CAPITULO 5

1. **ANÁLISIS MULTIVARIADO**
   1. **Técnicas para el Análisis Multivariado**

Para el análisis multivariado de cada una de las variables estudiadas, vamos a definir las técnicas que se utilizarán.

En este caso procederemos a definir lo que es una Tabla de Contingencia, Análisis de Correspondencia (ANACOR) y Análisis de Homogeneidad (HOMALS).

* + 1. **Análisis de las Tablas de Contingencia**

El Análisis de Contingencia es una técnica que permite determinar si dos variables o factores son independientes, para ello se construyen tablas de *r* filas y *c* columnas, en donde *c* es el número de niveles del *factor 1* y *r* es el número de niveles del *factor 2*.

A continuación se detalla un modelo de una tabla de contingencia:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Factor 1 | | | | |  |
| Factor 2 | Nivel 1 | Nivel 2 |  |  | Nivel c |  |
| Nivel 1 | X11 | X12 |  |  | X1c | X1. |
| E11 | E12 |  |  | E1c | E1. |
| Nivel 2 | X21 | X22 |  |  | X2c | X2. |
| E21 | E22 |  |  | E2c | E2. |
|  |  |  |  |  |  |  |
| Nivel r | Xr1 | Xr2 |  |  | Xrc | Xr. |
| Er1 | Er2 |  |  | Erc | Er. |
|  | X.1 | X.2 |  |  | X.c |  |
|  | E.1 | E.2 |  |  | E.c | X.. |

Donde Xij es la frecuencia observada de unidades de investigación sometidas al i-ésimo nivel del factor 2 y el j-ésimo nivel del factor 1.

y **** 

Xi. = es la suma de las frecuencias observadas en el i-ésimo reglón

X.j = es la suma de las frecuencias observadas en el j-ésimo columna

X.. = es la suma de todas las frecuencias observadas

Eij = es el número esperado de la frecuencia de la celda en el i-ésimo reglón y la j-ésima columna.

Ei. = es valor esperado para el i-ésimo reglón.

E.j = es valor esperado para la j-ésima columna.

Para el análisis de las tablas de contingencia se postula el siguiente contraste de hipótesis:

: Las variables 1 y 2 son independientes

Vs

: Existe dependencia entre las variables

El estadístico de prueba que se utilizará es:



Donde :

; 

Se probará que el estadístico  tiene una distribución Chi-Cuadrado con (r-1)(c-1) grados de libertad. <

De donde con (1-)100% de confianza se rechaza  a favor de si:

>

* + 1. **Análisis de Correspondencia (ANACOR)**

La finalidad del Análisis de correspondencia (ANACOR) es determinar la las relaciones existentes entre dos variables nominales, recogidas en una tabla de correspondencias, sobre un espacio de pocas dimensiones, mientras que al mismo tiempo se describen las relaciones entre las categorías de cada variable.

Para cada variable, las distancias sobre un gráfico entre los puntos de categorías reflejan las relaciones entre las categorías, con las categorías similares representadas próximas unas a otras. La proyección de los puntos de una variable sobre el vector desde el origen hasta un punto de categoría de la otra variable describe la relación entre ambas variables.

El análisis de las tablas de contingencia incluye a menudo los perfiles de fila y columna, así como contrastar la independencia a través del estadístico Chi-cuadrado. Sin embargo, el número de perfiles puede ser grande y la prueba Chi-cuadrado no determinaría la dependencia.

El análisis de correspondencia calculará, medidas de correspondencia, perfiles de fila y columna, valores propios, puntuaciones de fila y columna, inercia, estadísticos de confianza para valores propios, gráficos de transformación y diagramas de dispersión biespaciales, de las diversas filas y columnas de la tabla y además nos permitirá analizar la posible relación entre las variables, las diversas categorías de las variables estarán representadas en el gráfico más próximas o alejadas en las diversas dimensiones en función de su grado de similitud o diferencia.

* + 1. **Análisis de Homogeneidad (HOMALS)**

El análisis de homogeneidad, que se denomina HOMALS por: *HOM*ogeneity Análisis by jeans of Alternating *L*east *S*quares o también conocido como Análisis de Correspondencia Múltiple.

El HOMALS lo constituye una matriz de datos en la que las filas están formadas por objetos de diferente índole y las columnas por variables que en principio son tratadas como medidas en una escala nominal, y por tanto, con categorías meramente diferentes unas de otras.

Este procedimiento utiliza en la estimación de sus parámetros la alternancia de mínimos cuadrados. La estimación de mínimos cuadrados de los valores de los parámetros se realiza en dos fases, una de estimación del modelo propiamente dicha y otra de escalamiento óptimo. Estas dos fases se alternan iterativamente hasta conseguir una convergencia determinada. En la fase de escalamiento óptimo las variables son analizadas de acuerdo con una escala de medida especificada previamente.

La finalidad del análisis de homogeneidad (HOMALS), es conseguir cuantificaciones de los objetos y por tanto de las categorías de las variables que sean óptimas, en el sentido de que las categorías estén separadas unas de otras de la dimensión o dimensiones estudiadas tanto como sea posible, y a su vez dentro de cada categoría los sujetos estén lo más próximos unos a otros, es decir con puntuaciones cuanto más homogéneas entre sí.

Si el análisis de homogeneidad se realiza exclusivamente con dos variables, equivale a un análisis de correspondencia simple y aunque los resultados no sean exactamente los mismos, la interpretación con toda seguridad no va a diferir en ambos casos.

* 1. **Análisis Multivariado de las variables**
     1. **Análisis de Tablas de Contingencia para las variables de interés**

En esta sección se analizarán dos variables conjuntamente (X y Y), y a través de la prueba de independencia utilizando la distribución Chi-cuadrado analizaremos la dependencia entre estas variables que se consideran de importancia.

La mayoría de estas variables fueron recodificadas debido a que la validez de la prueba  habían muchas casillas cuya frecuencia esperada era menor a 5; las variables no independientes se les realizará un análisis de Homogeneidad.

* ***Sexo del estudiante vs. Tipo de música que les gusta escuchar***

: Sexo del estudiante es independiente al Tipo de música que le gusta escuchar

Vs

: Sexo y Tipo de música que le gusta escuchar no son independientes

En la Tabla 127 podemos observar que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, son 13 los que optan por la música clásica, donde la mayoría son mujeres. Se puede ver una notable diferencia en los gustos por la denominada música vieja, entre hombres y mujeres, ya que 14 de 18 son del sexo masculino.

El tipo de música de *Baladas* tuvo la preferencia entre los estudiantes ya que fue el más escogido, 23,5% (55 estudiantes) de donde 36 fueron del sexo femenino y 19 del masculino.

**GRAFICO 5.1**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



**TABLA 127**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



Como se observa en la Tabla 127, el p-valor es de 0,001, es decir que a nivel de significación de 0,05 se rechazará la hipótesis nula a favor de la alterna. Por lo tanto el tipo de música que escucha el estudiante depende del sexo.

A continuación se realizará el Análisis de Homogeneidad.

**TABLA 127A**



*Fuente y Elaboración:* *Lissette Flores Barcia*



**Frecuencias marginales**

Los autovalores representan la proporción de información categórica explicada por cada una de las dimensiones. En este caso, los autovalores de las dimensiones 1 y 2 son 0,663 y 0,500 respectivamente. El valor máximo para cada dimensión es igual a la unidad. Ver en la Tabla 127B.



**TABLA 127B**

*Fuente y Elaboración:* *Lissette Flores Barcia*



**TABLA 127C**

*Fuente y Elaboración:* *Lissette Flores Barcia*

En la Tabla 127C tenemos los valores de las medidas de discriminación para cada variable en cada dimensión y que equivalen a la variancia de las variables cuantificadas. Su valor máximo es igual a la unidad, que se consigue si todas las puntuaciones de los sujetos caen en grupos mutuamente excluyentes y a su vez dentro de cada grupo estas puntuaciones son idénticas.

En cuanto al gráfico de la Tabla 127C, confirma que la variable Sexo discrimina con la primera dimensión, mientras que la variable MSK\_PREF discrimina en ambas dimensiones.

En el caso de que estuviera alguna variable próxima al origen, no discriminaría en ninguna de las dos dimensiones.

En la Tabla 127D podemos ver el análisis de las frecuencias marginales y cuantificaciones de categorías para cada variable.

Podemos decir que -0,85 en la variable sexo, para la categoría 1, es la media de las puntuaciones en esta dimensión de las 112 mujeres de la población. La media de las puntuaciones en la dimensión 2 para la misma categoría 1, es de 0,001.

Para la Categoría 2, la media de las puntuaciones para la dimensión 1 es de 0,780, de los 122 hombres de la población. Así mismo para la dimensión 2 la media de las puntuaciones es de -0,001.

Para la variable MSK\_PREF, vemos que se detallan las medias de las puntuaciones de cada categoría y dimensión. Ver Tabla 127D.

La finalidad del HOMALS es encontrar cuantificaciones que sean óptimas en el sentido de que los centroides (medias) de cada categoría estén tan separados uno de otro como sean posibles; es decir, conseguir la máxima homogeneidad entre los sujetos dentro de cada categoría y en cambio que éstas sean lo más heterogéneas posibles, lo que se traducirá en una mayor separación en el gráfico.



**Frecuencias Marginales y Cuantificación de Categorías**

*Fuente y Elaboración:* *Lissette Flores Barcia*

**TABLA 127D**





En el gráfico de la TABLA D vemos las cuantificaciones, y la relación entre las variables. Notamos que sexo masculino está más relacionado con la música vieja y la música rock; mientras que el sexo femenino se relaciona más con las baladas y el regué.

El Gráfico 5.2 agrupa los individuos de la muestra de acuerdo con sus puntuaciones en ambas dimensiones.



**GRAFICO 5.2**

*Fuente y Elaboración:* *Lissette Flores Barcia*

Los girasoles son concentraciones de individuos y cada pétalo es un número determinado de individuos.

* ***Sexo del estudiante y Tipo de Programa de TV que prefiere ver***

: Sexo del estudiante es independiente al Tipo de programa que prefiere ver

Vs

: Sexo y Tipo de programa que prefiere ver no son independientes

Podemos ver en la Tabla 128 que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, son 53 los que optan por los programas de humor; mientras que los noticieros solamente optan 8 estudiantes.

Los programas culturales y deportivos han sido uno de los más escogidos, 60 estudiantes, donde 22 son del sexo femenino y 38 del sexo masculino.

Los programas musicales lo prefirieron 27 estudiantes, siendo 15 del sexo femenino y 12 del sexo masculino.

**TABLA 128**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



En la Tabla 128, el p-valor es de 0,001, es decir que a nivel de significación de 0,05 se va a rechazar la hipótesis nula de independencia entre estas variables.

Por lo tanto el tipo de programa que ve el estudiante con frecuencia depende del sexo del estudiante.

A continuación se realizará el Análisis de Homogeneidad.



*Fuente y Elaboración:* *Lissette Flores Barcia*



**TABLA 128A**

**Frecuencias marginales**

Los autovalores representan la proporción de información categórica explicada por cada una de las dimensiones. En este caso, los autovalores de las dimensiones 1 y 2 son 0,653 y 0,500 respectivamente. El valor máximo para cada dimensión es igual a la unidad.. Ver en la Tabla 128B.

**TABLA 128B**

*Fuente y Elaboración:* *Lissette Flores Barcia*



En la Tabla 128C tenemos los valores de las medidas de discriminación para cada variable en cada dimensión y que equivalen a la variancia de las variables cuantificadas. Su valor máximo es igual a la unidad, que se consigue si todas las puntuaciones de los sujetos caen en grupos mutuamente excluyentes y a su vez dentro de cada grupo estas puntuaciones son idénticas.



**TABLA 20C**

*Fuente y Elaboración:* *Lissette Flores Barcia*

En cuanto al gráfico de la Tabla 128C, confirma que la variable Sexo está relacionada con la primera dimensión, mientras que la variable PRG\_PREF discrimina en ambas dimensiones.

En el caso de que estuviera alguna variable próxima al origen, no discriminaría en ninguna de las dos dimensiones.

**Frecuencias Marginales y Cuantificación de Categorías**

*Fuente y Elaboración:* *Lissette Flores Barcia*

**TABLA 128D**



En la Tabla 128D podemos ver el análisis de las frecuencias marginales y cuantificaciones de categorías para cada variable.

Podemos decir que -0,844 en la variable sexo, para la categoría 1, es la media de las puntuaciones en esta dimensión de las 112 mujeres de la población. La media de las puntuaciones en la dimensión 2 para la misma categoría 1, es de -0,001.

Para la Categoría 2, la media de las puntuaciones para la dimensión 1 es de 0,775, de los 122 hombres de la población. Así mismo para la dimensión 2 la media de las puntuaciones es de 0,001.

Para la variable PRG\_PREF, vemos que se detallan las medias de las puntuaciones de cada categoría y dimensión. Ver Tabla 128D.

En el gráfico de la TABLA 128D vemos las cuantificaciones, y la relación entre las variables. Notamos que sexo masculino está más relacionado con los programas culturales y deportivos, y noticieros y periodísticos; mientras que el sexo femenino está más relacionado con los programas musicales y de concursos/variados.

El Gráfico 5.3 agrupa los individuos de la muestra de acuerdo con sus puntuaciones en ambas dimensiones.

**GRAFICO 5.3**

*Fuente y Elaboración:* *Lissette Flores Barcia*



Los girasoles representan la concentración de individuos y cada pétalo es un número determinado de individuos.

* ***Sexo del estudiante y Sección del periódico que lee con frecuencia***

: Sexo del estudiante es independiente a la sección del Periódico que lee con frecuencia

Vs

: Sexo y Sección del Periódico que lee con frecuencia no son independientes

En la Tabla 129 notamos que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, son 58 los que prefirieron la sección de Política, donde 25 son del sexo femenino y 33 son del sexo masculino; mientras que la sección Internacional solamente la prefieren 11 estudiantes, donde 3 son del sexo femenino y 8 del sexo masculino.

La sección Social la prefirieron 46 estudiantes, donde 36 fueron del sexo femenino y 10 del sexo masculino.

La sección Local tuvo 24 estudiantes, donde 15 fueron del sexo femenino y 9 del masculino.

La sección deportiva fue más escogida por los del sexo masculino, siendo una cantidad de 39, mientras que del sexo femenino solo hubieron 2.

Para los Avisos clasificados fuero 26 estudiantes los que escogieron esta opción, donde 16 fueron del sexo femenino y 20 del sexo masculino.

**TABLA 129**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



Vemos en la Tabla 129 que el p-valor es de 0,00, lo que vemos que a nivel de significación de 0,05 se rechazará la hipótesis nula de independencia entre estas variables. Por lo tanto la sección del periódico que el estudiante lee con frecuencia depende del sexo del estudiante.

**GRAFICO 5.4**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



A continuación se realizará el Análisis de Homogeneidad.

**TABLA 129A**



*Fuente y Elaboración:* *Lissette Flores Barcia*



**Frecuencias marginales**

Los autovalores representan la proporción de información categórica explicada por cada una de las dimensiones. En este caso, los autovalores de las dimensiones 1 y 2 son 0,740 y 0,500 respectivamente. El valor máximo para cada dimensión es igual a la unidad.. Ver en la Tabla 129B.

**TABLA 129B**

*Fuente y Elaboración:* *Lissette Flores Barcia*



En la Tabla 129C tenemos los valores de las medidas de discriminación para cada variable en cada dimensión y que equivalen a la variancia de las variables cuantificadas. Su valor máximo es igual a la unidad, que se consigue si todas las puntuaciones de los sujetos caen en grupos mutuamente excluyentes y a su vez dentro de cada grupo estas puntuaciones son idénticas.

En cuanto al gráfico de la Tabla 129C, confirma que la variable Sexo está relacionada con la primera dimensión, mientras que la variable SECC\_PRE discrimina en ambas dimensiones.

En el caso de que estuviera alguna variable próxima al origen, no discriminaría en ninguna de las dos dimensiones.

******

**TABLA 129C**

*Fuente y Elaboración:* *Lissette Flores Barcia*



En la Tabla 129D podemos ver el análisis de las frecuencias marginales y cuantificaciones de categorías para cada variable.

Podemos decir que -0,898 en la variable sexo, para la categoría 1, es la media de las puntuaciones en esta dimensión de las 112 mujeres de la población. La media de las puntuaciones en la dimensión 2 para la misma categoría 1, es de 0,00.

Para la Categoría 2, la media de las puntuaciones para la dimensión 1 es de 0,824, de los 122 hombres de la población. Así mismo para la dimensión 2 la media de las puntuaciones es de 0,00.

Para la variable SECC\_PRE, vemos que se detallan las medias de las puntuaciones de cada categoría y dimensión. Ver Tabla 129D.

En el gráfico de la TABLA D vemos las cuantificaciones, y la relación entre las variables. Notamos que sexo masculino está más relacionado con la sección de política y deportiva del periódico; mientras que el sexo femenino esta más relacionado con la sección local, social y económica.

**Frecuencias Marginales y Cuantificación de Categorías**

*Fuente y Elaboración:* *Lissette Flores Barcia*

**TABLA 129D**







**GRAFICO 5.5**

*Fuente y Elaboración:* *Lissette Flores Barcia*



Los girasoles son concentraciones de individuos y cada pétalo es un número determinado de individuos.

* ***Sexo del estudiante y Compra Joyas***

: Sexo del estudiante es independiente a Comprar Joyas

Vs

: Sexo y Comprar Joyas no son independientes

En la Tabla 130 vemos que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, 117 nunca compran joyas, donde 27 son del sexo femenino y 90 son del sexo masculino; mientras que 85 rara vez compran joyas, donde 55 son del sexo femenino y 30 del sexo masculino.

Vemos que 32 estudiantes compran joyas la mayoría de las veces, donde solamente 2 son del sexo masculino y 30 del femenino.

**TABLA 130**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



**GRAFICO 5.6**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

En la Tabla 130 notamos que el p-valor es de 0,00, es decir que a nivel de significación de 0,05 se va a rechazar la hipótesis nula de independencia entre estas variables. Por lo tanto la compra de joyas depende del sexo del estudiante.

A continuación se realizará el Análisis de Homogeneidad.



*Fuente y Elaboración:* *Lissette Flores Barcia*



**TABLA 130A**

**Frecuencias marginales**

Los autovalores representan la proporción de información categórica explicada por cada una de las dimensiones. En este caso, los autovalores de las dimensiones 1 y 2 son 0,764 y 0,500 respectivamente. El valor máximo para cada dimensión es igual a la unidad. Ver en la Tabla 130B.

**TABLA 130B**

*Fuente y Elaboración:* *Lissette Flores Barcia*



**TABLA 130C**

*Fuente y Elaboración:* *Lissette Flores Barcia*



En la Tabla 130C tenemos los valores de las medidas de discriminación para cada variable en cada dimensión y que equivalen a la variancia de las variables cuantificadas. Su valor máximo es igual a la unidad, que se consigue si todas las puntuaciones de los sujetos caen en grupos mutuamente excluyentes y a su vez dentro de cada grupo estas puntuaciones son idénticas.

En cuanto al gráfico de la Tabla 130C, confirma que la variable Sexo está relacionada con la primera dimensión, mientras que la variable JOYAS discrimina en ambas dimensiones.

En el caso de que estuviera alguna variable próxima al origen, no discriminaría en ninguna de las dos dimensiones.

En la Tabla 130D podemos ver el análisis de las frecuencias marginales y cuantificaciones de categorías para cada variable.

Podemos decir que -0,913 en la variable sexo, para la categoría 1, es la media de las puntuaciones en esta dimensión de las 112 mujeres de la población. La media de las puntuaciones en la dimensión 2 para la misma categoría 1, es de -0,001.

Para la Categoría 2, la media de las puntuaciones para la dimensión 1 es de 0,838, de los 122 hombres de la población. Así mismo para la dimensión 2 la media de las puntuaciones es de 0,001.

**Frecuencias Marginales y Cuantificación de Categorías**

*Fuente y Elaboración:* *Lissette Flores Barcia*

**TABLA 130D**







Para la variable JOYAS, vemos que se detallan las medias de las puntuaciones de cada categoría y dimensión. Ver Tabla 130D.

En el gráfico de la TABLA D vemos las cuantificaciones, y la relación entre las variables. El sexo masculino está más relacionado con nunca comprar joyas.

**GRAFICO 5.7**

*Fuente y Elaboración:* *Lissette Flores Barcia*



El Gráfico 5.7 agrupa los individuos de la muestra de acuerdo con sus puntuaciones en ambas dimensiones.

Los girasoles son concentraciones de individuos y cada pétalo es un número determinado de individuos.

* ***Sexo del estudiante y Compra Ropa de diario***

: Sexo es independiente a Comprar Ropa de diario

Vs

: Sexo y Comprar Ropa de diario no son independientes

Observamos en la Tabla 131, que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, 91 nunca o rara vez compran ropa de diario, donde 37 son del sexo femenino y 54 son del sexo masculino; mientras que 120 algunas veces compran ropa de diario, donde 32 son del sexo femenino y 58 del sexo masculino.

**TABLA 131**



*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

Vemos que 23 estudiantes compran ropa de diario la mayoría de las veces, 13 del sexo femenino y 10 del sexo masculino.



*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

**GRAFICO 5.8**



**TABLA 131A**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

En la Tabla 131A notamos que el p-valor es de 0,194, lo que vemos que a nivel de significación de 0,05, se va a aceptar la hipótesis nula de independencia entre estas variables. Por lo tanto que compre ropa de diario no depende del sexo del estudiante.

* ***Sexo del estudiante y Compra Discos, cassettes***

: Sexo del estudiante es independiente a Comprar Discos, cassettes

Vs

: Sexo y Comprar Discos, cassettes no son independientes

En la Tabla 132, vemos que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, 16 nunca o rara vez compran discos y cassettes, donde 9 son del sexo femenino y 7 son del sexo masculino; 102 estudiantes de 234 algunas veces compran discos y cassettes, donde 52 son de sexo femenino y 50 del sexo masculino.

Vemos que 116 estudiantes compran discos y cassettes la mayoría de las veces, 51 del sexo femenino y 65 del sexo masculino.



**TABLA 132**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

******

En la Tabla 132 notamos que el p-valor es de 0,460, lo que vemos que a nivel de significación de 0,05, se debe de aceptar la hipótesis nula de independencia entre estas variables. Por lo tanto comprar discos o cassettes no depende del sexo del estudiante.

* ***Sexo del estudiante y Compra artículos deportivos***

: Sexo del estudiante es independiente a Comprar artículos deportivos

Vs

: Sexo y Comprar artículos deportivos no son independientes

En la Tabla 133, observamos que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, 53 nunca compran artículos deportivos, donde 32 son del sexo femenino y 21 son del sexo masculino; 77 rara vez compran artículos deportivos, donde 38 son del sexo femenino y 39 del sexo masculino.

**TABLA 133**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



Vemos que 61 estudiantes de 234 algunas veces compran artículos deportivos, donde 26 son de sexo femenino y 35 del sexo masculino. Solo 14 estudiantes compran artículos deportivos la mayoría de las veces, 4 del sexo femenino y 10 del sexo masculino. 29 estudiantes siempre compra artículos deportivos, 12 del sexo femenino y 17 del sexo masculino.



**GRAFICO 5.9**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

**TABLA 133A**

En la Tabla 133A notamos que el p-valor es de 0,156, lo que vemos que es mayor que 0,05, es decir que a nivel de significación, se debe de aceptar la hipótesis nula de independencia entre estas variables. Por lo tanto comprar artículos deportivos no depende del sexo del estudiante.

* ***Sexo del estudiante y No debería haber tanto énfasis en el sexo hoy***

: Sexo del estudiante es independiente a No debería haber tanto énfasis en el sexo hoy

Vs

: Sexo y No debería haber tanto énfasis en el sexo hoy no son independientes

En la Tabla 26, observamos que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, 29 están en total desacuerdo en que no debería haber tanto énfasis en el sexo hoy, donde 12 son del sexo femenino y 17 son del sexo masculino; 33 están en parcial desacuerdo en que no debería haber tanto énfasis en el sexo hoy, donde 13 son del sexo femenino y 20 del sexo masculino. A 65 estudiantes de 234 les es indiferente en que no debería haber tanto énfasis en el sexo hoy, donde 27 son de sexo femenino y 38 del sexo masculino. Vemos que 60 estudiantes están en parcial acuerdo en que no debería haber tanto énfasis en el sexo hoy, 34 del sexo femenino y 26 del sexo masculino. A 47 estudiantes están en total acuerdo en que no debería haber tanto énfasis en el sexo hoy, donde 26 son del sexo femenino y 21 del sexo masculino.



**GRAFICO 5.10**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

**TABLA 134**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



En la Tabla 134 notamos que el p-valor es de 0,25, lo que vemos que es mayor que 0,05, es decir que a nivel de significación, se debe de aceptar la hipótesis nula de independencia entre estas variables. Por lo tanto que piensen que haya mucho énfasis en el sexo hoy no depende del sexo d los estudiantes.

* ***Sexo del estudiante y Me gustan los juegos de azar***

: Sexo del estudiante es independiente a Me gustan los juegos de azar

Vs

: Sexo y Me gustan los juegos de azar no son independientes

Podemos ver en la Tabla 135, que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, 53 están en total desacuerdo en que les gusta los juegos de azar, donde 27 son del sexo femenino y 26 son del sexo masculino; 26 están en parcial desacuerdo en que les gusta los juegos de azar, donde 12 son del sexo femenino y 14 del sexo masculino. 70 estudiantes de 234 les es indiferente en que les gusta los juegos de azar, donde 39 son de sexo femenino y 31 del sexo masculino.

Vemos que 51 estudiantes están en parcial acuerdo en que les gusta los juegos de azar, 21 del sexo femenino y 30 del sexo masculino. 34 estudiantes están en total acuerdo en que les gusta los juegos de azar, donde 13 son del sexo femenino y 21 del sexo masculino.

**TABLA 135**



*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



**GRAFICO 5.11**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



**TABLA 135A**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

Vemos en la Tabla 135A que el p-valor es de 0,388, es decir que a nivel de significación de 0,05, se debe de aceptar la hipótesis nula de independencia entre estas variables. Por lo tanto que les guste los juegos de azar no depende del sexo de los estudiantes.

* ***Sexo del estudiante y Veo TV los fines de semana***

: Sexo del estudiante es independiente a Veo TV los fines de semana

Vs

: Sexo y Veo TV los fines de semana no son independientes

Podemos ver en la Tabla 136, que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, 39 están en total o parcial desacuerdo en que ven TV los fines de semana, donde 19 son del sexo femenino y 20 son del sexo masculino; a 34 estudiantes de 234 les es indiferente en que ven TV los fines de semana, donde 14 son de sexo femenino y 20 del sexo masculino. Vemos que 58 estudiantes están en parcial acuerdo en que ven TV los fines de semana, 30 del sexo femenino y 28 del sexo masculino. 103 estudiantes están en total acuerdo en que ven TV los fines de semana, donde 49 son del sexo femenino y 54 del sexo masculino.

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

**TABLA 136**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



Vemos en la Tabla 28 que el p-valor es de 0,808, es mayor que 0,05 que es el nivel de significación, por lo que se debe de aceptar la hipótesis nula de independencia entre estas variables. Por lo tanto que vea TV los fines de semana no depende del sexo de los estudiantes.

**GRAFICO 5.12**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



* ***Edad del estudiante y Tipo de música que les gusta escuchar***

: Edad del estudiante es independiente a Tipo de música que les gusta escuchar

Vs

: Edad y Tipo de música que les gusta escuchar no son independientes

En la Tabla 137 podemos observar que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, son 31 los que optan por la música vieja, donde 8 son de edades de entre 17 a 19 años, 7 de 20 a 22 años y 16 de 23 años en adelante.

El tipo de música de Baladas tuvo un número significativo, ya que fueron 55 estudiantes, donde 11 fueron estudiantes de 17 a19 años, 16 estudiantes de 20 a 22 años, 28 estudiantes de 23 años en adelante.

**TABLA 137**

*Fuente y Elaboración: Lissette Flores Barcia*



La mayoría fue para el tipo de música bailable, que son la salsa, el regué, etc., donde 15 estudiantes de 17 a 19 años optaron por este tipo de música, 29 fueron estudiantes de 20 a 23 años y 32 de 23 años o más.

Vemos en la Tabla 137 que el p-valor es de 0,399, donde a nivel de significación de 0,05 se debe de aceptar la hipótesis nula. Por lo tanto el tipo de música que les gusta escuchar no depende de la edad de los estudiantes

**GRAFICO 5.13**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



* ***Edad del estudiante y Tipo de Programa que prefiere ver***

: Edad del estudiante es independiente a Tipo de programa que prefiere ver

Vs

: Edad y Tipo de programa que prefiere ver no son independientes

Podemos ver en la Tabla 138 que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, son 53 los que optan por los programas de humor, donde 14 son estudiantes de entre 17 y 19 años, 19 de entre 20 y 22 años, 18 de entre 23 y 25 años y solo 2 de estudiantes de más de 26 años; los noticieros solamente lo prefieren 8 estudiantes, donde 1 es de 17 a 19 años.

Los programas culturales han sido uno de los más escogidos, 42 estudiantes, donde 9 estudiantes fueron de edades entre 17 y 19 años, 10 de entre 20 y 22 años, 18 de entre 23 y 25 años, y 5 estudiantes de 26 años o más.

Los programas musicales también tuvieron acogida, donde lo prefirieron 27 estudiantes, siendo 10 los estudiantes de edades comprendidas entre 17 y 19 años, 10 estudiantes de 20 y 22 años, y 2 estudiantes de 23 y 25 años de edad que prefieren este tipo de programa.

**TABLA 138**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



Vemos en la Tabla 30 que el p-valor es de 0,343, donde a nivel de significación de 0,05 se debe de aceptar la hipótesis nula.

Por lo tanto el tipo de programas que prefieren ver no depende de la edad de los estudiantes.



**GRAFICO 5.14**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

* ***Edad del estudiante y Sección del periódico que lee con frecuencia***

: Edad del estudiante es independiente a Sección del periódico que lee con frecuencia

Vs

: Edad y Sección del periódico que lee con frecuencia no son independientes

En la Tabla 31 que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, son 58 los que prefirieron la sección de Política, donde 24 fueron estudiantes de entre 23 y 25 años de edad y solo 7 estudiantes de edades entre 17 y 19 años.; la sección Internacional solamente la prefieren 11 estudiantes, donde 5 fueron estudiantes de entre 23 y 25 años de edad y solo 2 estudiantes de edades entre 17 y 19 años.

La sección Social la prefirieron 46 estudiantes, donde 18 fueron estudiantes de edades entre 17 y 19 años, 15 estudiantes de entre 23 y 25 años y tan solo 2 de 26 años o más.

La sección Local tuvo 24 estudiantes, donde 10 fueron estudiantes de edades entre 20 y 22 años, y solo 2 estudiantes de 26 años o más. La sección deportiva fue más escogida por los estudiantes de edades entre 20 a 25 años, siendo una cantidad de 30.

Para los Avisos clasificados fuero 26 estudiantes los que escogieron esta opción, donde 10 fueron estudiantes de edades entre 20 y 22 años, y tan solo un estudiante d e 26 años o más.

**TABLA 139**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



Vemos en la Tabla 31 que el p-valor es de 0,171, donde a nivel de significación de 0,05 se debe de aceptar la hipótesis nula.

Por lo tanto la sección del periódico que leen con frecuencia no depende de la edad de los estudiantes.



**GRAFICO 5.15**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

* ***Edad del estudiante y Compra Joyas***

: Edad del estudiante es independiente a Comprar joyas

Vs

: Edad y Comprar joyas no son independientes

En la Tabla 32 que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, 117 nunca compran joyas, donde 44 son estudiantes de 20 a 22 años, y 12 de 26 años o más.; mientras que 85 rara vez compran joyas, donde 36 son estudiantes de 23 a 25 años y solo 4 de 26 años o más. Notamos que 32 estudiantes compran joyas la mayoría de las veces, donde 13 son estudiantes de 23 años o más.

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

**TABLA 140**



Vemos en la Tabla 140 que el p-valor es de 0,240, donde a nivel de significación de 0,05 se debe de aceptar la hipótesis nula. Por lo tanto el comprar joyas no depende de la edad de los estudiantes.

******

**GRAFICO 5.16**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

* ***Edad del estudiante y Compra Ropa de diario***

: Edad del estudiante es independiente a Comprar Ropa de diario

Vs

: Edad y Comprar Ropa de diario no son independientes

Observamos en la Tabla 141, que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, 91 nunca o rara vez compran ropa de diario, donde 16 fueron estudiantes de entre 17 y 19 años, 35 de entre 20 y 22 años y 40 de 23 años o más; mientras que 120 algunas veces compran ropa de diario, donde 32 son estudiantes de edades entre 17 y 19 años, 31 de entre 20 y 22 años y 57 de 23 años o más.

**GRAFICO 5.17**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



Notamos que solo 23 estudiantes de 234 la mayoría de veces compran ropa de diario, donde 9 son estudiantes de edades entre 17 y 19 años, 10 de entre 20 y 22 años y 4 de 23 años en adelante.

**TABLA 141**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



Vemos en la Tabla 141 que el p-valor es de 0,021, donde a nivel de significación de 0,05 se debe de rechazar la hipótesis nula a favor de la alterna. Por lo tanto el comprar ropa diario depende de la edad de los estudiantes. A continuación se realizará el Análisis de Homogeneidad.

**TABLA 141A**

*Fuente y Elaboración:* *Lissette Flores Barcia*



**Frecuencias marginales**

Los autovalores representan la proporción de información categórica explicada por cada una de las dimensiones. En este caso, los autovalores de las dimensiones 1 y 2 son 0,588 y 0,569 respectivamente. El valor máximo para cada dimensión es igual a la unidad.. Ver en la Tabla 141B.

**TABLA 141B**

*Fuente y Elaboración:* *Lissette Flores Barcia*



En la Tabla 141D podemos ver el análisis de las frecuencias marginales y cuantificaciones de categorías para cada variable.

Podemos decir que 0,775 en la variable edad, para la categoría 1, es la media de las puntuaciones en esta dimensión de los 57 estudiantes de entre 17 y 19 años. La media de las puntuaciones en la dimensión 2 para la misma categoría 1, es de 1,084.

Para la Categoría 2, la media de las puntuaciones para la dimensión 1 es de 0,577, de los 76 estudiantes de entre 20 y 22 años. Así mismo para la dimensión 2 la media de las puntuaciones es de -0,926.

Para la Categoría 3, la media de las puntuaciones para la dimensión 1 es de -0,871, de los 101 estudiantes de 23 años o más. Así mismo para la dimensión 2 la media de las puntuaciones es de 0,085.

Para la variable ROPA\_DIA, vemos que se detallan las medias de las puntuaciones de cada categoría y dimensión. Ver Tabla 141D.

Notamos en el gráfico de la Tabla 141D que los estudiantes que tienen edades entre 20 y 22 años están más relacionados a que nunca o rara vez compren ropa de diario. Mientras que los estudiantes de 23 años o más, están más relacionados con hacerlo la mayoría de las veces.

**Frecuencias Marginales y Cuantificación de Categorías**

*Fuente y Elaboración:* *Lissette Flores Barcia*

**TABLA 141D**



El Gráfico 5.18 agrupa los individuos de la muestra de acuerdo con sus puntuaciones en ambas dimensiones.

**GRAFICO 5.18**

*Fuente y Elaboración:* *Lissette Flores Barcia*



Los girasoles son concentraciones de individuos y cada pétalo es un número determinado de individuos.

* ***Edad del estudiante y Compra Discos, cassettes***

: Edad del estudiante es independiente a Comprar Ropa de diario

Vs

: Edad y Comprar Ropa de diario no son independientes

En la Tabla 142, observamos que 29 estudiantes de entre 17 y 19 años compran discos y cassettes, 36 son de 20 y 22 años y 53 son estudiantes de 23 años o más. Observamos que 116 estudiantes de 234 la mayoría de las veces compran discos y cassettes, donde 28 son estudiantes de edades entre 17 y 19 años, 40 estudiantes de 20 y 22 años y 48 de 23 años o más.

**TABLA 142**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia





**GRAFICO 5.19**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

Vemos en la Tabla 142 que el p-valor es de 0,795, donde a nivel de significación de 0,05 se debe de aceptar la hipótesis nula. Por lo tanto el comprar discos, cassettes no depende de la edad de los estudiantes.

* ***Edad del estudiante y Compra artículos deportivos***

: Edad es independiente a Comprar artículos deportivos

Vs

: Edad y Comprar Artículos deportivos no son independientes

Vemos en la Tabla 143, que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, 53 nunca compran artículos deportivos, donde 11 son estudiantes de edades entre 17 y 19 años, 19 de entre 20 y 22 años y 23 de 23 años más; 77 estudiantes rara vez compran artículos deportivos, donde 35 son de 23 años o más. 61 estudiantes algunas veces compran artículos deportivos, donde 25 son de 23 años o más. 43 estudiantes compra artículos deportivos la mayoría de las veces donde 18 son estudiantes de 23 años o más.

**GRAFICO 5.20**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



**TABLA 143**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



Vemos en la Tabla 143 que el p-valor es de 0,843, donde a nivel de significación de 0,05 se debe de aceptar la hipótesis nula. Por lo tanto el comprar artículos deportivos no depende de la edad de los estudiantes.

* ***Edad del estudiante y No debería haber tanto énfasis en el sexo hoy***

: Edad del estudiante es independiente a No debería haber tanto énfasis en el sexo hoy

Vs

: Edad y No debería haber tanto énfasis en el sexo hoy no son independientes

En la Tabla 144, observamos que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, 29 están en total desacuerdo en que no debería haber tanto énfasis en el sexo hoy, donde 11 son estudiantes de entre 17 y 19 años de edad; 33 están en parcial desacuerdo, donde 13 son estudiantes de 23 años o más.

A 65 estudiantes les es indiferente en que no debería haber tanto énfasis en el sexo hoy, donde 32 son estudiantes de 23 años o más. 60 estudiantes están en parcial acuerdo, donde 26 son estudiantes de 23 años o más. 47 están en total acuerdo en que no debería haber tanto énfasis en el sexo hoy, donde 20 son estudiantes de 23 años o más.

**TABLA 144**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia





**GRAFICO 5.21**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia

Vemos en la Tabla 144 que el p-valor es de 0,460, donde a nivel de significación de 0,05 se debe de aceptar la hipótesis nula. Por lo tanto el que piense que no debe haber tanto énfasis en el sexo hoy no depende de la edad de los estudiantes.

* ***Edad del estudiante y Me gustan los juegos de azar***

: Edad es independiente a Me gustan los juegos de azar

Vs

: Edad y Me gustan los juegos de azar no son independientes

Podemos ver en la Tabla 145, que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, que 53 están en total desacuerdo en que les gusta los juegos de azar, donde 26 son estudiantes de 23 años o más.

Son 26 estudiantes que están en parcial desacuerdo en que les gusta los juegos de azar, donde 10 son estudiantes de entre 20 y 22 años.

Para 70 estudiantes les es indiferente, donde 34 son estudiantes de 23 años o más.

Vemos que 51 están en parcial acuerdo en que les gusta los juegos de azar, son estudiantes de 23 años o más.

34 están en total acuerdo en que les gusta los juegos de azar, donde 12 son estudiantes de entre 17 y 19 años y 12 de entre 20 y 22 años.

**TABLA 145**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



**GRAFICO 5.22**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



Vemos en la Tabla 145 que el p-valor es de 0,638, donde a nivel de significación de 0,05 se debe de aceptar la hipótesis nula. Por lo tanto que le guste los juegos de azar no depende de la edad de los estudiantes.

* ***Edad del estudiante y Veo TV los fines de semana***

: Edad del estudiante es independiente a Veo TV los fines de semana

Vs

: Edad y Veo TV los fines de semana no son independientes

En la Tabla 146 tenemos que de los 234 estudiantes registrados en Ingeniería en Estadística Informática en el término 2 del año 2004, 39 están en total desacuerdo en que ven TV los fines de semana, donde 19 son estudiantes de 23 años o más.

A 34 estudiantes les es indiferente en que ven TV los fines de semana, donde 16 son estudiantes de entre 20 y 22 años.

58 estudiantes están en parcial acuerdo en que ven TV los fines de semana, donde 31 son estudiantes de 23 años o más.

Son 103 estudiantes los que están en total acuerdo en que ven TV los fines de semana, donde 40 son estudiantes de 23 años o más.

**GRAFICO 5.23**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



Podemos ver en la Tabla 146 que el p-valor es de 0,281, donde a nivel de significación de 0,05 se debe de aceptar la hipótesis nula. Por lo tanto el ver TV los fines de semana no depende de la edad de los estudiantes.

**TABLA 146**

*Fuente: Encuesta aplicada*

### Elaboración: Lissette Flores Barcia



* + 1. **Análisis de Componentes Principales**

El Análisis de Componentes Principales es una técnica multivariada que permite determinar la dimensionalidad del espacio en el cual caen los datos y encaminarlos en este espacio de dimensiones reducidas.

El Análisis de Componentes Principales comprende un procedimiento matemático que transforma un conjunto de variables correlacionadas de respuestas en un conjunto menor de variables no correlacionadas llamadas componentes principales.

Los objetivos para el análisis de componentes principales son de:

1. Reducir la dimensionalidad del conjunto de datos
2. Identificar nuevas variables subyacentes
3. Poder interpretar una matriz de correlaciones a partir del menor número posible de factores.

Las nuevas variables (componentes principales) presentan un orden decreciente de importancia:

* 1. No están correlacionadas
  2. La primera componente principal explica tanto de la variabilidad en los datos como sea posible
  3. Cada componente subyacente toma en cuenta de la variabilidad restante como sea posible.

El método de Componentes principales consiste en llevar a cabo una combinación lineal de todas las variables de modo que el primer componente principal sea una combinación que explique la mayor proporción de variancia de la muestra, el segundo la segunda mayor y que a su vez esté incorrelacionado con el primero, y así sucesivamente hasta tantos componentes como variables. Por lo tanto, reducir un número de variables incorrelacionadas a un número menor de factores no correlacionados.

El test de Barlett se utiliza para verificar si la matriz de correlaciones es una matriz de identidad, es decir si todos los coeficientes de la diagonal son iguales a 1 y los extremos a la diagonal iguales a 0.

Este estadístico se obtiene a partir de la transformación  del determinante de la matriz de correlaciones, cuanto mayor sea esta aproximación de la Chi-Cuadrado y menor sea el grado de significación es más improbable que la matriz sea una matriz de identidad. En este caso no se debería proceder con un análisis factorial.

El índice KMO compara los coeficientes de correlación de Pearson con los coeficientes de correlación parcial entre variables.

Se obtiene que:



Donde:

: es el coeficiente de correlación de Pearson entre las variables i y j

: es el coeficiente de correlación parcial entre las variables i y j

Si los coeficientes de correlación parcial entre las variables son muy pequeños, quiere decir que la relación entre cada par de las mismas se debe o puede ser explicada por el resto y por tanto llevar a cabo un análisis factorial de los datos no deja de de ser una buena solución. En este caso, si la suma de los coeficientes de correlación parcial al cuadrado es muy pequeña, KMO será un índice muy próximo a 1, y por tanto un análisis factorial es adecuado. Caso contrario, valores pequeños en este índice demuestra que no lo es. Según Kaiser (1974):

1≥KMO>0,90 son considerados excelentes

0,90≥KMO>0,80 son considerados buenos

0,80≥KMO>0,70 son considerados aceptables

0,70≥KMO>0,60 son considerados mediocres o regulares

0,60≥KMO>0,50 son considerados malos

KMO<0,50 son considerados inaceptables o muy malos

Interesan los coeficientes de las matrices anti-imagen de covarianzas y correlaciones de las variables, cuanto más pequeños sean éstos, mejor.

**TABLA 147**

*Fuente y Elaboración: Lissette Flores Barcia*



El resultado del test de Barlett con un =1570,581 y p=0,000. los valores de los coeficientes de matrices anti-imagen son muy bajos. Valores altos en los coeficientes de la diagonal (MSA) de la matriz de correlaciones anti-imagen.

El determinante de la matriz de correlaciones fue de 7,723E-04,que es un valor muy bajo.

Estos resultados nos dice que el análisis factorial puede ser pertinente para proporcionarnos conclusiones satisfactorias.

La varianza total explicada recoge en porcentajes individuales y acumulados, la proporción de varianza total explicada por cada factor, tanto para la solución no rotada como para la rotada.

Para saber cuantos componentes principales se deben utilizar el criterio de la varianza acumulada que con treinta y nueve variables de interés, se obtuvo como resultado 15 componentes principales (ver Tabla 149), los cuales en conjunto explican el 63,25% del total de la Varianza.

**TABLA 149**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Autovalores iniciales** | | | **Suma de saturaciones al cuadrado de la extracción** | | |
| **Componente** | **Total** | **% de la varianza** | **% acumulado** | **Total** | **% de varianza** | **% acumulado** |
| 1 | 3,1240266 | 8,0103246 | 8,0103246 | 3,1240266 | 8,0103246 | 8,0103246 |
| 2 | 2,5444425 | 6,5242116 | 14,5345361 | 2,5444425 | 6,5242116 | 14,5345361 |
| 3 | 2,2592784 | 5,7930215 | 20,3275577 | 2,2592784 | 5,7930215 | 20,3275577 |
| 4 | 1,9182442 | 4,9185749 | 25,2461326 | 1,9182442 | 4,9185749 | 25,2461326 |
| 5 | 1,7897571 | 4,5891208 | 29,8352534 | 1,7897571 | 4,5891208 | 29,8352534 |
| 6 | 1,6318539 | 4,1842407 | 34,0194940 | 1,6318539 | 4,1842407 | 34,0194940 |
| 7 | 1,5875429 | 4,0706229 | 38,0901170 | 1,5875429 | 4,0706229 | 38,0901170 |
| 8 | 1,4495884 | 3,7168935 | 41,8070104 | 1,4495884 | 3,7168935 | 41,8070104 |
| 9 | 1,3594481 | 3,4857645 | 45,2927749 | 1,3594481 | 3,4857645 | 45,2927749 |
| 10 | 1,3326298 | 3,4169995 | 48,7097744 | 1,3326298 | 3,4169995 | 48,7097744 |
| 11 | 1,2088900 | 3,0997181 | 51,8094924 | 1,2088900 | 3,0997181 | 51,8094924 |
| 12 | 1,1868417 | 3,0431838 | 54,8526762 | 1,1868417 | 3,0431838 | 54,8526762 |
| 13 | 1,1263613 | 2,8881059 | 57,7407821 | 1,1263613 | 2,8881059 | 57,7407821 |
| 14 | 1,0952445 | 2,8083192 | 60,5491013 | 1,0952445 | 2,8083192 | 60,5491013 |
| 15 | 1,0540490 | 2,7026898 | **63,2517911** | 1,0540490 | 2,7026898 | 63,2517911 |
| 16 | 0,9679523 | 2,4819290 | 65,7337201 |  |  |  |
| 17 | 0,9197390 | 2,3583050 | 68,0920251 |  |  |  |
| 18 | 0,8911433 | 2,2849828 | 70,3770079 |  |  |  |
| 19 | 0,8453683 | 2,1676111 | 72,5446190 |  |  |  |
| 20 | 0,8075714 | 2,0706959 | 74,6153149 |  |  |  |
| 21 | 0,7876189 | 2,0195357 | 76,6348506 |  |  |  |
| 22 | 0,7422273 | 1,9031470 | 78,5379976 |  |  |  |
| 23 | 0,7193999 | 1,8446152 | 80,3826128 |  |  |  |
| 24 | 0,6991238 | 1,7926252 | 82,1752380 |  |  |  |
| 25 | 0,6265711 | 1,6065927 | 83,7818307 |  |  |  |
| 26 | 0,6141802 | 1,5748211 | 85,3566518 |  |  |  |
| 27 | 0,5877237 | 1,5069840 | 86,8636357 |  |  |  |
| 28 | 0,5758604 | 1,4765650 | 88,3402008 |  |  |  |
| 29 | 0,5517161 | 1,4146567 | 89,7548575 |  |  |  |
| 30 | 0,5110122 | 1,3102876 | 91,0651451 |  |  |  |
| 31 | 0,4790567 | 1,2283505 | 92,2934956 |  |  |  |
| 32 | 0,4712366 | 1,2082991 | 93,5017947 |  |  |  |
| 33 | 0,4616759 | 1,1837844 | 94,6855791 |  |  |  |
| 34 | 0,3971689 | 1,0183819 | 95,7039610 |  |  |  |
| 35 | 0,3882736 | 0,9955734 | 96,6995344 |  |  |  |
| 36 | 0,3660170 | 0,9385052 | 97,6380396 |  |  |  |
| 37 | 0,3354105 | 0,8600269 | 98,4980665 |  |  |  |
| 38 | 0,3085502 | 0,7911544 | 99,2892210 |  |  |  |
| 39 | 0,2772038 | 0,7107790 | 100 |  |  |  |
| Fuente y Elaboración: Lissette Flores Barcia | | | | | | |

Se tiene una representación gráfica de estos resultados, donde las abscisas son los factores y las ordenadas el valor propio correspondiente a cada uno de ellos (ver Gráfico 5.24).

Los coeficientes utilizados para expresar cada variable estandarizada en término de los 21 factores del modelo, es decir la Matriz de Componentes Principales, lo podemos ver en el ANEXO A, llamado también ponderaciones factoriales ya que indica la carga de cada variable en cada factor, de modo que los factores con unos pesos factoriales más elevados en términos absolutos nos indican una relación estrecha con las variables. El ideal en el punto de vista del análisis factorial es encontrar un modelo en el que todas las variables saturen en algún factor, es decir, pesos factoriales altos en uno y bajos en el resto.

Si los factores no están correlacionados, es decir que no son ortogonales, los valores de estos coeficientes no dependen unos de otros.

**GRAFICO 5.24**

### Fuente y Elaboración: Lissette Flores Barcia



Representan por tanto la contribución única de cada factor en la variable y en ese sentido no dejan de ser sino los coeficientes de correlación de cada factor con la variable.

La correlación obtenida entre los distintos factores y las variables puede ser utilizada para estimar la correlación entre las variables, de modo que si los factores son ortogonales, el coeficiente de correlación entre las variables i y j será:

r

Donde:

k: es el número de factores comunes

: es el coeficiente de correlación entre el factor f y la variable i

De esta forma se calcularán los coeficientes para la matriz de Correlación reproducida, en donde aparecen además los residuales que es la diferencia entre estos coeficientes de correlación estimados y los coeficientes de correlación observados.

La magnitud y cuantía de los residuales con valores superiores a 0,05 son unos buenos indicadores del ajuste del modelo a los datos.

***ROTACION:*** La finalidad de la rotación es de ayudarnos a interpretar las componentes en el supuesto de que no quede clara la explicación de la varianza toral en el caso no rotado.

En este caso, utilizaremos la rotación *VARIMAX*, que es un procedimiento ortogonal que mantiene factores incorrelacionados y los ejes formando ángulos rectos. El objetivo principal es de minimizar el número de variables que tienen pesos en cada factor.

Utilizando el criterio de valores propios, el gráfico de sedimentación y rotación *Varimax*, se obtuvo que debe retenerse quince componentes principales, las que consiguen absorber el 63,25% del total de la Varianza.

* La primera componente resume a Frecuencia con la que estudia en laboratorios (0,408), Frecuencia con que estudia con exámenes pasados (0,439), Frecuencia con que estudia con apuntes personales (0,452), Frecuencia con que estudia con un texto guía (0,476), Debería haber una pistola en cada hogar (0,467), Soy Hogareño (0,371), Hay mucha violencia en la TV (0,416), Me gustan las apuestas (0,565), Mi tiempo libre lo comparto con mi familia (0,406), Los domingos almuerzo en mi casa (0,307). Esta componente se llamará *Forma de estudio y actitudes del estudiante*; ya que determina las actitudes del estudiante y su forma de estudiar.
* La segunda componente resume la información de Me divierte ir a conciertos (0,509), Las mujeres de hoy usan mucho maquillaje (0,432), La TV es fuente primaria de mi entretenimiento (0,414), Prefiero quedarme en casa antes que ir a una fiesta (0,499), Leo algún libro cuando tengo tiempo (0,429), Me gusta ir al cine frecuentemente (0,446), Veo TV los fines de semana (0,389). Esta componente se llamará *Preferencias del estudiante*; ya que se relacionan variables que determinan las preferencias que tienen los estudiantes ante determinadas circunstancias.
* La Tercera componente principal resume la información de Me gustan las historias de Guerra (0,458), Practico algún deporte (0,508), La lectura es mi pasatiempo favorito (0,440), Me gusta mucho la política (0,601). Esta componente se llamará *Gustos y* *Afinidades del estudiante*; porque determinan las afinidades en cuanto deportes o pasatiempos.
* La Cuarta componente principal resume la información de Con qué frecuencia estudia con apuntes de semestres anteriores (0,392), Frecuencia con que compra Ropa de diario (0,375), Me gusta los juegos de azar (0,386). Esta componente se llamará *Actividades del estudiante*; ya que se relacionan variables que determinan las actividades varias que acostumbra a tener el estudiante.
* La Quinta componente principal resume la información de Utiliza expresos públicos y/o de la ESPOL para trasladarse dentro y fuera de la ESPOL (0,428), Con qué frecuencia compra lociones y perfumes (0,649), Con qué frecuencia compra discos, cassettes (0,467). Esta componente se llamará *Preferencias en las compras*; ya que determina lo que el estudiante prefiere al momento de comprar.
* La Sexta componente principal resume la información de Me agrada distinguirme ante los demás (0,442). Esta componente se llamará *Distinción*; porque determina la satisfacción que tiene el estudiante al distinguirse ante los demás.
* La Séptima componente principal no se relaciona con ninguna variable.
* La Octava componente principal resume la información de Frecuencia que dispone de Internet en el lugar donde habita (0,348), Frecuencia con que estudia solo (0,481), Me considero un líder (0,442). Esta componente se llamará *Superación Personal;* porque se relacionan variables que determinan el empeño que tiene el estudiante para su superación.
* La Novena componente resume la información de No debería haber tanto énfasis en el sexo hoy (0,359). Esta componente se llamará *Énfasis en el sexo*; ya que determina la creencia de que deba haber énfasis en el sexo en la actualidad.
* La Décima componente principal resume la información de Frecuencia que compra el texto que sugiere el profesor (0,334), La religión es parte importante en mi vida (0,423). Esta componente se llamará *Responsabilidad y espiritualidad del estudiante.*
* La Décima primera componente principal resume la información de Me preocupo por la alimentación (0,366). Esta componente se llamará *Alimentación*; ya que determina la preocupación que tiene el estudiante por alimentarse bien.
* La Décima segunda componente principal resume la información de Frecuencia con que Estudia en Bibliotecas (0,432), Me gustan los quehaceres domésticos (0,318). Esta componente se llamará *Atracción a quehaceres del hogar.*
* La Décima tercera componente principal no se relaciona con ninguna variable.
* La Décima cuarta componente principal resume la información de Me gustan las actividades al aire libre (0,440). Esta componente se llamará *Actividades al aire libre*.
* La Décima quinta componente principal resume la información de Me siento atractivo ante el sexo opuesto (0,408). A esta componente se la llamará *Atractivo*.