CAPITULO 4

4. ANÁLISIS MULTIVARIADO

4.1 Introducción.

Este capítulo tratará sobre técnicas multivariadas como: Correlación, Tablas de Contingencia y Componentes Principales. Que nos servirán para tener una idea más clara de la información obtenida en nuestra muestra. Los cálculos correspondientes serán realizados por medio del software estadísticos como: SPPS y Systad.

4.2 Marco teórico

4.2.1 Matriz de correlación

La matriz de correlación es un arreglo de p filas por p columnas, donde p es el número de variables investigadas, que agrupa todas las relaciones de tipo lineal.

Para este estudio, la matriz de datos es una matriz de n filas que corresponden al número de estudiantes (432 estudiantes) y p columnas que son 14 variables que han sido investigadas en los estudiantes que utilizan computadora. Estas variables se presentan en la sección 4.3.

Se define ρ (matriz de correlación) como: $\rho = V^{-1/2} \Sigma V^{-1/2}$, donde Σ es la matriz de varianzas y covarianzas de X y V es la matriz cuya diagonal principal contiene las desviaciones estándar de cada variable, ρ resume los coeficientes de correlación de todos los pares de variables entre las p dadas $X_1, X_2, ..., X_p$.

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

$$= \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & \rho_{pp} \end{bmatrix}$$

El coeficiente de correlación $\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}}$ determina el grado de dependencia lineal entre la i-ésima y la j-ésima variable. Si $\rho = 1$ o $\rho = -1$ determina en ambos casos que existe una perfecta dependencia lineal entre las variables, la diferencia del signo determina únicamente el sentido de variación si es positivo ambas variables aumentan o disminuyen simultáneamente y si es negativo el sentido de la variación es opuesto lo que significa que al aumentar una disminuye la otra. El hecho de que el coeficiente de correlación sea igual a cero indica que entre ambas variables no existe relación lineal.

Como en este análisis se esta trabajando con muestras se desconoce los parámetros antes mencionados por lo que se hace necesario estimarlos; así que el estimador insesgado de la varianza poblacional es:

$$S_{ii} = (1/(n-1)) \times \sum_{i=1}^{n} (x_i - \bar{x})$$

Y el estimador de la covarianza entre dos variables X_i y X_k es:

$$S_{jk} = (1/(n-1)) \times \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)$$

4.2.2 Tablas de Contingencia

Una tabla de contingencia es un arreglo bidimensional en la que se detalla los factores a ser analizados con igual o diferente niveles de información que nos permitirá determinar si esos dos factores son independientes al realizar un contraste de hipótesis sobre independencia de los factores.

Sea A un factor con r niveles y B un factor con c niveles, se define el modelo de la tabla de contingencia como:

		Factor B				
		Nivel 1	Nivel 2	…..	Nivel c	$X_{i.}$
Factor A	Nivel 1	X_{11} E_{11}	X_{12} E_{12}	….. …..	X_{1c} E_{1c}	$X_{1.}$
	Nivel 2	X_{21} E_{11}	X_{22} E_{11}	….. …..	X_{2c} E_{11}	$X_{2.}$
		⋮ ⋮	⋮ ⋮		⋮ ⋮	⋮ ⋮
	Nivel r	X_{r1} E_{r1}	X_{r2} E_{r2}	…… …..	X_{rc} E_{rc}	$X_{r.}$
	$X_{.j}$	$X_{.1}$	$X_{.2}$	…..	$X_{.c}$	$X_{..}$ =n

Donde:

$n =$ Es el número de observaciones

$X_{ij} =$ Es el número de valores observados que simultáneamente poseen la i-ésima característica del factor A y la característica j-ésima del factor B.

$E_{ij} =$ Es el número de observaciones esperadas con la i-ésima característica del factor A y la característica j-ésima del factor B y se lo obtiene:

$$E_{ij} = \frac{X_{i.} * X_{.j}}{n} = \frac{\sum_{j=1}^{c} X_{ij} * \sum_{i=1}^{r} X_{ij}}{n}$$

$X_{i.} =$ Al número de observaciones que poseen la característica i-ésima del factor B.

$X_{.j} =$ Al número de observaciones que poseen la característica j-ésima del factor A.

Con los valores calculados procedemos a postular el siguiente contraste de hipótesis:

$H_0 =$ El factor A es independiente del factor B.

Vs.

$H_1 =$ No es cierto H_0

Siendo el estadístico de prueba utilizado $x^2 = \sum_{i=1}^{n} \sum_{j=1}^{k} (X_{ij} - E_{ij})^2 / E_{ij}$ el cual se distribuye según una ji-cuadrado con (r-1)*(c-1) grados de libertad, se rechaza la hipótesis nula a favor de la alterna con (1-α)100% de confianza si $\chi^2 > \chi_{\alpha(n-1)(k-1)}$.

4.2.3 Componentes Principales.

Es una técnica estadística multivariada que permite la reducción de datos, si las definimos de una forma algebraica, son una particular combinación lineal de las p variables aleatorias observadas $X_1, X_2, \ldots\ldots, X_p$. En forma geométrica, esta combinación lineal representa la elección de un nuevo sistema de coordenadas obtenidas al rotar el sistema original, con $X_1, X_2, \ldots\ldots, X_p$ como los ejes coordenados. Los nuevos ejes representan la dirección de máxima variabilidad.

Sea $X^t = \left[X_1, X_2, \ldots., X_p\right]$ un vector aleatorio p-variado, y cada una de las variables que lo componen son variables aleatorias observables. El vector p-variado X tiene como matriz de varianzas y covarianzas a Σ y sea $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$ los valores propios correspondiente a Σ.

Considere las siguientes combinaciones lineales:

$$
\begin{aligned}
Y_1 &= a_1^t = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p \\
Y_2 &= a_2^t = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p \\
&\vdots \\
Y_p &= a_p^t = a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pp}X_p
\end{aligned}
$$

Además se puede demostrar que:

$\mathrm{Var}(Y_i) = a_i^t \sum a_i$ i = 1,2,…..,p

$\mathrm{Cov}(Y_i, Y_j) = a_i^t \sum a_j$ i,j = 1,2,…..,p

De esta manera $Y_1, Y_{2,\ldots,} Y_k$ son las componentes principales, las cuales no están correlacionadas, son ortonormales entre ellas y además se tiene que $\mathrm{Var}(Y_1) \geq \mathrm{Var}(Y_1) \geq \ldots \geq \mathrm{Var}(Y_p) \geq 0$. Estas componentes deben cumplir con:

$\|a_i\| = 1$ para i = 1,2,.....,p

y

$\langle a_i, a_j \rangle = 0$ para i ≠ j

Donde $\|a_i\|$ es la norma del vector a_i y $\langle a_i, a_j \rangle$ es el productor interno entre los vectores a_i y a_j.

La primera componente principal es la combinación lineal $Y_1 = a_1^t X$ de máxima varianza, esto es que maximiza la varianza de Y_1, sujeta que la norma del vector a_1 sea unitaria.

La segunda componente principal es la combinación lineal $Y_2 = a_2^t X$ que maximiza la varianza de Y_2, sujeta a que la norma del vector a_2 sea unitaria y a que $\mathrm{Cov}(Y_1, Y_2) = 0$.

En general la i-ésima componente principal es la combinación lineal que maximiza la varianza de $Y_i = a_i^t X$, sujeta a que la norma del vector a_i sea unitaria y a que $\mathrm{Cov}(Y_i, Y_k) = 0$ para k<1.

Como resultado obtenemos que Σ es la matriz de covarianzas asociada con el vector aleatorio $X^t = [X_1\ X_2 \cdots X_p]$, Σ tiene los pares de valores y vectores propios (λ_1, e_1), $(\lambda_2, e_2), \ldots, (\lambda_p, e_p)$ donde $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$. Entonces la i-ésima componente principal viene dada por.

$Y_i = e_i^t X = e_{i1} X_1 + e_{i2} X_2 + \cdots + e_{ip} X_p$ i = 1,2,.....,p

y

$Var(Y_i) = e_i^t \Sigma e_i = \lambda_i$ i = 1,2,.....,p

$Cov(Y_i, Y_j) = e_i^t \Sigma e_j = 0$ i ≠ k

El porcentaje total de la varianza contenida por i-ésima componente principal, o su explicación viene dada por:

$$\frac{\lambda_i}{\sum_{i=1}^{p} \lambda_i}$$

Para poder conocer, si en principio, es procedente o no aplicar Componentes principales se usa el criterio de esfericidad de Bartlett, el cuál basado sobre un supuesto de normalidad de las variables aleatorias en discusión, implica en estas una independencia al decir que las covarianzas entre ellas son 0; así la hipótesis se plantea de la siguiente manera:

$$H_0 : \Sigma = \begin{bmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \sigma_{pp} \end{bmatrix}$$

Vs.

$H_1 :$ No es verdad H_0

Para verificar este contraste de hipótesis se debe calcular los estadísticos de tal forma que: u y u'

Donde, $u = \dfrac{\det(S)}{s_{11}s_{22}\ldots s_{pp}} = \det(R);$ y

R = matriz de correlación muestral; S = matriz de covarianzas muestrales y S_{ii} = son las desviaciones estándar muestrales.

El estadístico de prueba es:

$u' = -\left[v - \frac{2p+5}{6}\right]\ln(u)$ donde $v = n-1$; y n = tamaño de la muestra.

Este estadístico de prueba es aproximadamente $\chi^2(f)$; donde $f = {p(p-1)}\big/{2}$, siendo f los grados de libertad de la distribución Ji-cuadrado y p es el número de variables investigadas.

Entonces, con un (1-α)100% de confianza de rechaza H_0 a favor de H_1 si: $u' \geq \chi^2_{\alpha, f}$.

El número de componentes principales, dependerá del porcentaje de varianza que se desea que estas expliquen.

4.3 Análisis Multivariado

4.3.1 Análisis de Correlación

La matriz de datos para el análisis de correlación se encuentra en el Anexo 3 y esta conformada por las preguntas que respondieron los estudiantes que utilizan o han utilizado una computadora.

$X_1 =$ Número de horas promedio por semana que utiliza el computador.

$X_2 =$ Número de horas promedio por semana que utiliza el computador con fines educativos.

$X_3 =$ Edad desde que utiliza el computador.

$X_4 =$ Trabajaría mejor en mis tareas escolares si pudiera usar una computadora.

$X_5 =$ La computación ha influido favorablemente en mi desempeño estudiantil.

$X_6 =$ En el futuro podré conseguir un buen trabajo si aprendo a usar la computadora.

$X_7 =$ Es muy importante que aprenda a usar una computadora.

$X_8 =$ Me siento a gusto realizando mis tareas escolares con una computadora.

$X_9 =$ Es muy cansado usar una computadora.

$X_{10} =$ Me tardo más tiempo en terminar el trabajo escolar cuando uso una computadora.

$X_{11} =$ Me asusta que mis compañeros se rían de mí si cometo un error en el computador.

$X_{12} =$ Las computadoras son difíciles de usar.

$X_{13} =$ Trato de usar lo menor posible una computadora para estudiar y realizar mis tareas escolares.

a) La correlación lineal más fuerte que existió fue de las variables: X_1 (Número de horas promedio por semana que utiliza el computador) y X_2 (Número de horas promedio por semana que utiliza el computador con fines educativos), con una correlación de 0.844. Es decir que entre mas horas usan los estudiantes el computador mayor es el número de horas que lo utilizan para realizar trabajos escolares.

b) Otra relación lineal importante que se hallo fue la de X_6 (En el futuro podré conseguir un buen trabajo si aprendo a usar la computadora) y X_7 (Es muy importante que aprenda a usar una computadora), esta correlación fue de 0.509.

c) De la misma manera se puede decir que existe una relación importante entre las variables X_4 (Trabajaría mejor en mis tareas escolares si pudiera usar una computadora) y X_8 (Me siento a gusto realizando mis tareas escolares con una computadora), esta correlación es de 0.500.

Existen además correlaciones que pueden ser tomadas en cuenta, detallaremos a continuación alguna de ellas:

d) La correlación existen entre la variable X_9 (Es muy cansado usar una computadora) y la variable X_{10} (Me tardo más tiempo en terminar el trabajo escolar cuando uso una computadora) es de 0.440, como se puede observar esta correlación no es muy baja.

e) La correlación entre las variables X_{11} (Me asusta que mis compañeros se rían de mí si cometo un error en el computador) y X_{12} (Las computadoras son difíciles de usar) fue de 0.409.

f) La correlación entre las variables X_5 (La computación ha influido favorablemente en mi desempeño estudiantil) y X_8 (Me siento a gusto realizando mis tareas escolares con una computadora) fue de 0.376.

g) La correlación entre las variables X_9 (Es muy cansado usar una computadora) y X_{12} (Las computadoras son difíciles de usar) fue de 0.362.

h) La correlación entre las variables X_{10} (Me tardo más tiempo en terminar el trabajo escolar cuando uso una computadora) y X_{12} (Las computadoras son difíciles de usar) fue de 0.354.

Las demás correlaciones fueron bajas y se las puede revisar en el Anexo 3, que es donde se encuentra la matriz de correlación.

4.3.2 Análisis de Tabla de Contingencia

Este análisis permite establecer si existe algún tipo de dependencia entre dos variables; a continuación se presentan algunos contrastes que se consideran importantes.

4.3.2.1 Género vs. Usuario de Internet

$H_0 =$ El que un estudiante use Internet es independiente del sexo que este tenga.

Vs.

$H_1 =$ No es verdad H_0

Tabla 4.1

Tabla de contingencia para Género y Usuario de Internet

	Género		**Total**
Usuario de Internet	**Femenino**	**Masculino**	
No	90	118	208
Si	78	146	224
Total	168	264	432

Fuente y elaboración: H. Ruíz

El valor del estadístico χ^2 es 3.239

Numero de grados de libertad: 1

Valor p de la prueba: 0.072

Como el valor de la prueba es mayor que 0.05, entonces existe evidencia estadística para no rechazar H_0, por lo que se puede concluir que no existe algún tipo de relación entre el género del estudiante y si este se considera usuario de Internet.

4.3.2.2 Es muy cansado usar una computadora vs. Las computadoras son difíciles de usar.

$H_0 =$ Es muy cansado usar una computadora es independiente con que las computadoras son difíciles de usar

Vs.

$H_1 =$ No es verdad H_0

La tabla de contingencia de esta hipótesis se encuentra en el Anexo 4.

El valor del estadístico χ^2 es 95.919

Numero de grados de libertad: 16

Valor p de la prueba: 0.000

Como el valor de la prueba es menor a 0.01, entonces existe evidencia estadística para rechazar H_0, por lo que se puede concluir que existe algún tipo de relación no necesariamente lineal entre es muy cansado usar una computadora y las computadoras son difíciles de usar.

4.3.2.3 Trabajaría mejor en mis tareas escolares si pudiera usar una computadora vs. La computación ha influido favorablemente en mi desempeño estudiantil.

$H_0 =$ Trabajaría mejor en mis tareas escolares si pudiera usar una computadora es independiente de La computación ha influido favorablemente en mi desempeño estudiantil.

Vs.

$H_1 =$ No es verdad H_0

Factor A: Trabajaría mejor en mis tareas escolares si pudiera usar una computadora.

Factor B: La computación ha influido favorablemente en mi desempeño estudiantil.

La tabla de contingencia de esta hipótesis se encuentra en el Anexo 5.

El valor del estadístico χ^2 es 184.921

Numero de grados de libertad: 16

Valor p de la prueba: 0.000

Como el valor de la prueba es menor a 0.01, entonces existe evidencia estadística para rechazar H_0, por lo que se puede concluir que existe algún tipo de relación no necesariamente lineal entre Trabajaría mejor en mis tareas escolares si pudiera usar una computadora y La computación ha influido favorablemente en mi desempeño estudiantil.

4.3.2.3 Edad desde que empezó a utilizar el computador vs. Las computadoras son difíciles de usar.

$\mathrm{H}_0 =$ La edad desde que empezó a utilizar el computador es independiente de Las computadoras son difíciles de usar.

Vs.

$\mathrm{H}_1 =$ No es verdad H_0

La tabla de contingencia de esta hipótesis se encuentra en el Anexo 5.

El valor del estadístico χ^2 es 57.58

Numero de grados de libertad: 8

Valor p de la prueba: 0.082

Como el valor de la prueba es mayor que 0.05, entonces existe evidencia estadística para no rechazar H_0, por lo que se puede concluir que no existe algún tipo de relación entre la edad desde que empezó a utilizar la computadora y las computadoras son difíciles de usar.

A continuación se presenta los resultados de esta prueba en una tabla global. Para ello a las variables que sus nombres son más largos las vamos a llamar de la siguiente manera para poder realizar la tabla de una mejor manera:

$X_1 =$ Número de horas promedio por semana que utiliza el computador.

$X_2 =$ Número de horas promedio por semana que utiliza el computador con fines educativos.

$X_3 =$ Edad desde que utiliza el computador.

$X_4 =$ Trabajaría mejor en mis tareas escolares si pudiera usar una computadora.

$X_5 =$ La computación ha influido favorablemente en mi desempeño estudiantil.

$X_6 =$ En el futuro podré conseguir un buen trabajo si aprendo a usar la computadora.

$X_7 =$ Es muy importante que aprenda a usar una computadora.

$X_8 =$ Me siento a gusto realizando mis tareas escolares con una computadora.

$X_9 =$ Es muy cansado usar una computadora.

$X_{10} =$ Me tardo más tiempo en terminar el trabajo escolar cuando uso una computadora.

$X_{11} =$ Me asusta que mis compañeros se rían de mí si cometo un error en el computador.

$X_{12} =$ Las computadoras son difíciles de usar.

$X_{13} =$ Trato de usar lo menor posible una computadora para estudiar y realizar mis tareas escolares.

Tabla 4.2
Resultados de los análisis de las tablas de contingencia

FACTOR A	FACTOR B	Valor p	Resultado
X_4	X_{12}	0.204	Independientes
X_{11}	X_{12}	0.000	Dependientes
X_9	X_{11}	0.000	Dependientes
X_3	X_7	0.988	Independientes
X_1	X_3	0.005	Dependientes
Genero	X_7	0.537	Independientes
X_4	X_8	0.000	Dependientes
Genero	Si posee e-mail	0.680	Independientes
X_{11}	X_{13}	0.000	Dependientes
X_{12}	X_{13}	0.000	Dependientes
X_9	X_{13}	0.000	Dependientes
X_3	X_{11}	0.518	Independientes
X_5	X_{12}	0.009	Independientes
X_7	X_{13}	0.051	Independientes

Fuente y elaboración: H. Ruíz

4.3.3 Análisis de Componentes Principales.

El análisis de Componentes Principales es una técnica de reducción de datos que tiene como objeto explicar las varianzas observables a través de unas pocas combinaciones lineales de las mismas.

La matriz de datos para el análisis de componentes principales está formada por 13 variables observables:

$X_1 =$ Número de horas promedio por semana que utiliza el computador.

$X_2 =$ Número de horas promedio por semana que utiliza el computador con fines educativos.

$X_3 =$ Edad desde que utiliza el computador.

$X_4 =$ Trabajaría mejor en mis tareas escolares si pudiera usar una computadora.

$X_5 =$ La computación ha influido favorablemente en mi desempeño estudiantil.

$X_6 =$ En el futuro podré conseguir un buen trabajo si aprendo a usar la computadora.

$X_7 =$ Es muy importante que aprenda a usar una computadora.

$X_8 =$ Me siento a gusto realizando mis tareas escolares con una computadora.

$X_9 =$ Es muy cansado usar una computadora.

$X_{10} =$ Me tardo más tiempo en terminar el trabajo escolar cuando uso una computadora.

X_{11} = Me asusta que mis compañeros se rían de mí si cometo un error en el computador.

X_{12} = Las computadoras son difíciles de usar.

X_{13} = Trato de usar lo menor posible una computadora para estudiar y realizar mis tareas escolares.

Para conocer si es procedente o no aplicar esta técnica de reducción de datos, se utilizará el contraste de Barlett, que bajo supuesto de normalidad propone el siguiente contraste de hipótesis:

$$H_0 : \Sigma = \begin{bmatrix} \sigma_{11} & 0 & \cdots & 0 \\ 0 & \sigma_{22} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \sigma_{pp} \end{bmatrix}$$

Vs.

H_1 : No es verdad H_0

Estadístico de prueba: 1480.071

Grados de libertad: 78

Valor p: 0.000

Como el valor p es menor que 0.01, existe evidencia estadística para rechazar H_0, por lo que es posible utilizar la técnica estadística de componentes principales.

La Tabla 4.3 nos muestra los resultados obtenidos de la matriz de datos originales, así pues el 67.11% de la variabilidad de las 13 variables observables que han sido utilizadas en este estudio es explicada por 5 componentes principales. En la misma tabla se presentan los valores propios de cada componente, el porcentaje de la varianza que explica cada componente y el porcentaje acumulado de varianza.

Tabla 4.3
Análisis de Componentes Principales: Datos originales

Componentes	Valor Propio	% de Varianza	% Acumulado
Y_1	2.828	21.750	21.750
Y_2	2.370	18.231	39.981
Y_3	1.641	12.621	52.602
Y_4	0.977	7.514	60.116
Y_5	0.908	6.984	67.101
Y_6	0.822	6.324	73.425
Y_7	0.664	5.108	78.533
Y_8	0.611	4.702	83.236
Y_9	0.585	4.498	87.733
Y_{10}	0.540	4.155	91.888
Y_{11}	0.459	3.532	95.420
Y_{12}	0.453	3.481	98.902
Y_{13}	0.143	1.098	100.00

Fuente y elaboración: H. Ruíz

En la Tabla 4.4, se presentan las cinco componentes principales que explican el 67.11% de la variabilidad de los datos originales con su respectivo valor propio.

Tabla 4.4
Análisis de Componentes Principales: Datos originales

	Y_1	Y_2	Y_3	Y_4	Y_5
X_1	0,598	-0,310	0,621	0,016	0,128
X_2	0,602	-0,234	0,672	-0,013	0,133
X_3	-0,365	0,209	-0,291	-0,264	0,593
X_4	0,368	0,655	0,042	-0,324	0,038
X_5	0,465	0,509	0,152	-0,050	-0,082
X_6	0,337	0,612	-0,178	0,416	0,026
X_7	0,318	0,567	-0,140	0,527	0,123
X_8	0,429	0,553	0,024	-0,356	-0,006
X_9	-0,503	0,118	0,459	0,208	0,312
X_{10}	-0,520	0,217	0,392	0,244	0,202
X_{11}	-0,433	0,441	0,316	-0,272	-0,260
X_{12}	-0,544	0,412	0,333	-0,140	0,107
X_{13}	-0,465	0,243	0,261	0,158	-0,530

Fuente y elaboración: H. Ruíz

En la primera componente principal, la variable que más aporta es el número de horas con fines educativos que usted dedica con un peso de 0.602, seguida de la variable número de horas que utiliza el computador a la semana con un peso de 0.598. Pero vale recalcar que existen variables que tienen un peso importante para esta componente principal, como son las variables X_8 , X_9 y X_{12}.

En la segunda componente principal, la variable que mas peso tiene es trabajaría mejor en mis tareas escolares si pudiera usar una computadora con un valor de 0.655, la siguiente variable que tiene mas peso es en el futuro podré conseguir un buen trabajo si aprendo a usar la computadora. Como se puede apreciar que existen variables que tienen un gran peso también para esta variable y son las variables X_5, X_7 y X_8. Podemos entonces decir que esta componente principal explicaría la incidencia de la computación en los colegios fiscales de Guayaquil.

En la tercera componente principal, la variable que más aporta es el número de horas con fines educativos que usted dedica con un peso de 0.672, seguida de la variable número de horas que utiliza el computador a la semana con un peso de 6.21. Estas dos variables son las que mas aportan por lo que esta componente principal explicaría lo relacionado con este par de variables.

En la cuarta componente principal, la variable que mas aporta esta dada por es muy importante que aprenda a utilizar el computador con un peso de 0.527, seguido de la variable en el futuro podré conseguir un buen trabajo si aprendo a usar una computadora con un peso de 0.416.

Y por ultimo en la última componente principal, la variable que más pero tiene es la edad desde que el estudiante utiliza el computador con un peso de 0.593, seguido por la variable trato de usar lo menor posible una computadora para estudiar y realizar mis tareas escolares. Esta variable explicaría a este par de variables.

Como en este estudio de las 13 variables 10 de ellas están en la misma escala, por lo que no vamos a realizar el estudio de las componentes principales con los datos estandarizados, esto se lo realiza cuando existen muchas variables con diversas escales, ya que las variables que tienen mayor escala absorben mayor peso.

A continuación en la Tabla 4.5 presentamos el estudio de las componentes principales utilizando el método de Varimax para así distribuir la varianza a lo largo de las componentes, utilizando los datos originales.

Tabla 4.5
Análisis de Componentes Principales: Varimax

	Y_1	Y_2	Y_3	Y_4	Y_5
X_1	-0,090	0,917	0,055	-0,048	-0,019
X_2	-0,036	0,930	0,133	-0,039	-0,009
X_3	0,298	-0,421	0,089	-0,084	-0,634
X_4	0,009	0,008	0,794	0,196	-0,070
X_5	-0,051	0,201	0,588	0,329	0,108
X_6	-0,042	-0,078	0,297	0,772	0,031
X_7	0,017	-0,019	0,191	0,836	-0,026
X_8	-0,095	0,053	0,765	0,137	-0,056
X_9	0,754	0,076	-0,200	0,019	-0,049
X_{10}	0,732	-0,028	-0,157	0,089	0,051
X_{11}	0,515	-0,233	0,360	-0,216	0,358
X_{12}	0,714	-0,203	0,196	-0,119	0,059
X_{13}	0,390	-0,227	-0,029	-0,020	0,667

Fuente y elaboración: H. Ruíz

Los coeficientes de las componentes principales que poseen mayor peso utilizando el método Varimax son:

Primera componente principal: El mayor peso lo posee la variable es muy canso usar una computadora, esta componente explica la frustración de los estudiantes frente a las computadoras.

Segunda componente principal: El mayor peso lo tiene la variable número de horas dedicadas a fines educativos, esta componente explica el número de horas a la semana que utiliza el computador y el número de horas que lo utiliza con fines educativos.

Tercera componente principal: El mayor peso lo tiene la variable trabajaría mejor en mis tareas escolares si pudiera usar una computadora, esta componente podría explicar la incidencia de la computación en los estudiantes.

Cuarta componente principal: El mayor pero lo tiene la variable es muy importante que aprenda a usar una computadora.

Quinta componente principal: El mayor peso lo tiene la variable trato de usar lo menos posible una computadora para estudiar y realizar las tareas escolares.