4.4.4 Análisis de las componentes principales de los Directores o Rectores

El primer paso en el análisis de componentes principales es determinar si es aplicable el método de componentes principales a las variables aleatorias de la matriz de datos de los directores o rectores, mediante el estadístico de prueba de Bartlett, en la cual se plantearán supuestos de normalidad sobre las variables aleatorias al decir que sus covarianzas son 0 implicando dependencia entre ellas, de esta forma el contraste de hipótesis esta dado por:

$$H_o: \quad \Sigma = \begin{bmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{pp} \end{bmatrix} \qquad \text{vs.} \qquad H_1 \qquad H_0$$

Tabla IV.XXIII
Tabla de estadístico de prueba de Bartlett

	Val□□	□.l.	Val□□ □
□□□□ba □□□□a□□a□□	353342.15	1540	0.000
□□□□ □a□□□	5457		

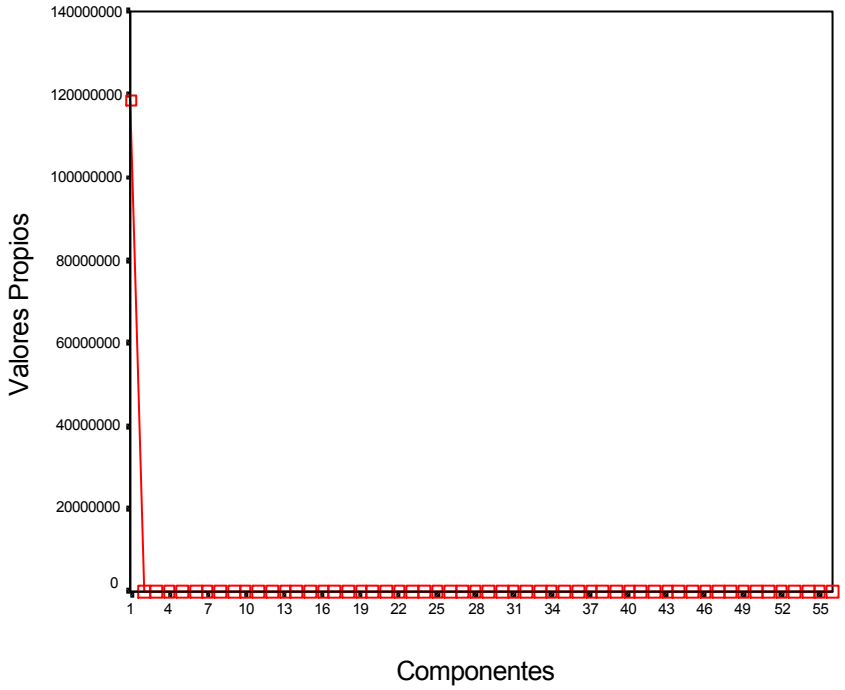
□□□□□□ [Base de datos del Censo del Magisterio Fiscal. 2000

Como podemos observar en la tabla IV.XXIII el valor del estadístico de prueba de Bartlett es de 353342.15, con un valor p de 0.000, por lo que podemos concluir que existe evidencia estadística para rechazar la hipótesis nula planteada, es decir los valores de las covarianzas son diferentes de 0, y podemos afirmar que no existe independencia entre las variables de la matriz de datos de los directores o rectores; una vez rechazada la hipótesis nula se procede a realizar el análisis de las componentes principales.

Con la ayuda del paquete estadístico SPSS obtuvimos los valores propios de la matriz de covarianzas, los cuales representan las varianzas de cada componente, con su respectivo porcentaje de explicación, y el porcentaje de explicación acumulado para cada componente principal, como se detallan en el Anexo 5

En el gráfico 4.1 se ilustran los valores propios obtenidos de la matriz Σ de los directivos, con respecto a las 56 componentes principales respectivas, en este gráfico se pueden apreciar que a partir de la segunda componente sus valores propios toman un valor de 0 ($Var(Y_i) = 0$), de esta forma para nuestro análisis se tomamos solo la primera componente principal.

Figura 4.
Valores propios de las componentes principales calculada a partir de los datos originales



Fuente: Base de datos del Censo del Magisterio Fiscal. 2000
Elaborado: Julio Veloz Serrano

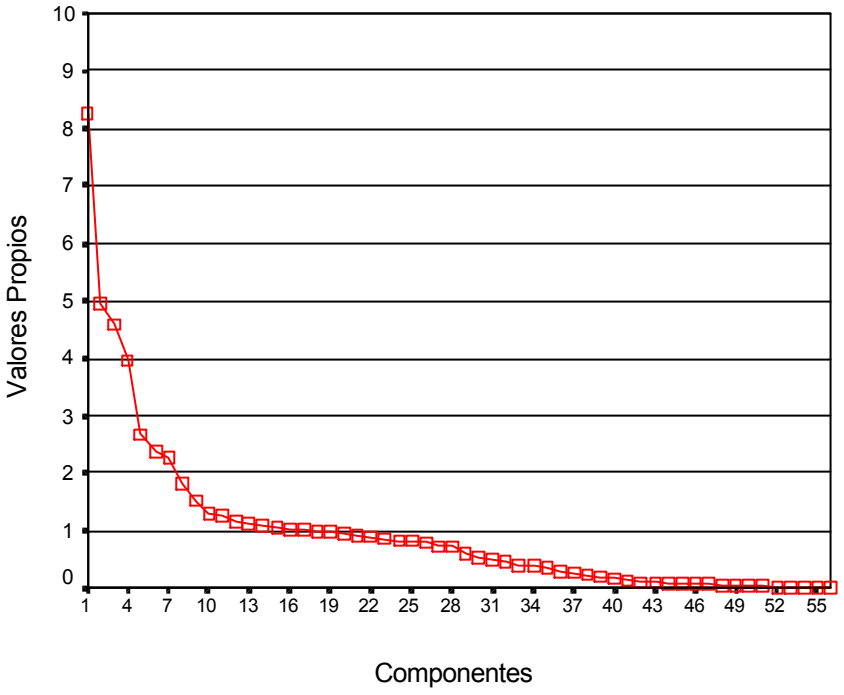
En la Anexo 6 se muestran los coeficientes de la primera componente principal, debido a que esta explica en conjunto el 99.967% de la varianza, a su vez se puede observar en esta tabla que el valor del coeficiente de la variable Alumnos del ciclo básico (ID23) es de 10878.4, la cual absorbe a los demás coeficientes debido a que esta variable contiene valores grandes, sin saber si esta variable realmente tiene una gran relevancia sobre las demás,

por este motivo surge el problema al utilizar el análisis de componentes principales con diferentes tipos de escala al tomar la matriz de datos originales.

Para evitar este tipo de problema, se procede a llevar a todas las variables a una misma escala, este procedimiento se lo realiza estandarizando los valores de cada una de las variables. A partir de la matriz de correlación de los directores o rectores (Ver anexo 1) y con la ayuda del software estadístico SPSS, se obtuvo 18 componentes principales, las cuales explican el 75.696% de la varianza. Se puede observar en la Anexo 7 los valores propios para la matriz de correlación que son la varianza de cada componente, el porcentaje de explicación de cada varianza de las componentes y su acumulado, en el Anexo 8 se pueden apreciar las 18 primeras componentes principales con sus respectivos coeficientes.

En el gráfico 4.2 se ilustran los valores propios obtenidos a partir de la matriz de correlación, con respecto a las 56 componentes principales, en este gráfico se puede apreciar que a partir de la decimonovena componente principal toma un valor menor de 1 (varianza de la componente), por lo que para nuestro análisis se tomará solo las primeras 18 primeras componentes principales.

□□□□□ 4.□
Valores propios de las componentes principales calculada a partir de la matriz de correlación



□□□□□ □Base de datos del Censo del Magisterio Fiscal. 2000

En las tablas anteriores podemos observar la diferencia que existe entre usar los datos originales o los estandarizados; como se observa en el Anexo 7 que representa las varianza de explicación de las componentes principales con respecto a los datos estandarizados, como esta representa la varianza de una distribución normal, por esta razón se utiliza el criterio de escoger las componentes cuyos valores propios sean mayores a 1, el

número de componentes a elegir serán las primeras 18 componentes, con un porcentaje de explicación de 76.696%, es decir que el porcentaje de la variación total de la población se explica a través de las 18 componentes

Analizando la tabla del Anexo 8 podemos notar que la primera componente estaría compuesta por 14 variables, mientras que la segunda componente estaría compuesta por 7 variables, mientras que para el resto de componentes, las cargas o coeficientes de participación son pequeños, repitiendo incluso variables entre factores, quitando de esta manera representatividad a las demás componentes. Con estos resultados se puede apreciar un patrón que ocurre regularmente al efectuar este método: la primera componente tiende a ser muy general agrupando variables significativas y acumula el mayor porcentaje individual de explicación, mientras que los restantes a partir de la cuarta componente en nuestro caso se basan en cantidades ínfimas de participación en la varianza y el determinar o nombrar a las componentes se torna difícil por la confusión surgida entre los aportes que da una misma variable a diferentes factores.

Por tal razón, resulta beneficioso una rotación de factores, útil para esta clase de situaciones ya que permite a través de este procedimiento en la matriz de factores de las componentes principales redistribuir la varianza a lo largo de todas las componentes simplificando el modelo y obteniendo patrones más claros para identificar los factores en cada componente, para este objetivo se utilizará el método VARIMAX, este método permite realizar una rotación ortogonal.

Una vez efectuada la rotación de las componentes por el método de Varimax, dio como resultado una mejor redistribución de la varianza, pero la varianza total determinada a partir de las 18 componentes se mantiene en 75.696%. (Véase la Tabla XXVIII).

Tabla IV.XXVIII
LITORAL ECUATORIANO: Censo del Magisterio Fiscal. Valores propios de las componentes principales rotada por el método VARIMAX (matriz estandarizada de los profesores).

Componente	λ_j	% de la varianza total	% Acumulado
1	5.9683	10.6577	10.658
2	3.0321	5.4144	16.072
3	4.4980	8.0322	24.104
4	3.6446	6.5083	30.613
5	2.9366	5.2440	35.857
6	2.2216	3.9672	39.824
7	2.1751	3.8841	43.708
8	3.9128	6.9872	50.695
9	1.8925	3.3795	54.075
10	2.3731	4.2376	58.312
11	1.4022	2.5039	60.816
12	1.7074	3.0489	63.865
13	1.1530	2.0588	65.924
14	1.0971	1.9591	67.883
15	1.1466	2.0476	69.931
16	1.1483	2.0505	71.981
17	1.0590	1.8911	73.872
18	1.0212	1.8236	75.696

□□□□□□ □Base de datos del Censo del Magisterio Fiscal. 2000
□lab□□a□□□□ Julio Veloz Serrano

En la tabla del Anexo 9 se presentan los coeficientes de las componentes rotadas por el método de Varimax calculada a partir de la matriz de datos estandarizados de los directivos. En la tabla del Anexo 10 se presentan los vectores propios, a partir de los cuales se forman los coeficientes de las combinaciones lineales en cada componente.

Una vez analizada los coeficientes de mayor incidencia en cada componente rotada (Véase Anexo 10), dando como resultado

mejores patrones de agrupación, en donde prevalecen variables que definen nuestro factor, de esta manera se pueden obtener las siguientes combinaciones lineales, en nuestro caso se detallarán las primeras 7 componentes:

- En la primera componente Y_1 prevalecen las siguientes variables:

  IL17: Lugar donde habita (Sólo rural)

  ID29: Vivienda para docentes (solo rural)

  ID30: Servicio de agua

  ID31: Servicio de luz

  ID32: Alcantarillado

  $Y_1 = 0.2707\ IL17 + 2.765\ ID29 + 2.772\ ID30 + 0.2676\ ID31 + 0.2823\ ID32 + \ldots$

  A esta componente Ψ_1 se le podría denominar "Servicios básicos que poseen los directores que viven en el sector rural", esta componente explica el 10.66% del total de la variación de la población.

- En la segunda componente Y_2 prevalecen las siguientes variables:

  IP1: Provincia de Nacimiento

  IP8: Provincia donde habita

  IL3: Provincia donde labora

  IL6: Provincia donde pertenece presupuestariamente

  ID3: Provincia del plantel educativo

  $Y_2 = 0.0681\ IP1 + 0.0929\ IP8 + 0.0895\ IL3 + 0.0957\ IL6 + 0.0895\ ID3 + \ldots$

  A esta componente Y_2 se le podría denominar “Provincia donde nació y laboran los directores”, esta componente explica el 5.41% del total de la variación de la población.

- En la tercera componente Y_3 prevalecen las siguientes variables:

  ID14: Personal Docente que labora en el Plantel

  ID15: Personal Administrativo y de Servicio que labora en el Plantel

  ID17: Relación Laboral por nombramiento

  ID18: Relación laboral por contrato

ID24: N° de alumnos en el diversificado

Y_3 = 0.0817 ID14 + 0.1299 ID15 + 0.0747 ID17 + 0.1149 ID18 + 0.1048 ID24 + . . .

A esta componente Y_3 se le podría denominar "N° de personal docente y administrativo, con su respectivo N° de alumnos del diversificado", esta componente explica el 8.03% del total de la variación de la población.

- En la cuarta componente Y_4 prevalecen las siguientes variables:

  IP2: Edad

  IE6: Tipo de nombramiento

  IE7: Años de experiencia

  IE8: Escala nominal

  IE9: Escala económica

  IL10: Relación laboral

Y_4 = 0.1219 IP2 + 0.2357 IE6 + 0.1564 IE7 + 0.2468 IE8 + 0.2426 IE9 - 0.2205 IL10 + . . .

A esta componente Y_4 se le podría denominar “Forma de medir los años de experiencia del director”, esta componente explica el 6.51% del total de la variación de la población.

- En la quinta componente Y_5 prevalecen las siguientes variables:

  IL2: Nivel del plantel educativo donde labora

  IL7: Nivel del plantel educativo donde pertenece presupuestariamente

  ID4: Nivel del plantel educativo

  $Y_5 = 0.0632\ IL2 + 0.0665\ IL7 + 0.0637\ ID4 + \ldots$

  A esta componente Y_5 se le podría denominar “Nivel del plantel educativo”, esta componente explica el 5.24% del total de la variación de la población.

- En la sexta componente Y_6 prevalecen las siguientes variables:

  IL4: Sostenimiento del plantel educativo donde labora

  IL8: Sostenimiento del plantel educativo donde pertenece presupuestariamente

  ID6: Sostenimiento del plantel educativo

Y_6 = 0.0153 IL4+ 0.0161 IL8 + 0.0156 ID6 + . . .

A esta componente Y_6 se le podría denominar “Sostenimiento del plantel educativo”, esta componente explica el 3.97% del total de la variación de la población.

- En la séptima componente Y_7 prevalecen las siguientes variables:

  IL5: Zona del plantel educativo donde labora

  IL9: Zona del plantel educativo donde pertenece presupuestariamente

  ID7: Zona del plantel educativo

  Y_7 = 0.0153 IL5+ 0.0161 IL9 + 0.0156 ID7 + . . .

  A esta componente Y_7 se le podría denominar “Zona del plantel educativo”, esta componente explica el 3.88% del total de la variación de la población.