaciones) y 10 columnas (variables).

Exportaciones de banano en TM.: Parte del total de la producción bananera, en miles de toneladas métricas, que se destina para la exportación.

Exportaciones de banano en dólares FOB : Ingreso por concepto de las exportaciones bananeras, expresado en miles de dólares FOB.

Producción bananera en TM.: Producción total de banano.

Porcentaje de participación de las exportaciones bananeras en el PIB: Relación entre las exportaciones bananeras y el producto interno bruto (PIB), expresada como porcentaje.

Porcentaje de participación de las exportaciones bananeras en las exportaciones totales: Relación entre las exportaciones bananeras y las exportaciones totales (petróleo, camarón, cacao, café, etc...) expresada como porcentaje.

Índice de precios al consumidor (total) : Es un indicador que mide la evolución de los precios correspondientes al conjunto de artículos que conforman la *Canasta Básica Familiar* de los hogares bajos y medios del área urbana del país. La canasta básica familiar es el conjunto de

artículos más representativos que han consumido o usado los hogares con niveles de ingreso bajos y medios en un período determinado.

La canasta se conforma de 197 artículos a nivel nacional, que se clasifican en diez agrupaciones: alimentos, bebidas y tabaco; vestido y calzado; alquiler, agua, electricidad, gas y otros; muebles, equipamiento y mantenimiento de la vivienda; salud; transporte; esparcimiento y cultura; educación; hoteles, cafeterías y restaurantes; y bienes y servicios misceláneos.

Se calcula con el *Índice de Laspeyres* de la siguiente forma:

$$IPC = \frac{\sum_{i=1}^n P_{it} Q_{i0}}{\sum_{i=1}^n P_{i0} Q_{i0}}$$

donde

P_{i0} : Precio del ítem i en el período base

P_{it} : Precio del ítem i en el tiempo t

Q_{i0} : Cantidad del ítem i en el período base

Se considera el año en que se ha realizado la encuesta de hogares como el período o año base. A partir del índice de Laspeyres se determinará

la evolución de los precios de los ítems en el tiempo, es decir si han registrado incrementos (inflación) o decrementos (deflación), por ello se mantienen constantes las cantidades de los ítems tanto para el año base como para el año actual.

Por ejemplo, consideremos la siguiente tabla de cantidades y precios de productos para los años 1998 y 1999:

Producto	Cantidad	Precio (1998)	Precio (1999)
arroz	5 libras	s./ 1.200 (libra)	s./ 1.500 (libra)
azúcar	3 libras	s./ 1.250 (libra)	s./ 1.600 (libra)
aceite	2 litros	s./ 8.000 (litro)	s./ 10.000 (litro)
pollo	4 libras	s./ 6.000 (libra)	s./ 8.000 (libra)

El año 1998 es el periodo base y 1999 el periodo actual. Aplicando el índice de Las peyres se obtiene que:

$$Indice = \frac{(1500)5 + (1600)3 + (10000)2 + (8000)4}{(1200)5 + (1250)3 + (8000)2 + (6000)4} = 1.29$$

el índice registró un incremento del 29% entre los años 1998 y 1999.

Índice de precios al consumidor del rubro alimentos: Es un indicador que mide la evolución de los precios correspondientes al grupo de alimentos, en el área urbana del país.

Cotización del dólar : Equivalencia anual del dólar en sucres, medida al final del periodo.

Superficie cultivada : Area en hectáreas destinada a la cosecha bananera.

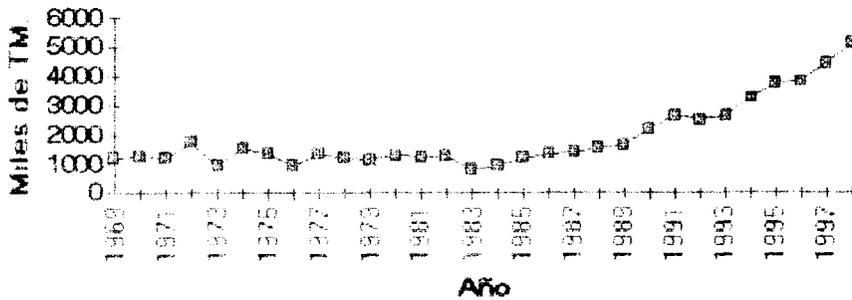
Población económicamente activa: Grupo de la población, que se clasifica en ocupados, subempleados (subempleados urbanos, agrícolas y de servicio doméstico) y los desocupados.

3.1 Análisis univariado

El análisis univariado de las variables descritas consistirá en la obtención de las estadísticas básicas: Mínimo, máximo, media, mediana, desviación estándar, varianza, sesgo y kurtosis. Se emplearán ilustraciones gráficas tales como: histogramas de frecuencia, diagramas de cajas y ojivas. Se realizarán pruebas de bondad de ajuste utilizando el método Kolmogorov-Smirnov. Iniciaremos el análisis con la variable: Exportaciones bananeras del Ecuador en TM.

Figura 3.1

Exportaciones de Banano en el Ecuador



Fuente: El Banano en el Ecuador, Larrea Carlos, Corporación Editora Nacional, 1997

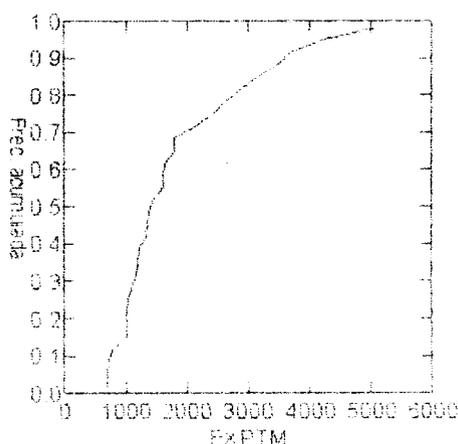
Estadísticas básicas (en miles de toneladas métricas)

Mínimo	800.15
Máximo	5154.600
Media	1897.725
Mediana	1355.15
Desviación estándar	1137.683
Varianza	1294323.354
Sesgo	1.533
Kurtosis	1.527

El coeficiente de sesgo indica que los datos se encuentran concentrados en el lado izquierdo, es decir que la media es mayor que la mediana. El coeficiente de kurtosis indica que la distribución es platicúrtica, ya que tiene menor concentración de datos que la normal en sus colas.

Figura 3.2

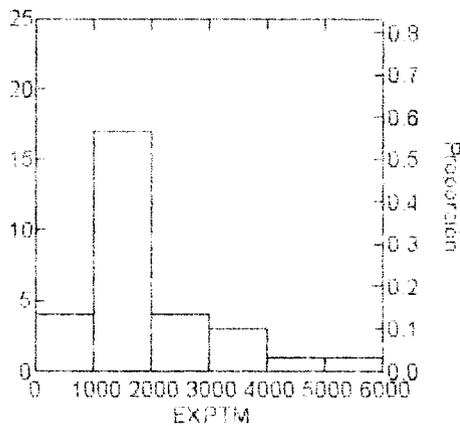
Ojiva y diagrama de cajas de las exportaciones bananeras en TM



Si $Q_1 = 1'200.000$, esto significa que el 25 % de los valores anuales de las exportaciones bananeras registrados entre los años 1969 y 1998, han sido menores o iguales que 1'200.000 toneladas métricas. Si $Q_2 = 1'355.150$, quiere decir que la mediana de las observaciones es 1'355.150 toneladas métricas. Si $Q_3 = 2'600.000$, implica que el 75% de las observaciones son menores o iguales a 2'600.000 toneladas métricas. Existe un valor aberrante, 5'154.600 toneladas métricas (año 1998). Así lo podemos observar en la figura 3.2.

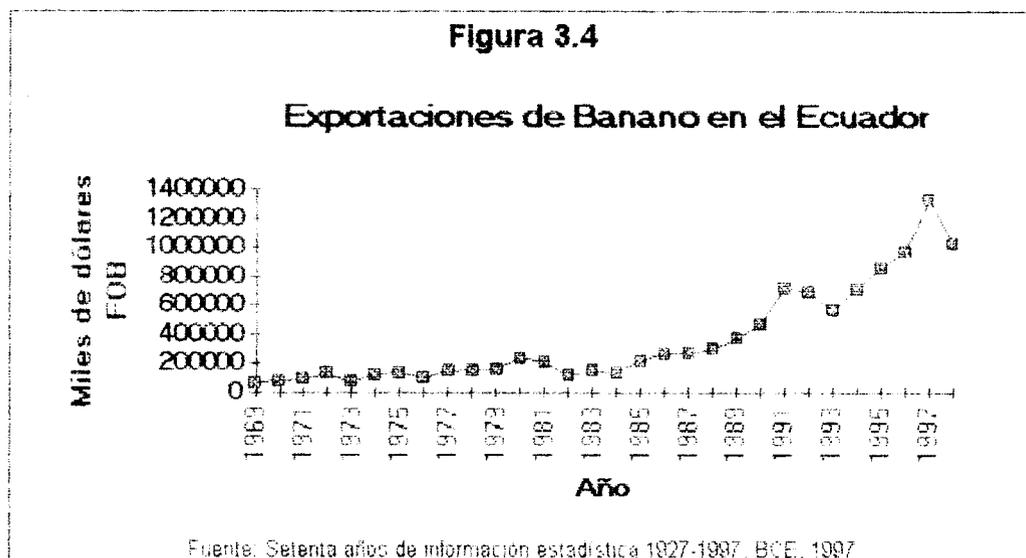
Figura 3.3

Histograma de frecuencias de las exportaciones bananeras (miles de TM)



Aproximadamente un 60% de las exportaciones bananeras anuales (1969-1998), se encuentran entre 1'000.000 y 2'000.000 de TM. No más de un 5% de las exportaciones han sobrepasado la barrera de 4'000.000 de TM. Así lo podemos observar en la figura 3.3 .

A continuación se presentará el análisis univariado de la variable: exportaciones bananeras del Ecuador en miles de dólares FOB.



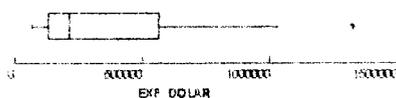
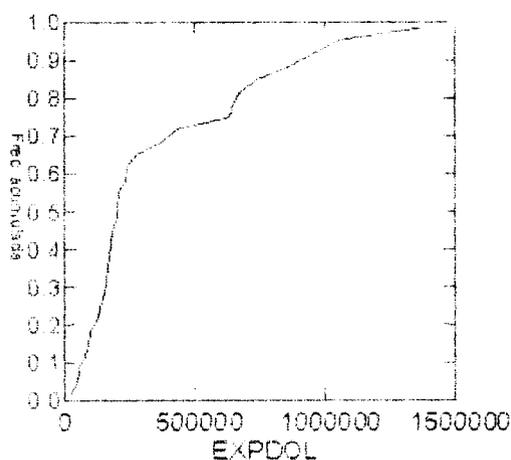
Estadísticas básicas (en miles de dólares FOB)

Mínimo	68175
Máximo	1326892
Media	362096.167
Mediana	213931
Desviación estándar	338273.603
Varianza	1.14429E+11
Sesgo	1.386
Kurtosis	1.068

El coeficiente de sesgo indica que los datos se encuentran concentrados en el lado izquierdo, es decir que la media es mayor que la mediana. El coeficiente de kurtosis indica que la distribución es platicúrtica, ya que tiene menor concentración de datos que la normal, en sus colas.

Figura 3.5

Ojiva y diagrama de cajas de las exportaciones bananeras en miles de dólares FOB

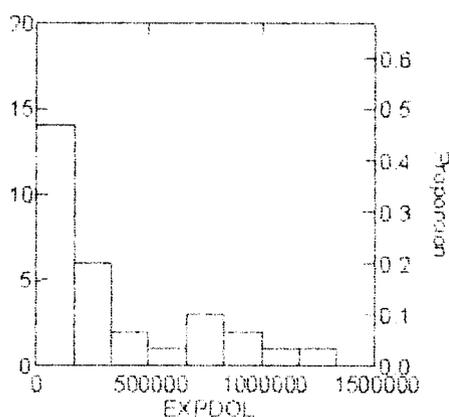


Si $Q_1 = 125'000.000$, esto significa que el 25 % de los valores anuales de las exportaciones bananeras registrados entre los años 1969 y 1998, han sido menores o iguales que 125'000.000 dólares FOB. Si $Q_2 = 213'931.000$, quiere decir que la mediana de las observaciones es 213'931.000 dólares FOB. Si $Q_3 = 600'000.000$, implica que el 75% de las observaciones son menores o iguales a

600'000.000 dólares FOB. Existe un valor aberrante, 1.326'892.000 dólares FOB (año 1997). Así lo podemos apreciar en la figura 3.5 .

Figura 3.6

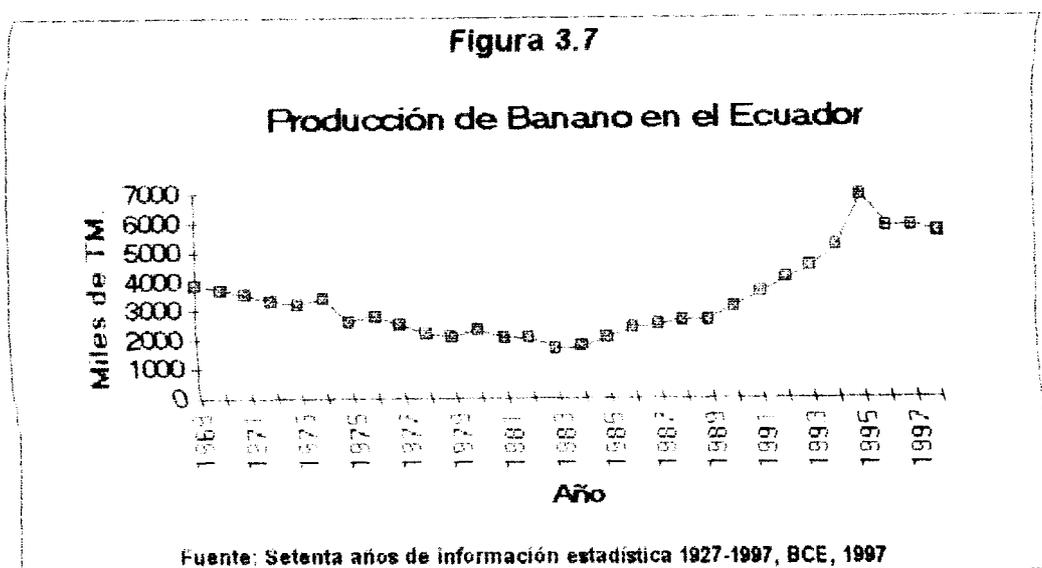
Histograma de frecuencias de las exportaciones bananeras en miles de dólares FOB



Aproximadamente, un 70% de las exportaciones bananeras anuales, (1969-1998), no han sobrepasado la barrera de los 500'000.000 dólares FOB. El 30% restante, han registrado valores entre 500'000.000 y 1.400'000.000 dólares FOB. Así lo podemos observar en la figura 3.6. Los patrones entre las figuras 3.3 y 3.6 no coinciden, debido a que el precio del banano presenta continuos incrementos y decrementos, y en ciertos casos se incrementa la exportación de banano en miles de TM. , pero la exportación de banano en miles de dólares FOB decrece y viceversa, o también puede establecerse una

relación directa entre ellas, es decir si las exportaciones en miles de TM aumentan (disminuyen), las exportaciones en miles de dólares FOB también aumentan (disminuyen).

A continuación se presentará el análisis univariado de la variable: producción bananera del Ecuador en miles de TM.



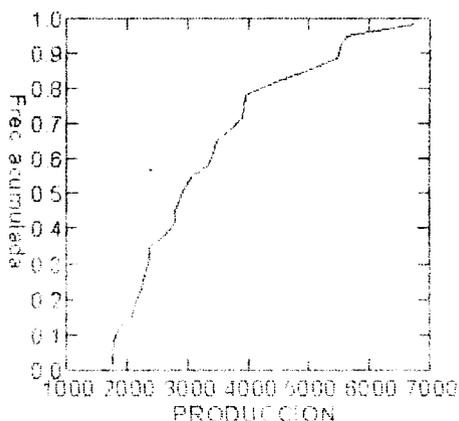
Estadísticas básicas (en miles de toneladas métricas)

Mínimo	1642.1
Máximo	6809.2
Media	3276.79
Mediana	2921.25
Desviación estándar	1361.043
Varianza	1852437.075
Sesgo	1.045
Kurtosis	0.354

El coeficiente de sesgo indica que los datos se encuentran concentrados en el lado izquierdo, es decir que la media es mayor que la mediana. El coeficiente de kurtosis indica que la distribución es platicúrtica, ya que tiene menor concentración de datos que la normal en sus colas.

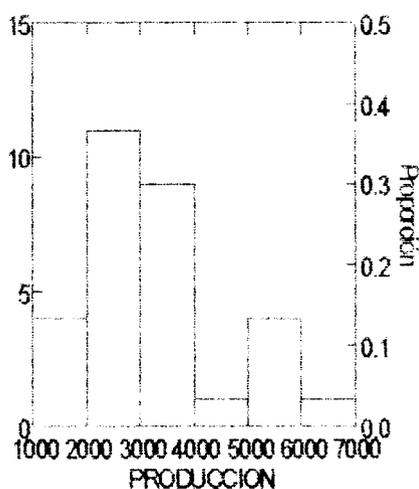
Figura 3.8

Ojiva y diagrama de cajas de la producción bananera en TM



Si $Q_1 = 2'160.000$, esto significa que el 25 % de los valores anuales de producción bananera registrados entre los años 1969 y 1998, han sido menores o iguales que 2'160.000 toneladas métricas. Si $Q_2 = 2'921.250$, quiere decir que la mediana de las observaciones es 2'921.250 toneladas métricas. Si $Q_3 = 3'880.000$, implica que el 75% de las observaciones son menores o iguales a 3'880.000 toneladas métricas. Existe un valor aberrante, 6'809.200 toneladas métricas (1995). Así lo podemos apreciar en la figura 3.8.

Histograma de frecuencias de la producción bananera en miles de TM

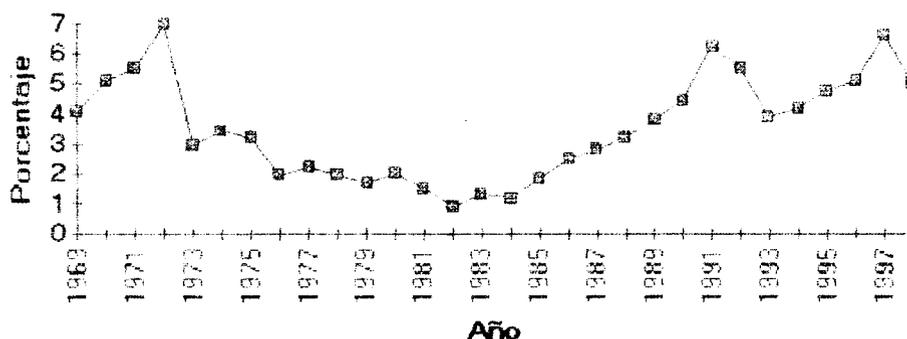


Aproximadamente, el 40% de la producción bananera anual (1969-1998), se encuentra entre 2'000.000 y 3'000.000 TM. El 30% está entre 3'000.000 y 4'000.000 TM. Tan solo un 4% sobrepasa la barrera de 6'000.000 TM. Así lo podemos notar en la figura 3.9 .

A continuación se presentará el análisis univariado de la variable: Participación de las exportaciones bananeras en el PIB del Ecuador.

Figura 3.10

Participación de las Exportaciones Bananeras en el PIB del Ecuador



Fuente: Setenta años de información estadística 1927-1997, BCE, 1997

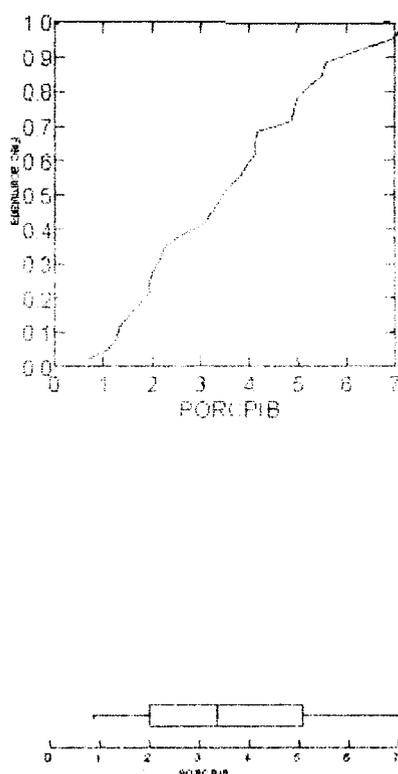
Estadísticas básicas (en porcentajes)

Mínimo	0.871
Máximo	6.990
Media	3.534
Mediana	3.338
Desviación estándar	1.724
Varianza	2.972
Sesgo	0.305
Kurtosis	-0.916

El coeficiente de sesgo indica que los datos se encuentran concentrados en el lado izquierdo, es decir que la media es mayor que la mediana. El coeficiente de kurtosis indica que la distribución es platicúrtica, ya que tiene menor concentración de datos que la normal en sus colas.

Figura 3.11

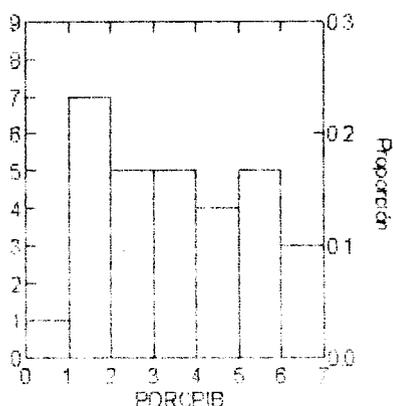
Ojiva y diagrama de cajas de la participación de las exportaciones bananeras en el PIB (%)



Si $Q_1 = 1.9$, esto significa que el 25 % de los valores anuales de participación de las exportaciones bananeras en el Producto Interno Bruto (PIB) registrados entre los años 1969 y 1998, han sido menores o iguales que 1.9%. Si $Q_2 = 3.338$, quiere decir que la mediana de las observaciones es 3.338%. Si $Q_3 = 5$, implica que el 75% de las observaciones son menores o iguales a 5%. Así lo podemos notar en la figura 3.11.

Figura 3.12

Histograma de frecuencias de la participación de las exportaciones bananeras en el PIB (%)

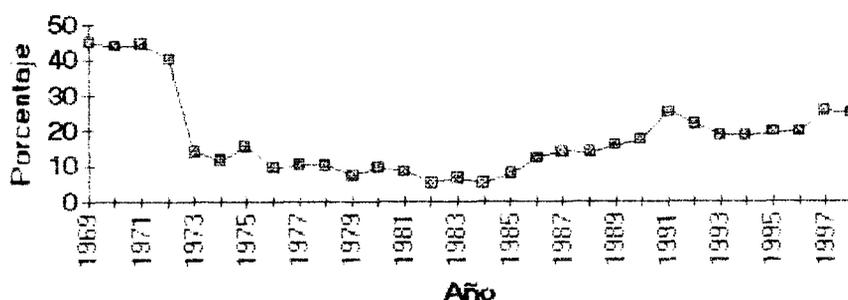


Aproximadamente, un 25% de las exportaciones bananeras anuales (1969-1998), han tenido una participación en el PIB comprendida entre el 1 y 2 %. Un 34% tuvieron una participación entre 2 y 4 %. Sólo un 10% de las exportaciones registraron una participación superior al 6%. Así lo podemos apreciar en la figura 3.12.

A continuación presentaremos el análisis univariado de la variable: Participación de las exportaciones bananeras en las exportaciones totales del Ecuador.

Figura 3.13

**Participación de las Exportaciones de
Banano en las Exportaciones Totales del
Ecuador**



Setenta años de información estadística 1927-1997, BCE, 1997

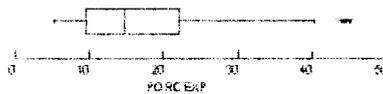
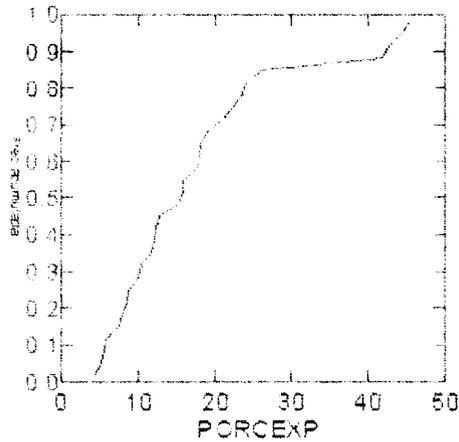
Estadísticas básicas (en porcentajes)

Mínimo	5.159
Máximo	44.886
Media	17.969
Mediana	14.694
Desviación estándar	11.67
Varianza	136.198
Sesgo	1.282
Kurtosis	0.845

El coeficiente de sesgo indica que los datos se encuentran concentrados en el lado izquierdo, es decir que la media es mayor que la mediana. El coeficiente de kurtosis indica que la distribución es platicúrtica, ya que tiene menor concentración de datos que la normal en sus colas.

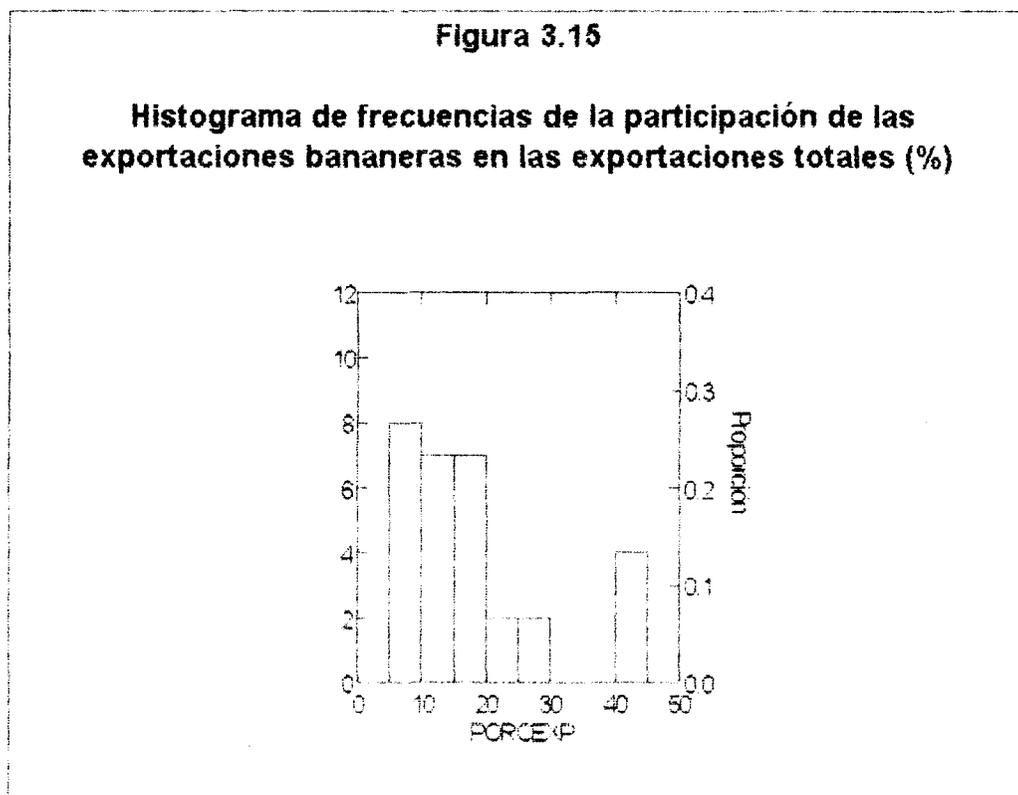
Figura 3.14

Ojiva y diagrama de cajas de la participación de las exportaciones bananeras en las exportaciones totales (%)



Si $Q_1 = 9$, esto significa que el 25 % de los valores anuales de participación de las exportaciones bananeras en las exportaciones totales registrados entre los años 1969 y 1998, han sido menores o iguales que 9%. Si $Q_2 = 14.694$, quiere decir que la mediana de las observaciones es 14.694%. Si $Q_3 = 22$, implica que el 75% de las observaciones son menores o iguales a 22%. Existen tres valores

aberrantes, 43.82% (año 1970), 44.28% (año 1971) y 44.88% (año 1969). Así lo podemos apreciar en la figura 3.14 .

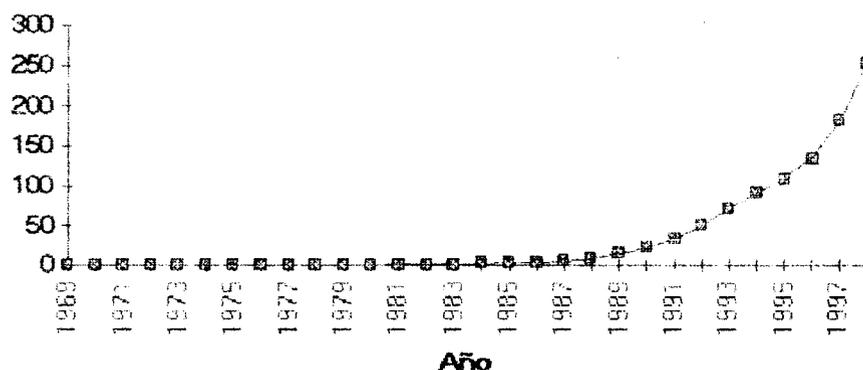


Aproximadamente un 30% de las exportaciones bananeras anuales (1969-1998), tuvieron una participación en las exportaciones totales de entre 5 y 10%. Un 50% tienen una participación de entre 10 y 20 %. Sólo un 15% registra una participación que supera el 40% de las exportaciones totales. Así lo podemos observar en la figura 3.15.

A continuación presentaremos el análisis univariado de la variable: índice de precios al consumidor (alimentos).

Figura 3.16

Índice de precios al consumidor (IPC) del rubro alimentos



Fuente: Setenta años de información estadística 1927-1997, BCE, 1997

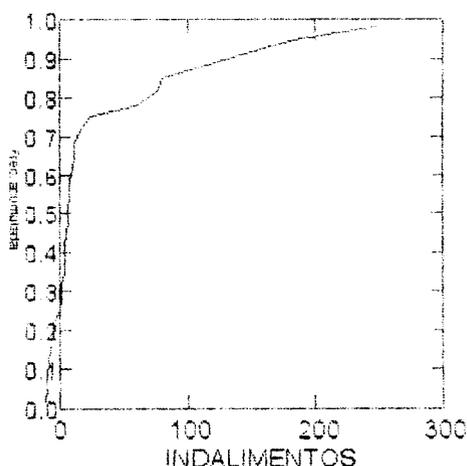
Estadísticas básicas

Mínimo	0.190
Máximo	253
Media	33.181
Mediana	2.05
Desviación estándar	62.233
Varianza	3872.975
Sesgo	2.299
Kurtosis	5.128

El coeficiente de sesgo indica que los datos se encuentran concentrados en el lado izquierdo, es decir que la media es mayor que la mediana. El coeficiente de kurtosis indica que la distribución es leptocúrtica, ya que tiene mayor concentración de datos que la normal en sus colas.

Figura 3.17

**Ojiva y diagrama de cajas del índice de precios al consumidor
(alimentos)**

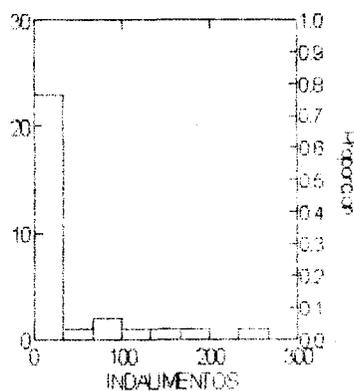


Si $Q_1 = 0.4$, esto significa que el 25 % de los valores anuales del índice de precios al consumidor del rubro alimentos registrados entre los años 1969 y 1998, han sido menores o iguales que 0.4. Si $Q_2 = 2.05$, quiere decir que la mediana de las observaciones es 2.05. Si $Q_3 = 33$, implica que el 75% de las observaciones son menores o iguales a 33. Existen dos valores aberrantes, 90.61 (año

1994) y 109.67 (año 1995) y existen tres valores mucho mayores que los considerados como aberrantes, 135.7 (año 1996), 181.2 (año 1997) y 253 (año 1998). Así lo podemos observar en la figura 3.17. Se puede deducir de estos valores, que el precio de los alimentos y por ende su influencia en la inflación, ha presentado un lento crecimiento entre los años 1969 y 1988 (0.19 a 8.06), pero a partir del año de 1989 hasta 1998, ha registrado un rápido y elevado incremento de 15.16 a 253, es decir que ha crecido un 1669%. Siendo el alto rango de los datos la causa de que el diagrama de cajas correspondiente sea tan atípico.

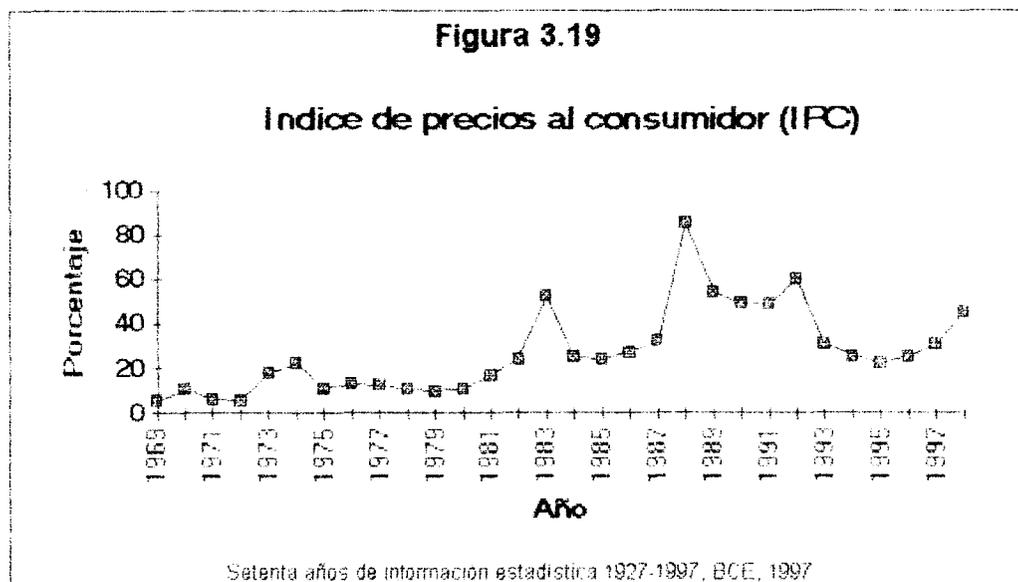
Figura 3.18

Histograma de frecuencias del índice de precios al consumidor (alimentos)



Casi el 80% de los IPC (alimentos) anuales (1969-1998), no son mayores de 30. Mientras que el 20% restante fluctúan entre 30 y 260, aproximadamente. Así lo podemos apreciar en la figura 3.18.

A continuación presentamos el análisis univariado de la variable: Índice de precios al consumidor.



Estadísticas básicas

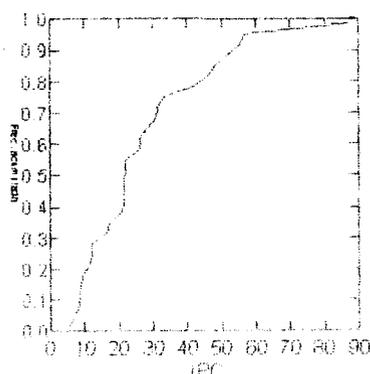
Mínimo	5.22
Máximo	85.7
Media	27.291
Mediana	24.4
Desviación estándar	19.177
Varianza	367.742
Sesgo	1.237
Kurtosis	1.511

El coeficiente de sesgo indica que los datos se encuentran concentrados en el lado izquierdo, es decir que la media es mayor que la mediana. El coeficiente de kurtosis indica que la distribución es

platicúrtica, ya que tiene menor concentración de datos que la normal en sus colas.

Figura 3.20

Ojiva y diagrama de cajas del índice de precios al consumidor (IPC)

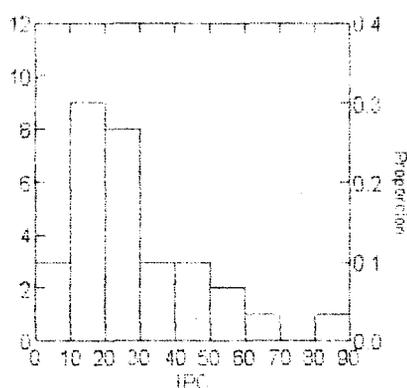


Si $Q_1 = 10.7$, esto significa que el 25 % de los valores anuales del índice de precios al consumidor (IPC) registrados entre los años 1969 y 1998, han sido menores o iguales que 10.7%. Si $Q_2 = 24.4$, quiere decir que la mediana de las observaciones es 24.4%. Si $Q_3 = 33$, implica que el 75% de las observaciones son menores o iguales a

33%. Existe un valor aberrante, 85.7% (año 1988). Así lo podemos notar en la figura 3.20.

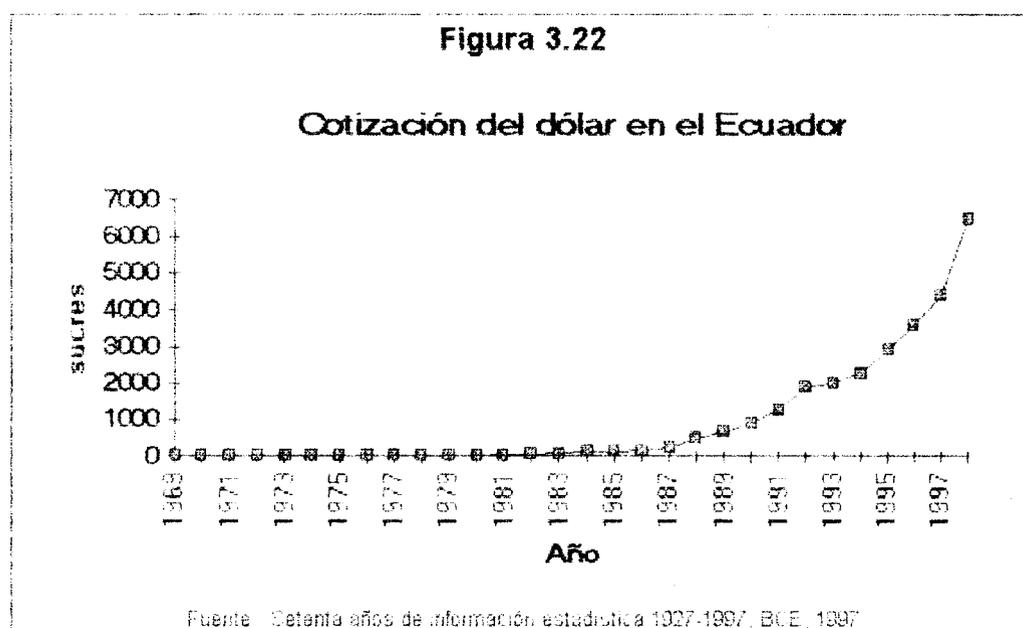
Figura 3.21

Histograma de frecuencias del índice de precios al consumidor



Más de un 30% de los índices de precios al consumidor (1969-1998), han fluctuado entre un 10 y 20%. Aproximadamente un 28% fluctuaron entre el 20 y 30%. Menos del 3% registraron una tasa mayor al 80%. Así lo podemos apreciar en la figura 3.21.

A continuación presentamos el análisis univariado de la variable: Cotización del dólar en el Ecuador.



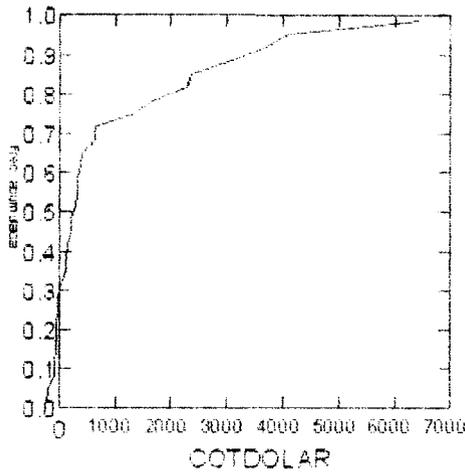
Estadísticas básicas (en sucres)

Mínimo	18.18
Máximo	6480
Media	934.493
Mediana	103.35
Desviación estándar	1582.573
Varianza	2504538.075
Sesgo	2.146
Kurtosis	4.585

El coeficiente de sesgo indica que los datos se encuentran concentrados en el lado izquierdo, es decir que la media es mayor que la mediana. El coeficiente de kurtosis indica que la distribución es leptocúrtica, ya que tiene mayor concentración de datos que la normal en sus colas.

Figura 3.23

Ojiva y diagrama de cajas de la cotización del dólar (sucres/dólar)

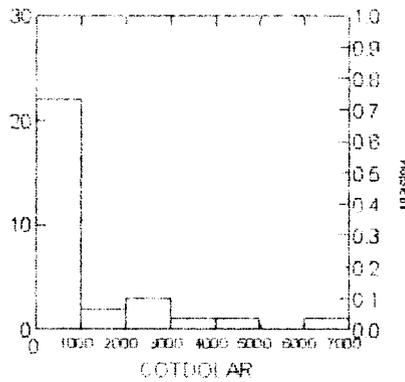


Si $Q_1 = 26.7$, esto significa que el 25 % de los valores anuales de la cotización del dólar registrados entre los años 1969 y 1998, han sido menores o iguales que 26.7 sucres. Si $Q_2 = 103.35$, quiere decir que la mediana de las observaciones es 103.35 sucres. Si $Q_3 = 1200$, implica que el 75% de las observaciones son menores o iguales a 1200 sucres. Existen dos valores aberrantes, 3596.2 sucres (año 1996) y 4393 sucres (año 1997), y un valor mucho mayor que los

considerados aberrantes, 6480 sucres (año 1998). Así lo podemos apreciar en la figura 3.23. Recordemos que este análisis no incluye observaciones de los años 1999 y 2000. En diciembre de 1999, el dólar alcanzó una cotización de 19260 sucres, es decir que entre 1998 y 1999, tuvo un incremento de aproximadamente 200%. En el 2000, se pone en práctica la dolarización en el Ecuador, por lo que la cotización del dólar se fija en 25000 sucres, es decir que entre 1999 y el 2000, hay un 30% de incremento, aproximadamente. Si estos cambios económicos hubieran sido considerados, el análisis correspondiente al período 1969-1998, registraría notables variaciones.

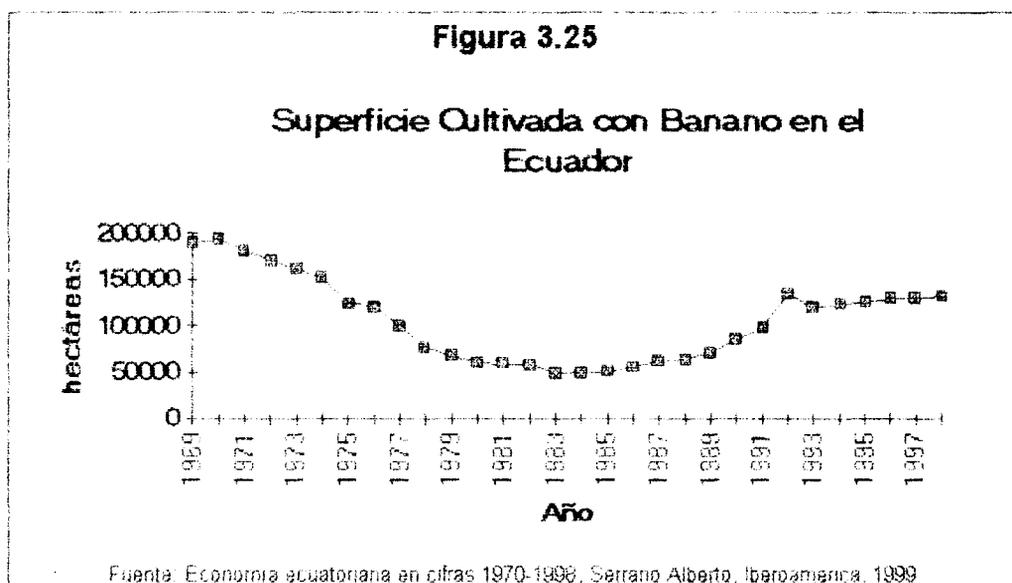
Figura 3.24

Histograma de frecuencias de la cotización del dólar (sucres/dólar)



Más de un 70% de las cotizaciones anuales del dólar (1969-1998) se encuentran entre 18 y 1000 sucres . Un 10% han fluctuado entre 2000 y 3000 sucres. Menos de un 5% han presentado una cotización del dólar mayor a 6000 sucres. Así lo podemos notar en la figura 3.24.

A continuación presentaremos el análisis univariado de la variable:
Superficie cultivada con banano en el Ecuador



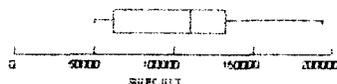
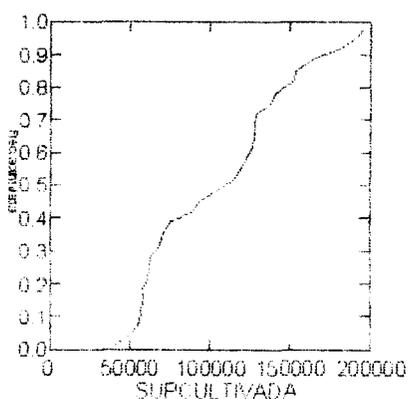
Estadísticas básicas (en hectáreas)

Mínimo	49562
Máximo	193560
Media	106669.6
Mediana	109982.5
Desviación estándar	45439.084
Varianza	2.06471E+09
Sesgo	0.378
Kurtosis	-1.020

El coeficiente de sesgo indica que los datos se encuentran concentrados en el lado izquierdo, es decir que la media es mayor que la mediana. El coeficiente de kurtosis indica que la distribución es platicúrtica, ya que tiene menor concentración de datos que la normal en sus colas.

Figura 3.26

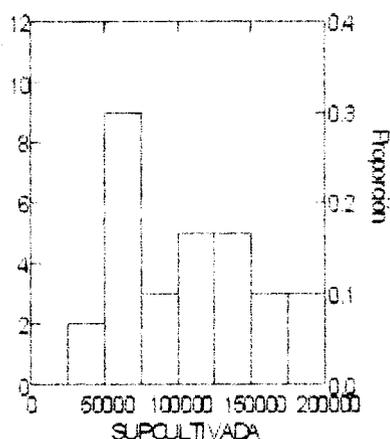
Ojiva y diagrama de cajas de la superficie cultivada (hectáreas)



Si $Q_1 = 60568$, esto significa que el 25 % de los valores anuales de superficie cultivada con banano registrados entre los años 1969 y 1998, han sido menores o iguales que 60568 hectáreas. Si $Q_2 = 109982.5$, quiere decir que la mediana de las observaciones es 109982.5 hectáreas. Si $Q_3 = 134000$, implica que el 75% de las observaciones son menores o iguales a 134000 hectáreas. Así lo podemos observar en la figura 3.26.

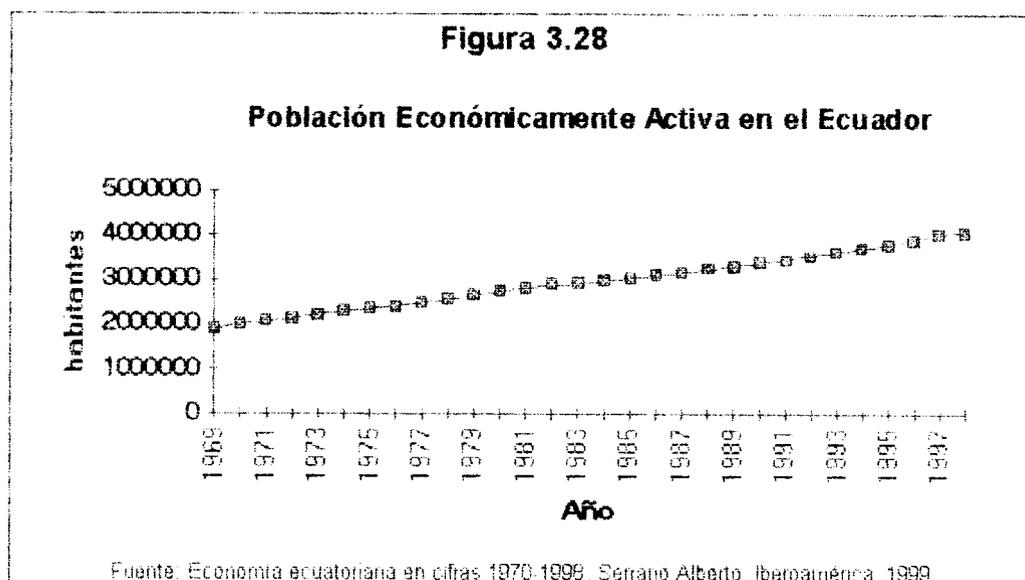
Figura 3.27

Histograma de frecuencias de la superficie cultivada (hectáreas)



Más de un 30% de los registros anuales de superficie cultivada con banano (1969-1998), han fluctuado entre 50000 y 75000 hectáreas. Aproximadamente el 36% fluctuaron entre 100000 y 150000 hectáreas. Un 20% presentan superficies cultivadas entre 150000 y 200000 hectáreas. Así lo podemos observar en la figura 3.27.

A continuación presentaremos el análisis univariado de la variable: Población económicamente activa del Ecuador.



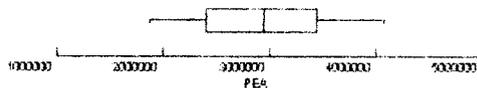
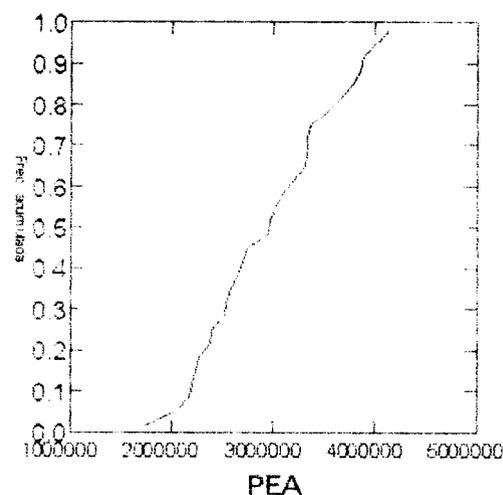
Estadísticas básicas (en número de personas)

Mínimo	1879000
Máximo	4073581
Media	2950798.533
Mediana	2950824
Desviación estándar	636571.899
Varianza	4.052E+11
Sesgo	0.074
Kurtosis	-1.035

El coeficiente de sesgo indica que los datos se encuentran concentrados en el lado izquierdo, es decir que la media es mayor que la mediana. El coeficiente de kurtosis indica que la distribución es platicúrtica, ya que tiene menor concentración de datos que la normal en sus colas.

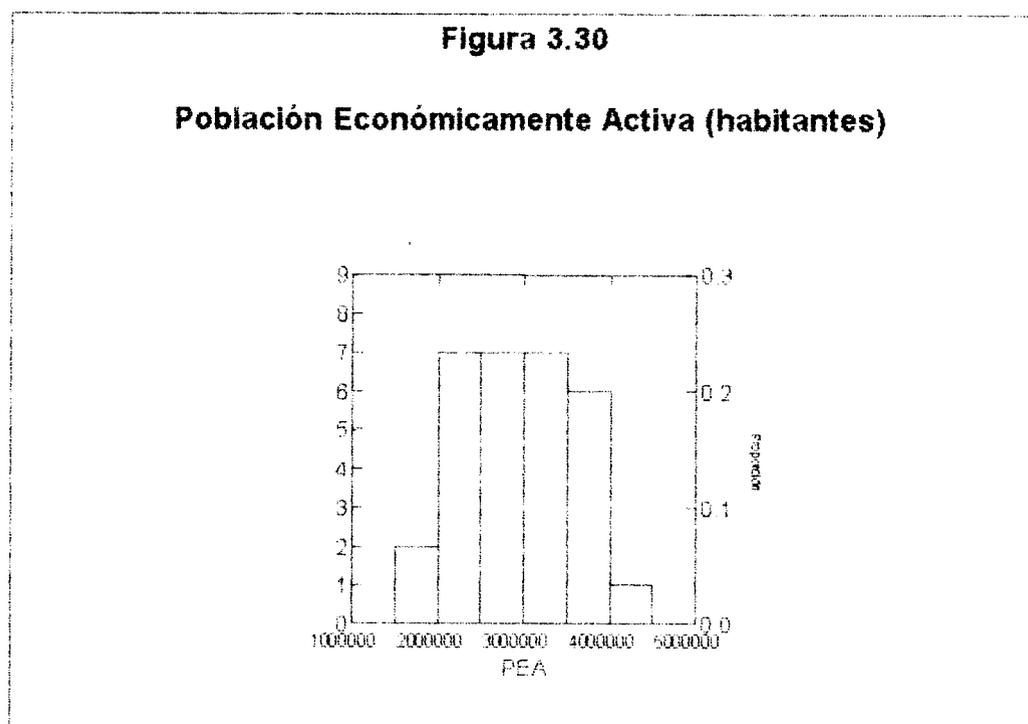
Figura 3.29

Ojiva y diagrama de cajas de la Población Económicamente Activa (en número de personas)



Si $Q_1 = 2'400.000$, esto significa que el 25 % de los valores anuales de la población económicamente activa (PEA) registrados entre los años 1969 y 1998, han sido menores o iguales que 2'400.000 personas. Si $Q_2 = 2'950.824$, quiere decir que la mediana de las observaciones es 2'950.824 personas. Si $Q_3 = 3'250.000$, implica

que el 75% de las observaciones son menores o iguales a 3'250.000 personas. Así lo podemos observar en la figura 3.29.



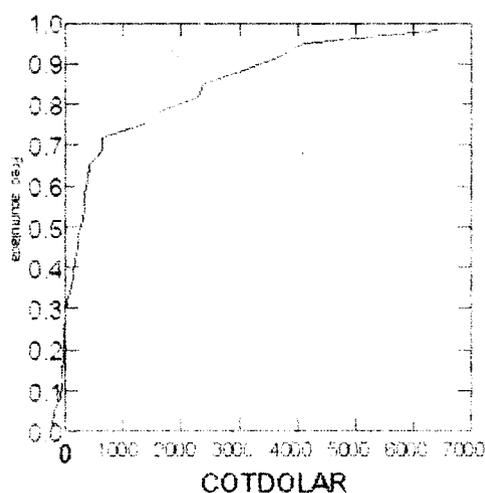
Aproximadamente, un 70% de los registros anuales de la población económicamente activa (1969-1998), fluctúan entre 2 y 3 millones de habitantes del Ecuador. Así lo podemos observar en la figura 3.30.

Contraste de las ojivas correspondientes a las variables :
Cotización del dólar, superficie cultivada y población
económicamente activa

Hay un rápido crecimiento de la frecuencia acumulada de las observaciones correspondientes a la cotización del dólar, ya que aproximadamente el 75% de los valores se registran en el primer intervalo (0-1000) sucres, mientras que el 25% restante se encuentran distribuidos en los demás intervalos (1000-7000) sucres, lo cual podemos apreciarlo en la figura 3.31: ojiva de la variable *cotización del dólar*.

Figura 3.31

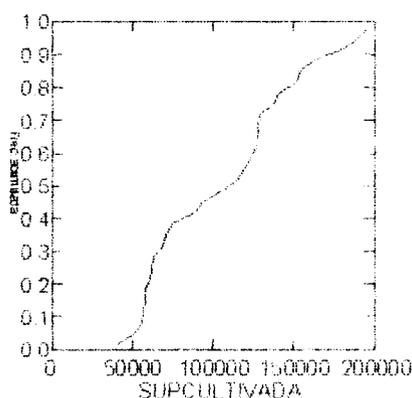
Ojiva de la cotización del dólar (sucres/dólar)



El crecimiento de la frecuencia acumulada de los valores correspondientes a la superficie cultivada, es más lento que en la anterior (figura 3.31). El 45% de los valores de superficie cultivada con banano, se encuentran en los dos primeros intervalos (0-100000) hectáreas. En el tercer intervalo (100000-150000) hectáreas se registran aproximadamente el 37% de las observaciones, mientras que el 18% restante se encuentran en el último intervalo (150000-200000) hectáreas, lo cual podemos apreciarlo en la figura 3.32: ojiva de la variable *superficie cultivada*.

Figura 3.32

Ojiva de la superficie cultivada (hectáreas)

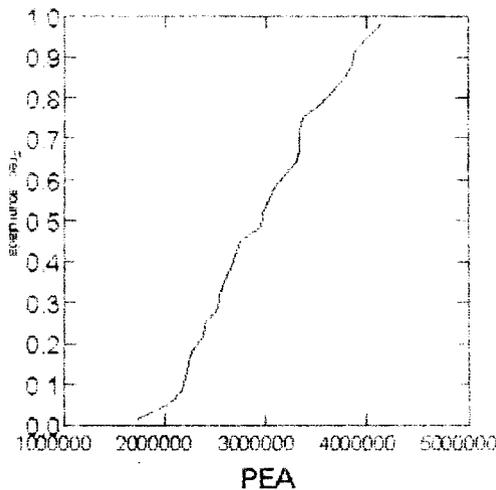


El crecimiento de la frecuencia acumulada de los valores correspondientes a la población económicamente activa, es más

rápido que en la anterior (figura 3.32), ya que el 60% de los valores registrados de la población económicamente activa, se encuentran en los dos primeros intervalos (1'000.000-3'000.000) habitantes, mientras que el 37 y 3% restante se encuentran distribuidos en el tercer (3'000.000-4'000.000) habitantes y cuarto intervalo (4'000.000-5'000.000) respectivamente, lo cual podemos apreciarlo en la figura 3.33 ojiva de la variable *población económicamente activa*.

Figura 3.33

Población Económicamente Activa (en número de personas)



Por lo que podemos concluir que de los tres modelos contrastados, la ojiva correspondiente a la variable: *cotización del dólar*, tiene el más

rápido crecimiento de frecuencia acumulada, mientras que la ojiva correspondiente a la variable: *superficie cultivada*, registra el crecimiento más lento.

Prueba de Bondad de Ajuste

Para realizar una prueba de bondad de ajuste, es necesario conocer los valores de los parámetros correspondientes a la media y la varianza poblacional, lo cual sería ideal, ya que estos valores exactos son desconocidos. Por ello, se ha calculado para cada variable de estudio, la media y la varianza muestral, debido a que son los estimadores más adecuados de la media y la varianza poblacional.

Los valores propuestos en cada prueba son sumamente cercanos a las correspondientes media y varianza muestral de la variable de estudio analizada.

Prueba de Kolmogorov-Smirnov para la variable: Exportaciones bananeras en miles de TM.

H_0 : Las exportaciones bananeras en miles de TM. se ajustan a una distribución Normal con media 1900 y desviación estándar 1140

vs.

$H_1: \neg H_0$

MaxDif = 0.261

Valor p (2-colas) = 0.034

El valor p resultante se encuentra en el intervalo de indiferencia, es decir que la hipótesis nula H_0 propuesta, puede ser aceptada o rechazada, tendiendo más a ser rechazada.

Prueba de Kolmogorov-Smirnov para la variable: Exportaciones bananeras en millones de dólares FOB

H_0 : Las exportaciones bananeras en millones de dólares FOB se ajustan a una distribución Normal con media 363 y desviación estándar 340

vs.

$H_1: \neg H_0$

MaxDif = 0.245

Valor p (2-colas) = 0.055

El valor p resultante se encuentra en el intervalo de indiferencia, es decir que la hipótesis nula H_0 propuesta, puede ser aceptada o rechazada, tendiendo más a ser rechazada.

Prueba de Kolmogorov-Smirnov para la variable: Producción bananera en miles de TM.

H_0 : La producción bananera en miles de TM. se ajusta a una distribución Normal con media 3300 y desviación estándar 1360

vs.

$H_1: \bar{\neq} H_0$

MaxDif = 0.169

Valor p (2-colas) = 0.356

Hay evidencia estadística para no rechazar la hipótesis nula H_0 . En la figura 3.34 se presenta su función de densidad.

Prueba de Kolmogorov-Smirnov para la variable: Participación de las exportaciones bananeras en el PIB

H_0 : La participación de las exportaciones bananeras en el PIB se ajusta a una distribución Normal con media 4 y desviación estándar 2

vs.

$H_1: \bar{\neq} H_0$

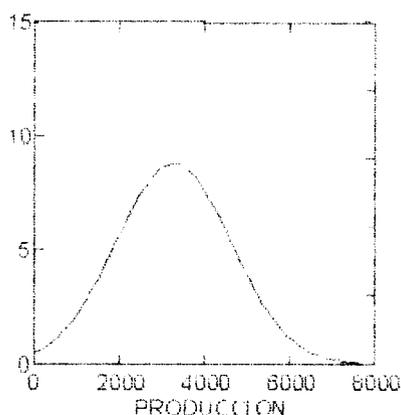
MaxDif = 0.148

Valor p (2-colas) = 0.523

Hay evidencia estadística para no rechazar la hipótesis nula H_0 . En la figura 3.35 se presenta su función de densidad.

Figura 3.34

Producción bananera en el Ecuador



Prueba de Kolmogorov-Smirnov para la variable: Participación de las exportaciones bananeras en las exportaciones totales

H_0 : La participación de las exportaciones bananeras en las exportaciones totales se ajusta a una distribución Normal con media 18 y desviación estándar 12

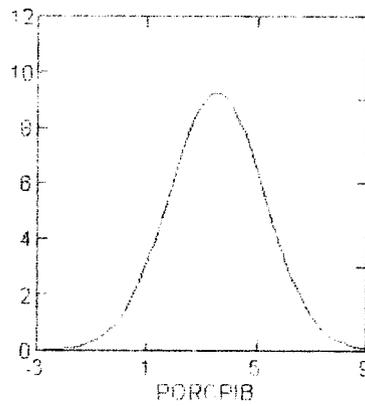
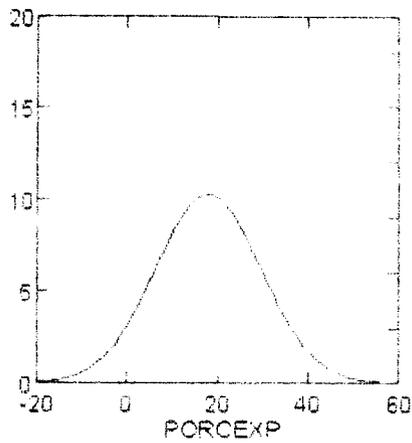
vs.

$H_1: \neq H_0$

MaxDif = 0.170

Valor p (2-colas) = 0.348

Hay evidencia estadística para no rechazar la hipótesis nula H_0 . En la figura 3.36 se presenta su función de densidad.

Figura 3.35**Participación de las exportaciones bananeras en el PIB****Figura 3.36****Participación de las exportaciones bananeras en las exportaciones totales**

Prueba de Kolmogorov-Smirnov para la variable: Índice de precios al consumidor (alimentos)

H_0 : El índice de precios al consumidor (alimentos) se ajusta a una distribución normal con media 33 y desviación estándar 63

vs.

$H_1: \neq H_0$

MaxDif = 0.321

Valor p (2-colas) = 0.004

Hay evidencia estadística para rechazar la hipótesis nula H_0 .

Prueba de Kolmogorov-Smirnov para la variable: Índice de precios al consumidor (IPC total)

H_0 : La variable IPC se ajusta a una distribución Normal con media 23 y desviación estándar 20

vs.

$H_1: \neq H_0$

MaxDif = 0.187

Valor p (2-colas) = 0.245

Hay evidencia estadística para no rechazar la hipótesis nula H_0 . En la figura 3.37 se presenta su función de densidad.

Prueba de Kolmogorov-Smirnov para la variable: Cotización del dólar

H_0 : La cotización del dólar tiene distribución Normal con media 935 y desviación estándar 1600

vs.

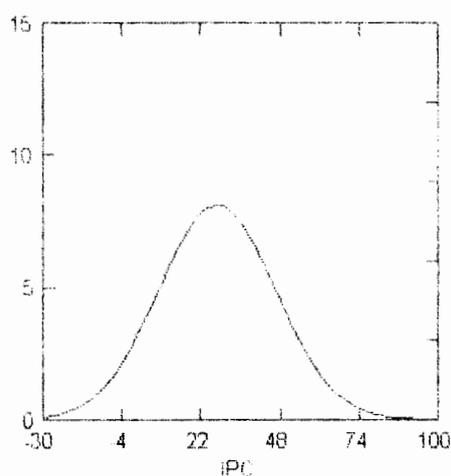
$H_1: \neg H_0$

MaxDif = 0.3 Valor p (2-colas) = 0.009

Hay evidencia estadística para rechazar la hipótesis nula H_0 .

Figura 3.37

Indice de precios al consumidor



Prueba de Kolmogorov-Smirnov para la variable: Superficie cultivada de banano

H_0 : La superficie cultivada de banano se ajusta a una distribución Normal con media 107000 y desviación estándar 45500

vs.

$H_1: \bar{\neq} H_0$

MaxDif = 0.148

Valor p (2-colas) = 0.524

Hay evidencia estadística para no rechazar la hipótesis nula H_0 . En la figura 3.38 se presenta su función de densidad.

Prueba de Kolmogorov-Smirnov para la variable: Población económicamente activa

H_0 : La variable población económicamente activa se ajusta a una distribución Normal con media 2950000 y desviación estándar 640000

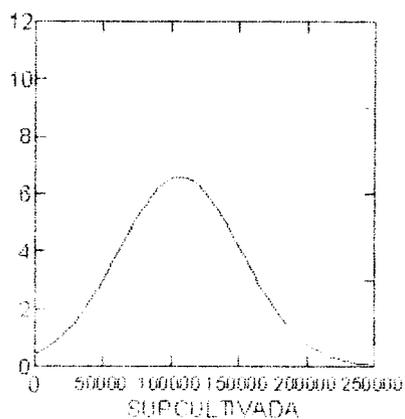
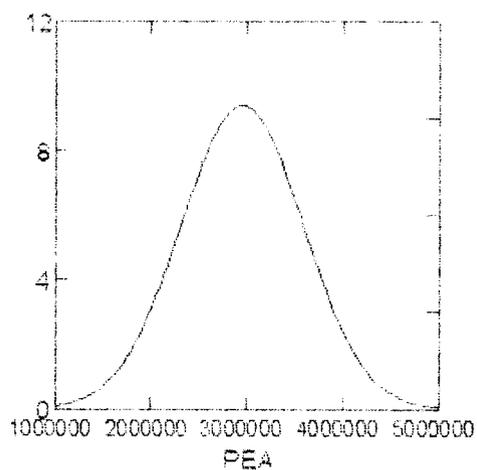
vs.

$H_1: \bar{\neq} H_0$

MaxDif = 0.068

Valor p (2-colas) = 0.999

Hay evidencia estadística para no rechazar la hipótesis nula H_0 . En la figura 3.39 se presenta su función de densidad.

Figura 3.38**Superficie cultivada con banano****Figura 3.39****Población económicamente activa**

3.2 Análisis Multivariado

3.2.1 Matriz de correlación de las variables de estudio

En el APENDICE C se presenta la matriz de correlación de las variables de estudio, mediante la cual se determinará la relación lineal existente entre ellas.

La variable X_1 (Exportaciones bananeras en miles de TM) está altamente correlacionada con las variables: X_2 (Exportaciones bananeras en miles de dólares FOB), X_3 (Producción bananera en miles de TM), X_7 (Índice de precios al consumidor, alimentos), X_8 (Cotización del dólar) y X_{10} (Población económicamente activa). La variable X_2 (Exportaciones bananeras en miles de dólares FOB) está altamente correlacionada con las variables: X_3 (Producción bananera en miles de TM), X_7 (Índice de precios al consumidor, alimentos), X_8 (Cotización del dólar) y X_{10} (Población económicamente activa). La variable X_3 (Producción bananera en miles de TM) está altamente correlacionada con las variables: X_4 (Participación de las exportaciones bananeras en el PIB), X_7 (Índice de precios al consumidor, alimentos), X_8 (Cotización del dólar). La variable X_4 (Participación de las exportaciones bananeras en el PIB) está altamente correlacionada con la variable X_5 (Participación de las exportaciones bananeras en las exportaciones totales). La variable X_5 (Participación de las exportaciones bananeras en las

exportaciones totales) está altamente correlacionada con la variable X_9 (Superficie cultivada con banano). La variable X_6 (Índice de precios al consumidor) no presenta una correlación representativa con alguna de las variables. La variable X_7 (Índice de precios al consumidor, alimentos) está altamente correlacionada con las variables: X_8 (Cotización del dólar) y X_{10} (Población económicamente activa). La variable X_8 (Cotización del dólar) está altamente correlacionada con la variable X_{10} (Población económicamente activa).

Se aplicará la técnica multivariada de Componentes Principales.

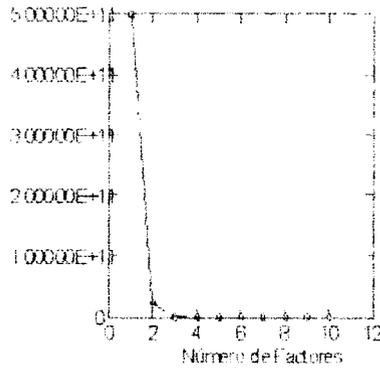
3.2.2 Análisis de componentes principales

Se realizará el análisis de componentes principales utilizando los datos originales. Los valores propios obtenidos a partir de la matriz de covarianza de la muestra son:

$$\begin{array}{ll}
 \lambda_1 = 4.98315E+11 & \lambda_6 = 24256.032 \\
 \lambda_2 = 2.28127E+10 & \lambda_7 = 147.000 \\
 \lambda_3 = 5.95232E+08 & \lambda_8 = 40.539 \\
 \lambda_4 = 366331.637 & \lambda_9 = 11.920 \\
 \lambda_5 = 162291.928 & \lambda_{10} = 0.139
 \end{array}$$

Figura 3.40

Valores propios



La varianza de la primera componente es $4.98315E+11$, lo que representa un porcentaje de explicación de la varianza total de 95.51%, la varianza de la segunda componente es $2.28127E+10$, lo que representa un porcentaje de explicación de la varianza total de 4.37% y la varianza de la tercera componente es $5.95232E+08$, lo que representa un porcentaje de explicación de la varianza total de 0.11%. Lo que significa que las tres componentes explican casi la totalidad de la varianza.

La matriz de covarianzas de las componentes con las variables de estudio se presenta en el APENDICE D. Las variables que presentan los mayores valores de covarianza con la primera componente son: Exportaciones bananeras en miles de TM.,

exportaciones bananeras en miles de dólares FOB, cotización del dólar e índice de precios al consumidor (alimentos). Con la segunda componente son: Población económicamente activa, superficie cultivada, producción bananera en miles de TM., participación de las exportaciones bananeras en el PIB y participación de las exportaciones bananeras en las exportaciones totales. El índice de precios al consumidor presenta el mayor valor de covarianza con la séptima componente.

La matriz de carga (vectores propios) obtenida a partir de la matriz de covarianza es:

	Y ₁	Y ₂	Y ₃
X ₁₀	0.897	-0.424	0.122
X ₂	0.441	0.871	-0.215
X ₉	-0.015	0.247	0.969
X ₈	0.002	0.004	0.008
X ₁	0.001	0.003	0.006
X ₃	0.001	0.006	0.02
X ₇	0	0	0
X ₆	0	0	0
X ₅	0	0	0
X ₄	0	0	0

En la primera componente, la variable X₁₀ presenta el mayor coeficiente. El coeficiente de la variable X₂ difiere mucho de la

anterior. Los demás coeficientes son sumamente pequeños. En la segunda componente, la variable X_2 presenta el mayor coeficiente, el cual contrasta con el coeficiente de la variable X_{10} . El coeficiente de la variable X_9 es mayor que los demás, ya que estos tienden a cero. En la tercera componente, el mayor coeficiente lo presenta la variable X_9 .

Las componentes resultantes son:

$$Y_1 = 0.897 \cdot X_{10} + 0.441 \cdot X_2 - 0.015 \cdot X_9 + 0.002 \cdot X_8 + 0.001 \cdot X_1 + 0.001 \cdot X_3$$

$$Y_2 = -0.424 \cdot X_{10} + 0.871 \cdot X_2 + 0.247 \cdot X_9 + 0.004 \cdot X_8 + 0.003 \cdot X_1 + 0.006 \cdot X_3$$

$$Y_3 = 0.122 \cdot X_{10} - 0.215 \cdot X_2 + 0.969 \cdot X_9 + 0.008 \cdot X_8 + 0.006 \cdot X_1 + 0.02 \cdot X_3$$

A pesar de haber obtenido nueve componentes principales, tan solo con las tres primeras componentes se explica casi la totalidad de la varianza, debido a que los valores propios correspondientes son sumamente grandes en comparación con los demás, cumpliéndose el objetivo de las componentes principales, que es considerar la variabilidad de todo el sistema

en el menor número de componentes, logrando reducir los datos, en virtud de las diferentes escalas en las que se presentan X_1, X_2, \dots, X_{10} , trataremos ahora a estas variables, pero estandarizadas.

Se realizará el análisis de componentes principales utilizando los datos estandarizados. Los valores propios obtenidos a partir de la matriz de correlación de la muestra son:

$$\lambda_1 = 5.930$$

$$\lambda_2 = 2.682$$

$$\lambda_3 = 0.831$$

$$\lambda_4 = 0.233$$

$$\lambda_5 = 0.177$$

$$\lambda_6 = 0.091$$

$$\lambda_7 = 0.026$$

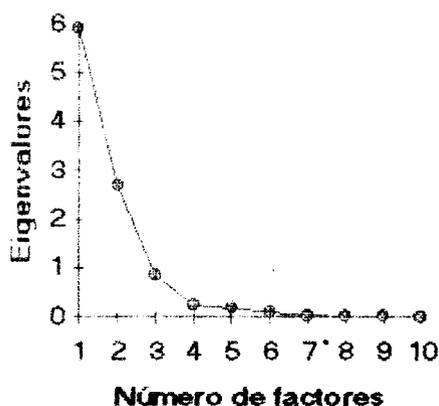
$$\lambda_8 = 0.020$$

$$\lambda_9 = 0.008$$

$$\lambda_{10} = 0.001$$

Figura 3.41

Valores propios (eigenvalores)



La varianza de la primera componente es 5.93, lo que representa un porcentaje de explicación de la varianza total de 59.297%, mientras que la varianza de la segunda componente es 2.682, lo que representa un porcentaje de explicación de la varianza total de 26.819%. Lo que significa que las dos componentes explican el 86.1% de la varianza total.

El análisis de factores determinó que la información de las diez variables de estudio propuestas, pueden ser explicadas mediante dos componentes principales, cuyas correlaciones con las variables de estudio se presentan a continuación:

	1	2
X ₁	0.982	0.094
X ₈	0.956	0.158
X ₂	0.951	0.230
X ₇	0.942	0.136
X ₃	0.905	-0.274
X ₁₀	0.762	0.613
X ₄	0.720	-0.534
X ₉	0.338	-0.895
X ₅	0.357	-0.845
X ₆	0.315	0.571

Se puede observar que en la primera componente, las variables con altas correlaciones son : Exportaciones bananeras en miles de TM, cotización del dólar , exportaciones bananeras en miles de dólares FOB, índice de precios al consumidor del rubro alimentos, producción bananera en miles de TM, población económicamente activa, porcentaje de participación de las exportaciones bananeras en el PIB. Mientras que en la segunda componente, las variables con altas correlaciones son: Superficie cultivada de banano en hectáreas, porcentaje de participación de las exportaciones bananeras en las exportaciones totales, índice de precios al consumidor.

La matriz de cargas (vectores propios) obtenida a partir de la matriz de correlación es:

	Y ₁	Y ₂
X ₁	0.403	0.057
X ₈	0.393	0.096
X ₂	0.390	0.140
X ₇	0.387	0.083
X ₃	0.372	-0.167
X ₁₀	0.313	0.374
X ₄	0.296	-0.326
X ₉	0.139	-0.547
X ₅	0.147	-0.516
X ₆	0.129	0.349

En la primera componente, la variable X₁ tiene el mayor coeficiente. Las variables X₈, X₂, X₇, X₃ tienen coeficientes similares. En la segunda componente, la variable X₉ tiene el mayor coeficiente. Las variables X₄, X₅ y X₉ contrastan con las variables X₁₀ y X₆.

Las componentes principales resultantes son:

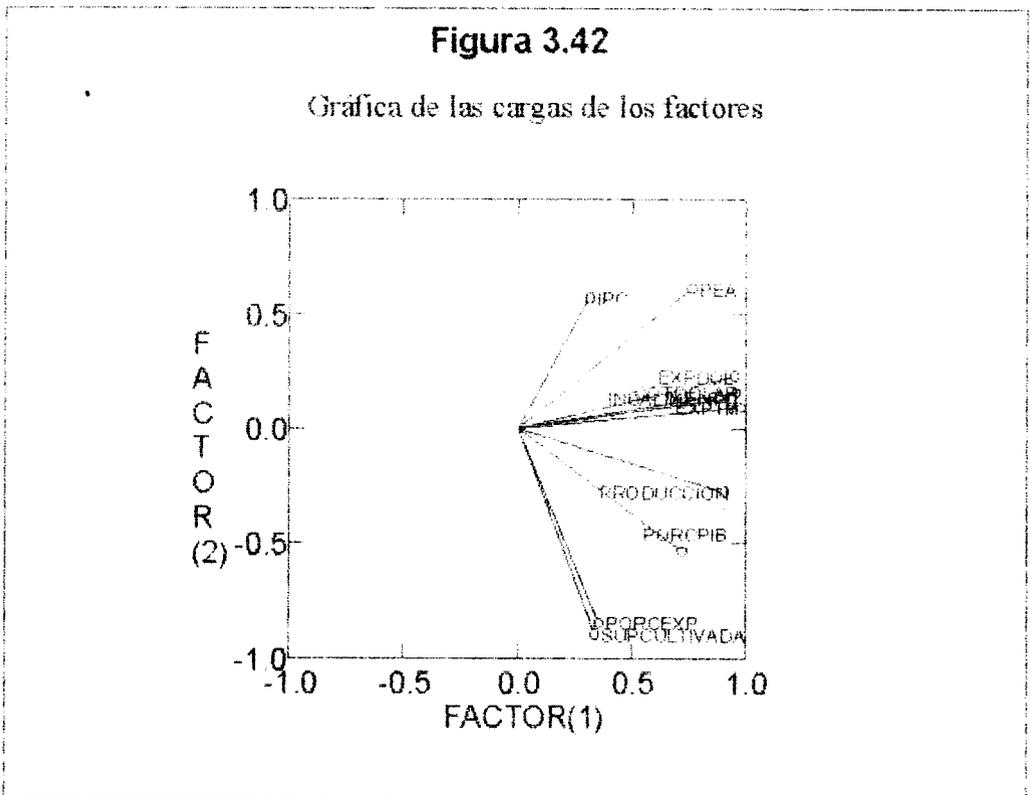
Exportaciones

$$Y_1 = 0.403 \cdot X_1 + 0.393 \cdot X_8 + 0.39 \cdot X_2 + 0.387 \cdot X_7 + 0.372 \cdot X_3 + 0.313 \cdot X_{10} + 0.296 \cdot X_4 + 0.139 \cdot X_9 + 0.147 \cdot X_5 + 0.129 \cdot X_6$$

Rendimiento económico de los cultivos bananeros

$$Y_2 = 0.057 \cdot X_1 + 0.096 \cdot X_8 + 0.14 \cdot X_2 + 0.083 \cdot X_7 - 0.167 \cdot X_3 + 0.374 \cdot X_{10} - 0.326 \cdot X_4 - 0.547 \cdot X_9 - 0.516 \cdot X_5 + 0.349 \cdot X_6$$

En la figura 3.42 se puede observar que las variables altamente correlacionadas con la primera componente, se encuentran a mayor distancia del eje correspondiente al Factor (1) y más cercanas al eje correspondiente al Factor (2), mientras que con las variables altamente correlacionadas con la segunda componente ocurre lo contrario.



También se puede realizar el análisis de componentes principales utilizando los datos rotados. Aplicando el método de rotación de factores (Varimax) a los datos originales, se obtienen ocho

componentes que explican el 100% de la varianza total. Las tres primeras componentes explican el 79.9%. Si se aplica el método de rotación a los datos estandarizados, se obtienen dos componentes que explican el 84.5% de la varianza total.

En conclusión, si se toma como criterio de elección al porcentaje de explicación de la varianza, el análisis de componentes principales obtenido utilizando los datos originales es el más adecuado, ya que explica con tres componentes casi el 100% de la varianza total, pero se debe recordar que las variables de estudio tienen diferentes escalas.

Componentes principales obtenidas utilizando los datos originales

$$Y_1 = 0.897 \cdot X_{10} + 0.441 \cdot X_2 - 0.015 \cdot X_9 + 0.002 \cdot X_8 + 0.001 \cdot X_1 + 0.001 \cdot X_3$$

$$Y_2 = -0.424 \cdot X_{10} + 0.871 \cdot X_2 + 0.247 \cdot X_9 + 0.004 \cdot X_8 + 0.003 \cdot X_1 + 0.006 \cdot X_3$$

$$Y_3 = 0.122 \cdot X_{10} - 0.215 \cdot X_2 + 0.969 \cdot X_9 + 0.008 \cdot X_8 + 0.006 \cdot X_1 + 0.02 \cdot X_3$$

Aquellas variables cuyas varianzas son grandes, pueden influir notablemente en el porcentaje de explicación de la varianza de la componente a la que pertenecen, mientras que las variables cuya varianza sea pequeña, influirán muy poco. Por ello, es preferible tomar como base del análisis de componentes principales a la matriz de correlación (datos estandarizados), ya que sus escalas no pueden incidir en el porcentaje de varianza total explicada.

Componentes principales obtenidas utilizando los datos estandarizados

Exportaciones

$$Y_1 = 0.403 \cdot X_1 + 0.393 \cdot X_8 + 0.39 \cdot X_2 + 0.387 \cdot X_7 + 0.372 \cdot X_3 + 0.313 \cdot X_{10} + 0.296 \cdot X_4 + 0.139 \cdot X_9 + 0.147 \cdot X_5 + 0.129 \cdot X_6$$

Rendimiento económico de los cultivos bananeros

$$Y_2 = 0.057 \cdot X_1 + 0.096 \cdot X_8 + 0.14 \cdot X_2 + 0.083 \cdot X_7 - 0.167 \cdot X_3 + 0.374 \cdot X_{10} - 0.326 \cdot X_4 - 0.547 \cdot X_9 - 0.516 \cdot X_5 + 0.349 \cdot X_6$$

Por lo tanto, se selecciona el análisis de componentes principales obtenido utilizando los datos estandarizados.

CAPITULO 4

4. ANALISIS DE DATOS UTILIZANDO SERIES TEMPORALES Y MODELOS DE REGRESION

En este capítulo, se realizará el análisis de datos utilizando series temporales y modelos de regresión. Para el análisis, mediante series temporales, se consideraron las siguientes variables: exportaciones de banano en miles de toneladas métricas, exportaciones de banano en miles de dólares FOB, producción de banano en miles de toneladas métricas, porcentaje de participación de las exportaciones bananeras en el PIB, el índice de precios al consumidor (alimentos), cotización del dólar y población económicamente activa. Se presentarán 49 observaciones de cada variable, las cuales representan los valores anuales registrados entre los años 1950-1998. Para el análisis, mediante modelos de regresión, se considerarán, además de las variables mencionadas, las siguientes: índice de precios al consumidor, participación de las exportaciones bananeras en las exportaciones totales y superficie cultivada. Se presentarán 30 observaciones anuales de cada variable, correspondientes al periodo 1969-1998.

4.1 Series temporales

Para cada serie, se probarán modelos estacionarios ARMA, no estacionarios ARIMA y estacionales SARIMA, de acuerdo a las características que presente cada una.

Exportaciones bananeras en miles de TM

La serie de las exportaciones bananeras en miles de TM. se presenta en la figura 4.1. El número de parámetros media móvil q es 3, ya que los tres primeros valores de autocorrelaciones superan el límite de confianza superior (LCS), y el número de parámetros autoregresivos p es 1, ya que el primer valor de las autocorrelaciones parciales supera el límite de confianza superior (LCS). Así lo podemos observar en las figuras 4.2 y 4.3 que presentan las gráficas de las autocorrelaciones y las autocorrelaciones parciales, con sus respectivos intervalos de confianza superior (LCS) e inferior (LCI).

Se probará el modelo ARMA(1,3). Los resultados indican que el modelo no es estacionario, ya que no cumple la propiedad de que su distribución conjunta sea invariante en el tiempo, y que además los

Figura 4.1

Exportaciones bananeras del Ecuador en miles de TM.

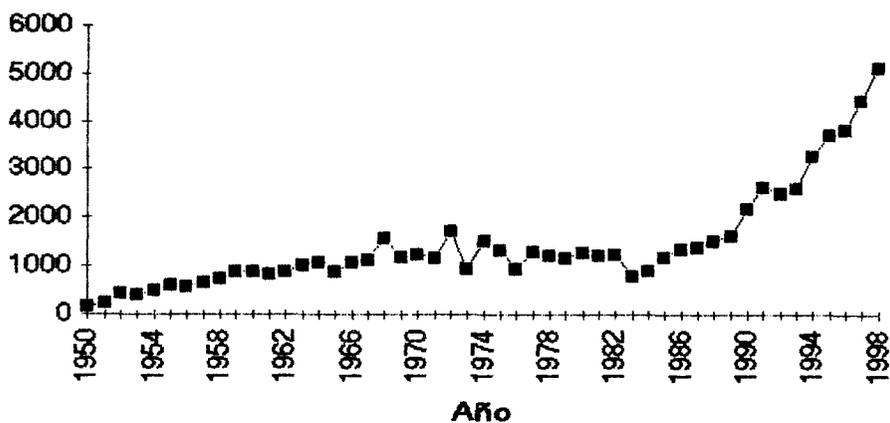
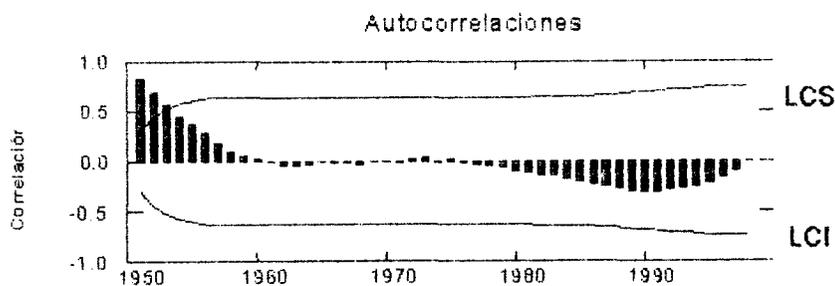


Figura 4.2

Autocorrelaciones de las exportaciones bananeras (TM)



parámetros media móvil no son necesarios, ya que sus intervalos de confianza incluyen al cero.

Se aplicará una diferenciación d de orden 1, para transformar la serie no estacionaria a estacionaria, y se probará el modelo ARIMA(1,1,0). Los resultados indican que el modelo no pasa la prueba de parámetros, es decir que el intervalo de confianza del parámetro autoregresivo incluye al cero.

ARIMA (1,1,0)

La varianza del modelo es 98375.902

Indice	Tipo	Estimador	Inferior <95%>	Superior
1	AR	-0.033	-0.343	0.277

La serie no tiene un comportamiento estacional, por lo que no es posible aplicar modelos estacionales SARIMA. Así lo indican las funciones de autocorrelaciones y autocorrelaciones parciales.

Exportaciones bananeras en miles de dólares FOB

La serie de las exportaciones bananeras en miles de dólares FOB, se presenta en la figura 4.4. El número de parámetros media móvil q es 3, ya que los tres primeros valores de las autocorrelaciones superan

Figura 4.3

Autocorrelaciones parciales de las exportaciones bananeras (TM)

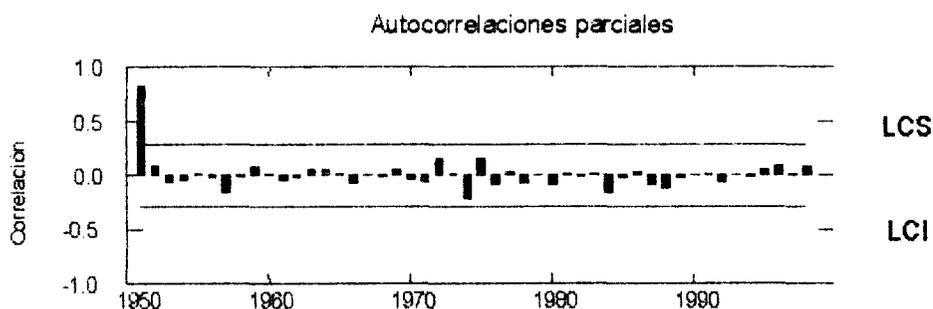
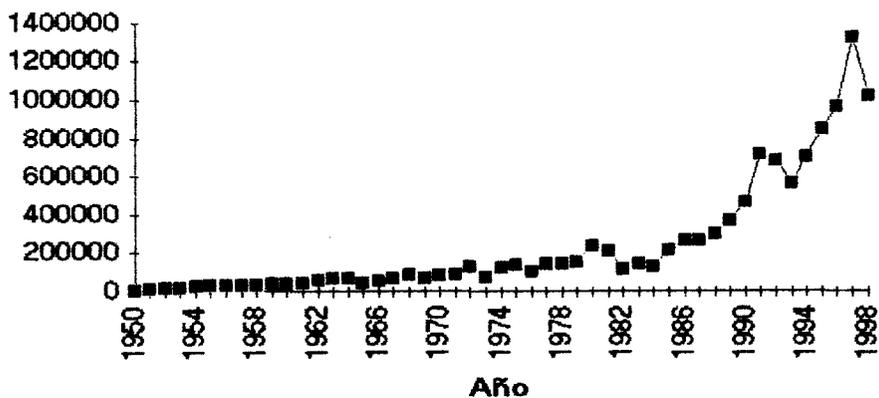


Figura 4.4

Exportaciones bananeras del Ecuador
en miles de dólares FOB

el límite de confianza superior (LCS), y el número de parámetros autoregresivos p es 1, ya que el primer valor de las autocorrelaciones parciales supera el límite de confianza superior (LCS). Así lo podemos constatar en las figuras 4.5 y 4.6 que presentan las gráficas de las autocorrelaciones y las autocorrelaciones parciales con sus respectivos intervalos de confianza superior (LCS) e inferior (LCI).

Los modelos ARMA(1,2) y ARMA(1,1) no son estacionarios, pues no cumplen con la propiedad de que sus distribuciones conjuntas sean invariantes en el tiempo. Para transformar la serie no estacionaria a estacionaria, se le aplicará una diferenciación d de orden 1 y se probará el modelo ARIMA(1,1,1). Los resultados indican que el modelo no pasa la prueba de parámetros, ya que ambos parámetros, el media móvil y el autoregresivo, incluye en sus respectivos intervalos de confianza al cero. Esto significa que el parámetro objetivo puede tomar el valor de cero, lo cual implicaría que no es necesario en el modelo.

ARIMA (1,1,1)

La varianza del modelo es 8.72776E+09

Indice	Tipo	Estimador	Inferior 95%	Superior 95%
1	AR	-0.111	-2.949	2.728
2	MA	-0.048	-2.824	2.727

Figura 4.5

Autocorrelaciones de las exportaciones bananeras (dólares FOB)

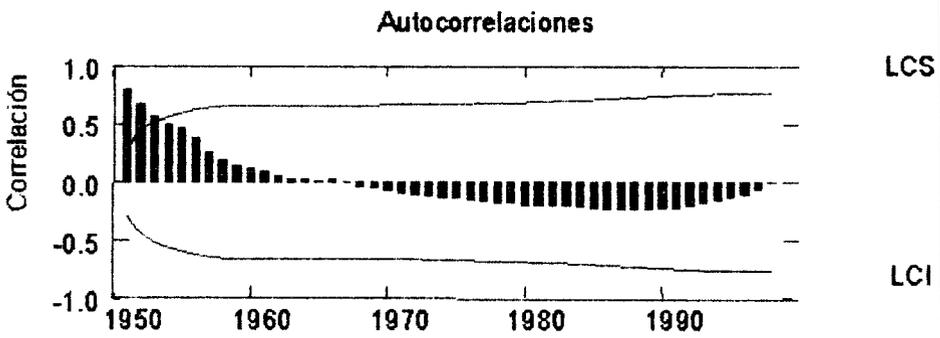
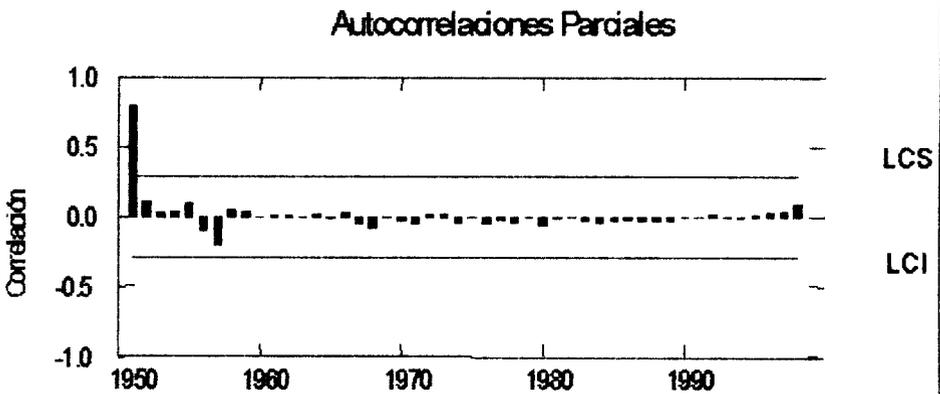


Figura 4.6

Autocorrelaciones parciales de las exportaciones bananeras (dólares FOB)



Matriz de correlación de los estimadores de los parámetros

	1	2
1	1.000	
2	0.993	1.000

La serie no tiene un comportamiento estacional, por lo que no es posible aplicar modelos estacionales SARIMA. Así lo indican las funciones de autocorrelaciones y autocorrelaciones parciales.

Producción bananera en miles de TM

La serie de la producción bananera en miles de TM presenta ciertas fluctuaciones. Así lo podemos notar en la figura 4.7. El número de parámetros media móvil q es 2, ya que los dos primeros valores de autocorrelaciones superan el límite de confianza superior (LCS), y el número de parámetros autoregresivos p es 1, ya que el primer valor de las autocorrelaciones parciales supera el límite de confianza superior (LCS). Así lo podemos observar en las figuras 4.8 y 4.9 que presentan las gráficas de las autocorrelaciones y las autocorrelaciones parciales con sus respectivos intervalos de confianza superior (LCS) e inferior (LCI).

Se probará el modelo ARMA(1,2). Los resultados indican que el modelo no es estacionario, ya que no cumple con la propiedad de que

Figura 4.7

Producción bananera del Ecuador en miles de TM.

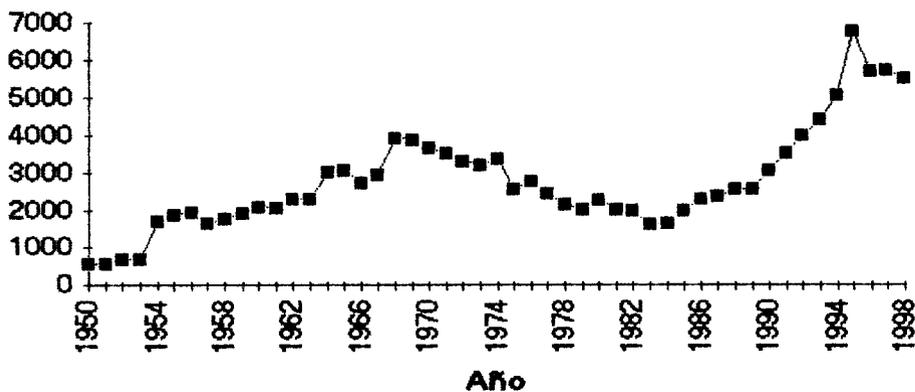
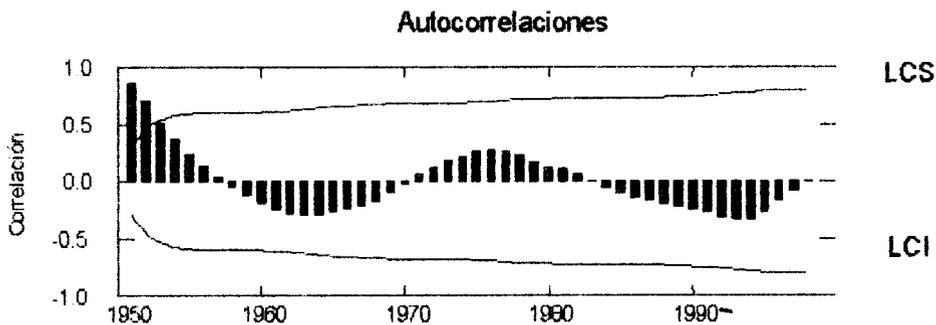


Figura 4.8

Autocorrelaciones de la producción bananera (TM)



su distribución conjunta sea invariante en el tiempo, y que además los parámetros media móvil no son necesarios, ya que sus intervalos de confianza incluyen al cero, es decir que el parámetro objetivo puede tomar el valor de cero.

Para transformar la serie no estacionaria a estacionaria, se le aplicará una diferenciación d de orden 1 y se probará el modelo ARIMA(1,1,0). Los resultados indican que el modelo no pasa la prueba de parámetros, es decir que el intervalo de confianza del parámetro autoregresivo incluye al cero, lo que implicaría que puede tomar el valor de cero y por lo tanto, no ser necesario en el modelo.

La serie tiene un comportamiento estacional, pero no lo suficientemente relevante para aplicar modelos estacionales SARIMA, ya que los valores de las autocorrelaciones y autocorrelaciones parciales, correspondientes a los periodos estacionales, no superan los límites de confianza respectivos. Así lo indican las funciones de autocorrelaciones y autocorrelaciones parciales.

Participación de las exportaciones bananeras en el PIB

La serie de la participación de las exportaciones bananeras en el PIB, presenta varias fluctuaciones. Así lo podemos observar en la figura 4.10. El número de parámetros media móvil q es 3, ya que los tres primeros valores de las autocorrelaciones superan el límite de confianza superior (LCS) y el número de parámetros autoregresivos p

Figura 4.9

Autocorrelaciones parciales de la producción bananera (TM)

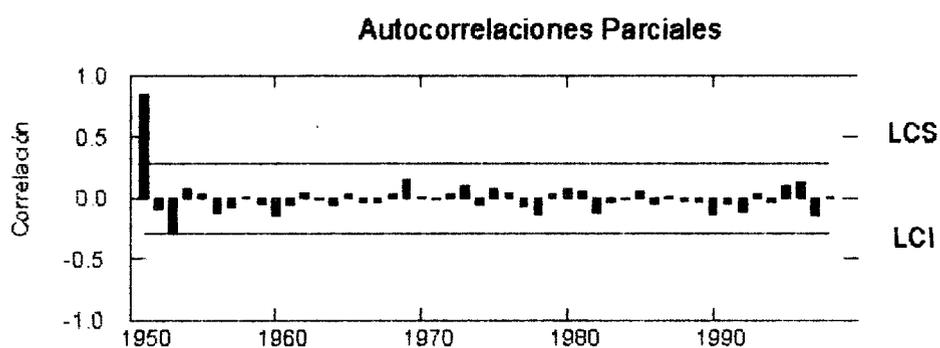
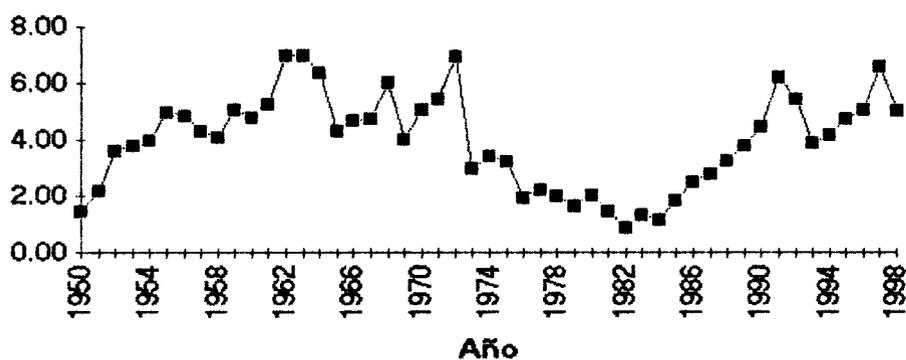


Figura 4.10

Participación de las exportaciones bananeras en el PIB



es 1, ya que el primer valor de las autocorrelaciones parciales supera el límite de confianza superior (LCS). Así lo podemos constatar en la figura 4.11 y 4.12 que presentan las gráficas de las autocorrelaciones y las autocorrelaciones parciales con sus respectivos intervalos de confianza superior (LCS) e inferior (LCI).

Se probará el modelo ARMA(1,3). Los resultados indican que el modelo no pasa la prueba de parámetros, porque los parámetros media móvil no son necesarios, ya que sus intervalos de confianza incluyen al cero, es decir que pueden tomar el valor de cero. El modelo ARMA (1,0) sí pasa la prueba de parámetros. Tanto los valores reales como los estimados presentan un descenso entre 1997 y 1998, pero difieren notablemente. Las predicciones para los años 1997 y 1998 se muestran en la figura 4.13.

ARMA(1,3)

La varianza del modelo es 1.238

Indice	Tipo	Estimador	Inferior 95%	Superior 95%
1	AR	0.994	0.928	1.061
2	MA	0.134	-0.192	0.459
3	MA	0.002	-0.278	0.281
4	MA	0.048	-0.262	0.357

Figura 4.11

Autocorrelaciones de la participación de las exportaciones bananeras en el PIB

Autocorrelaciones

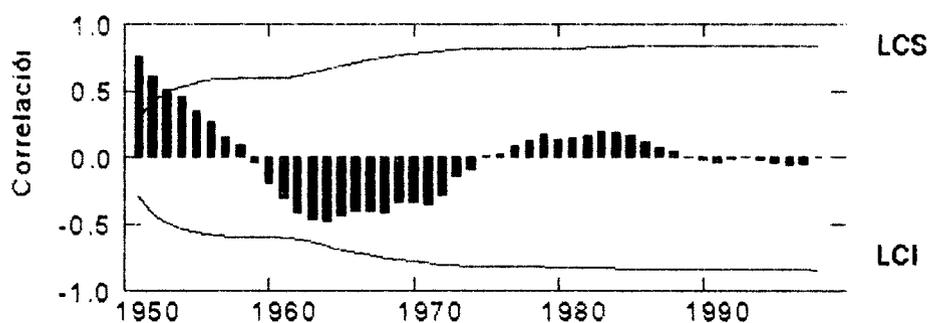
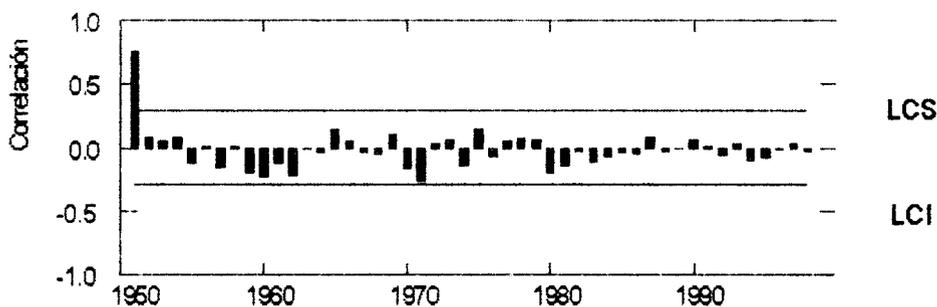


Figura 4.12

Autocorrelaciones parciales de la participación de las exportaciones bananeras en el PIB

Autocorrelaciones parciales



Matriz de correlación de los estimadores de los parámetros

	1	2	3	4
1	1.000			
2	0.205	1.000		
3	0.156	-0.107	1.000	
4	0.126	-0.261	-0.101	1.000

ARMA(1,0)

La varianza del modelo es 1.186

Indice	Tipo	Estimador	Inferior <95%>	Superior
1	AR	0.982	0.909	1.056

Período	Inferior 95%	Predicción	Superior 95%
1997	2.831	4.965	7.100
1998	1.886	4.877	7.869
1999	1.159	4.790	8.422
2000	0.687	4.758	8.830
2001	0.165	4.681	9.197
2002	-0.303	4.605	9.514

La serie tiene, aproximadamente, un comportamiento estacional, pero no lo suficientemente relevante para aplicar modelos estacionales SARIMA, ya que los valores de las autocorrelaciones y autocorrelaciones parciales, correspondientes a los periodos estacionales, no superan los límites de confianza respectivos. Así lo

Figura 4.13

Predicciones del modelo ARMA(1,0)
para la participación de las
exportaciones bananeras en el PIB

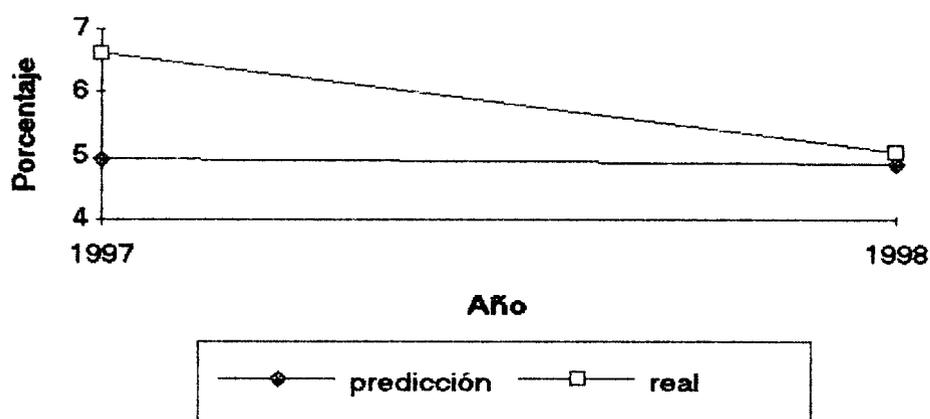
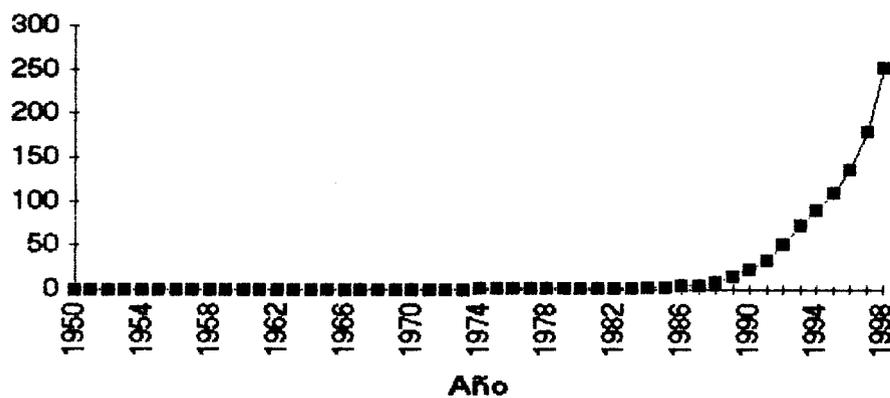


Figura 4.14

Cotización del dólar en el Ecuador



indican las funciones de autocorrelaciones y autocorrelaciones parciales.

Cotización del dólar

La serie de la cotización del dólar se presenta en la figura 4.14. El número de parámetros media móvil q es 2, ya que los dos primeros valores de las autocorrelaciones superan el límite de confianza superior (LCS) y el número de parámetros autoregresivos p es 1, ya que el primer valor de las autocorrelaciones parciales supera el límite de confianza superior (LCS). Así lo podemos notar en las figuras 4.15 y 4.16 que presentan las gráficas de las autocorrelaciones y las autocorrelaciones parciales con sus respectivos intervalos de confianza superior (LCS) e inferior (LCI).

Se probará el modelo ARMA(1,2). Los resultados indican que el modelo no es estacionario, ya que no cumple con la propiedad de que su distribución conjunta sea invariante en el tiempo. Para transformar la serie no estacionaria a estacionaria, se le aplicará una diferenciación d de orden 1 y se probará el modelo ARIMA(1,1,2). Los resultados indican que este modelo, al igual que el modelo ARMA(1,2), tampoco es estacionario. La serie no muestra un comportamiento estacional, por lo que no es posible aplicar modelos estacionales SARIMA. Así lo indican las funciones de autocorrelaciones y autocorrelaciones parciales.

Figura 4.15

Autocorrelaciones de la cotización del dólar

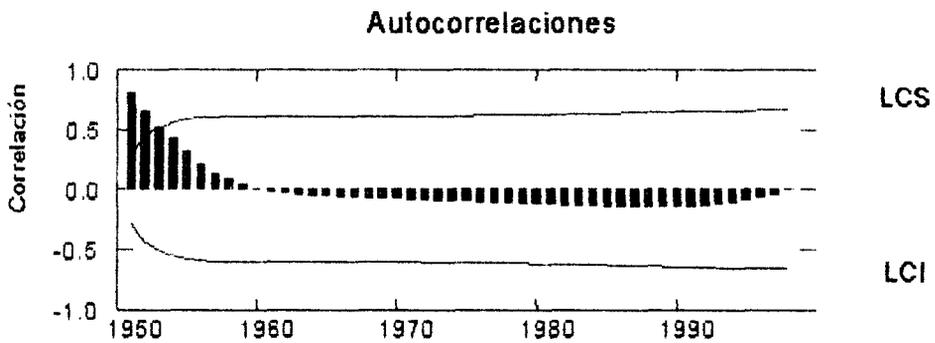
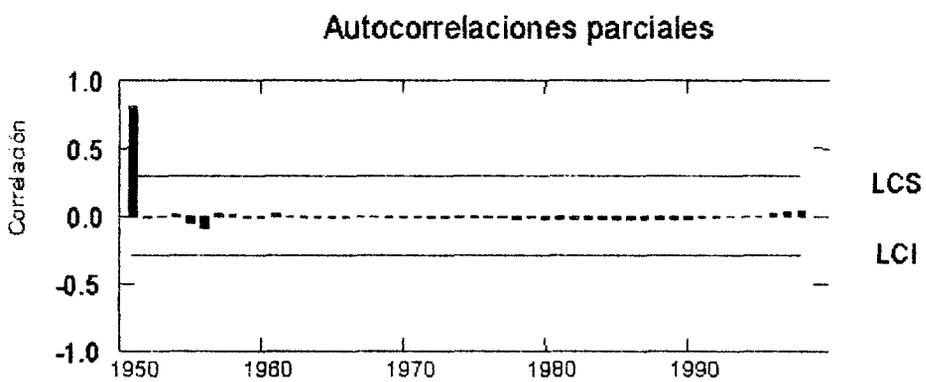


Figura 4.16

Autocorrelaciones parciales de la cotización del dólar



Indice de precios al consumidor (alimentos)

La serie de los índices de precios al consumidor (alimentos) se presenta en la figura 4.17. El número de parámetros media móvil q es 2, ya que los dos primeros valores de las autocorrelaciones superan el límite de confianza superior (LCS) y el número de parámetros autoregresivos p es 1, ya que el primer valor de las autocorrelaciones parciales supera el límite de confianza superior (LCS). Así lo podemos notar en las figuras 4.18 y la 4.19 que presentan las gráficas de las autocorrelaciones y las autocorrelaciones parciales con sus respectivos intervalos de confianza superior (LCS) e inferior (LCI).

Se probará el modelo ARMA(1,2). Los resultados indican que el modelo no es estacionario, ya que no cumple con la propiedad de que su distribución conjunta sea invariante en el tiempo. Para transformar la serie no estacionaria a estacionaria, se le aplicará una diferenciación d de orden 1 y se probará el modelo ARIMA(1,1,2). Los resultados indican que este modelo tampoco es estacionario. La serie no muestra un comportamiento estacional, por lo que no es posible aplicar modelos estacionales SARIMA.

Figura 4.17

Indice de precios al consumidor (alimentos)

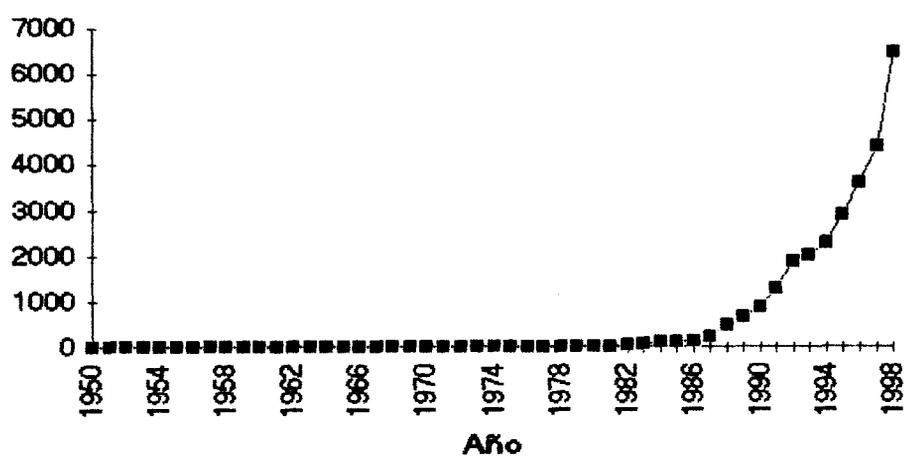
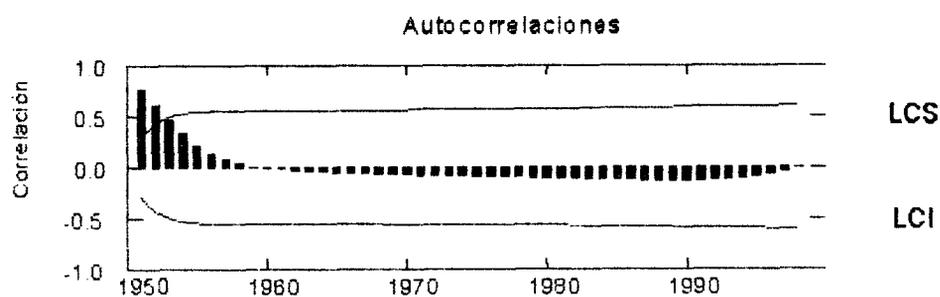


Figura 4.18

Autocorrelaciones del índice de precios al consumidor (alimentos)



Población económicamente activa

La serie de la población económicamente activa del Ecuador, presenta una forma lineal. Así lo podemos ver en la figura 4.20. El número de parámetros media móvil q es 4, ya que los primeros cuatro valores de autocorrelaciones superan el límite de confianza superior (LCS) y el número de parámetros autoregresivos p es 1, ya que el primer valor de las autocorrelaciones parciales supera el límite de confianza superior (LCS). Así lo podemos observar en las figuras 4.21 y la 4.22 que presentan las gráficas de las autocorrelaciones y las autocorrelaciones parciales con sus respectivos intervalos de confianza superior (LCS) e inferior (LCI).

Se probará el modelo ARMA(1,4). Los resultados indican que el modelo no es estacionario, ya que no cumple con la propiedad de que su distribución conjunta sea invariante en el tiempo. Para transformar la serie no estacionaria a estacionaria, se le aplicará una diferenciación d de orden 1. Al aplicar la diferenciación, tanto las autocorrelaciones como las autocorrelaciones parciales, no sobrepasaban los intervalos de confianza correspondientes, por lo que no fue posible probar modelos ARIMA. Además, la serie no muestra un comportamiento estacional, por lo que no es posible aplicar modelos estacionales SARIMA.

Figura 4.19

Autocorrelaciones parciales del índice de precios al consumidor (alimentos)

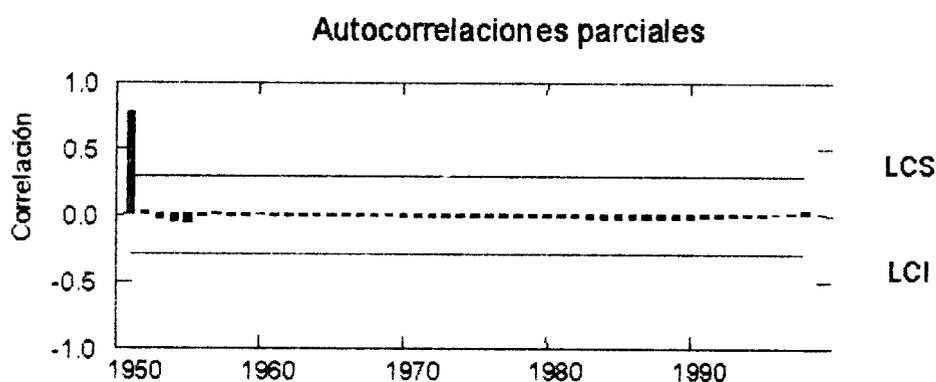


Figura 4.20

Población económicamente activa

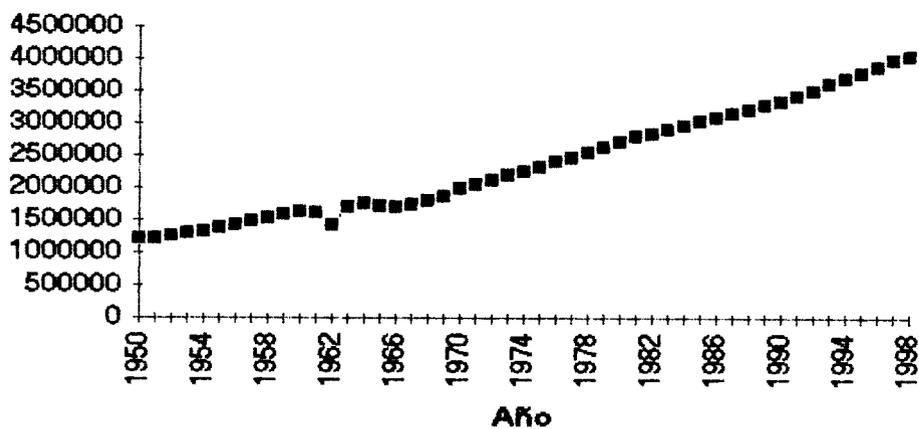


Figura 4.21

Autocorrelaciones de la población económicamente activa

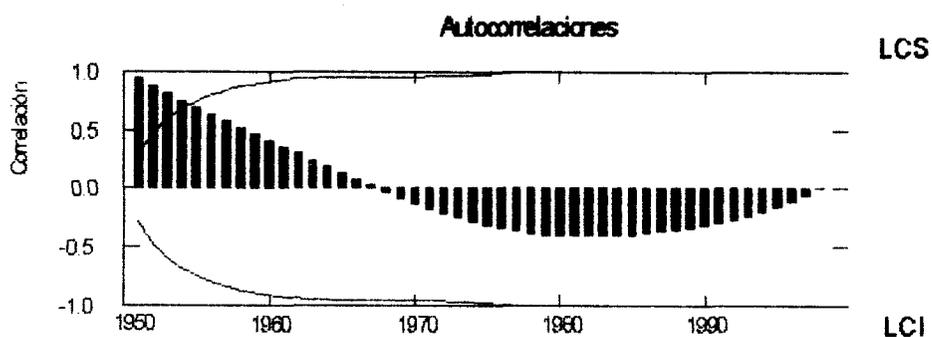
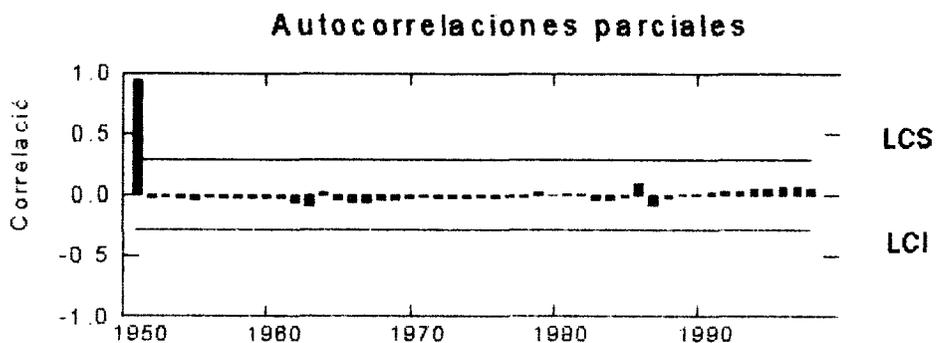


Figura 4.22

Autocorrelaciones parciales de la población económicamente activa



En conclusión, la técnica de series temporales no dió buenos resultados, ya que no se obtuvieron modelos de predicción para las variables de estudio propuestas, a excepción del modelo estacionario, ARMA(1,0), para la participación de las exportaciones bananeras en el PIB, cuyas predicciones carecen de importancia, si no pueden ser relacionadas con las predicciones de las variables fundamentales, tales como, exportaciones y producción bananera.

4.2 Modelos de regresión

Con el análisis de series temporales presentado anteriormente, se pudo concluir que no era la técnica más adecuada para modelar los datos correspondientes a las variables de estudio, especialmente las exportaciones bananeras. Por esta razón, se emplearán modelos de regresión, en los que la variable dependiente o de respuesta será la exportación bananera en miles de TM, y los regresores o variables independientes serán: la producción bananera, la participación de las exportaciones bananeras en el PIB, la participación de las exportaciones bananeras en las exportaciones totales, la cotización del dólar, el índice de precios al consumidor, el índice de precios al consumidor (alimentos), la superficie cultivada y la población económicamente activa. A continuación, se presentarán los diferentes modelos de regresión.

Modelo 1

El primer modelo es de regresión lineal simple:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2), \text{ cov}(\epsilon_i, \epsilon_j) = 0 \text{ para } i \neq j$$

donde: Y representa las exportaciones bananeras en miles de TM; X_1 , el índice de precios al consumidor (alimentos) y ϵ_i es un término de error aleatorio normal con media cero y varianza σ^2 . En la figura 4.23 se presenta la gráfica del modelo junto con los datos de la muestra.

Fuente	Suma de cuadrados	gl	Media cuadrática
Regresión	3.47219E+07	1	3.47219E+07
Residuo	2813464.745	28	100480.884
Total	3.75354E+07	29	

Coefficiente de determinación R^2 = 0.925

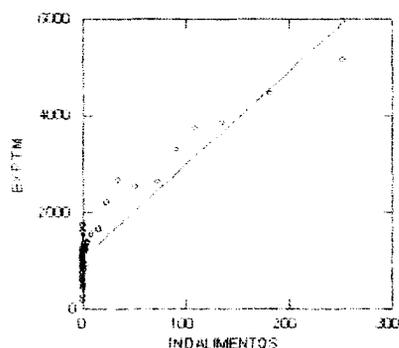
Intervalo de confianza

Parámetro	Estimador	Inferior 95%	Superior 95%
β_0	1314.320	1179.462	1449.178
β_1	17.582	15.645	19.520

El modelo considera el 92.5% de la variabilidad que presentan los datos.

Figura 4.23

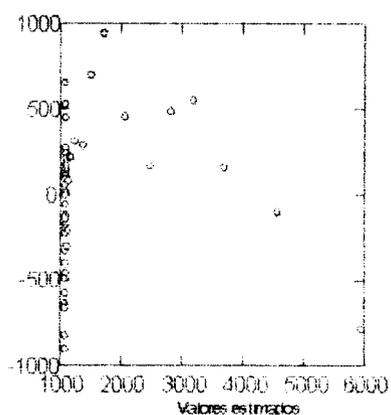
Gráfica de dispersión de las exportaciones bananeras contra el índice de precios al consumidor (alimentos)



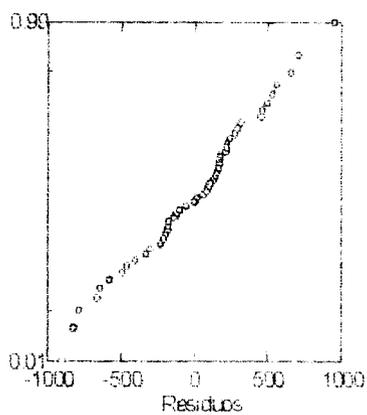
La gráfica de dispersión de los residuos contra los valores estimados, revela un patrón distante del ideal, ya que los datos no se encuentran en forma homogénea alrededor de la media, sino que están sumamente concentrados al principio, y a medida que aumenta x o y , la mayoría de los valores están sobre el eje central, lo cual se puede observar en la figura 4.24. Los residuos se ajustan adecuadamente a la distribución normal, lo cual se observa en la figura 4.25.

Figura 4.24

Gráfica de la dispersión de los residuos contra los valores estimados del modelo 1

**Figura 4.25**

Gráfica de probabilidad normal del modelo 1



A continuación se realizará la prueba de Kolmogorov-Smirnov, para probar si los residuos son normales con media cero y varianza s^2 . Dicha varianza es igual a la media cuadrática del residuo.

H_0 : Los residuos son normales con media cero y desviación estándar 316.99

vs.

$H_1: \neg H_0$

Max Dif. = 0.139

Valor p = 0.605

Por lo tanto, hay evidencia estadística para no rechazar la hipótesis nula H_0 .

Se empleará el estadístico de Durbin-Watson para comprobar si los residuos son independientes.

$C_1: \rho = 0$

vs.

$C_2: \rho > 0$

Durbin- Watson = 1.099

El estadístico de Durbin-Watson es menor que el límite de confianza inferior (1.35) correspondiente a $n=30$ y $p-1=1$, donde n representa el

número de observaciones y $(p-1)$ es el número de variables independientes. Por lo tanto, hay evidencia estadística para afirmar que ρ es mayor que cero, es decir que los residuos son dependientes (correlacionados positivamente).

Modelo 2

El segundo modelo es de regresión lineal simple:

$$y_i = \beta_0 + \beta_1 x_{1i} + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2), \text{ cov}(\epsilon_i, \epsilon_j) = 0 \text{ para } i \neq j$$

donde Y representa las exportaciones bananeras en miles de TM; X_1 , la cotización del dólar y ϵ_i es un término de error aleatorio normal con media cero y varianza σ^2 . En la figura 4.26 se presenta la gráfica del modelo junto con los datos de la muestra.

Fuente	Suma de cuadrados	gl	Media cuadrática
Regresión	3.53615E+07	1	3.53615E+07
Residuo	2173924.582	28	77640.164
Total	3.75354E+07	29	

Coefficiente de determinación R^2 = 0.942

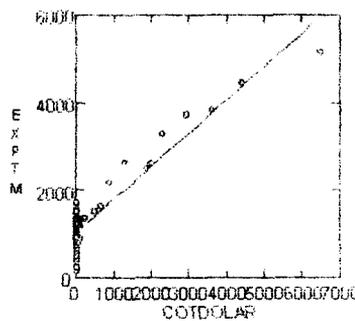
Intervalo de confianza

Parámetro	Estimador	Inferior 95%	Superior 95%
β_0	1245.679	1124.122	1367.236
β_1	0.698	0.631	0.765

El modelo considera el 94.2% de la variabilidad que presentan los datos.

Figura 4.26

Gráfica de la dispersión de las exportaciones bananeras contra la cotización del dólar



La dispersión de los residuos contra los valores estimados revela un patrón muy distante del ideal, ya que los datos no se encuentran en forma homogénea alrededor de la media, sino que están sumamente concentrados al principio, y a medida que aumenta x o y , los valores se sitúan sobre el eje central, lo cual se puede ver en la figura 4.27.

Los residuos se ajustan adecuadamente a la distribución normal, así lo podemos ver en la figura 4.28 que presenta la gráfica de probabilidad normal.

Figura 4.27

Gráfica de la dispersión de los residuos contra los valores estimados del modelo 2

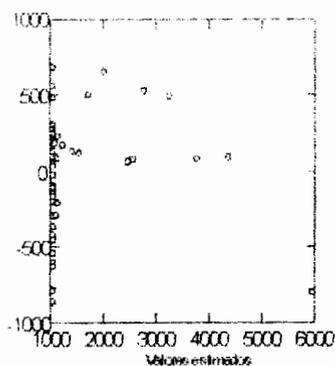
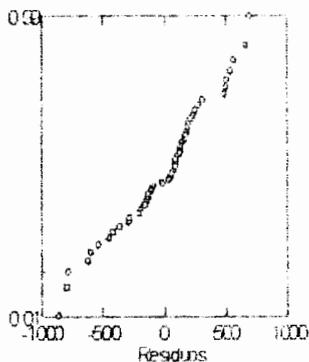


Figura 4.28

Gráfica de probabilidad normal del modelo 2



A continuación se realizará la prueba de Kolmogorov-Smirnov, para probar si los residuos son normales con media cero y varianza σ^2 . Dicha varianza es igual a la media cuadrática del residuo.

H_0 : Los residuos son normales con media cero y desviación estándar 278.64

vs.

$H_1: \neg H_0$

Max Dif. = 0.167

Valor p = 0.372

Por lo tanto, hay evidencia estadística para no rechazar la hipótesis nula H_0 .

Se empleará el estadístico de Durbin-Watson para comprobar si los residuos son independientes.

$C_1: \rho = 0$

vs.

$C_2: \rho > 0$

Durbin- Watson = 1.557

El estadístico de Durbin-Watson es mayor que el límite de confianza superior (1.49) correspondiente a $n=30$ y $p-1=1$, donde n representa el número de observaciones y $(p-1)$ es el número de variables

independientes. Por lo tanto, hay evidencia estadística para afirmar que ρ es igual a cero, es decir que los residuos son independientes..

Modelo 3

El tercer modelo es de regresión lineal múltiple con dos regresores:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2), \text{ cov}(\epsilon_i, \epsilon_j) = 0 \text{ para } i \neq j$$

donde Y representa las exportaciones bananeras en miles de TM; X_1 , la producción bananera en miles de TM, X_2 , la cotización del dólar, y ϵ_i es un término de error aleatorio normal con media cero y varianza σ^2 .

Fuente	Suma de cuadrados	gl	Media cuadrática
Regresión	3.60947E+07	2	1.80474E+07
Residuo	1440672.861	27	53358.254
Total	3.75354E+07	29	

Coefficiente de determinación R^2 = 0.962

Intervalo de confianza

Parámetro	Estimador	Inferior 95%	Superior 95%
β_0	719.754	411.651	1027.856
β_1	0.200	0.090	0.311
β_2	0.558	0.462	0.653

El modelo considera el 96.2% de la variabilidad que presentan los datos.

La dispersión de los residuos contra los valores estimados revela una mayor concentración de datos para los valores iniciales, pero a medida que aumenta x o y , los valores se encuentran en forma homogénea alrededor de la media, lo cual se puede ver en la figura 4.29. Los residuos se ajustan adecuadamente a la distribución normal, así lo podemos ver en la figura 4.30 que presenta la gráfica de probabilidad normal.

Figura 4.29

Gráfica de la dispersión de los residuos contra los valores estimados del modelo 3

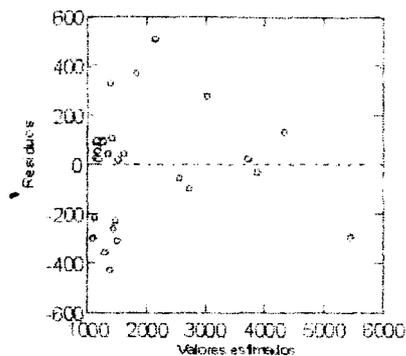
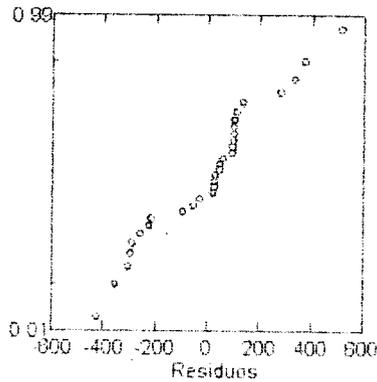


Figura 4.30
Gráfica de probabilidad normal del modelo 3



A continuación se realizará la prueba de Kolmogorov-Smirnov, para probar si los residuos son normales con media cero y varianza σ^2 . Dicha varianza es igual a la media cuadrática del residuo.

H_0 : Los residuos son normales con media cero y desviación estándar 230.99

vs.

$H_1: \neg H_0$

Max Dif. = 0.169

Valor p = 0.362

Por lo tanto, hay evidencia estadística para no rechazar la hipótesis nula H_0 .

Se empleará el estadístico de Durbin-Watson para comprobar si los residuos son independientes.

$$C_1: \rho = 0$$

vs.

$$C_2: \rho > 0$$

$$\text{Durbin-Watson} = 1.898$$

El estadístico de Durbin-Watson es mayor que el límite de confianza superior (1.57) correspondiente a $n=30$ y $p-1=2$, donde n representa el número de observaciones y $(p-1)$ es el número de variables independientes. Por lo tanto, hay evidencia estadística para afirmar que ρ es igual a cero, es decir que los residuos son independientes..

Modelo 4

El cuarto modelo es de regresión lineal múltiple con tres regresores:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2), \text{ cov}(\epsilon_i, \epsilon_j) = 0 \text{ para } i \neq j$$

donde Y representa las exportaciones bananeras en miles de TM; X_1 , el índice de precios al consumidor (alimentos); X_2 , el índice de precios al consumidor; X_3 , la producción bananera en miles de TM., y ϵ_i es un término de error aleatorio normal con media cero y varianza σ^2 .

Fuente	Suma de cuadrados	gl	Media cuadrática
Regresión	3.59236E+07	3	1.19745E+07
Residuo	1611731.199	26	61989.661
Total	3.75354E+07	29	

Coefficiente de determinación R^2 = 0.957

Intervalo de confianza

Parámetro	Estimador	Inferior 95%	Superior 95%
β_0	485.367	83.157	887.576
β_1	12.631	9.850	15.411
β_2	5.718	0.430	11.007
β_3	0.255	0.132	0.370

El modelo considera el 95.7% de la variabilidad que presentan los datos.

La dispersión de los residuos contra los valores estimados revela una alta concentración de datos para los valores iniciales, pero a medida que aumenta x o y , los datos se encuentran en forma homogénea alrededor de la media, lo cual se puede ver en la figura 4.31. Los residuos se ajustan adecuadamente a la distribución normal, así lo podemos ver en la figura 4.32 que presenta la gráfica de probabilidad normal.

Figura 4.31

Gráfica de la dispersión de los residuos contra los valores estimados del modelo 4

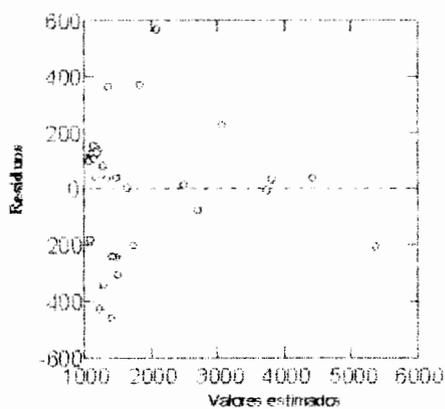
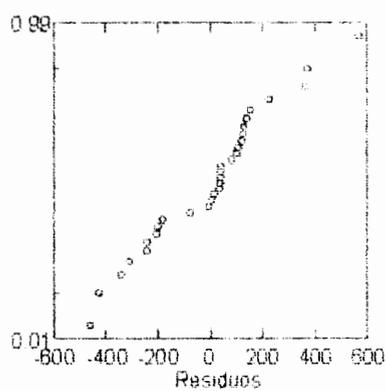


Figura 4.32

Gráfica de probabilidad normal del modelo 4



A continuación se realizará la prueba de Kolmogorov-Smirnov, para probar si los residuos son normales con media cero y varianza σ^2 . Dicha varianza es igual a la media cuadrática del residuo.

H_0 : Los residuos son normales con media cero y desviación estándar 248.98

vs.

$H_1: \neg H_0$

Max Dif. = 0.158

Valor p = 0.439

Por lo tanto, hay evidencia estadística para no rechazar la hipótesis nula H_0 .

Se empleará el estadístico de Durbin-Watson para comprobar si los residuos son independientes.

$C_1: \rho = 0$

vs.

$C_2: \rho > 0$

Durbin- Watson = 1.832

El estadístico de Durbin-Watson es mayor que el límite de confianza superior (1.65) correspondiente a $n=30$ y $p-1=3$, donde n representa el número de observaciones y $(p-1)$ es el número de variables independientes. Por lo tanto, hay evidencia estadística para afirmar que ρ es igual a cero, es decir que los residuos son independientes.

Modelo 5

El quinto modelo es de regresión lineal múltiple con cuatro regresores:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2), \text{ cov}(\epsilon_i, \epsilon_j) = 0 \text{ para } i \neq j$$

donde Y representa las exportaciones bananeras en miles de TM; X_1 , la producción bananera en miles de TM.; X_2 , la participación de las exportaciones bananeras en el PIB; X_3 , la cotización del dólar ; X_4 , la superficie cultivada; y ϵ_i es un término de error aleatorio.

Fuente	Suma de cuadrados	gl	Media cuadrática
Regresión	3.68802E+07	4	9220044.816
Residuo	655197.990	25	26207.920
Total	3.75354E+07	29	

Coefficiente de determinación R^2 = 0.983

Intervalo de confianza

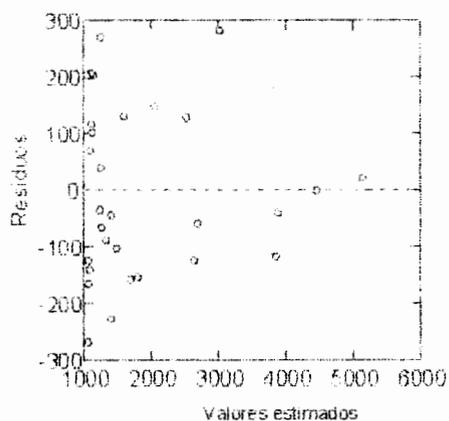
Parámetro	Estimador	Inferior 95%	Superior 95%
β_0	662.623	444.752	880.494
β_1	0.281	0.163	0.399
β_2	132.164	72.792	191.535
β_3	0.456	0.372	0.540
β_4	-0.005	-0.008	-0.003

El modelo considera el 98.3% de la variabilidad que presentan los datos.

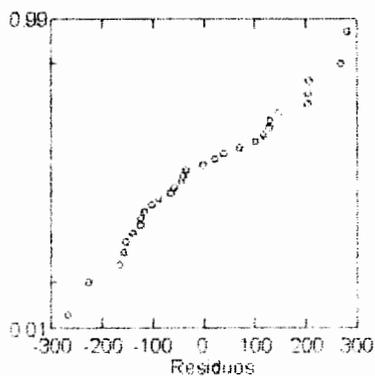
La dispersión de los residuos contra los valores estimados revela una alta concentración de datos para los valores iniciales, pero a medida que aumenta x o y , los datos se encuentran en forma homogénea alrededor de la media, lo cual se puede ver en la figura 4.33. Los residuos se ajustan adecuadamente a la distribución normal, así lo podemos ver en la figura 4.34 que presenta la gráfica de probabilidad normal.

Figura 4.33

Gráfica de la dispersión de los residuos contra los valores estimados del modelo 5

**Figura 4.34**

Gráfica de probabilidad normal del modelo 5



A continuación se realizará la prueba de Kolmogorov-Smirnov, para probar si los residuos son normales con media cero y varianza σ^2 . Dicha varianza es igual a la media cuadrática del residuo.

H_0 : Los residuos son normales con media cero y desviación estándar 161.89

vs.

$H_1: \neg H_0$

Max Dif. = 0.119

Valor p = 0.788

Por lo tanto, hay evidencia estadística para no rechazar la hipótesis nula H_0 .

Se empleará el estadístico de Durbin-Watson para comprobar si los residuos son independientes.

$C_1: \rho = 0$

vs.

$C_2: \rho > 0$

Durbin- Watson = 2.026

El estadístico de Durbin-Watson es mayor que el límite de confianza superior (1.74) correspondiente a $n=30$ y $p-1=4$, donde n representa

el número de observaciones y $(p-1)$ es el número de variables independientes. Por lo tanto, hay evidencia estadística para afirmar que ρ es igual a cero, es decir que los residuos son independientes..

Modelo 6

El sexto modelo de regresión es múltiple con cinco regresores:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2), \text{ cov}(\epsilon_i, \epsilon_j) = 0 \text{ para } i \neq j$$

donde Y representa las exportaciones bananeras en miles de TM; X_1 , la producción bananera en miles de TM., X_2 , la participación de las exportaciones bananeras en el PIB, X_3 , el índice de precios al consumidor, X_4 , la cotización del dólar, X_5 , la superficie cultivada, y ϵ_i es un término de error aleatorio.

Fuente	Suma de cuadrados	gl	Media cuadrática
Regresión	3.69647E+07	5	7392947.334
Residuo	570640.584	24	23776.691
Total	3.75354E+07	29	

Coefficiente de determinación R^2 = 0.985

Intervalo de confianza

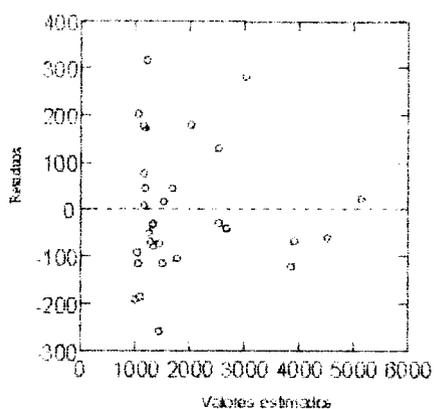
Parámetro	Estimador	Inferior 95%	Superior 95%
β_0	820.286	550.063	1090.510
β_1	0.266	0.152	0.379
β_2	157.144	94.223	220.065
β_3	-3.669	-7.684	0.346
β_4	0.472	0.390	0.554
β_5	-0.006	-0.009	-0.004

El modelo considera el 98.5% de la variabilidad que presentan los datos, pero β_3 , no supera la prueba de parámetros, ya que su intervalo de confianza contiene al cero.

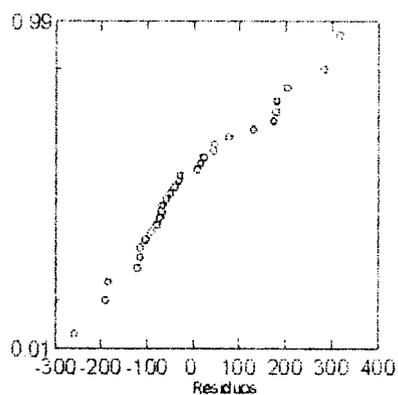
La dispersión de los residuos contra los valores estimados revela una alta concentración de datos para los valores iniciales, pero a medida que aumenta x o y , los valores se encuentran en forma homogénea alrededor de la media, lo cual se puede ver en la figura 4.35. Los residuos se ajustan adecuadamente a la distribución normal, así lo podemos ver en la figura 4.36 que presenta la gráfica de probabilidad normal.

Figura 4.35

Gráfica de la dispersión de los residuos contra los valores estimados del modelo 6

**Figura 4.36**

Gráfica de probabilidad normal del modelo 6



A continuación se realizará la prueba de Kolmogorov-Smirnov, para probar si los residuos son normales con media cero y varianza σ^2 . Dicha varianza es igual a la media cuadrática del residuo.

H_0 : Los residuos son normales con media cero y desviación estándar 154.2

vs.

$H_1: \neg H_0$

Max Dif. = 0.140

Valor p = 0.597

Por lo tanto, hay evidencia estadística para no rechazar la hipótesis nula H_0 .

Se empleará el estadístico de Durbin-Watson para comprobar si los residuos son independientes.

$C_1: \rho = 0$

vs.

$C_2: \rho > 0$

Durbin- Watson = 1.988

El estadístico de Durbin-Watson es mayor que el límite de confianza superior (1.83) correspondiente a $n=30$ y $p-1=5$, donde n representa

el número de observaciones y $(p-1)$ es el número de variables independientes. Por lo tanto, hay evidencia estadística para afirmar que ρ es igual a cero, es decir que los residuos son independientes..

Modelo 7

El séptimo modelo de regresión es múltiple con seis regresores:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2), \text{ cov}(\epsilon_i, \epsilon_j) = 0 \text{ para } i \neq j$$

donde Y representa las exportaciones bananeras en miles de TM; X_1 , la producción bananera en miles de TM., X_2 , la participación de las exportaciones bananeras en el PIB, X_3 , la participación de las exportaciones bananeras en las exportaciones totales, X_4 , el índice de precios al consumidor, X_5 , la cotización del dólar, X_6 , la superficie cultivada, y ϵ_i es un término de error aleatorio.

Fuente	Suma de cuadrados	gl	Media cuadrática
Regresión	3.70667E+07	6	6177775.633
Residuo	468723.458	23	20379.281
Total	3.75354E+07	29	

Coefficiente de determinación R^2 = 0.988

Intervalo de confianza

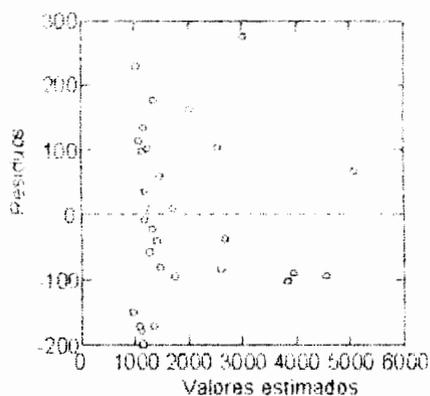
Parámetro	Estimador	Inferior 95%	Superior 95%
β_0	793.074	541.063	1045.085
β_1	0.239	0.131	0.348
β_2	210.599	134.087	287.111
β_3	-11.749	-22.617	-0.881
β_4	-4.330	-8.106	-0.554
β_5	0.470	0.394	0.546
β_6	-0.005	-0.008	-0.002

El modelo considera el 98.8% de la variabilidad que presentan los datos.

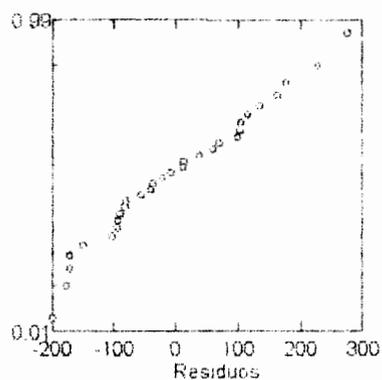
La dispersión de los residuos contra los valores estimados revela una alta concentración de datos para los valores iniciales, pero a medida que aumenta x o y , los datos se encuentran en forma homogénea alrededor de la media, lo cual se puede ver en la figura 4.37. Los residuos se ajustan adecuadamente a la distribución normal, así lo podemos ver en la figura 4.38 que presenta la gráfica de probabilidad normal.

Figura 4.37

Gráfica de la dispersión de los residuos contra los valores estimados del modelo 7

**Figura 4.38**

Gráfica de probabilidad normal del modelo 7



A continuación se realizará la prueba de Kolmogorov-Smirnov, para probar si los residuos son normales con media cero y varianza σ^2 . Dicha varianza es igual a la media cuadrática del residuo.

H_0 : Los residuos son normales con media cero y desviación estándar 142.76

vs.

$H_1: \neg H_0$

Max Dif. = 0.082

Valor p = 0.987

Por lo tanto, hay evidencia estadística para no rechazar la hipótesis nula H_0 .

Se empleará el estadístico de Durbin-Watson para comprobar si los residuos son independientes.

$C_1: \rho = 0$

vs.

$C_2: \rho > 0$

Durbin- Watson = 2.277

El estadístico de Durbin-Watson es mayor que el límite de confianza superior (1.92) correspondiente a $n=30$ y $p-1=6$, donde n representa

el número de observaciones y $(p-1)$ es el número de variables independientes. Por lo tanto, hay evidencia estadística para afirmar que ρ es igual a cero, es decir que los residuos son independientes..

Modelo 8

El octavo modelo de regresión es múltiple con nueve regresores:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \beta_4 X_{4i} + \beta_5 X_{5i} + \beta_6 X_{6i} + \beta_7 X_{7i} + \beta_8 X_{8i} + \beta_9 X_{9i} + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2), \text{cov}(\epsilon_i, \epsilon_j) = 0 \text{ para } i \neq j$$

donde Y representa las exportaciones bananeras en miles de TM, X_1 , las exportaciones bananeras en miles de dólares FOB, X_2 , la producción bananera en miles de TM., X_3 , la participación de las exportaciones bananeras en el PIB, X_4 , la participación de las exportaciones bananeras en las exportaciones totales, X_5 , el índice de precios al consumidor , X_6 , el índice de precios al consumidor (alimentos), X_7 , la cotización del dólar, X_8 , la superficie cultivada, X_9 , la población económicamente activa, y ϵ_i es un término de error aleatorio.

Fuente	Suma de cuadrados	gl	Medla cuadrática
Regresión	3.70679E+07	9	4118659.190
Residuo	467444.544	20	20379.281
Total	3.75354E+07	29	

Coefficiente de determinación R^2 = 0.988

Intervalo de confianza

Parámetro	Estimador	Inferior 95%	Superior 95%
β_0	937.136	-640.033	2514.305
β_1	-0.000	-0.001	0.001
β_2	0.252	0.070	0.434
β_3	214.168	91.073	337.264
β_4	-12.050	-24.999	0.898
β_5	-4.354	-10.030	1.321
β_6	-1.305	-18.309	15.699
β_7	0.532	-0.194	1.258
β_8	-0.006	-0.011	0.000
β_9	-0.000	-0.001	0.001

El modelo considera el 98.8% de la variabilidad que presentan los datos, pero solo β_2 y β_3 , superan la prueba de parámetros, ya que sus intervalos no contienen al cero.

La dispersión de los residuos contra los valores estimados revela una alta concentración de datos para los valores iniciales, pero a medida que aumenta x o y , los datos se encuentran en forma homogénea alrededor de la media, lo cual se puede ver en la figura 4.39. Los residuos se ajustan adecuadamente a la distribución normal, así lo podemos ver en la figura 4.40 que presenta la gráfica de probabilidad normal.

Figura 4.39

Gráfica de la dispersión de los residuos contra los valores estimados del modelo 8

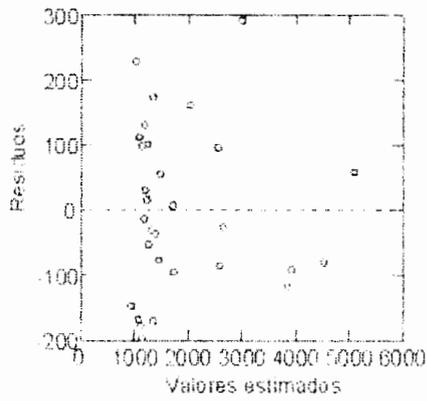
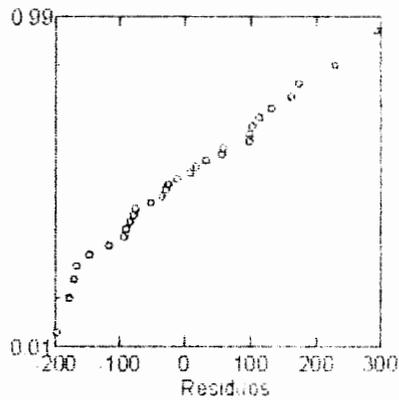


Figura 4.40

Gráfica de probabilidad normal del modelo 8



A continuación se realizará la prueba de Kolmogorov-Smirnov, para probar si los residuos son normales con media cero y varianza σ^2 . Dicha varianza es igual a la media cuadrática del residuo.

H_0 : Los residuos son normales con media cero y desviación estándar 152.88

vs.

$H_1: \neg H_0$

Max Dif. = 0.099

Valor p = 0.932

Por lo tanto, hay evidencia estadística para no rechazar la hipótesis nula H_0 .

Se empleará el estadístico de Durbin-Watson para comprobar si los residuos son independientes.

$C_1: \rho = 0$

vs.

$C_2: \rho > 0$

Durbin- Watson = 2.296

El estadístico de Durbin-Watson es mayor que el límite de confianza superior (2.19) correspondiente a $n=30$ y $p-1=9$, donde n representa el número de observaciones y $(p-1)$ es el número de variables

independientes. Por lo tanto, hay evidencia estadística para afirmar que ρ es igual a cero, es decir que los residuos son independientes..

En conclusión, se han obtenido seis modelos de regresión lineales: dos simples y cuatro múltiples, que cumplen las pruebas de parámetros y el supuesto de normalidad de los residuos. El modelo 7 tiene la mayor potencia de explicación, ya que considera el 98.8% de la variabilidad de los datos. En lo que respecta a la prueba de independencia de los residuos, mediante el estadístico de Durbin-Watson, todos los modelos cumplen con el supuesto de independencia de residuos, excepto el modelo 1, que explica las exportaciones bananeras en miles de TM mediante la variable correspondiente al índice de precios al consumidor (alimentos).

CONCLUSIONES

1. El análisis de componentes principales seleccionado para explicar la relación existente entre las diez variables de estudio, es el obtenido utilizando los datos estandarizados, ya que de esa forma sus diferentes escalas no pueden incidir en el porcentaje de varianza total explicada. Las componentes principales, denominadas *Exportaciones* y *Rendimiento económico de los cultivos bananeros*, son:

Exportaciones

$$Y_1 = 0.403*X_1 + 0.393*X_8 + 0.39*X_2 + 0.387*X_7 + 0.372*X_3 + 0.313*X_{10} + 0.296*X_4 + 0.139*X_9 + 0.147*X_5 + 0.129*X_6$$

Rendimiento económico de los cultivos bananeros

$$Y_2 = 0.057*X_1 + 0.096*X_8 + 0.14*X_2 + 0.083*X_7 - 0.167*X_3 + 0.374*X_{10} - 0.326*X_4 - 0.547*X_9 - 0.516*X_5 + 0.349*X_6$$

2. El análisis de factores obtenido utilizando los datos estandarizados determinó que la información de las diez variables de estudio propuestas, puede ser explicada mediante dos componentes principales. En la primera componente, las variables con mayores correlaciones son:

Exportaciones bananeras en miles de TM, cotización del dólar , exportaciones bananeras en miles de dólares FOB, índice de precios al consumidor del rubro alimentos, producción bananera en miles de TM, población económicamente activa, porcentaje de participación de las exportaciones bananeras en el PIB. Mientras que en la segunda componente, las variables con mayores correlaciones son: Superficie cultivada de banano en hectáreas, porcentaje de participación de las exportaciones bananeras en las exportaciones totales, índice de precios al consumidor. Las dos componentes explican el 86.1% de la varianza total.

3. El análisis de componentes principales obtenido utilizando los datos originales de las diez variables, con tan solo tres componentes explica casi el 100% de la varianza total de las diez variables de estudio, las cuales tienen distintas escalas. La primera componente explica el 95.5%. Además, determinó dos grupos de variables altamente relacionadas, el primer grupo está conformado por: Exportaciones bananeras en miles de TM, exportaciones bananeras en miles de dólares FOB, cotización del dólar e índice de precios al consumidor (alimentos), y el segundo grupo: Población económicamente activa, producción bananera, participación de las exportaciones bananeras en el PIB, participación de las exportaciones bananeras en las exportaciones totales y la superficie cultivada con banano. El índice de precios al consumidor, no se relaciona con las exportaciones bananeras, en ninguna de las dos componentes.

4. Aplicando el método de rotación de factores (Varimax) al análisis de factores obtenido utilizando los datos originales, se determinó que ocho componentes explican el 100% de la varianza total. Las tres primeras componentes explican el 79.9%. Y aplicando el método de rotación al análisis de factores obtenido utilizando los datos estandarizados, se determinó que dos componentes explican el 84.5% de la varianza total.

5. Se han obtenido seis modelos de regresión lineal, dos simples y cuatro múltiples, que cumplen las pruebas de parámetros y el supuesto de normalidad de los residuos. La prueba de independencia de los residuos, mediante el estadístico de Durbin-Watson, revela que todos los modelos cumplen con el supuesto de independencia de residuos, excepto el modelo 1, que explica las exportaciones bananeras en miles de TM mediante la variable correspondiente al índice de precios al consumidor (alimentos).

6. El séptimo modelo de regresión, que explica las exportaciones bananeras en miles de TM., mediante las variables: producción bananera en miles de TM., participación de las exportaciones bananeras en el PIB, participación de las exportaciones bananeras en las exportaciones totales, índice de precios al consumidor, cotización del dólar y superficie cultivada, tiene el mayor coeficiente de determinación, válido para una potencia de explicación del 98.8%.

7. El primer modelo de regresión, que explica las exportaciones bananeras en miles de TM. mediante la variable: índice de precios al consumidor (alimentos), tiene una potencia de explicación del 92.5%.
8. El segundo modelo de regresión, que explica las exportaciones bananeras en miles de TM. mediante la variable: cotización del dólar, tiene una potencia de explicación del 94.2%.
9. El tercer modelo, que explica las exportaciones bananeras en miles de TM. mediante las variables: producción bananera en miles de TM y cotización del dólar, tiene una potencia de explicación del 96.2%.
10. El cuarto modelo, que explica las exportaciones bananeras en miles de TM. mediante las variables: índice de precios al consumidor (alimentos), índice de precios al consumidor y producción bananera en miles de TM., tiene una potencia de explicación del 95.7%.
11. El quinto modelo, que explica las exportaciones bananeras en miles de TM. mediante las variables: producción bananera en miles de TM., participación de las exportaciones bananeras en el PIB, cotización del dólar y superficie cultivada, tiene una potencia de explicación del 98.3%.
12. El sexto modelo, que explica las exportaciones bananeras en miles de TM. mediante las variables: producción bananera en miles de TM., participación

de las exportaciones bananeras en el PIB, índice de precios al consumidor, cotización del dólar y superficie cultivada, tiene una potencia de explicación del 98.5%, pero el coeficiente de regresión que corresponde a la variable índice de precios al consumidor, no supera la prueba de parámetros, ya que su intervalo de confianza contiene al cero.

13. El octavo modelo, que explica las exportaciones bananeras en miles de TM. mediante las variables: exportaciones bananeras en miles de dólares FOB, producción bananera en miles de TM., participación de las exportaciones bananeras en el PIB, participación de las exportaciones bananeras en las exportaciones totales, el índice de precios al consumidor, índice de precios al consumidor (alimentos), cotización del dólar, superficie cultivada y población económicamente activa, tiene una potencia de explicación del 98.8%, pero sólo los coeficientes de regresión que corresponden a las variables producción bananera y participación de las exportaciones bananeras en el PIB, superan la prueba de parámetros, ya que sus intervalos no contienen al cero.

14. La técnica de series temporales no dió buenos resultados, ya que no se obtuvieron modelos de predicción para las variables de estudio propuestas, a excepción del modelo estacionario, ARMA(1,0), para la participación de las exportaciones bananeras en el PIB, cuyas predicciones carecen de importancia, si no pueden ser relacionadas con las predicciones de las variables fundamentales, tales como, exportaciones y producción bananera.

15. Aproximadamente un 60% de las exportaciones bananeras anuales registradas entre los años 1969-1998, se encuentran entre 1'000.000 y 2'000.000 de TM. No más de un 5% de las exportaciones han sobrepasado la barrera de 4'000.000 de TM.
16. De los valores anuales de las exportaciones bananeras registrados entre los años 1969 y 1998, el 25 % se encuentran entre 800.15 (mínimo) y 1'200.000 toneladas métricas (Q_1). El segundo cuartil Q_2 (Mediana) es 1'355.150 toneladas métricas. El 50% de los valores registrados se encuentran entre 1'200.000 y 2'600.000 toneladas métricas (Q_3). El restante 25% se encuentran entre 2'600.000 y 4'456.300 toneladas métricas (máximo). Existe un valor aberrante, 5'154.600 toneladas métricas registrado en el año 1998.
17. Aproximadamente, un 75% de las exportaciones bananeras anuales, no han sobrepasado la barrera de los 500'000.000 dólares FOB. El 25% restante, han registrado valores entre 500'000.000 y 1.400'000.000 dólares FOB. Los patrones entre las figuras 3.3 y 3.6 no coinciden, debido a que el precio del banano presenta continuos incrementos y decrementos, y en ciertos casos se incrementa la exportación de banano en miles de TM. , pero la exportación de banano en miles de dólares FOB decrece y viceversa, o también puede establecerse una relación directa entre ellas, es decir si las exportaciones en miles de TM aumentan (disminuyen), las exportaciones en miles de dólares FOB también aumentan (disminuyen).

18. De los valores anuales de las exportaciones bananeras en miles de dólares FOB, registrados entre los años 1969 y 1998, el 25% se encuentran entre 68'175.000 (mínimo) y 125'000.000 dólares FOB (Q_1). El segundo cuartil Q_2 (Mediana) es 213'931.000 dólares FOB. El 50% de los valores registrados se encuentran entre 125'000.000 y 600'000.000 dólares FOB (Q_3). El restante 25% se encuentran entre 600'000.000 y 1.027'631.000 dólares FOB (máximo). Existe un valor aberrante, 1.326'892.000 dólares FOB, que fue el récord de ventas en el año 1997.

19. Aproximadamente, el 40% de la producción bananera anual, registrada en el período 1969-1998, se encuentra entre 2'000.000 y 3'000.000 TM. El 30% de los valores registrados está entre 3'000.000 y 4'000.000 TM. Tan solo un 4% sobrepasa la barrera de 6'000.000 TM.

20. De los valores anuales de producción bananera en miles de toneladas métricas, registrados entre los años 1969 y 1998, el 25% se encuentran entre 1'642.100 (mínimo) y 2'160.000 toneladas métricas (Q_1). El segundo cuartil Q_2 (Mediana) es 2'921.250 toneladas métricas. El 50% de los valores registrados se encuentran entre 2'160.000 y 3'880.000 toneladas métricas (Q_3). El restante 25% se encuentran entre 3'880.000 y 5'750.300 toneladas métricas (máximo). Existe un valor aberrante, 6'809.200 toneladas métricas, ya que en el año de 1995, se registró el mayor valor de productividad que fue de 37.05 TM./ hectárea.

21. Aproximadamente, un 25% de las exportaciones bananeras anuales han tenido una participación en el PIB comprendida entre el 1 y 2 %. Un 34% tuvieron una participación entre 2 y 4 %. Sólo un 10% de las exportaciones tuvieron una participación superior al 6%.

22. De los valores anuales de participación de las exportaciones bananeras en el Producto Interno Bruto (PIB), registrados entre los años 1969 y 1998, el 25% se encuentran entre 0.871 (mínimo) y 1.9% (Q_1). El segundo cuartil Q_2 (Mediana) es 3.338 %. El 50% de los valores registrados se encuentran entre 1.9 y 5% (Q_3). El restante 25% se encuentran entre 5 y 6.990% (máximo).

23. Aproximadamente un 30% de las exportaciones bananeras anuales, tuvieron una participación en las exportaciones totales de entre 5 y 10%. Un 50% tienen una participación de entre 10 y 20 %. Sólo un 15%, aproximadamente, tiene una participación que supera el 40% de las exportaciones totales.

24. De los valores anuales de participación de las exportaciones bananeras en las exportaciones totales, registrados entre los años 1969 y 1998, el 25 % se encuentran entre 5.159 (mínimo) y 9% (Q_1). El segundo cuartil Q_2 (Mediana) es 14.694%. El 50% de los valores registrados se encuentran entre 9 y 22% (Q_3). El restante 25% se encuentran entre 22 y 40.14% (máximo). Existen tres valores aberrantes, 43.82% (año 1970), 44.28%

(año 1971) y 44.88% (año 1969). Podemos notar que estos valores aberrantes se registraron antes de 1972, año del Boom petrolero, porque a partir de este año, las exportaciones petroleras desplazaron a las exportaciones bananeras y por ende, registraron mayores porcentajes de participación en las exportaciones totales.

25. Casi el 80% del índice de precios al consumidor (alimentos) anuales, registrados entre los años 1969-1998, no presentan incrementos mayores de 30%. Mientras que el 20% restante fluctúan entre 30 y 260%, aproximadamente.

26. De los valores anuales del índice de precios al consumidor del rubro alimentos, registrados entre los años 1969 y 1998, el 25% se encuentran entre 0.19 (mínimo) y 0.4 (Q_1). El segundo cuartil Q_2 (Mediana) es 2.05. El 50% de los valores registrados se encuentran entre 0.4 y 33 (Q_3). El restante 25% se encuentran entre 33 y 72.6 (máximo). Existen dos valores aberrantes, 90.61 (año 1994) y 109.67 (año 1995) y existen tres valores mucho mayores que los considerados como aberrantes, 135.7 (año 1996), 181.2 (año 1997) y 253 (año 1998). Se puede deducir de estos valores, que el precio de los alimentos y por ende su influencia en la inflación, ha presentado un lento crecimiento entre los años 1969 y 1988 (0.19 a 8.06), pero a partir del año de 1989 hasta 1998, ha registrado un rápido y elevado incremento de 15.16 a 253, es decir que ha crecido en 1669%. Siendo el alto rango de los datos la causa de que el diagrama de cajas correspondiente sea tan atípico.

27. Más de un 30% del índice de precios al consumidor (1969-1998), han registrado incrementos comprendidos entre 10 y 20%. Aproximadamente un 28% fluctuaron entre el 20 y 30% y menos del 3% registraron incrementos superiores al 80%.

28. De los valores anuales del índice de precios al consumidor (IPC), registrados entre los años 1969 y 1998, el 25% se encuentran entre 5.22 (mínimo) y 10.7% (Q_1). El segundo cuartil Q_2 (Mediana) es 24.4%. El 50% de los valores registrados se encuentran entre 10.7 y 33% (Q_3). El restante 25% se encuentran entre 33 y 60.2% (máximo). Existe un valor aberrante, 85.7% (año 1988).

29. Más de un 70% de las cotizaciones anuales del dólar (1969-1998), se encuentran entre 18 y 1000 sucres. Un 10% han fluctuado entre 2000 y 3000 sucres. Menos de un 5% han presentado una cotización del dólar mayor a 6000 sucres.

30. De los valores anuales de la cotización del dólar, registrados entre los años 1969 y 1998, el 25% se encuentran entre 18.18 (mínimo) y 26.7 sucres (Q_1). El segundo cuartil Q_2 (Mediana) es 103.35 sucres. El 50% de los valores registrados se encuentran entre 26.7 y 1200 sucres (Q_3). El restante 25% se encuentran entre 1200 y 2914.8 sucres (máximo). Existen dos valores aberrantes, 3596.2 sucres (año 1996) y 4393 sucres (año 1997), y un valor mucho mayor que los considerados

aberrantes, 6480 sucres (año 1998). Esto se debe al alza inesperada de la cotización del dólar. Recordemos que este análisis no incluye observaciones de los años 1999 y 2000. En diciembre de 1999, el dólar alcanzó una cotización de 19260 sucres, es decir que entre 1998 y 1999, tuvo un incremento de aproximadamente 200%. En el 2000, se pone en práctica la dolarización en el Ecuador, por lo que la cotización del dólar se fija en 25000 sucres, es decir que entre 1999 y el 2000, hay un 30% de incremento, aproximadamente. Si estos cambios económicos hubieran sido considerados, el análisis correspondiente al período 1969-1998, registraría notables variaciones.

31. Más de un 30% de los registros anuales de superficie cultivada con banano (1969-1998), han fluctuado entre 50000 y 75000 hectáreas. Aproximadamente el 36% fluctuaron entre 100000 y 150000 hectáreas. Un 20% de las observaciones, registran superficies cultivadas de entre 150000 y 200000 hectáreas.

32. De los valores anuales de superficie cultivada con banano, registrados entre los años 1969 y 1998, el 25% se encuentran entre 49562 (mínimo) y 60568 hectáreas (Q_1). El segundo cuartil Q_2 (Mediana) es 109982.5 hectáreas. El 50% de los valores registrados se encuentran entre 60568 y 134000 hectáreas (Q_3). El restante 25% se encuentran entre 134000 y 193560 (máximo).

33. Aproximadamente, un 70% de los registros anuales de la población económicamente activa del Ecuador (1969-1998), fluctúan entre 2 y 3 millones de personas.

34. De los valores anuales de la población económicamente activa del Ecuador (PEA), registrados entre los años 1969 y 1998, el 25% se encuentran entre 1'879.000 (mínimo) y 2'400.000 personas (Q_1). El segundo cuartil Q_2 (Mediana) es 2'950.824 personas. El 50% de los valores registrados se encuentran entre 2'400.000 y 3'250.000 habitantes (Q_3). El restante 25% se encuentran entre 3'250.000 y 4'073.581 (máximo).

35. Los datos empleados para la realización del análisis de la producción bananera del Ecuador, no incluyen los registrados en el año 1999, debido a su falta de disponibilidad, de allí que no se realice mención alguna sobre el repunte de la cotización del dólar y su influencia en las exportaciones bananeras, como tampoco se hace referencia a la dolarización y consecuente estabilización del valor del dólar frente al sucre (25000 sucres por dólar).

36. Aunque entre 1997 y 1998, las exportaciones bananeras aumentaron un 16% aproximadamente, el ingreso que generaron en dólares FOB disminuyó un 22.5% aproximadamente, debido a que el precio referencial depende del mercado, así en la Unión Europea el precio es más elevado

que en Estados Unidos, y por lo tanto, los ingresos dependerán de las proporciones de ventas realizadas en cada país comprador.

37. A partir de 1996, la producción bananera ha registrado continuos descensos. La principal causa de este hecho, ha sido el continuo desacuerdo entre productores y exportadores, acerca del costo de la caja de banano, ya que los productores afirman que los exportadores desean adquirir el banano a precios inferiores al costo, provocándoles importantes pérdidas. Esto ocasiona que los productores disminuyan la producción o en casos más extremos, dejen de producir por completo.

38. El aumento que viene registrando la producción bananera es mayor que el ritmo de crecimiento que registran las exportaciones. Este comportamiento ha dado como resultado que una buena parte de la producción bananera, aproximadamente entre el 15 y 40% , se destine para el consumo interno y para la fabricación de productos industrializados (elaborados de banano) que son exportados a precios superiores, debido a su valor agregado.

RECOMENDACIONES

1. Todos los modelos, ya sean series temporales o regresiones, planteados en base a registros de años anteriores, presentarán predicciones que seguirán la tendencia otorgada por dichos datos, sin considerar la situación actual del Ecuador y las consecuentes fluctuaciones inesperadas, que definitivamente anula cualquier modelo propuesto. Por ello, es recomendable esperar que el país adopte un plan económico, que establezca su situación, y a partir de la nueva información que se registre, elaborar modelos aplicando cualquiera de las dos técnicas: series temporales o modelos de regresión.
2. El Gobierno debe establecer conversaciones con los productores y exportadores, para plantear soluciones definitivas al desacuerdo que existe sobre el precio de la caja de banano, ya que de esta forma se evitará que gran cantidad de productores abandonen sus cultivos o prefieran perder la fruta, lo cual se refleja en que el 30% de la producción se comercializa y el 70% restante se pierde en las plantaciones.

3. Las autoridades deben ser más enérgicas y poner un alto a la siembra indiscriminada de miles de hectáreas ilegales (hectáreas no consideradas en las superficies de cultivo de banano, establecidas oficialmente), ya que ha dado origen a la caída de los precios de la fruta y a la reducción de la calidad de las siembras.
4. Se debe realizar un estudio de la capacidad de adquisición de los países compradores, actuales y potenciales, para poder proyectar, adecuadamente, la cantidad a producir, evitando la sobreproducción y por ende la pérdida del producto.
5. Este estudio se ha enfocado en el banano como fruto, pero sería interesante, realizar un análisis similar de la información referente a los elaborados del banano y su exportación, ya que es una excelente alternativa de generación de ingresos para el país y una posible solución al problema de las continuas pérdidas de los productores, ya que toda la fruta no seleccionada para la exportación, se la emplearía en la producción de elaborados.

APENDICES

APENDICE A
PARTICIPACION DE LAS PRINCIPALES EMPRESAS EN LAS
EXPORTACIONES DE BANANO DEL ECUADOR
1959-1964

Año	Noboa TM	%	Standard F. TM	%	United F. TM	%	UBESA TM	%	TOTAL TM
1959	80489	9.09	106365	12.01	188013	21.2	179793	20.3	885600
1960	86770	9.69	94780	10.59	220664	24.7	146499	16.37	895100
1961	90550	10.75	101107	12.07	181613	21.6	156769	18.61	842300
1962	118236	13.17	127286	14.18	186263	20.8	160360	17.86	897800
1963	170841	16.84	136140	13.42	149979	14.8	254243	25.07	1014300
1964	172068	15.83	180462	16.6	151927	14	248242	22.84	1086800

Fuente: El Banano en el Ecuador, Larrea, M. Carlos. Corporación Editora Nacional, 1987

APENDICE B
PARTICIPACION DE LAS PRINCIPALES EMPRESAS EN LAS
EXPORTACIONES DE BANANO DEL ECUADOR
1965-1976

Año	Noboa		Standard F.		United F.		UBESA		TOTAL TM
	TM	%	TM	%	TM	%	TM	%	
1965	175926	20.12	230903	26.4	—	—	232009	26.53	874600
1966	210117	19.62	294368	27.5	—	—	320722	29.95	1070700
1967	234731	20.74	310955	27.5	—	—	311151	27.49	1131800
1968	346218	21.67	127286	25.5	—	—	407041	25.48	1597700
1969	255696	21.33	136140	23	—	—	287253	23.97	1198600
1970	286462	22.98	228087	18.3	—	—	236040	18.94	1246300
1971	295362	25.04	202708	17.2	—	—	179493	15.22	1179700
1972	606231	35.12	265244	15.4	—	—	258471	14.97	1726100
1973	394392	41.58	173443	18.3	—	—	141223	14.89	948500
1974	593952	38.95	411808	27	107208	7.03	181996	11.93	1525100
1975	569564	41.14	350278	25.3	111210	8.03	169282	12.23	1384500
1976	417870	44.58	192160	20.5	93804	10.01	124678	13.3	937300

Fuente: El Banano en el Ecuador, Larrea M. Carlos, Corporación Editora Nacional, 1987

APENDICE C

MATRIZ DE CORRELACION

	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	X ₉	X ₁₀
X ₁	1.000									
X ₂	0.955	1.000								
X ₃	0.870	0.798	1.000							
X ₄	0.629	0.585	0.722	1.000						
X ₅	0.239	0.134	0.482	0.788	1.000					
X ₆	0.276	0.382	0.045	0.131	-0.195	1.000				
X ₇	0.962	0.906	0.811	0.501	0.181	0.239	1.000			
X ₈	0.971	0.923	0.813	0.525	0.183	0.306	0.995	1.000		
X ₉	0.240	0.094	0.583	0.681	0.817	-0.353	0.208	0.196	1.000	
X ₁₀	0.796	0.877	0.529	0.256	-0.217	0.598	0.760	0.792	-0.315	1.000

Número de observaciones: 30

APENDICE D**MATRIZ DE COVARIANZAS DE LAS COMPONENTES Y LAS VARIABLES DE ESTUDIO**

	1	2	3	4	5	6	7	8	9
X ₁₀	633329.04	-64103.14	2982.57	0.697	0.345	0.016	-0.001	0	0
X ₂	311593.4	131571.34	-5250.47	2.14	0.41	-0.15	0	0	0
X ₉	-10737.03	37295.59	23631.87	8.18	7.85	0.217	-0.007	-0.001	0
X ₈	1323.37	631.23	202.15	-551.9	78.68	-55.77	0.103	0.032	-0.15
X ₁	962.16	521.39	139.4	-233.44	-50.154	142.36	0.096	-0.043	0.012
X ₃	811.88	888.23	493.4	-80.83	-391.76	-29.44	0.207	0.016	-0.008
X ₇	50.19	26.4	6.04	-24.36	1.43	-2.92	-2.21	-0.23	3.39
X ₆	10.84	-6.72	2.5	3.45	6.71	-0.83	11.85	-0.71	0.61
X ₅	-1.73	8.29	4.48	0.91	1.04	1.1	1.24	6.31	0.196
X ₄	0.569	1.27	0.517	0.4	0.26	0.42	0.25	0.42	-0.03

BIBLIOGRAFIA

BCE, (1997) Setenta Años de Información Estadística 1927-1997, Dirección Gral. de Estudios, Guayaquil, Ecuador

BCE, (1998) Informe Estadístico Anual, Dirección Gral. de Estudios, Guayaquil, Ecuador

BCE, (1999) Informe Estadístico Mensual, Vol. 11, Dirección Gral. de Estudios, Guayaquil, Ecuador

JOHNSON R., WICHERN, D. (1998), Applied Multivariate Statistical Analysis, cuarta edición, Prentice Hall, Upper Saddle River, New Jersey.

LARREA, C. (1987) El Banano en el Ecuador, Corporación Editora Nacional, Quito, Ecuador

MENDNHALL W., WACKERLY D., SCHEAFFER, R. (1994) Estadística Matemática con Aplicaciones, Editorial Iberoamérica, México D.F., México

MONTGOMERY C., RUNGER G. (1996) Probabilidad y estadística aplicadas a la Ingeniería, Mc Graw Hill editores, México D.F., México

NELSON, R. (1973) Applied Time Series Analysis For Managerial Forecasting, Editora Holden-Day, San Francisco, California, USA

NETER J., WASSERMAN W. (1974) Applied Linear Statistical Models, Richard D. Irwin, Inc. Homewood, Illinois, USA

RIOFRIO, J. (1997) Banano Ecuatoriano, Perspectivas, Producciones Agropecuarias, Guayaquil, Ecuador

SOTO, M. (1987) Bananos, Cultivo y Comercialización , Editora Iberoamérica, San José, Costa Rica

INTERNET

www.statsoft.com

www.conabanecuador.com

www.yahoo.com