CAPITULO 5

# **ANÁLISIS ESTADÍSTICO MULTIVARIADO DE LA POBLACIÓN INVESTIGADA**

##  Introducción

En este capítulo se efectuará un análisis estadístico multivariado, de las respuestas de los cuestionarios, aplicados a 593 estudiantes, de los colegios de sostenimiento Fiscal y Particular del cantón Guayaquil. Entre uno de los objetivos del estudio es determinar la independencia entre variables a través de la matriz de correlación y tablas de contingencia.

##  Análisis de Correlación entre pares de variables (5)

5.2.1 Coeficiente de correlación

El coeficiente de correlación , definido en términos del cociente entre la covarianza  y el producto de varianzas  y . El coeficiente de correlación entre Xj y Xk se denota por y se define por:



En general, si dos variables se distribuyen de manera independiente, entonces  y .

El coeficiente de correlación lineal simple, mide el grado de asociación lineal entre dos variables medidas en escala de intervalo o de razón, tomando valores entre -1 y 1.

Valores de próximos a 1 indicarán una fuerte asociación lineal positiva ( a medida que aumentan los valores de una de las dos variables, aumentan los de la otra); valores de próximos a -1 indicarán una fuerte asociación lineal negativa ( a medida que aumentan los valores de una delas dos variables, disminuyen los de la otra) y valores  próximos a 0 indicarán no asociación lineal ( lo que no significa que no pueda existir otro tipo de asociación).

Para determinar si la asociación es estadísticamente significativa, se puede plantear la hipótesis nula de que el coeficiente de correlación lineal es igual a cero:



El estadístico de contraste se construye a partir del coeficiente de correlación y del tamaño de la muestra , n. Si el p –valor asociado es menor que α, se rechazará la hipótesis nula al nivel de significación α.

5.2.2 Matriz de Correlación

Es un arreglo de p filas y p columnas que agrupa todas las medidas de las relaciones de tipo lineal que existen entre las p variables investigadas, la Matriz de Correlación () , está definida por:



  =

##  Tablas de contingencia (8)

Las tablas de contingencia constituyen una técnica que permite determinar si dos variables o factores son independientes. A continuación se presenta un modelo de una tabla de contingencia.

Tabla de Contingencia del factor 1 vs. factor 2

|  |
| --- |
| FACTOR 1 |
| FACTOR2 |  | Nivel 1 | Nivel 2 | ... | Nivel c | Xi. |
| 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.j | X.1E.1 | X.2E.2 | ... | X.cE.c | X..=n |

Como podemos ver, la tabla de contingencia es un arreglo matricial de r filas y c columnas, en donde existen dos factores en el que r es el número de niveles del factor 1 o de la variable Xi y c el número de niveles del factor 2 o de la variable Xj , en cada variable debe tener al menos dos niveles los cuales deben ser exhaustivos y mutuamente excluyentes.

Donde:

n= es el número de observaciones

Xi j : es el número de unidades de investigación sometidas al i-ésimo nivel del factor 2 y el j-ésimo nivel del factor 1.

Se definen las siguientes ecuaciones:





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

X.j es la suma de las frecuencias observadas en la j-ésima columna.

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

Eij es el valor esperado de la frecuencia de la celda en el i-ésimo renglón y la j-ésima columna.

Ei. es el valor esperado para el i-ésimo renglón.

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

La hipótesis planteada es:



Y el estadístico de prueba a utilizar es:



Se puede probar que el estadístico *X2* tiene una distribución ji-cuadrado con (r-1)(c-1) grados de libertad, *X2* ~ .

De donde con (1-α)100% de confianza se rechaza Ho en favor de H1 si:

*X2* > .

5.4 Análisis de Correspondencia (7)

El análisis de correspondencias simple es un técnica que permite analizar la homogeneidad entre las categorías de cada una de los dos variables respecto a las categorías de la otra, 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 a menudo incluye examinar los perfiles de fila y de columna, así como contrastar la independencia a través del estadístico de chi-cuadrado. Sin embargo, el número de perfiles puede ser bastante grande y la prueba de chi-cuadrado no revelará la estructura de la dependencia. El procedimiento Tablas de contingencia ofrece varias medidas y pruebas de asociación pero no puede representar gráficamente ninguna relación entre las variables.

Para la medida de distancia entre las filas y columnas se puede seleccionar el estadístico Chi-Cuadrado. Utiliza una distancia ponderada entre los perfiles, donde la ponderación es la masa de las filas o de las columnas. Esta distancia es necesaria para el análisis de correspondencias típico.

### 5.5 Análisis de Homogeneidad (5) (6)

El Análisis de Homogeneidad conocido también como HOMALS ( Análisis de Homogeneidad mediante Mínimos Cuadrados Alternantes) . El fin de este análisis es cuantificar los datos categóricos nominales mediante la asignación de valores numéricos a los casos “objetos” y a las categorías.

Es similar al Análisis de Correspondencias pero no está limitado a dos variables, es por esta razón que se lo denomina también análisis de Correspondencia múltiple.

La interpretación de los resultados del Análisis de Homogeneidad es mucho más práctico que otro tipo de técnicas categóricas que se emplean, puesto que las categorías de las variables son cuantificables, sobre las cuantificaciones técnicas que necesitan de datos numéricos en análisis posteriores.

* 1. Análisis de las variables
		1. Correlaciones

En este punto se analizará la relación lineal existente entre las variables, haciendo uso de la matriz de correlaciones, ( Ver anexo 2 ).

De la matriz de correlación se extraen los siguientes resultados mostrados en la siguiente Tabla XLIV, que corresponden a las variables investigadas, se puede observar que existen 37 pares de variables que tienen una correlación significativa en comparación a los demás pares de variables.

 **TABLA XLIV**

 **Correlaciones más significativas**

|  |  |  |  |
| --- | --- | --- | --- |
| Variable 1 | Variable 2 | Correlación de Pearson | Sig. (bilateral) |
| con quien vive | de que forma se entero | -0,11 | 0,03 |
| grado de parentesco | 0,13 | 0,01 |
| grado de parentesco | país al que emigro | -0,13 | 0,01 |
| de que forma se entero | que sintió cuando se entero la noticia | 0,12 | 0,02 |
| como le hubiese gustado enterarse | sostenimiento | -0,11 | 0,03 |
| sintió desanimo y tristeza | 0,13 | 0,01 |
| que sintió cuando se entero la noticia | edad | -0,11 | 0,03 |
| año | -0,12 | 0,02 |
| acepta que haya viajado para cambiar | edad | 0,13 | 0,01 |
| grado de parentesco | -0,11 | 0,09 |
| situaciones que se dieron | grado de parentesco | 0,14 | 0,01 |
| sintió desanimo y tristeza | sin esperanzas en la vida | 0,12 | 0,02 |
| sintió deseos de morir | grado de parentesco | -0,11 | 0,04 |
| sin esperanzas en la vida | con quien vive | 0,11 | 0,04 |
| de que forma se entero | 0,12 | 0,02 |
| ganas frecuentes de llorar | como le hubiese gustado enterarse | 0,14 | 0,01 |
| sintió angustia | nombre colegio | -0,11 | 0,03 |
| que sintió cuando se entero la noticia | -0,13 | 0,01 |
| sintió confusión | como lo trata la persona que lo cuida | -0,12 | 0,02 |
| antes de que se fuera Ud. hacia | -0,12 | 0,02 |
| sintió irritabilidad | como le hubiese gustado enterarse | 0,11 | 0,04 |
| sentimientos de abandono | nombre colegio | -0,12 | 0,02 |
| que sintió cuando se entero la noticia | -0,11 | 0,04 |
| frustración | edad | 0,11 | 0,03 |
| que sintió cuando se entero la noticia | -0,13 | 0,02 |
| con quien se quedo viviendo | frustración | 0,12 | 0,03 |
| antes de que se fuera Ud. hacia | -0,12 | 0,03 |
| como lo trata la persona que lo cuida | sintió desanimo y tristeza | -0,13 | 0,01 |
| sentimientos de abandono | -0,14 | 0,01 |
| sintió deseos de morir | -0,11 | 0,04 |
| piensa q fue lo mejor q te dejara a su | con quien vive | -0,13 | 0,01 |
| sentimientos de abandono | -0,13 | 0,01 |
| sintió deseos de morir | -0,12 | 0,02 |
| antes de que se fuera Ud. hacia | sintió desanimo y tristeza | -0,11 | 0,04 |
| sentimientos de abandono | -0,13 | 0,01 |
| después de que se fue Ud. hacia | genero | 0,13 | 0,01 |
| sintió angustia | -0,11 | 0,04 |

* + 1. Tablas de Contingencia, Análisis de Correspondencia y Homogeneidad.

En esta sección vamos analizar la dependencia lineal entre las variables, a través del uso de las tablas de contingencia, y del estadístico de prueba de independencia Chi - Cuadrado para algunos pares de variables. En caso de existir variables dependientes se realizará un análisis de correspondencia.

TABLA XLV

Tabla de contingencia de que forma se entero \* que sintió cuando se entero la noticia

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | que sintió cuando se entero la noticia | Total |
|  |  | Confusión | Tristeza | Alegría |  |
| de que forma se entero | inesperadamente | 19 | 76 | 25 | 120 |
|  | ya lo esperaba | 16 | 138 | 60 | 214 |
| Total |  | 35 | 214 | 85 | 334 |

Fuente: Encuesta Aplicada a colegios fiscales y particulares de Guayaquil

Elaboración: Lorena Álvarez

 Pruebas de chi-cuadrado

|  |  |  |  |
| --- | --- | --- | --- |
|  | Valor | gl | Sig. Asintótica (bilateral) |
| Chi-cuadrado de Pearson | 6,708 | 2 | ,035 |

 a 0 casillas (,0%) tienen una frecuencia esperada inferior a 5.

 La frecuencia mínima esperada es 12,57.

: De que forma se entero vs. que sintió cuando se enteró la noticia , son independientes.

Vs.

 : No es verdad 

El estadístico de prueba Chi- Cuadrado es igual 6,708 y el valor p es igual 0,035 por lo tanto existe evidencia estadística para rechazar la Hipótesis nula, y las variables de que forma se enteró la noticia y que sintió cuando se enteró la noticia son variables dependientes.

Análisis de Homogeneidad para las variables de que forma se enteró vs. que sintió cuando se enteró de la noticia.

En la Tabla XLVI se presentan los autovalores y las medidas de discriminación. Dentro de las dos dimensiones, el autovalor más grande es 0.57 y en la segunda 0.50 lo que nos indica que en la solución general tendrá más importancia el autovalor 1 que el autovalor 2.

TABLA XLVI

Resultado del análisis de homogeneidad de las variables

De la forma como se enteró vs. que sintió cuando se enteró de la noticia

 Autovalores

|  |  |
| --- | --- |
| Dimensión | Autovalores |
| 1 | 0,57 |
| 2 | 0,50 |

 Medidas de discriminación

|  |  |  |
| --- | --- | --- |
|   | Dimensión |   |
|  Variables  | 1 | 2 |   |
| de que forma se entero | 0,57 | 0,00 |   |
| que sintió cuando se entero la noticia | 0,57 | 1,00 |   |

 Fuente: Encuesta aplicada a los estudiantes de

 Colegios fiscales y particulares de Guayaquil.

 Elaboración: Lorena Álvarez

Las dos variables discriminan en igual forma en la primera dimensión, mientras que en la segunda dimensión discrimina mejor la variable que sintió cuando se enteró de la noticia.

En la Gráfico 33 de cuantificaciones se puede ver, que los estudiantes que ya esperaban la noticia sintieron tristeza porque su familiar emigraría a otro lugar.

GRAFICO 33

Que sintió cuando se enteró la noticia vs.

de que forma se enteró



TABLA XLVII

 Tabla de contingencia como le hubiese

gustado enterarse \* sostenimiento

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | sostenimiento | Total |
|  |  | Fiscal | Particular |  |
| como le hubiese gustado enterarse | que nadie le contara | 39 | 61 | 100 |
|  | que hubiese contado con su opinión | 121 | 113 | 234 |
| Total |  | 160 | 174 | 334 |

 Fuente: Encuesta aplicada

 Elaboración: Lorena Álvarez

Pruebas de chi-cuadrado

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Valor | gl | Sig. asintótica (bilateral) | Sig. exacta (bilateral) | Sig. exacta (unilateral) |
| Chi-cuadrado de Pearson | 4,535 | 1 | 0,033 |  |  |

 a Calculado sólo para una tabla de 2x2.

 b 0 casillas (,0%) tienen una frecuencia esperada inferior a 5.

 La frecuencia mínima esperada es 47,90.

: como le hubiese gustado enterarse vs. sostenimiento, son independientes.

Vs.

 : No es verdad 

El estadístico de prueba Chi- Cuadrado es igual 4,535 y el valor p es igual 0,033 por lo tanto existe evidencia estadística para rechazar la hipótesis nula, y las variables como le hubiese gustado enterarse y sostenimiento son variables dependientes.

Análisis de Homogeneidad para las variables como le hubiese gustado enterarse y sostenimiento.

En la dimensión 1 tiene el autovalor con mayor valor que el de la dimensión 2. Ambas variables se discriminan mejor en en la dimensión 1.

TABLA XLVIII

Autovalores

|  |  |  |
| --- | --- | --- |
| Dimensión | Autovalores |   |
| 1 | 0,558 |   |
| 2 | 0,442 |   |

Medidas de discriminación

|  |  |
| --- | --- |
|   | Dimensión |
| Variables | 1 | 2 |
| sostenimiento | 0,558 | 0,442 |
| como le hubiese gustado enterarse | 0,558 | 0,442 |

 Fuente: Encuesta aplicada a los colegios fiscales y particulares

 Elaboración: Lorena Álvarez

En el gráfico 34 se observa que los estudiantes de colegio particular hubiesen preferido que su familiar contara con su opinión, antes de tomar la decisión de viajar.

GRAFICO 34

Como le hubiese gustado enterarse vs. Sostenimiento



TABLA XLVIX

Tabla de contingencia acepta que haya viajado para cambiar \* edad

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | edad | Total |
|  |  | [11,12] | [13,14] | [15,16] |  |
| acepta que haya viajado para cambiar | Total desacuerdo | 17 | 30 | 47 | 94 |
|  | Parcial desacuerdo | 6 | 10 | 35 | 51 |
|  | Parcial acuerdo | 9 | 28 | 63 | 100 |
|  | Total Acuerdo | 10 | 16 | 63 | 89 |
| Total |  | 42 | 84 | 208 | 334 |

Fuente: Encuesta aplicada

Elaboración: Lorena Álvarez

Pruebas de chi-cuadrado

|  |  |  |  |
| --- | --- | --- | --- |
|  | Valor | gl | Sig. asintótica (bilateral) |
| Chi-cuadrado de Pearson | 11,565 | 6 | 0,072 |

 a 0 casillas (,0%) tienen una frecuencia esperada inferior a 5.

 La frecuencia mínima esperada es 6,41.

: acepta que haya viajado para cambiar vs. edad , son independientes.

Vs.

 : No es verdad 

El estadístico de prueba Chi- Cuadrado es igual 11,565 y el valor p es igual 0,072 por lo tanto no se rechaza la hipótesis nula, y las variables acepta que haya viajado para cambiar y la edad , son variables independientes bajo todo nivel de significancia estadística.

TABLA L

Tabla de contingencia ganas frecuentes de llorar \* como le hubiese gustado enterarse

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | como le hubiese gustado enterarse | Total |
|  |  | que nadie le contara | que hubiese contado con su opinión |  |
| ganas frecuentes de llorar | Total desacuerdo | 91 | 186 | 277 |
|  | Total Acuerdo | 9 | 48 | 57 |
| Total |  | 100 | 234 | 334 |

 Fuente: Encuesta aplicada

 Elaboración: Lorena Álvarez

Pruebas de chi-cuadrado

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Valor | gl | Sig. asintótica (bilateral) | Sig. exacta (bilateral) | Sig. exacta (unilateral) |
| Chi-cuadrado de Pearson | 6,561 | 1 | 0,010 |  |  |

 a Calculado sólo para una tabla de 2x2.

 b 0 casillas (,0%) tienen una frecuencia esperada inferior a 5.

 La frecuencia mínima esperada es 17,07.

: ganas frecuentes de llorar vs. como le hubiese gustado enterarse de la noticia , son independientes.

Vs.

 : No es verdad 

El estadístico de prueba Chi- Cuadrado es igual 6,561 y el valor p es igual 0,010 por lo tanto se rechaza la hipótesis nula, y las variables ganas frecuentes de llorar y como le hubiese gustado enterarse de la noticia, son dependientes.

Análisis de Homogeneidad de la variable ganas frecuentes de llorar y como le hubiese gustado enterarse de la noticia

 TABLA LI

Resultados del análisis de homogeneidad de la variable ganas de llorar vs. como le hubiese gustado enterarse de la noticia.

Autovalores

|  |  |  |
| --- | --- | --- |
| Dimensión | Autovalores |   |
| 1 | 0,570 |   |
| 2 | 0,430 |   |

Medidas de discriminación

|  |  |  |
| --- | --- | --- |
|  | Dimensión |   |
| Variables | 1 | 2 |   |
| como le hubiese gustado enterarse | 0,570 | 0,430 |   |
| ganas frecuentes de llorar | 0,570 | 0,430 |   |

En el gráfico 35 se puede ver que los estudiantes que hay estudiantes que no sintieron ganas de llorar, al preferir que sus familiares le cuenten del viaje.

GRAFICO 35

Ganas frecuentes de llorar vs.

como le hubiese gustado enterarse



TABLA LII

Tabla de contingencia de género\* acepta que haya viajado para cambiar.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | acepta que haya viajado para cambiar | Total |   |
|  |  | Total desacuerdo | Parcial desacuerdo | Parcial acuerdo | Total Acuerdo |  |
| genero | Masculino | 46 | 33 | 41 | 28 | 148 |
|  | Femenino | 48 | 18 | 59 | 61 | 186 |
| Total |  | 94 | 51 | 100 | 89 | 334 |

 Fuente : Encuesta aplicada

 Elaboración: Lorena Álvarez

######  Pruebas de chi-cuadrado

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Valor | gl | Sig. asintótica (bilateral) |   |
| Chi-cuadrado de Pearson | 15,812 | 3 | 0,001 |   |

 a 0 casillas (,0%) tienen una frecuencia esperada inferior a 5.

 La frecuencia mínima esperada es 22,60.

: El género vs. acepta que haya tenido que viajar para cambiar el estilo de vida son independientes.

Vs.

 : No es verdad 

El estadístico de prueba Chi- Cuadrado es igual 15,812 y el valor p es igual 0,001 por lo tanto existe evidencia estadística para rechazar la hipótesis nula, y las variables género vs. acepta que haya tenido que viajar son variables dependientes.

Análisis de Homogeneidad entre el género y si acepta que haya tenido que viajar .

En el gráfico 36 se puede observar, que los estudiantes de género masculino están en parcial desacuerdo con que su familiar haya tenido que viajar, caso contrario a lo que sucede con los estudiantes de género femenino que se muestran indiferentes ante este hecho.

GRAFICO 36



TABLA LIII

Tabla de contingencia como le hubiese gustado enterarse \* sintió desanimo y tristeza

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | sintió desanimo y tristeza | Total |
|  |  | Total desacuerdo | Indiferente | Parcial acuerdo |  |
| como le hubiese gustado enterarse | que nadie le contara | 70 | 7 | 23 | 100 |
|  | que hubiese contado con su opinión | 138 | 14 | 82 | 234 |
| Total |  | 208 | 21 | 105 | 334 |

 Fuente: Encuesta aplicada

 Elaboración: Lorena Álvarez

 Pruebas de chi-cuadrado

|  |  |  |  |
| --- | --- | --- | --- |
|  | Valor | gl | Sig. asintótica (bilateral) |
| Chi-cuadrado de Pearson | 4,715 | 2 | 0,095 |
| Razón de verosimilitud | 4,883 | 2 | 0,087 |
| Asociación lineal por lineal | 4,189 | 1 | 0,041 |
| N de casos válidos | 334 |  |  |

 a 0 casillas (,0%) tienen una frecuencia esperada inferior a 5.

 La frecuencia mínima esperada es 6,29.

: Como le hubiese gustado enterarse y sintió desánimo y tristeza, son independientes.

Vs.

 : No es verdad 

El estadístico de prueba Chi- Cuadrado es igual 4,715 y el valor p es igual 0,095 por lo tanto no se rechaza la hipótesis nula, y las variables como le hubiese gustado enterarse y sintió desánimo y tristeza son independientes.

Análisis de Homogeneidad para las variables género, como se sintió y de que forma se enteró

La variable que discrimina mejor en la primera dimensión es la forma como se enteró seguida de lo que sintió cuando se enteró mientras que la segunda dimensión estará mejor representada por el género del estudiante.

TABLA LIV

Resultado del análisis de homogeneidad para

las variables género, lo que sintió y la forma

 como se enteró

Autovalores

|  |  |
| --- | --- |
| Dimensión | Autovalores |
| 1 | 0,410 |
| 2 | 0,386 |

Medidas de discriminación

|  |  |  |
| --- | --- | --- |
|   | Dimensión |   |
|   | 1 | 2 |   |
| genero | ,016 | ,712 |   |
| que sintió cuando se entero la noticia | ,535 | ,396 |   |
| de que forma se entero | ,679 | ,049 |   |

Fuente: Encuesta aplicada a los estudiantes de colegios fiscales

Y particulares de Guayaquil. Elaboración: Lorena Álvarez

En el gráfico 37 se puede notar que los estudiantes de género femenino, que recibieron la noticia inesperadamente sintieron tristeza, mientras que los estudiantes de género femenino que ya esperaban la noticia se sentían un poco culpables por la partida de su familiar.

GRÁFICO 37

Medidas discriminantes para las variables genero,

 lo que sintió y la forma como se enteró



Análisis de Homogeneidad para las variables género, como se sintió y de que forma se enteró

En la tabla LV podemos observar los autovalores de las variables el autovalor de la dimensión 1 es mayor que el autovalor de la dimensión 2.

TABLA LV

Resultado del análisis de homogeneidad para las variables género, lo que sintió y la edad

Autovalores

|  |  |
| --- | --- |
| Dimensión | Autovalores |
| 1 | 0,403 |
| 2 | 0,379 |

Medidas de discriminación

|  |  |  |
| --- | --- | --- |
|   | Dimensión |   |
| Variables | 1 | 2 |   |
| genero | 0,663 | 0,021 |   |
| que sintió cuando se entero la noticia | 0,212 | 0,259 |   |
| edad | 0,333 | 0,857 |   |

Fuente: Encuesta aplicada a los estudiantes de colegios fiscales

Y particulares de Guayaquil. Elaboración: Lorena Álvarez

Los valores que discriminan mejor en la dimensión uno son los de la variable género seguido de la edad, en la dimensión dos está la variable edad y lo que sintió.

En el gráfico 38 se ven las cuantificaciones de cada categoría, se puede observar que los estudiantes de sexo femenino y que tienen entre 15 y 16 años sintieron tristeza y culpabilidad al saber que su familiar se iría al exterior, mientras que los estudiantes de género masculino sintieron confusión.

GRÁFICO 38

Medidas discriminantes para las variables genero,

lo que sintió y la edad.



TABLA LVI

Resultados de los análisis de correspondencia de las variables Acepta que haya viajado para cambiar vs. situaciones que se dieron

######  Examen de los puntos de fila

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|   | Masa | Puntuación en la dimensión | Inercia | Contribución |
|   |   | 1 | 2 |   | De los puntos a la inercia de la dimensión | De la dimensión a la inercia del punto  |
| acepta que haya viajado para cambiar |  |  |  |  | 1 | 2 | 1 | 2 | Total |
| Total desacuerdo | 0,281 | 0,178 | -0,358 | 0,008 | 0,049 | 0,314 | 0,214 | 0,542 | ,0756 |
| Parcial desacuerdo | 0,153 | -0,277 | -0,437 | 0,008 | 0,064 | 0,254 | 0,259 | 0,406 | 0,664 |
| Parcial acuerdo | 0,299 | 0,491 | 0,302 | 0,017 | 0,395 | 0,237 | 0,791 | 0,187 | 0,979 |
| Total Acuerdo | 0,266 | -0,582 | 0,290 | 0,019 | 0,493 | 0,194 | 0,855 | 0,133 | 0,987 |
| Total activo | 1,000 |  |  | 0,052 | 1,000 | 1,000 |  |  |  |

 a Normalización Simétrica

Se puede ver en la tabla de examen de los puntos fila , que la calificación de Parcial tiene el mayor peso que las demás y que la contribución de la dimensión a la inercia del punto está representada por la calificación de Total acuerdo.

En la tabla de examen de puntos columna la variable situaciones que se dieron , la que mejor está representada es no cambio nada , mientras que la mayor contribución de los puntos de inercia de la dimensión es la calificación de Cambiaste domicilio.

######  Examen de los puntos columna

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|   | Masa | Puntuación en la dimensión | Inercia | Contribución |   |
|   |   | 1 | 2 |   | De los puntos a la inercia de la dimensión | De la dimensión a la inercia del punto |   |
| Situacionesque se dieron |   |   |   |   | 1 | 2 | 1 | 2 | Total |
| Cambiastede colegio | ,135 | -,584 | ,070 | ,011 | ,251 | ,006 | ,760 | ,007 | ,767 |
| Cambiastede domicilio | ,045 | 1,443 | ,277 | ,018 | ,510 | ,030 | ,976 | ,023 | ,999 |
| cambiaste de amistades | ,030 | ,583 | 1,446 | ,009 | ,056 | ,545 | ,198 | ,763 | ,961 |
| cambiaste de forma de ser | ,186 | -,377 | ,338 | ,010 | ,144 | ,184 | ,507 | ,256 | ,763 |
| no cambio nada | ,605 | ,110 | -,211 | ,004 | ,040 | ,235 | ,300 | ,698 | ,998 |
| Total activo | 1,000 |   |   | ,052 | 1,000 | 1,000 |   |   |   |

 a Normalización Simétrica

En el Gráfico 39 se puede ver como interactúan las características de las variables estudiadas.

La calificación de total acuerdo se encuentra muy cercana a la calificación de cambiaste de colegio , se podría interpretar que los estudiantes que los estudiantes que están de acuerdo con que sus familiares viajaron por mejorar su estilo de vida, se vea reflejado en el cambio de lugar de estudio.

GRAFICO 39

Situaciones que se dieron vs.

acepta que haya tenido que viajar



5.7 ANALISIS DE COMPONENTES PRINCIPALES CATEGÓRICO

**5.7.1 Comparación entre el CATPCA y el ACP**

El análisis típico de componentes principales ( ACP) asume relaciones lineales entre las variables numéricas, mientras que el (CATPCA), permite escalar las variables a diferentes niveles. Las variables categóricas se cuantifican de forma óptima en la dimensionalidad especificada. Como resultado, se pueden modelar relaciones no lineales entre las variables.

5.7.2 Análisis de componentes principales categórico (CATPCA).

El análisis de componentes principales categóricos es una técnica estadística multivariada utilizada para la reducción de datos, que no hace supuestos de normalidad, en la que se estudia p variables observables, *X1, X2,..., Xp,* a través de las cuales se generarán otras k variables no observables, k<p.

Sea el vector aleatorio p variado,  con matriz de varianzas y covarianzas Σdonde 1 2 p y a1, a2, ...,ap son los valores y vectores propios correspondientes a Σconsidere las siguientes combinaciones lineales:



Se puede probar que:



Las componentes principales son las combinaciones lineales Y1, Y2,......,Yp que no están correlacionadas entre sí y cuyas varianzas cumplen que Var(Y1)Var(Y2)Var(Yp)0.

Se define entonces a las componentes principales como sigue:

La primera componente principal es la combinación lineal Y1=a1tX que maximiza  sujeto a .

La segunda componente principal es la combinación lineal Y2=a2tX

que maximiza  sujeto a  y Cov(Y1,Y2) = 0.

En general la i-ésima componente principal es la combinación lineal Yi=aitX que maximiza sujeto a  y Cov(Yi,Yk)=0 para k < i.

Para este método se tiene el siguiente resultado: Sea Σ la matriz de varianzas y covarianzas asociada al vector   Rp y sean (1, a1), (2, a2),.......,(p, ap) los valores y vectores propios correspondientes a la matriz Σ donde 1 2 p , entonces la i-ésima componente principal está dada por:



sujeto a las siguientes condiciones:



La mayor proporción del total de varianza de la población explicada por las componentes principales está dado por:



El número de componentes principales categóricos escogidos dependerá del porcentaje de varianza que se desee explicar**.**

Este procedimiento cuantifica simultáneamente las variables categóricas a la vez que reduce la dimensionalidad de los datos. El análisis de componentes principales categórico se conoce también por el acrónimo CATPCA, del inglés Categorical Principal Components Analysis.

La técnica es más útil cuando un extenso número de variables impide una interpretación eficaz de las relaciones entre los objetos (sujetos y unidades). Al reducir la dimensionalidad, se interpreta un pequeño número de componentes en lugar de un extenso número de variables.

* + 1. **Escalamiento Optimo.**

**Definir escala y ponderación en CATPCA**

Se puede definir el nivel de escalamiento óptimo de las variables del análisis y de las suplementarias. Por defecto, se escalan como líneasSp (ordinales) monotónicas de segundo orden con dos nudos interiores. Asimismo, se puede definir la ponderación para las variables del análisis.

**Ponderación de la variable.** Se puede definir una ponderación para cada variable. El valor especificado debe ser un entero positivo. El valor por defecto es 1.

**Nivel de escalamiento óptimo.** Asimismo, se puede seleccionar el nivel de escalamiento que se utilizará para cuantificar cada variable. Estos pueden ser: LíneaSp nominal, Nominal múltiple, Ordinal, Nominal, Numérico. En este estudio se escogió el nivel Ordinal.

**Ordinal.** El orden de las categorías de la variable observada se conserva en la variable escalada óptimamente. Los puntos de categoría estarán sobre una recta (vector) que pasa por el origen. La transformación resultante se ajusta mejor que la transformación de líneaSp ordinal pero la suavidad es menor.

**Discretización de los componentes principales categóricos**

El cuadro de diálogo Discretización permite seleccionar un método para recodificar las variables. Las variables con valores fraccionarios se agrupan en siete categorías (o en el número de valores diferentes de la variable si dicho número es inferior a siete) con una distribución aproximadamente normal, si no se especifica lo contrario. Las variables de cadena se convierten siempre en enteros positivos mediante la asignación de indicadores de categoría en función del orden alfanumérico ascendente. La discretización de las variables de cadena se aplica a estos enteros resultantes. Por defecto, las variables restantes se dejan inalteradas. A partir de ese momento, se utilizan en el análisis las variables discretizadas.

**Método:** Seleccione entre agrupación, asignación de rangos y multiplicación.

**Asignación de rangos**. La variable se discretiza mediante la asignación de rangos a los casos.

* + 1. **CATPCA de las variables investigadas.**

Para realizar el análisis de componentes categóricos se tomaron en cuenta a todas la variables de estudio.

Los autovalores de la tabla LV siguiente expresan la cantidad de la varianza total que está explicada por cada factor. En este estudio existen 10 autovalores mayores que uno por lo tanto el método de Componentes Principales Categóricos muestra que estos valores explican el 69.23 % de varianza total .

**TABLA LVII**

**Resumen del Modelo**

**Componentes Principales Categóricos**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Alfa de Cronbach | Varianza explicada |   |
| Dimensión |  | Total (Autovalores) | % de la varianza |   |
| 1 | 0,850 | 5,591 | 19,280 |   |
| 2 | 0,709 | 3,167 | 10,922 |   |
| 3 | 0,598 | 2,367 | 8,162 |   |
| 4 | 0,461 | 1,803 | 6,217 |   |
| 5 | 0,302 | 1,411 | 4,866 |   |
| 6 | 0,265 | 1,345 | 4,637 |   |
| 7 | 0,230 | 1,284 | 4,429 |   |
| 8 | 0,094 | 1,100 | 3,793 |   |
| 9 | 0,014 | 1,014 | 3,495 |   |
| 10 | -0,006 | 0,994 | 3,428 |   |
| Total | 0,984 | 20,077 | 69,230 |   |

 a El Alfa de Cronbach Total está basado en los autovalores totales.

Se agruparán las variables de acuerdo a su relación con cada componente de la siguiente forma:

La primera componente principal , identificada con el nombre de lo que sintió el estudiante, estará compuesta por las siguientes variables:

* Sintió desánimo y tristeza
* Sintió deseos de morir
* Sin esperanzas en la vida
* Ganas frecuentes de llorar
* Sintió miedo
* Sintió angustia
* Sintió confusión
* Sintió irritabilidad
* Sentimientos de abandono
* Frustración

La segunda componente principal, definida está formada por las siguientes variables:

* Piensa que fue lo mejor
* Género
* Nombre del Colegio
* Sostenimiento

La tercera componente principal, definida con el nombre de Familiaridad, estará definida con el nombre de y está compuesta por :

* Con quien vive
* Grado de parentesco
* Con quién se quedó viviendo

La cuarta componente principal , denominada como información del estudiante, formada por dos variables:

* Edad
* Año o curso

La quinta componente principal, denominada con el nombre de Tiempo conformada por la variable:

* Hace que tiempo se fue su familiar

La sexta componente principal, definida con el nombre de Preferencia, conformada por :

* Como le hubiese gustado enterarse de que el familiar emigraría.

La séptima componente principal, denominada como Vicios y está compuesta por las siguientes variables:

* Antes de que se fuera el familiar que hacía
* Después de que se fue el familiar que hace

La octava componente principal, definida con el nombre de Reacción a la noticia, está compuesta por las siguientes variables:

* De que forma se enteró
* Que sintió cuando se fue
* Acepta que haya viajado su familiar

La novena componente principal, definida con el nombre de Cambios de acuerdo al trato, está compuesta por :

* Situaciones que se presentaron
* Como lo trata la persona que lo cuida