

519.88
-7230
A678



ESPECIALIDAD EN MÉTODOS ESTADÍSTICOS
FACULTAD DE ESTADÍSTICA E INFORMÁTICA
UNIVERSIDAD VERACRUZANA

ESTUDIO SOBRE EL EFECTO DE LAS
CONDICIONES SOCIOECONÓMICAS EN LA
DESERCIÓN, REPROBACIÓN Y EFICIENCIA
TERMINAL DEL NIVEL MEDIO SUPERIOR, EN
LOS MUNICIPIOS DEL ESTADO DE VERACRUZ

BIBLIOTECA
UNIDAD DOCENTE INTERDISCIPLINARIA
DE ECONOMÍA Y ESTADÍSTICA

Trabajo recepcional que como requisito
parcial para obtener el diploma de esta
Especialidad presenta:

Josefina Arellano Reducindo

9 ENE. 1996

Tutor Académico:
Dr. Manuel Martínez Morales

Xalapa, Ver. Diciembre de 1995

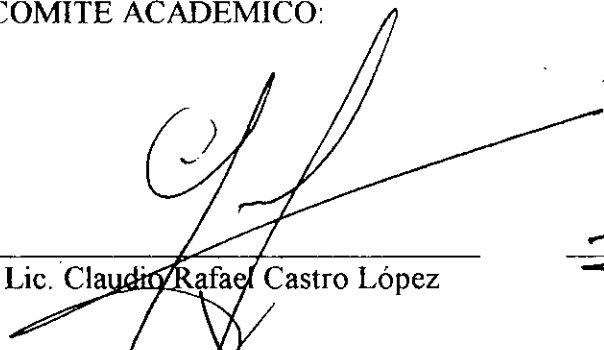
T-334

DATOS DEL AUTOR: Josefina Arellano Reducindo, originaria de la Ciudad de Xalapa, Ver., efectuó estudios básicos de primaria, secundaria y bachillerato en esta Ciudad Capital, realizó estudios profesionales en la Universidad Veracruzana de 1987-1991, en la carrera de Contador Público y Auditor de la cual es titulada, continuó estudios en el Instituto de Contaduría de la U.V. dentro de la Especialidad en Administración Fiscal. Desde 1982 se desempeña laborando en esta Máxima Casa de Estudios, actualmente está adscrita en el Depto. de Supervisión y Desarrollo Escolar de la U.V., en esta ocasión se obtendrá la Especialidad en Métodos Estadísticos.


DEDICATORIA: A mi pareja, porque sus ejemplos y consejos motivan mis deseos de superación y con su paciencia permite terminar una meta más para nosotros.

El Comité Académico de la Especialidad en Métodos Estadísticos y el respectivo Tutor Académico del trabajo recepcional "Estudio sobre el efecto de las condiciones socioeconómicas en la deserción, reprobación y eficiencia terminal del nivel medio superior, en el Estado de Veracruz", una vez cubiertos todos los requisitos académicos y administrativos establecidos, autorizan la impresión y la constitución del jurado para la defensa del mismo.

POR EL
COMITÉ ACADÉMICO:



Lic. Claudio Rafael Castro López

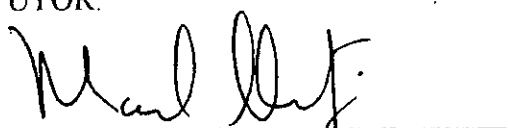


Dr. Mario Miguel Ojeda Ramirez



Lic. Sergio Hernández González

TUTOR:



Dr. Manuel Martínez Morales

INDICE

	PAGS.
INTRODUCCION	
CAPITULO I FUNDAMENTO.....	1
1.1 Antecedentes.....	1
1.2 Justificación.....	2
1.3 Objetivo.....	2
CAPITULO II Metodología.....	3
2.1 Obtención de los datos.....	3
2.2. Estructura	4
2.3 Transformación de datos.....	5
2.4 Descripción de las varaibales.....	6
CAPITULO III Análisis Descriptivo.....	7
3.1 Estadísticas descriptivas.....	7
3.2 Gráficas comparativas.....	8
CAPITULO IV Técnicas Estadísticas.....	11
4.1 Definición del Análisis Canónico.....	11
4.2 Planteamiento de Hipótesis.....	12
4.3 Programa y salida de SAS.....	14
4.4 Comentarios de Inferencia.....	21
CAPITULO V Técnicas de Agrupación en Cúmulos.....	23
5.1 Definición de Cúmulo.....	23
5.2 Programa y salida de Systat.....	23
5.3 Análisis de Varianza Múltiple.....	29
5.4 Planteamiento de Hipótesis.....	29
5.5 Comentarios de Inferencia.....	34
5.6 Análisis de gráfico de cúmulos.....	35
CAPITULO VI Conclusiones.....	39
Bibliografía.....	42

9 ENE. 1996

BIBLIOTECA
UNIDAD DOCENTE INTERDISCIPLINAR
DE ECONOMIA Y ESTADISTICA

INTRODUCCION

Actualmente las sociedades humanas se organizan en razón de los retos para su supervivencia. En los últimos años el país ha venido sufriendo una crisis económica que se expresa entre otras manifestaciones en: disminución del poder adquisitivo, desempleo abierto y descenso de los niveles de vida en la mayoría de la población. Ante esta crisis los sectores productivos, económicos y educativos, también se han visto afectados. Por ello pretendemos estudiar dos aspectos :

- a) En lo económico las condiciones socioeconómicas que prevalecen en la sociedad y
- b) En lo educativo, los indicadores como deserción, reprobación y eficiencia terminal desarrollados durante el ciclo escolar del nivel medio superior en la generación 1990-1993, en el Estado de Veracruz.

Teniendo como objetivo conocer la posible dependencia entre éstos dos aspectos, es decir, encontrar la correlación que permita explicar el comportamiento que ambos guardan, considerando que el progreso educativo y cultural son decisivos para avanzar hacia el bienestar económico y social en toda sociedad. El trabajo se encuentra estructurado de la siguiente manera:

- Capítulo I, descripción breve de los antecedentes que para efectos del estudio fue necesario incluir como la trayectoria que ha tenido la educación hasta llegar a la consolidación educativa en el estado, y retomar así el nivel medio superior; en la justificación se detallan los motivos que originaron el estudio; en el objetivo se describe lo que pretendemos alcanzar como es, constatar la dependencia o independencia de las variables arriba descritas.
- Capítulo II, incluye la metodología en cuanto a la obtención de los datos, su estructura la transformación matemática para comparar los datos y la descripción de las variables en cuanto al significado de cada una de ellas.
- Capítulo III, análisis descriptivo de estadísticas de tendencia central.

|- Capítulo IV, Técnica de Análisis de Correlación Canónica, en la cual se pretende relacionar las variables socioeconómicas y educativas a fin de reducir la dimensión de éstas para obtener las correlaciones más significativas entre ellas e interpretar la contribución relativa a las funciones canónicas.

- Capítulo V, Técnica de Agrupación en Cúmulos, la cual nos proporcionó tres grandes grupos de municipios con características similares entre sí, respecto a los indicadores de educación.

- Capítulo VI, Conclusiones, se describen las variables socioeconómicas que resultaron más significativas, respecto a la relación que guardan con el indicador educativo reprobación, el cual está correlacionado con las condiciones de vida actual; además se conocen los niveles de clase en que se agrupan los municipios del Estado de Veracruz; reconociendo los cúmulos de municipios en los cuales se desarrollan la deserción, reprobación y eficiencia terminal.

Esperamos que el estudio logre adentrarlos al interés por conocer más acerca del comportamiento y la relación que guardan estos dos grandes sectores en la comunidad Veracruzana.

CAPITULO I

FUNDAMENTO

1.1 ANTECEDENTES

De acuerdo a la trayectoria que ha tenido la educación en México a través de las diversas políticas que han gobernado al país, podemos señalar como relevantes la organización del sistema educativo nacional en la implementación del nivel primaria, la enseñanza rural e indígena así como la enseñanza técnica, la cual abrió las modernas especialidades que exigía el desarrollo de un país hacia la industrialización. Sin embargo a lo largo de este recorrido surge la figura de la educación pública, como un reto más para la enseñanza libre, al crearse instituciones dedicadas al servicio de la educación. Por la gran historia que trae consigo la educación, incluyendo la creación de la Secretaría de Educación Pública, y la consolidación educativa del Estado establecida en el artículo 3o. Constitucional que dice “ La educación que imparta el Estado será democrática, considerada a la democracia, no solamente como una estructura jurídica y un régimen político, sino como un sistema de vida fundado en el constante mejoramiento económico, social y cultural del pueblo”, por todo ello es que llama nuestra atención saber el comportamiento que tienen los niveles de educación hoy en día, en relación con nuestro desarrollo económico. Cabe señalar que sólo nos abocaremos a la trayectoria que presentan los jóvenes del nivel medio superior (bachillerato) para un período determinado.

1.2 JUSTIFICACION

Debemos señalar que los motivos para estudiar sólo el nivel medio superior, estriba en el interés de reconocer al bachillerato como una posición de educación importante por las atribuciones que tiene el aspirante en este nivel, como la edad, los hábitos de estudio, la forma en que se organiza, su posición económica, los deseos de superación, y hasta la posibilidad de adquirir con este nivel un trabajo mejor remunerado, lo que significa que al llegar a este nivel el alumno prácticamente estaría al inicio de una carrera profesional. De manera que nuestro propósito al igual que el compartido por la sociedad y algunos sectores que se preocupan por la cuestión escolar, es conocer los avances de forma cualitativa y cuantitativa, en términos de que la educación y el medio que lo rodean cumplan la finalidad de prepararlos para resolver los problemas de siempre y además de los nuevos problemas que trae aparejados la civilización científico-tecnológica en la que nos desarrollamos, y de la cual no podemos escapar, de manera que se hace indispensable contar con una preparación básica. Considero además que las condiciones socioeconómicas tienen contribuciones significativas en el desenvolvimiento del educando durante el ciclo del bachillerato, es por ello que efectuaremos el análisis correspondiente.

1.3 OBJETIVO

Constatar la dependencia o independencia de los indicadores de educación como deserción, reprobación y eficiencia terminal en el nivel medio superior con condiciones socioeconómicas, para los municipios del Estado de Veracruz.

FINACIADA DE REP.

CAPITULO II

METODOLOGIA

2.1.OBTENCION DE DATOS

Para facilitar el manejo de la información en el aspecto educativo se dispuso de datos estadísticos Oficiales de la S.E.C. (Secretaría de Educación y Cultura), correspondientes al inicio de cursos 1994-1995, clasificados por nivel, modalidad y sostenimiento; sin embargo adicional a esta información para el Estado de Veracruz, se tenía disponible un recuadro de los indicadores por nivel educativo del fin de cursos 1993-1994 respecto a la deserción, reprobación y eficiencia terminal por municipio el cual retomamos.

Para llevar a cabo nuestro estudio , se validaron los 207 municipios del Estado de Veracruz, a fin de considerar aquellos municipios en los cuales se ofrece educación a nivel medio superior (bachillerato) y conocer los indicadores antes descritos para conforman con ellos 3 variables dependientes. Al efectuar la validación se detectaron 7 localidades, las cuales excenden del 100% calculable en la eficiencia terminal, 1 descrita como no calculable y 9 más descritas como no egresión; por lo que finalmente se conformó una base con 126 municipios del Estado de Veracruz, dado que se excluyeron los 17 casos antes descritos, por considerar que es una información errónea y que podría llevarnos a una impresión descriptiva del análisis; cabe señalar que existen 64 municipios que no cuentan con el nivel educativo en estudio y por consiguiente también fueron excluidos. Para cumplir nuestro objetivo requerimos de variables de condición social, las cuales fueron obtenidas del Anuario Estadístico del Estado de Veracruz edición 1994 (INEGI), éstas condiciones nos van a permiten saber la proporción de disponibilidad en los servicios, adquisición de la vivienda, alfabetismo, ocupación e ingreso promedio por municipio, de esta manera se conformaron 6 variables independientes.



La importancia para elegir el servicios como energía eléctrica, se debe en gran parte, a la posibilidad de medir el grado de infraestructura adquirido por municipio, así como el nivel sociodemográfico encuancto a las proporciones de vivienda propia, de personas alfabetas, ingreso promedio por salario mínimo y por último, en lo económico conocer la proporción de ocupación por sector de actividad primario y secundario, ya que estos nos permiten agrupar entre otros, un indicador de bienestar y desarrollo económico, en las cuales estarían creadas condiciones propicias para la convivencia humana.

Posiblemente para otros estudios hubieran sido otras las condiciones de vida actual; sin embargo creemos que los factores que elegimos nos van a permitir conocer cuales son las condiciones de vida que están más correlacionadas con los indicadores de educación.

2.2 ESTRUCTURA

Podemos decir que los datos se agrupan en dos tipos de variables, es decir:

Dependiente. las respuestas que tienen los municipios en los indicadores de educación como la deserción, reprobación y eficiencia terminal.

Independiente. las condiciones socioeconómicas que permiten agrupar un indicador de bienestar social por municipio.

La característica de los datos es de tipo cuantitativo, dado que me permite cuantificar a una escala continua las obtenidas en cada variable de estudio, de tal manera que la estructura de éste análisis se medirá en la misma escala. Debemos recordar que la información fue obtenida directamente, dado que se consideraron municipios con bachillerato, y estos vendrían siendo nuestra población universo:

$$U = \{ m_1, m_2, m_3, \dots, m_n \}$$

Es una población finita, homogénea de acuerdo a ciertas características, digamos nivel medio superior, pero heterogénea sobre otras características, digamos que los bachilleres pertenecen a modalidades distintas como Estatal, Federal y Particular, las cuales no se consideraron para el estudio.

Existen otras características que pueden concurrir en los elementos de U, digamos W_1, W_2, \dots, W_i , las cuales se conocen por covariables, tampoco consideradas, pero de haberlas incluido podrían o no estar influyendo a las respuestas conocidas como Y_1, Y_2, Y_3 , que son las variables de interés central. Y finalmente hay una serie de variables X_1, X_2, \dots, X_p , que son las variables causa o explicatorias de la respuesta.

2.3 TRANSFORMACION DE LOS DATOS

Debido a la forma con que se encontraron los datos, se realizaron algunos procesos matemáticos sencillos, a fin de obtener una proporción porcentual en: servicio, ocupación por actividad, disposición de energía eléctrica, viviendas propias, y personas alfabetizadas; por ejemplo:

$$(\text{dato de interés} / \text{total de personas según el dato de interés}) * 100$$

sin embargo en la obtención del ingreso en función del salario mínimo, fue necesario aplicar un valor promedio por habitante en cada municipio, por ejemplo:

Municipio	Ingreso Salario Mínimo	Punto medio de la clase	Frecuencia	Frecuencia por punto medio
Acajete	menos de 1	.5	632	311.5
	de 1 a -2	1.5	479	718.5
	2 a -3	2.5	55	137.5
	3 a -5	4.0	29	116
	más de 5	7.5	24	180

$$N=1210 \quad T=1463.5$$

$$\text{Ingreso promedio es} = \frac{1463.5}{1210} = 1.21$$

2.4 DESCRIPCION DE LAS VARIABLES

VAR	NOMBRE	DESCRIBE
X ₁	ENERGIA ELECTRICA	Porcentaje de viviendas que disponen del servicio de energía eléctrica en zonas urbanas y rurales.
X ₂	VIVIENDA PROPIA	Porcentaje de viviendas propias en zona urbana o rural
X ₃	PERSONAS ALFABETIZADAS	Porcentaje de Población en condiciones de leer y escribir
X ₄	SECTOR PRIMARIO	Porcentaje de personas ocupadas con actividad en agricultura ganadería, caza y pesca en área rural.
X ₅	SECTOR SECUNDARIO	Porcentaje de personas ocupadas con actividad en minería, extracc.de petróleo, industria manufacturera, sector electricidad, agua y construcción.
X ₆	INGRESO PROMEDIO	Proporción del ingreso en función del salario mínimo según los rangos de percepción.

VAR	NOMBRE	DESCRIBE
Y ₁	DESERCION	Porcentaje de alumnos que abandona sus estudios de manera temporal o definitiva en cualquier periodo del ciclo que conforma el bachillerato.
Y ₂	REPROBACION	Porcentaje de alumnos que no aprueba el grado cursado en algún periodo escolar del ciclo que conforma el bachillerato.
Y ₃	EFICIENCIA TERMINAL	Porcentaje de alumnos que promueve regularmente en el bachillerato y que logran egresar de éste en el ciclo escolar establecido.

CAPITULO III

ANALISIS DESCRIPTIVO

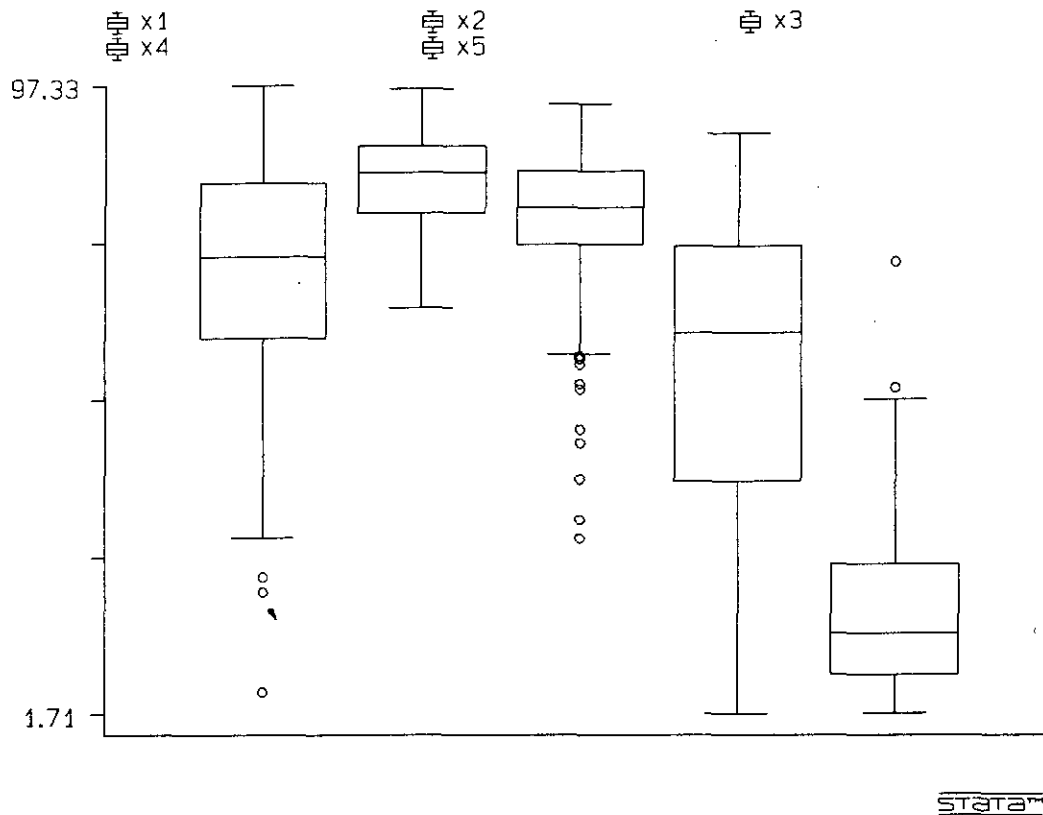
3.1 Estadísticas descriptivas

Las estadísticas básicas de las variables consideradas se muestran a continuación:

Medida	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	Y ₁	Y ₂	Y ₃
MIN	4.960	63.500	28.420	1.710	1.840	0.790	0.000	0.000	7.780
MAX	97.330	96.840	94.450	89.910	70.440	2.910	27.530	73.330	96.670
RANGO	92.370	33.340	66.030	88.200	68.600	2.120	27.530	73.330	88.890
MEDIA	69.297	83.128	76.538	53.863	17.334	1.643	5.492	32.684	62.782
S.D.	18.253	7.238	12.475	25.124	12.295	0.436	4.435	17.080	14.712
MEDIANA	71.345	84.110	78.660	59.630	13.950	1.685	4.560	33.280	64.090

Con apoyo del Gráfico No.1 podremos visualizar mejor el comportamiento que guardan las medianas en las variables de condición socioeconómica, Debemos aclarar que debido a la escala que guarda el ingreso promedio se presenta por separado en el Gráfico No. 2, a fin de conocer mejor la mediana para esta variable, también presentamos las medianas para los indicadores de educación en el Gráfico No. 3.

GRAFICO No. 1
COMPARACION DE VARIABLES SOCIOECONOMICAS

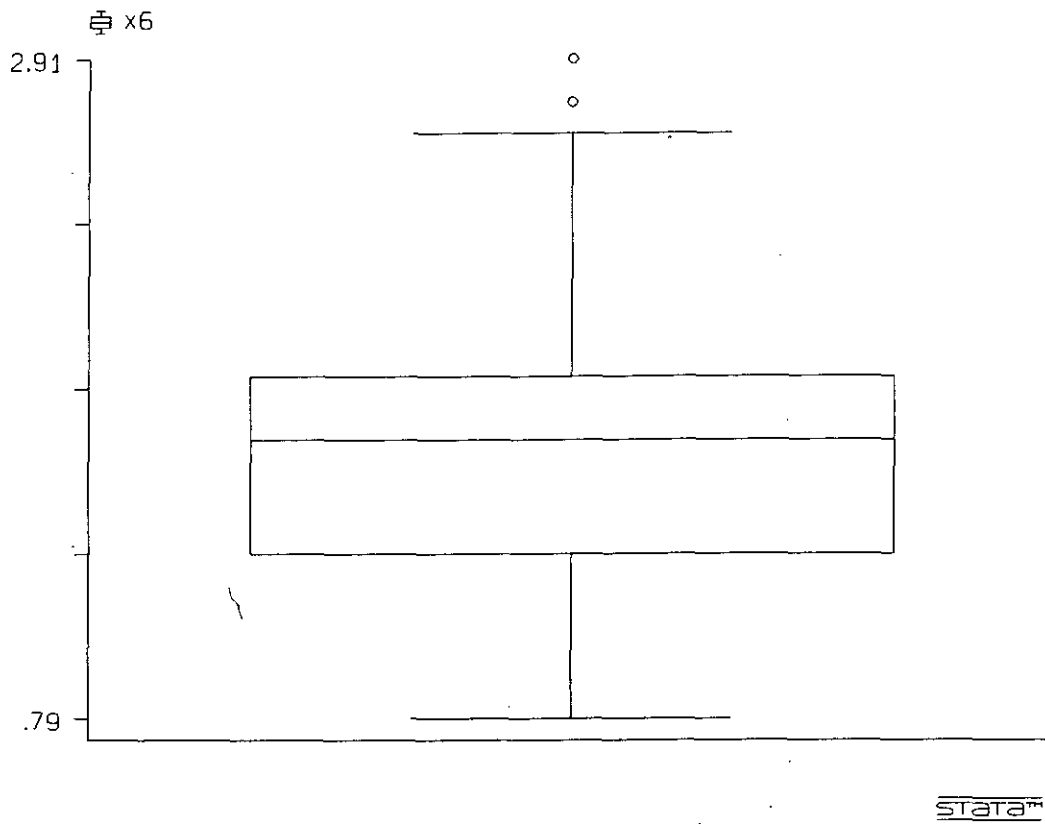


%Disp.Energía %Viv.Propia %Pnas.Alfab. %Primario %Secundario

Podemos observar que las variables con mayor mediana guardan el siguiente orden:

- 1° La proporción de viviendas propias.
- 2° La proporción de personas alfabetizadas.
- 3° La proporción de viviendas que disponen de energía eléctrica.
- 4° La proporción de actividades en ocupación sector primario y
- 5° La proporción de actividades en ocupación sector secundario.

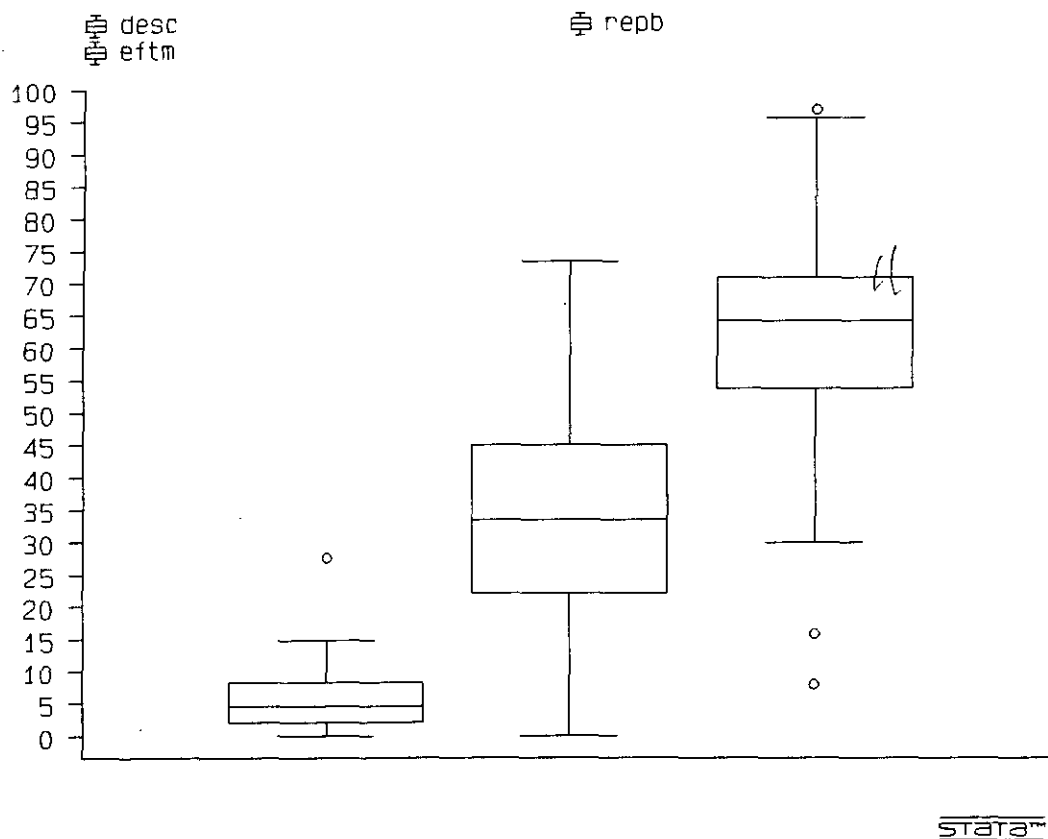
GRAFICO No. 2
VARIABLE DE CONDICION SOCIOECONOMICA



INGRESO PROMEDIO

La mediana que presenta esta variable es de 1.685, podemos ver que se aprecia una asimetría hacia la izquierda, podriamos decir que por debajo de la mediana se concentran el mayor porcentaje del ingreso promedio en razón a la mediana.

GRAFICO No. 3
COMPARACION DE VARIABLES DE EDUCACION



% DESERCION % REPROBACION % EFIC. TERMINAL

En relación a las medianas de los indicadores de educación, el orden que guardan es el siguiente:

- 1º Porcentaje de eficiencia terminal
- 2º Porcentaje de reprobación y
- 3º Porcentaje de deserción

CAPITULO IV

TECNICAS ESTADISTICAS APLICADAS

4.1 Definición de Análisis de Correlación Canónica

La técnica de Análisis de Correlación Canónica (ACC), es un modelo estadístico multivariado, es particularmente útil en situaciones donde pretendemos relacionar un conjunto de variables predictoras múltiples independientes y otras variables de criterio múltiple dependientes.

Objetivos

Podemos mencionar como objetivo principal, encontrar dos variables compuestas llamadas canónicas, que permitan reducir la dimensionalidad de las variables originales, y obtener las correlaciones más significativas entre ellas. A partir de esta información podemos señalar otros objetivos que pretende también esta técnica como:

- a) Obtener la relación entre variables
- b) Determinar si dos conjuntos de variables son independientes una de otra o determinar la magnitud de la asociación que puede existir entre los 2 conjuntos.
- c) Derivar un conjunto de pesos para cada conjunto variables de criterio o predictores, de manera que las combinaciones lineales estén correlacionadas al máximo.
- d) Explicar la naturaleza de cualquier conjunto de variables predictoras al medir la contribución relativa de cada variable a las funciones canónicas, entre otras.

4.2 PLANTEAMIENTO DE HIPOTESIS

Ho: Existe una correlación entre la deserción, reprobación y eficiencia terminal y los niveles socioeconómicos en los municipios del Estado de Veracruz.

Ha: No existe una correlación entre la deserción, reprobación y eficiencia terminal y los niveles socioeconómicos en los municipios del Estado de Veracruz.

Por el interés que mostramos para ejemplificar la técnica Canónica en un análisis de tipo Socio-Económico, asumimos el criterio de explicar en el estudio sólo aquel componente de variables que puedan ofrecer la máxima correlación entre ellas. Por lo que se decidió tomar sólo el primer componente de pesos canónicos estandarizados.

Identificaremos dos grupos de variables (X_1, \dots, X_p) y (Y_1, \dots, Y_q), para valores $p=6$ y $q=3$, aún cuando normalmente se aplica ($p < q$), podemos decir que no es impedimento alguno para realizar lo que nos interesa, por lo que la matriz quedaría representada así:

$$[nXp, nYq]$$

para obtener posteriormente la matriz simétrica de correlaciones asociadas :

$$R = \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} \text{ donde } R_{yx} = R'_{xy}$$

(p+q) (p+q)

${}_pR_{xxp}$ Tiene como entradas a las correlaciones entre las variables del primer grupo

${}_pR_{xyq}$ Tiene como entradas a las correlaciones entre las variables del primer y segundo

${}_pR_{yxq}$ Es la transpuesta de R_2 ($q \times p$)

${}_qR_{yyq}$ Tiene como entradas a las correlaciones entre las variables del segundo grupo

Como en otros procedimientos multivariados esta matriz muestra información sobre la asociación entre parejas, sin posibilidad de establecer redes de asociación, por considerarse compleja, por lo que tendremos que conocer las variables canónicas, debiendo transformar las variables originales para facilitar la interpretación de éstas, se llamarían: variables transformadas Z_1, Z_2, \dots, Z_p , en el caso de las variables X variables transformadas U_1, U_2, \dots, U_q , en el caso de las variables Y

Las construiremos según (Anderson Rolph E.)²

a) Combinación lineal entre X y Y

$$Z_i = a^t X = a_{i1} X_1 + a_{i2} X_2 + \dots + a_{ip} X_p$$

$$U_j = b^t Y = b_{j1} Y_1 + b_{j2} Y_2 + \dots + b_{jq} Y_q$$

donde a, a, a, \dots, a_p y b, b, b, \dots, b_q , se les llama pesos canónicos.

las Z_i son variables no correlacionadas entre sí cuando, $Cov(z_i, z_{i'}) = 0$ para $i \neq i'$

las U_j son variables no correlacionadas entre sí cuando, $Cov(u_i, u_{i'}) = 0$ para $i \neq i'$

las Z_i y U_j no son correlacionadas entre sí cuando $Cov(Z_i, U_j) = 0$ para $i \neq j$

sin embargo cuando Z_i y U_j tienen $i = j$ entonces si tienen las máximas correlaciones posible, como:

1a. Máx. correlación $r_{c1} = cov [z, u]$

2a. Máx. Correlación

3a. Máx. correlación etc: $r_{cs} = cov [z, u]$ la j-ésima máx. Correlación posible

es decir $r_{c1} \geq r_{c2} \geq r_{c3} \geq \dots \geq r_{cs}$ donde las r_k se les llama correlaciones canónicas, por

lo que la matriz de correlaciones canónicas quedaría así:

$$R = \begin{vmatrix} I & Z^t \\ Z^t & I \end{vmatrix} +$$

² Anderson Rolph E., et al. -Multivariate data Analysis. Mcmillan Publishing Company Inc. USA. 1987.

¿Cómo podríamos encontrar las a's y b's que producen lo que deseamos ?

- 1) Que Z_i y U_j produzcan como ya se dijo las máximas correlaciones
- 2) Que las Z's sean incorrelacionadas entre si al igual que las U's
- 3) Que solo existan las correlaciones (Z_i, U_j) y que las restantes sean nulas.

$$a) \text{Corr}(Z_i, U_j) = \begin{cases} 0 & \text{si } i \neq j \\ \rho_i & \text{si } i = j \end{cases}$$

$$b) \text{Corr}(Z_i, Z_j) = \text{Corr}(U_i, U_j) = 0 \text{ para todo } i \neq j$$

En relación con la forma más detallada respecto a las matrices, procedencia de valores característicos nulos y vectores característicos se sugiere revisar la bibliografía anexa en relación con ésta técnica.

4.3 Programa y Salida de SAS

A continuación presentamos las intrucciones del programa SAS.

```
Data a;  
input name $ x1....xp y1....yq;  
cards;  
Acayucan 75.43 78.60 15.7.....8.21 41.98 64.83  
Actopan 89.36 84.07 6.23....4.63 26.05 94.56  
.  
.; (cierre del listado)  
proc cancorr all out = salida  
vp = condic vn = 'Condición Social'  
wp = educac wn = 'Indicadores de Educación';  
var x1....xp;  
with y1 y2 y3 ; run;
```

Identificamos las variables de condición Social (X) y las de Educación (Y)

Recordando que las proporciones son

X₁ = Viviendas que disponen de energía eléctrica

X₂ = Viviendas propias

X₃ = Personas Alfabetizadas

X₄ = Ocupacion sector primario

X₅ = Ocupación sector secundario

X₆ = Ingreso promedio

Y₁ = Porcentaje de deserción

Y₂ = Porcentaje de reprobación

Y₃ = Porcentaje de eficiencia terminal

Presentamos las salidas considerando como primer caso a la correlación de X₁,...X₆ vs Y₁, Y₂, Y₃ y como segundo caso todas las X's vs únicamente Y₃, dentro de los cuales pretendemos conocer el grado de significancia y de correlación que presentan las variables.

SALIDA

(primer caso)

CANCORR PROCEDURE
Correlations Among the Original Variables
Correlations Among the Condición Socioeconómica

	X1	X2	X3
X1	1.0000	-0.5732	0.6110
X2	-0.5732	1.0000	-0.4498
X3	0.6110	-0.4498	1.0000
X4	-0.6218	0.6764	-0.5876
X5	0.5319	-0.5285	0.4422
X6	0.6595	-0.6445	0.6171

	X4	X5	X6
X1	-0.6218	0.5319	0.6595
X2	0.6764	-0.5285	-0.6445
X3	-0.5876	0.4422	0.6171
X4	1.0000	-0.8536	-0.8139
X5	-0.8536	1.0000	0.7176
X6	-0.8139	0.7176	1.0000

Correlations Among the Indicador Escolar

	Y1	Y2	Y3
Y1	1.0000	0.4698	-0.0480
Y2	0.4698	1.0000	-0.1404
Y3	-0.0480	-0.1404	1.0000

Correlations Between the Condición Socioeconómica and the Indicador Escolar

	Y1	Y2	Y3
X1	0.2172	0.4007	0.1065
X2	-0.2243	-0.4583	0.0795
X3	0.3354	0.4758	-0.0196
X4	-0.3458	-0.5513	-0.0153
X5	0.1841	0.4221	-0.0098
X6	0.2609	0.4206	-0.0402

CANCORR PROCEDURE
Canonical Correlation Analysis

	Adjusted Canonical Correlation	Approx Canonical Correlation	Squared Standard Error	Squared Canonical Correlation
--	--------------------------------------	------------------------------------	------------------------------	-------------------------------------

1	0.627288	0.600188	0.054248	0.393490
2	0.260955	0.149988	0.083352	0.068098
3	0.206103		0.085643	0.042478

Likelihood					
	Ratio	Approx F	Num DF	Den DF	Pr > F
1	0.54119887	4.4636	18	331.4113	0.0001
2	0.89231659	1.3835	10	236	0.1886
3	0.95752155	1.3198	4	119	0.2666

CANCORR PROCEDURE

Canonical Correlation Analysis

Raw Canonical Coefficients for the Condición Socioeconómica

CONDIC1

X1	0.00423761881
X2	-0.0207505253
X3	0.03276357176
X4	-0.0431856544
X5	-0.0227794428
X6	-0.935884789

Raw Canonical Coefficients for the Indicador Escolar

EDUCA1

Y1	.05520901514
Y2	.05092136498
Y3	.01170870696

Standardized Canonical Coefficients for the Condición Socioeconómica

CONDIC1

X1	0.0773
X2	-0.1502
X3	0.4087
X4	-1.0850
X5	-0.2801
X6	-0.4082

Standardized Canonical Coefficients for the Indicador Escolar

EDUCA1

Y1	0.2449
Y2	0.8697
Y3	0.1723

CANCORR PROCEDURE

Canonical Structure

Correlations Between the Condición Socioeconómica and Their Canonical Variables

CONDIC1

X1	0.6696
X2	-0.7011
X3	0.7853
X4	-0.9035
X5	0.6544
X6	0.6740

Correlations Between the Indicador Escolar and Their Canonical Variables

EDUCA1

Y1	0.6452
Y2	0.9606
Y3	0.0384

Correlations Between the Condición Socioeconómica and the Canonical Variables of the Indicador Escolar

EDUCA1

X1	0.4201
X2	-0.4398
X3	0.4926
X4	-0.5668
X5	0.4105
X6	0.4228

**Correlations Between the Indicador Escolar and the
Canonical Variables of the Condición Socioeconómica**

CONDIC1

Y1	0.4047
Y2	0.6025
Y3	0.0241□

SALIDA
(segundo caso)

Correlations Among the Original Variables

Correlations Among the Condición Social

	X1	X2	X3
X1	1.0000	-0.5732	0.6110
X2	-0.5732	1.0000	-0.4498
X3	0.6110	-0.4498	1.0000
X4	-0.6218	0.6764	-0.5876
X5	0.5319	-0.5285	0.4422
X6	0.6595	-0.6445	0.6171

	X4	X5	X6
X1	-0.6218	0.5319	0.6595
X2	0.6764	-0.5285	-0.6445
X3	-0.5876	0.4422	0.6171
X4	1.0000	-0.8536	-0.8139
X5	-0.8536	1.0000	0.7176
X6	-0.8139	0.7176	1.0000

**Correlations Among the Indicador Escolar
Y3**

Y3 1.0000

Correlations Between the Condición Social and the Indicador Escolar

Y3
X1 0.1065
X2 0.0795
X3 -0.0196
X4 -0.0153
X5 -0.0098
X6 -0.0402

Canonical Correlation Analysis

	Adjusted	Approx	Squared
Canonical	Canonical	Standard	Canonical
Correlation	Correlation	Error	Correlation

1	0.255409	0.183166	0.083608	0.065234
---	----------	----------	----------	----------

**Eigenvalues of INV(E)*H
= CanRsqr/(1-CanRsqr)**

	Eigenvalue	Difference	Proportion	Cumulative
1	0.0698		1.0000	1.0000

**Test of H0: The canonical correlations in the current row
and all that follow are zero**

Likelihood

	Ratio	F	Num DF	Den DF	Pr > F
--	-------	---	--------	--------	--------

1	0.93476638	1.3841	6	119	0.2265
---	------------	--------	---	-----	--------

Canonical Correlation Analysis
Raw Canonical Coefficients for the Condición Social

CONDII	
X1	0.0651084992
X2	0.1198709966
X3	-0.036021416
X4	-0.047207436
X5	-0.043177299
X6	-1.580013884

Raw Canonical Coefficients for the Indicador Escolar
EDUCA1

Y3	0.0679730703
----	--------------

Standardized Canonical Coefficients for the Condición Social

CONDII	
X1	1.1884
X2	0.8677
X3	-0.4494
X4	-1.1860
X5	-0.5309
X6	-0.6891

Standardized Canonical Coefficients for the Indicador Escolar

EDUCA1	
Y3	1.0000

4.4 Comentarios de Inferencia

De acuerdo a la salida obtenida en SAS, para el primer caso podemos observar el grado de significancia que corresponde al primer par de variables donde sus valores característicos son:

$$\text{corr}(Z_1 U_1) = r_1 = 0.627288 \text{ con } Pr > F = 0.0001$$

$$\text{corr}(Z_2 U_2) = r_2 = 0.260955 \text{ con } Pr > F = 0.1886$$

$$\text{corr}(Z_3 U_3) = r_3 = 0.206103 \text{ con } Pr > F = 0.2666$$

Y respecto al segundo caso encontramos el siguiente valor característico:

$$\text{corr}(V_1, W_1) = r_1 = 0.255409 \text{ con } \text{Pr}>F = 0.2265$$

Entonces de acuerdo a los resultados, podemos decir que en el primer caso, las variables de condición social con los indicadores de educación, nos aportan una mejor y significativa correlación, por lo que nuestros coeficientes canónicos de las variables a escoger están en CONDIC1 y EDUCA1; sin embargo para el segundo caso observamos que las variables socioeconómicas resultaron ser no significativas para explicar la eficiencia terminal. de manera que la combinación lineal con los pesos canónicos es:

(sólo primer caso)

$$Z1 = (-0.1502)X2 + (0.4087)X3 - (1.0850)X4 - (0.2801)X5 - (0.4082)X6$$

$$U1 = (0.2449)Y1 + (0.8697)Y2 + (0.1723)Y3$$

De acuerdo a las hipótesis planteadas podemos decir, que no se rechaza la hipótesis nula, es decir existe una correlación entre las variables; y respecto al segundo caso, podemos decir que se rechaza la hipótesis nula, es decir no existe una correlación significativa entre las variables. se puede deducir que las variables de condición social, no son lo suficientemente dependientes como para explicar el indicador de eficiencia terminal, lo que significa que existirían otras variables por ejemplo (número de hijos por familia, tipo de casa, hábitos de estudio, tipo de escuela, criterios de escuela, etc.) que posiblemente expliquen más a ésta variable.

Podríamos comentar respecto al análisis de Redundancia canónica, que para interpretarlo se debe tener cuidado, por que las combinaciones R^2 representan la varianza compartida por los compuestos lineales de los 2 conjuntos de variables criterio y predictores. al respecto no se considero.

CAPITULO V

TECNICAS DE AGRUPACION EN CÚMULOS

5.1. DEFINICION DE CÚMULO

Definición:

EL análisis de agrupación por cúmulos, llamado también (Cluster), es el nombre de un grupo de técnicas multivariadas, las cuales su principal propósito es identificar conjuntos similares de elementos de acuerdo a las características que ellos presenten.

Objetivo:

Reducir la dimensión de los datos a través del número de cúmulos deseados, con el fin de agruparlos de acuerdo a sus similitudes.

5.2 PROGRAMA Y SALIDA DE SYSTAT

Con el uso de este paquete, obtuvimos la salida de CLUSTER para k=3 means. debemos señalar que aquí nuestra variable de agrupamiento son las x's; entendiendo por variables de agrupamiento a las de condición socioeconómicas, que van a permitir conocer al respecto, que municipios tienen mayor similitud.

Las instrucciones para la ejecución del análisis:

POR MENÚ

FILE	USE A:	STATISTICS	CLUSTER	K MEANS	VARIABLES
------	--------	------------	---------	---------	-----------

SALIDA

SUMMARY STATISTICS FOR 3 CLUSTERS

BIBLIOTECA
 UNIDAD DOCENTE INTERDISCIPLINARIA
 DE ECONOMIA Y ESTADISTICA

VARIABLE	BETWEEN SS	DF	WITHIN SS	DF	F-RATIO	PROB
X1	27551.356	2	14094.239	123	120.220	0.000
X2	2446.758	2	4102.373	123	36.680	0.000
X3	9483.669	2	9969.016	123	58.506	0.000
X4	60858.838	2	18042.185	123	207.448	0.000
X5	12217.169	2	6679.418	123	112.488	0.000
X6	14.152	2	9.628	123	90.394	0.000

CLUSTER NUMBER: 1

MEMBERS		STATISTICS				
CASE	DISTANCE	VARIABLE	MINIMUM	MEAN	MAXIMUM	ST.DEV.
2	7.93	X1	52.01	70.95	94.43	9.64
3	8.21	X2	66.68	84.36	94.52	5.59
6	8.54	X3	28.42	76.87	88.09	8.04
8	10.02	X4	41.21	64.20	89.91	11.73
9	9.84	X5	4.29	12.55	27.95	5.52
11	2.41	X6	0.87	1.58	2.12	0.26
13	13.69					
15	5.90					
19	6.93					
21	10.29					
22	5.41					
29	8.13					
32	8.38					
34	5.61					
38	2.43					
40	8.13					
41	11.74					
42	4.44					
43	8.35					
47	21.44					
50	6.00					
51	8.12					
53	3.42					
55	8.10					
58	10.23					
59	7.32					
61	4.02					
62	4.51					
63	6.99					
65	5.15					
70	3.18					
73	9.34					
74	5.81					
77	4.27					
79	5.46					
80	10.33					

81	7.47
82	2.40
85	5.05
89	9.98
90	6.17
91	3.38
92	6.19
93	3.77
94	3.88
96	5.31
97	9.97
98	5.92
100	5.87
101	2.76
103	7.57
104	11.60
105	5.63
107	7.07
108	9.42
109	6.88
110	7.08
112	10.68
113	3.30
114	7.75
115	8.89
117	5.00
119	8.55
121	6.09
123	5.21

CLUSTER NUMBER: 2

MEMBERS		STATISTICS				
CASE	DISTANCE	VARIABLE	MINIMUM	MEAN	MAXIMUM	ST.DEV.
1	8.31	X1	56.42	84.54	97.33	10.54
4	7.93	X2	63.50	76.93	87.41	6.02
10	2.05	X3	44.96	86.22	94.45	7.98
12	4.35	X4	1.71	20.61	46.72	13.71
14	9.23	X5	7.83	32.29	70.44	10.81
17	6.98	X6	1.50	2.09	2.91	0.32
18	9.39					
20	6.72					
24	11.04					
26	5.52					
27	9.25					
28	3.80					
30	8.46					
31	4.23					
33	8.66					
37	10.65					

46	4.20
56	13.06
57	10.55
60	11.55
64	7.30
66	9.91
68	11.07
69	4.19
71	13.29
72	24.67
75	3.72
76	3.67
78	10.73
83	6.37
86	9.35
87	5.58
88	9.73
116	6.05
118	8.53
120	10.30
122	10.37

CLUSTER NUMBER: 3

MEMBERS		STATISTICS				
CASE	DISTANCE	VARIABLE	MINIMUM	MEAN	MAXIMUM	ST.DEV.
5	6.73	X1	4.96	41.33	64.78	12.82
7	7.61	X2	75.50	89.35	96.84	5.50
16	4.43	X3	31.30	60.72	77.67	11.92
23	7.32	X4	58.87	77.14	89.47	9.45
25	5.25	X5	1.84	7.23	15.12	3.97
35	4.70	X6	0.79	1.12	1.68	0.24
36	4.03					
39	6.62					
44	5.79					
45	15.89					
48	8.25					
49	5.97					
52	17.74					
54	3.93					
67	4.34					
84	8.09					
95	10.14					
99	12.89					
102	9.82					
106	5.10					
111	6.42					
124	9.95					
125	8.87					
126	8.17					

NUMBER OF OBSERVATIONS: 126

Para identificar mejor a los municipios que se agrupan en los cluster se presentan las siguientes tablas, conservando el mismo orden como se presentan en la salida.

TABLA QUE CONTIENE LOS MUNICIPIOS DEL CÚMULO 1

ACTOPAN	OMEALCA
ACULA	OTATITLAN
ALTO LUCERO	PANUCO
ALVARADO	PAPANTLA
AMATLAN DE LOS REYES	PASO DEL MACHO
ANGEL R.CABADA	PLAYA VICENTE ♡
APAZAPAN	SALTABARRANCA
ATZACAN	SAN ANDRES TUXTLA
CAMARON DE TEJEDA	SAN JUAN EVANGELISTA
CASTILLO DE TEAYO	SANTIAGO TUXTLA
CATEMACO	SAYULA DE ALEMAN
COLIPA	SOLEDAD DE DOBLADO
COSAUTLAN DE CARBAJAL	TAMALIN
COTAXTLA	TAMIAHUA
CHACALTIANGUIS	TANCOCO
CHICONTEPEC	TECOLUTLA
LAS CHOAPAS	TEMAPACHE
CHOCAMAN	TEOCELO
EMILIANO ZAPATA	TEPATLAXCO
GUTIERREZ ZAMORA	TEPETZINTLA
HUEYAPAN DE OCAMPO	TEZONAPAN
IGNACIO DE LA LLAVE	TIERRA BLANCA
ISLA	TLACOJALPAN
IXHUATLAN DEL CAFE	TLACOTALPAN
JAMAPA	TLAPACOYAN
JESUS CARRANZA	TOMATLAN
JOSE AZUETA	TOTUTLA
JUAN RODRIGUEZ CLARA	TRES VALLES
JUCHIQUE DE FERRER	TUXTILLA
MANLIO FABIO ALTAMIRANO	VEGA DE LA TORRE
MISANTLA	TUXTILLA
NAOLINCO DE VICTORIA	VEGA DE LA TORRE
NAUTLA	VIGAS DE RAMIREZ
	XICO

TABLA QUE CONTIENE LOS MUNICIPIOS DEL CÚMULO 2

ACAYUCAN	MARTINEZ DE LA TORRE
AGUA DULCE	MEDELLIN DE BRAVO
LA ANTIGUA	MINTITLAN
ATOYAC	MOLOACAN
BANDERILLA	NANCHITAL
BOCA DEL RIO	NOGALES
CAMERINO Z.MENDOZA	OLUTA
CERRO AZUL	ORIZABA
COATEPEC	PEROTE
COATZACOALCOS ♡	POZA RICA
COATZINTLA	PUEBLO VIEJO
CORDOBA	RIO BLANCO
COSAMALOAPAN	TUXPAN
COSOLEACAQUE	URSULO GALVAN
CUITLAHUAC	VERACRUZ
FORTIN	XALAPA
IXHUATLAN DEL SURESTE	
IXTACZOQUITLAN	
JILOTEPEC	
LERDO DE TEJADA	

TABLA QUE CONTIENE LOS MUNICIPIOS DEL CÚMULO 3

ALPATLAHUAC	ILAMATLAN
ALTOTONGA	IXHUATLAN DE MADERO
ATZALAN	MECAYAPAN
CAZONES DE HERRERA	PLATON SANCHEZ
CITLALTEPEC	SOTEAPAN
COXQUIHUI	TANTOYUCA
COYUTLA	TEMPOAL
CHICONQUIACO	TEXHUACAN
ESPINAL	TLACHICHILCO
FILOMENO MATA	YECUATLA
HIDALGOTTILAN	ZONGOLICA
HUAYACOCOTLA	ZOZOCOLCO DE HIDALGO

5.3. ANALISIS DE VARIANZA MULTIPLE

Definición:

Es una técnica que nos permite probar ciertas hipótesis importantes mediante procedimientos relativamente fáciles.

Se hizo uso de esta técnica a partir del agrupamiento por cúmulos, a fin de probar el grado de significancia que presentan éstos respecto a cada variable dependiente.

El modelo aplicado para las variables dependientes es el siguiente:

$$y_i = CL + e_{ij} \quad \begin{array}{l} i=1, \dots, 3 \\ j=1, \dots, 3 \end{array}$$

En la comparación múltiple de medias nos interesa probar la igualdad de estas, en cada uno de los cúmulos, es decir :

5.4 PLANTEAMIENTO DE HIPOTESIS:

$$H_0: \mu_1 = \mu_2 = \mu_3 \quad H_a: \mu_i \neq \mu_j \text{ par al menos un } i \text{ distinto de } j$$

Si H_0 es verdadera, todos los cúmulos tienen la media común μ

Para obtener las técnicas de Anova y comparación múltiple de medias, se utilizó SYSTAT con el siguiente programa :

```
>MGLH
>fedit a: .cmd
>use a: .sys
>output a:salida.dat
>print long
>category variable
>model y1=constant + C1
>estimate
>hipótesis
>post C1/Tuckey
>tes
>submit
SALIDA
```

LEVELS ENCOUNTERED DURING PROCESSING ARE:

CL

1.000 2.000 3.000

DEP VAR: Y1 N: 126 MULTIPLE R: 0.262 SQUARED MULTIPLE R: 0.069

-1

ESTIMATES OF EFFECTS $B = (X'X)^{-1} X'Y$

	Y1
CONSTANT	5.404
CL	1.000 -0.285
CL	2.000 1.749

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
--------	----------------	----	-------------	---------	---

CL	168.895	2	84.448	4.536	0.013
ERROR	2289.762	123	18.616		

LEAST SQUARES MEANS.

		LS MEAN	SE	N
CL =	1.000	5.120	0.535	65
CL =	2.000	7.153	0.709	37
CL =	3.000	3.940	0.881	24

COL/

ROW	CL
1	1.000
2	2.000
3	3.000

USING LEAST SQUARES MEANS.

POST HOC TEST OF Y1

USING MODEL MSE OF 18.616 WITH 123. DF.

MATRIX OF PAIRWISE MEAN DIFFERENCES:

	1	2	3
1	0.000		
2	2.033	0.000	
3	-1.180	-3.213	0.000

TUKEY HSD MULTIPLE COMPARISONS.

MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

	1	2	3
1	1.000		
2	0.057	1.000	
3	0.487	0.012	1.000

LEVELS ENCOUNTERED DURING PROCESSING ARE:

CL	1.000	2.000	3.000
----	-------	-------	-------

DEP VAR: Y2 N: 126 MULTIPLE R: 0.501 SQUARED MULTIPLE R: 0.251

-1

ESTIMATES OF EFFECTS B = (X'X)⁻¹X'Y

	Y2
CONSTANT	31.315
CL 1.000	0.312
CL 2.000	12.285

ANALYSIS OF VARIANCE

SOURCE	SUM-OF-SQUARES	DF	MEAN-SQUARE	F-RATIO	P
CL	9162.816	2	4581.408	20.641	0.000
ERROR	27301.310	123	221.962		

LEAST SQUARES MEANS.

		LS MEAN	SE	N
CL =	1.000	31.626	1.848	65
CL =	2.000	43.600	2.449	37
CL =	3.000	18.718	3.041	24

COL/ ROW	CL
1	1.000
2	2.000
3	3.000

USING LEAST SQUARES MEANS.

POST HOC TEST OF Y2

USING MODEL MSE OF 221.962 WITH 123. DF.
MATRIX OF PAIRWISE MEAN DIFFERENCES:

	1	2	3
1	0.000		
2	11.974	0.000	
3	-12.909	-24.882	0.000

TUKEY HSD MULTIPLE COMPARISONS.
MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

	1	2	3
1	1.000		
2	0.000	1.000	
3	0.001	0.000	1.000

LEVELS ENCOUNTERED DURING PROCESSING ARE:

CL	1.000	2.000	3.000
----	-------	-------	-------

DEP VAR Y3 N 126 MULTIPLE R 0.096 SQUARED MULTIPLE R 0.009

-1

ESTIMATES OF EFFECTS B = (X'X)⁻¹ X'Y

Y3

CONSTANT 63.206

CL 1.000 -1.741

CL 2.000 1.384

ANALYSIS OF VARIANCE

SOURCE SUM-OF-SQUARES DF MEAN-SQUARE F-RATIO P

CL 248.365 2 124.183 0.570 0.567

ERROR 26805.931 123 217.934

LEAST SQUARES MEANS.

		LS MEAN	SE	N
CL =	1.000	61.465	1.831	65
CL =	2.000	64.590	2.427	37
CL =	3.000	63.563	3.013	24

COL/ ROW	CL
1	1.000
2	2.000
3	3.000

USING LEAST SQUARES MEANS.

POST HOC TEST OF Y3

USING MODEL MSE OF 217.934 WITH 123. DF.

MATRIX OF PAIRWISE MEAN DIFFERENCES:

	1	2	3
1	0.000		
2	3.125	0.000	
3	2.098	-1.027	0.000

TUKEY HSD MULTIPLE COMPARISONS.

MATRIX OF PAIRWISE COMPARISON PROBABILITIES:

	1	2	3
1	1.000		
2	0.559	1.000	
3	0.823	0.962	1.000

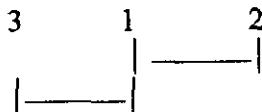
5.5 COMENTARIOS DE INFERENCIA

De acuerdo a la salida, el Anova nos indica una $P=0.013$ lo que muestra que sí son significativos los cúmulos bajo un nivel confiabilidad del 95 % para la variable dependiente Y_1 , conocida como deserción y encunto a su comparación múltiple de medias ésta nos indica que hay diferencias entre las medias de los cúmulos 2 y 3, en contraste con las medias de los cúmulos 1 y 2, 1 y 3 que son semejantes.

Para la variable reprobación, el Anova nos muestra una $P=0.000$ lo que indica que existe significancia en los cúmulos bajo el mismo nivel de confiabilidad. Respecto a la comparación múltiple, se observa diferencia entre las tres medias de los cúmulos.

En cuanto a la eficiencia terminal , podemos decir que el Anova muestra una $P=0.567$ lo que indica que no resultaron significativos los cúmulos bajo el mismo grado de confiabilidad. Y en cuanto a la comparación múltiple se observa cierta semejanza entre las medias de los cúmulos.

Una manera de ejemplificar la comparación múltiple de medias Tuckey de la salida, sería de la siguiente forma en el caso de la variable deserción con un $\alpha=0.05$, donde podemos observar como los cúmulos 1 y 2 no difieren; sin embargo el cúmulo 3 difiere del cúmulo 2, pero no del cúmulo 1.

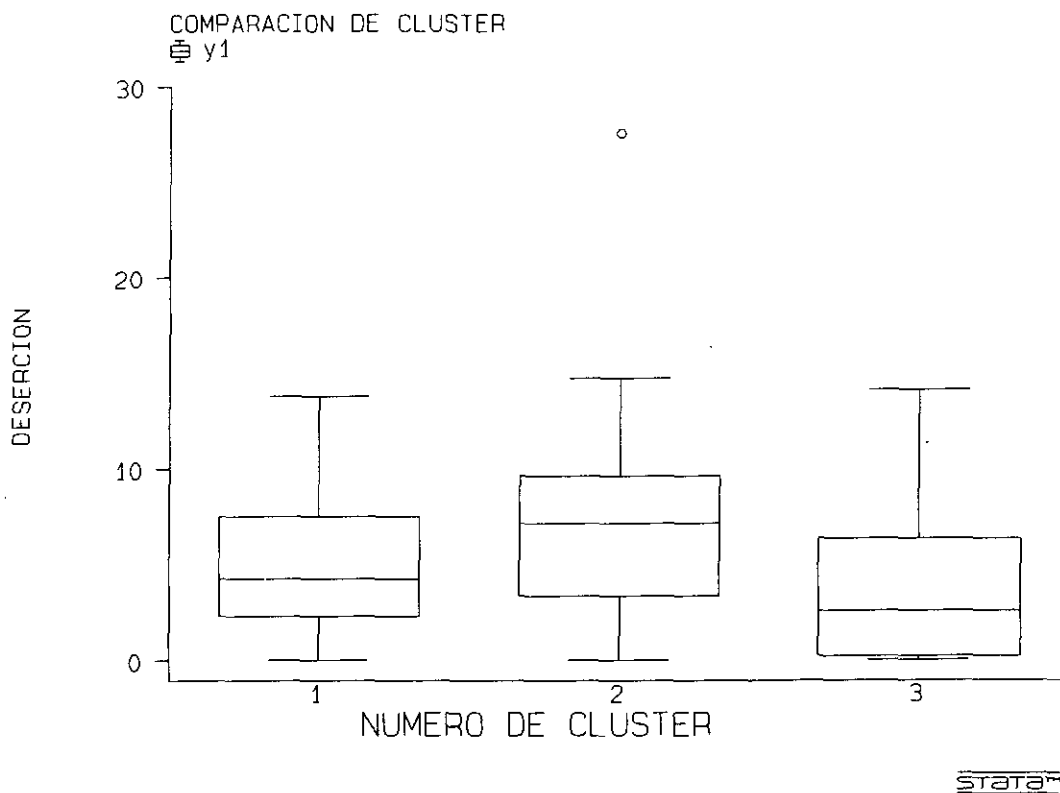


Respecto a Y_2 (reprobación) con el mismo alpha, los cúmulos difieren todos entre sí y para Y_3 (eficiencia terminal), no presentan diferencias los cúmulos con respecto a esta variable.

5.6 ANALISIS GRAFICO DE CÚMULOS

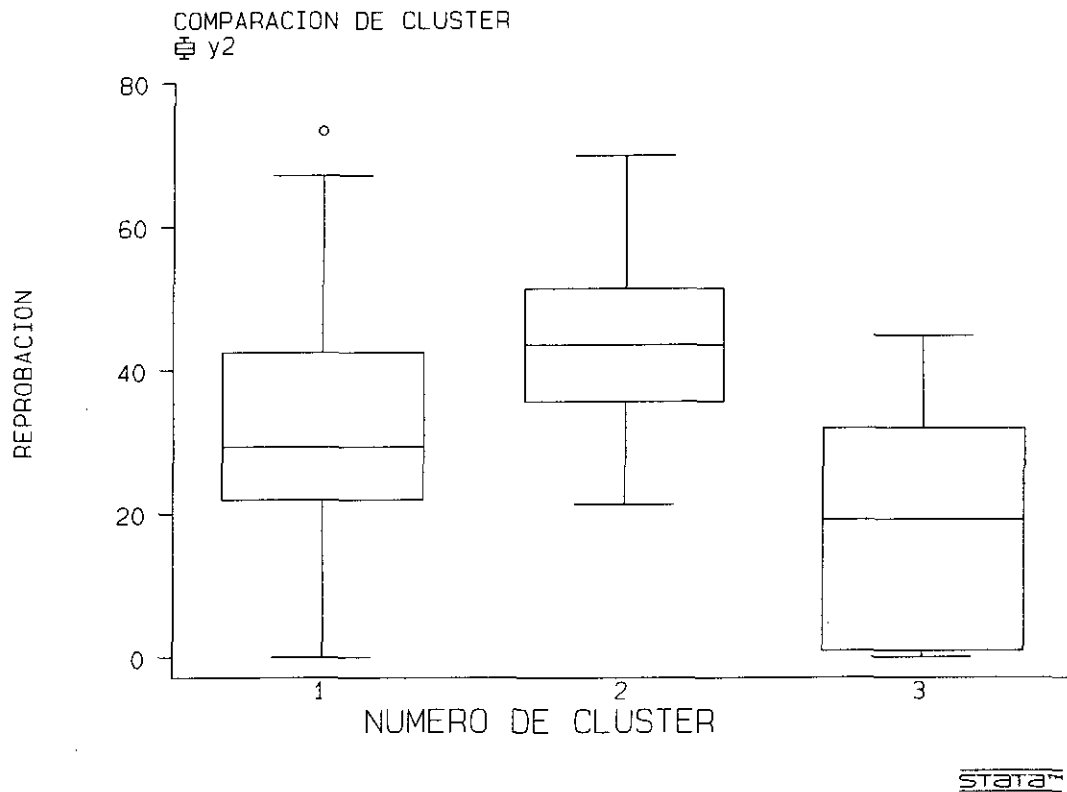
Una manera de contrastar lo antes señalado es através de los gráficos de cajas y alambres, en donde se contemplan los tres cúmulos de los indicadores de educación como deserción, reprobación y eficiencia terminal.

GRAFICO No. 4



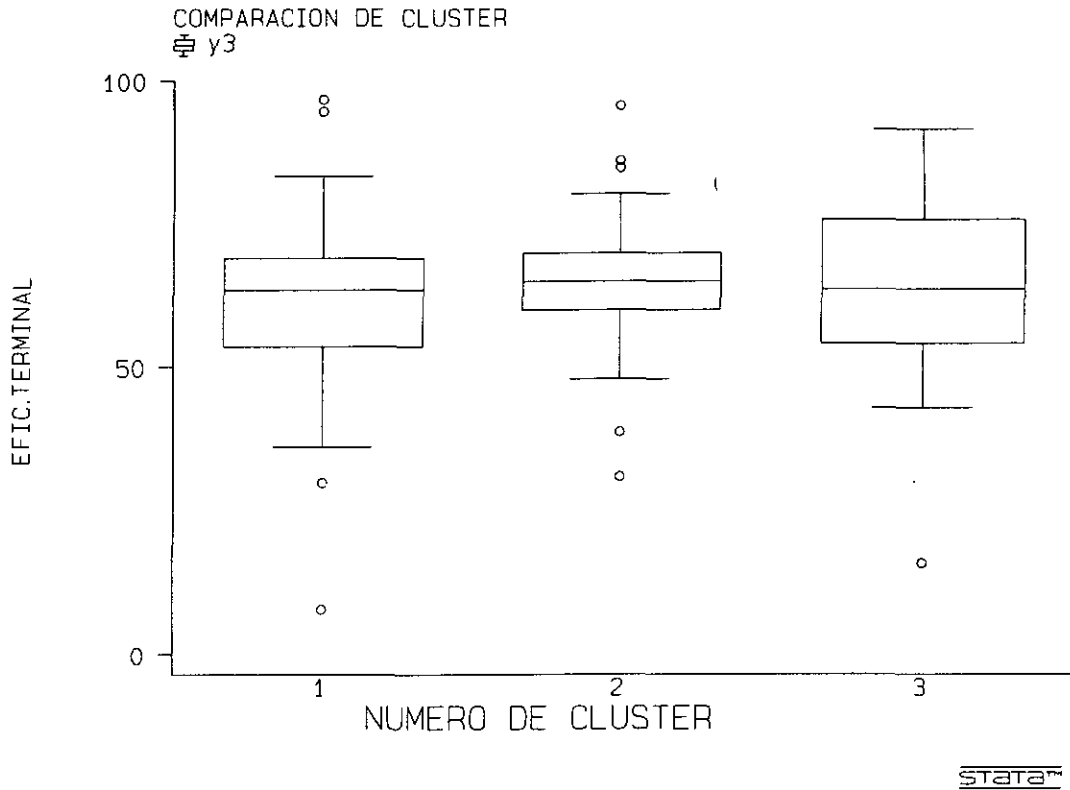
En cuanto a la interpretación para este gráfico podemos decir que la composición de los cúmulos respecto a la proporción de deserción, está representada por el grupo 2 el cual presenta la mayor de las medianas, por lo que se entiende que dicho grupo concentra el mayor porcentaje de deserción; además observamos un valor alejado que corresponde al municipio que presenta mayor porcentaje de deserción, a saber Cerro Azul con el 27.53%. Otro aspecto relevante en el análisis de esta variable lo son aquellos municipios agrupados en el cúmulo 3, en el cual no presentan deserción los municipios, dado que su rango mínimo inicia en porcentajes de cero. Respecto al cúmulo 1, podemos decir que ocupa el segundo lugar en cuanto a su mediana.

GRAFICO No. 5



Podemos señalar que en el presente gráfico para la variable reprobación, también el grupo 2 presenta la mayor de las medianas, se observa cierta simetría en éste; respecto al cúmulo 1, podemos apreciar un valor alejado que corresponde al municipio de Otatitlán con un 73.33% de reprobación, siendo este el valor más alto; sin embargo es relevante señalar que la agrupación de municipios del cúmulo 3 conservan un porcentaje bajo de reprobación

GRAFICO No. 6



Nuevamente se observa en este gráfico que el cúmulo 2 posee la mayor de las medianas; aún cuando existe proximidad entre las medianas de los tres grupos .

Por otro lado se aprecian valores lejanos en los tres grupos, pero solo analizaremos éstos valores para el agrupamiento 1, dado que es el que refleja los porcentajes mínimos y máximos en la eficiencia terminal, los cuales pertenecen al municipio de Isla con el 7.78% y el municipio de Actopan con un 94.56% respectivamente.

Podríamos mencionar además sobre la simetría que presenta el cúmulo 2, entendiendola como una proporción equilibrada entre el número de alumnos que ingresan respecto al numero de egresados.

CAPITULO VI

CONCLUSION

Durante el desarrollo del trabajo, fuimos encontrando a través del uso de estadística descriptivas, el comportamiento de las variables en estudio como la mediana, media, valores máximos y mínimos entre otros; lo que nos permitió adentrarnos a la aplicación de las técnicas ya conocidas como Análisis de Correlación Canónica y Análisis de Cúmulos; las cuales nos permitieron llegar a conocer el objetivo planteado, como observar, si mantenían una correlación entre ellas, a fin de obtener una mejor explicación del comportamiento que reflejan los porcentajes de deserción, reprobación y eficiencia terminal emitidos por la Secretaría de Educación y Cultura, en relación con las condiciones socioeconómicas. Después de aplicar la primera técnica pudimos encontrar para el primer caso, que sí existe una correlación significativa entre ellas es decir, que los indicadores de educación tienen una dependencia con las variables de condición socioeconómicas, las cuales quedaron representadas en forma canónica por los porcentajes de la población alfabetizada, población ocupada en el sector primario así como por el ingreso promedio en función del salario mínimo que percibe la población. Con esto podríamos decir que, cuanto más personas alfabetizadas estén en la familia que rodea al joven de bachillerato, podría no darse la reprobación, ésta variable es la más representativa de los indicadores de educación, en relación con el sector primario podemos decir que, a menor gente ocupada en agricultura, ganadería, caza y pesca menor podría ser la reprobación y respecto al ingreso promedio podemos decir que, a menor ingreso promedio podría ser menor la reprobación es decir que, cuando hay menos ingreso se distribuye mejor el dinero y se podría reducir el fenómeno de la reprobación. Para el segundo caso encontramos cómo las variables socioeconómicas no fueron tan signifi-

cativas con la eficiencia terminal, lo que nos lleva a pensar que posiblemente existen otros aspectos que determinan la eficiencia, que no necesariamente son de tipo socioeconómicas, como podrían ser la existencia de criterios bajo los cuales la S.E.C. mantiene porcentajes determinados de egresión y hasta de reprobación. Cabe señalar que el simple hecho de retomar los datos emitidos por ésta Secretaría, representó en algunos indicadores incongruencias como decir, que egresó el 300% de alumnos del bachillerato, sin embargo como se mencionó en el trabajo bajo estas circunstancias no fueron considerados aquellos municipios.

La segunda técnica aplicada, nos permitió conocer la forma en que se agrupan los municipios en relación con los indicadores de educación. De manera que se observó como predomina la mediana del grupo 2 en la deserción, concentrando así el mayor porcentaje en éste indicador; respecto a la reprobación y a la eficiencia terminal se encontró que también los municipios del grupo 2 predominan en estos indicadores, sólo que en la eficiencia terminal los grupos 1 y 3 se encuentran prácticamente a niveles muy similares al cúmulo 2. Al respecto podríamos decir que los municipios del cúmulo 2 cuenta con una posición en términos relativos económicamente alta, es decir que deberían conservar cifras menores en la deserción y reprobación por la infraestructura que guardan éstos respecto a los otros municipios. Por su parte el grupo 3 es el que conserva menores porcentajes en deserción y reprobación, aún cuando no se cumple en algunos de sus municipios satisfactoriamente el egreso de alumnos del nivel medio superior, de acuerdo a los municipios que se agrupan en él su infraestructura no es igual como el resto de los municipios, de manera que podríamos decir que, pertenecen a un nivel de clase baja; por lo que los municipios del grupo 1, se podrían considerar como un nivel de clase media en el que se presentan los tres indicadores de manera moderada, sin embargo también existe algún municipio en el que se dá mucho la reprobación así como exceso de alumnos egresados del bachillerato.

A través de estos análisis pudimos conocer, la relación que guardan las condiciones socioeconómicas con los indicadores de educación del nivel medio superior en los municipios del Estado de Veracruz; conocer y dar a conocer como avanzamos en el aspecto educativo, es preocupación de muchos que nos interesa saber respecto a ello. Podemos decir que aún queda mucho por investigar en relación ahora, con los fenómenos que presentan los municipios bajo los tres niveles de clase en que se agrupan, además de considerar por otro lado la modalidad del plantel educativo a decir, Estatal, Federal y Particular; para visualizar mejor donde ocurren los hechos e investigar a que obede el comportamiento de los indicadores que en cada generación se presentan.

BIBLIOGRAFIA

- 1.- Cuadras Avellana, Carles M. Fac.Biología. "Métodos de Análisis Multivariantes", Univ. Barcelona.
- 2.- Johnson, Richard A., Univ. of Wisconsin-Madison, "Applied Multivariate Statistical Analysis.
- 3.- Ato, Manuel/López, Juan José. "Fundamentos de estadística con Systat, Addison-Wesley Iberoamericana, Edic. Ra-ma, 1994.
- 4.- Infante, gil Said/Zárate, Lara Guillermo P. "Métodos Estadísticos" Edicc. Trillas. Imp.1990
- 5.- Lawrence, C. Hamilton. "Statistics with stata.3" Univ.New Hampshire. Edicc. Duxbury. Publis. Company Belmont Calif.
- 6.- Solana, Fernando/Cardiel, Reyes Raúl/Bolaños, Martínez Raúl."Historia de la Educación en México".SEP. Edit.Fondo de Cultura Económica, México.Imp.1981.
- 7.- Anuario Estadístico del Estado de Veracruz. INEGI. Edición 1994.
- 8.- Estadísticas de inicio y fin de cursos 1994-95 y 1993 en el nivel educativo del Estado de Veracruz. Secretaría de Educación y Cultura.