

APLICACIÓN DE MODELOS MULTIVARIADOS EN EL PLANTEAMIENTO DE UN  
MODELO DE PRONÓSTICO PARA LA CLASIFICACION DE ESTUDIANTES  
CON RIESGO DE MORTALIDAD EN ASIGNATURAS CRÍTICAS DE LAS  
FACULTADES DE INGENIERIAS.

PRESENTADO POR:

VIVIANA CAROLINA SALDARRIAGA TORO

CARLOS ARTURO QUITIAN URREGO

TRABAJO DE GRADO PRESENTADO COMO REQUISITO PARCIAL PARA  
OPTAR AL TÍTULO DE INGENIERO INDUSTRIAL

DIRECTOR

M.Sc. SANTIAGO VASQUEZ ARTUNDUAGA

UNIVERSIDAD TECNOLÓGICA DE PEREIRA

FACULTAD DE INGENIERIA INDUSTRIAL

PEREIRA, FEBRERO DE 2014.

**NOTA DE ACEPTACIÓN**

---

---

---

---

---

---

---

**JURADO**

---

**JURADO**

**Pereira FEBRERO de 2014**

## **AGRADECIMIENTOS**

Le agradecemos a Dios por guiarnos en este proceso iluminando nuestros caminos, por dejarnos luchar por nuestros sueños y por permitirnos culminar esta meta tan importante en nuestras vidas.

Especial agradecimientos a nuestros padres por ser nuestro apoyo en todo este proceso, por confiar, creer y tener fe en nosotros y por ser ellos una gran fuente de inspiración para nosotros.

A nuestro director que con gran paciencia siempre nos guio de la mejor manera.

A otras personas que estuvieron a nuestro lado.

## **ÍNDICE GENERAL**

1. INTRODUCCION	5
1.1. PLANTEAMIENTO DEL PROBLEMA	8
1.2. FORMULACION DEL PROBLEMA	9
1.3. DELIMITACIÓN	9
1.4. JUSTIFICACIÓN	10
2. OBJETIVO GENERAL	12
2.1. OBJETIVOS ESPECÍFICOS	13
3. MARCO CONCEPTUAL	13.
3.1.1. ANALISIS PREVIO DE LOS DATOS	14
3.1.2. ANALISIS EXPLORATORIO Y GRAFICO DE LOS DATOS	18
3.1.2.1. Histogramas	18
3.1.2.2. Gráficos de Normalidad	19
3.1.2.3. Gráfico de Caja y Bigotes	20
3.1.2.4. Gráfico de dispersión	21
3.1.3. ANALISIS DE DATOS AUSENTES	22
3.1.3.1. Detección y Diagnóstico de los Datos Ausentes	22
3.1.4. ANALISIS Y DETECCION DE VALORES ATÍPICOS	23
3.1.4.1. Detección Univariante de Valores Atípicos	24
3.1.4.2. Detección Bivariante de Valores Atípicos	24
3.1.4.3. Detección Multivariante de Valores Atípicos	24
3.2. MARCO TEORICO	25
3.2.1. DESERCIÓN ESTUDIANTIL	25
3.2.2. REPITENCIA	27
3.2.3. REZAGO	28
3.2.4. MUESTREO	28

3.2.4.1. TIPOS DE MUESTREO	29
3.2.5. ANALISIS MULTIVARIADOS	32
3.2.5.1. MODELOS DE CLASIFICACION. MODELOS LOGIT Y PROBIT	33
3.2.6. REGRESION LOGISTICA	34
3.2.7. REGRESION LOGISTICA MULTINOMIAL	38
4. ESTADO DEL ARTE TRABAJO DE GRADO	40
5. TIPO DE INVESTIGACIÓN	51
5.1. RECURSOS DE LA INVESTIGACIÓN	53
6. MORTALIDAD ACADÉMICA	57
6.1. ANALISIS DESCRIPTIVO	58
7. APLICACIÓN DE MODELOS MULTIVARIADOS EN EL PLANTEAMIENTO DE UN MODELO DE PRONÓSTICO PARA LA CLASIFICACION DE ESTUDIANTES CON RIESGO DE DESERCIÓN ESTUDIANTIL CRÍTICAS DE LAS FACULTADES DE INGENIERIAS	72
8. GRAFICAS DE CADA UNA DE LAS PREGUNTAS	80
9. VARIABLES DE IMPORTANCIA EN LA MORTALIDAD ACADEMICA Y EN EL DESEMPEÑO DE LOS ESTUDIANTES EN LA UNIVERSIDAD	94
9.1 Variables de tipo socioeconómico	96
9.2 Variables de tipo académico	97
9.3 Variables de Interpretación personal	97
10. MUESTREO	99
10.1 TIPO DE MUESTREO Y CLASIFICACIÓN	99
10.2 METODOLOGIA DEL MUESTREO	101
11. ANÁLISIS MULTIVARIADO	105
11.1 TÉCNICA DE ANÁLISIS MULTIVARIADO	106
11.2 METODOLOGIA DEL ANALISIS MULTIVARIADO	107
12. RESULTADOS	109
13. ANALISIS REGRESION LOGISITICA BINARIA	134

14. RESULTADOS	138
15. CREACION DE UN MODELO DE PRONOSTICO	148
15.1 MODELO LOGÍSTICO	150
16. RESULTADOS DE NUESTRO MODELO	151
17. CONCLUSIONES	155
18. RECOMENDACIONES	157
19. BIBLIOGRAFIA Y WEBGRAFIA	159

## LISTADO DE GRAFICAS

<b>Figura 1.</b> Histograma con curva de probabilidad	19
<b>Figura 2.</b> Gráfico normal de probabilidad	19
<b>Figura 3.</b> Gráfico de cajas y bigotes	21
<b>Figura 4.</b> Gráfico de dispersión	22
<b>Figura 5.</b> Gráfica de valores atípicos	24
<b>Figura 6: Grafica Matemáticas 1 Ingeniería Eléctrica</b>	61
<b>Figura 7: Grafica Matemáticas 1 Ingeniería Física</b>	62
<b>Figura 8: Grafica Matemáticas 2 Ingeniería Física</b>	64
<b>Figura 9: Grafica Matemáticas 1 Ingeniería en Sistemas y Computación Matemáticas</b>	66
<b>Figura 10: Grafica Informática 1 Ingeniería en Sistemas y Computación</b>	67
<b>Figura 11: Grafica Matemáticas 1 Ingeniería Mecánica</b>	68
<b>Figura 12: Grafica Matemáticas 2 Ingeniería Mecánica</b>	69
<b>Figura 13: Grafica Matemáticas 1 Facultad de Ingeniería</b>	71
<b>Figura 14: Grafica Resultado encuesta Género</b>	80
<b>Figura 15: Grafica Resultado encuesta Rango de edad</b>	80
<b>Figura 16: Grafica Resultado encuesta Estrato socioeconómico</b>	80
<b>Figura 17: Grafica Resultado encuesta Departamento procedencia</b>	81
<b>Figura 18: Grafica Resultado encuesta Naturaleza colegio</b>	81
<b>Figura 19: Grafica Resultado encuesta característica de colegio</b>	81
<b>Figura 20: Grafica Resultado encuesta ha tenido dificultad para pagar la universidad</b>	82
<b>Figura 21: Grafica Resultado encuesta se le presenta cruces con el horario de trabajo</b>	82
<b>Figura 22: Grafica Resultado encuesta nivel educativo del padre</b>	82

<b>Figura 23: Grafica Resultado encuesta nivel educativo de la madre</b>	<b>83</b>
<b>Figura 24: Grafica Resultado encuesta en que se desempeña el padre</b>	<b>83</b>
<b>Figura 25: Grafica Resultado encuesta en que se desempeña la madre</b>	<b>83</b>
<b>Figura 26: Grafica Resultado encuesta presenta alguna discapacidad</b>	<b>84</b>
<b>Figura 27: Grafica Resultado encuesta tiene personas a cargo (Hijos)</b>	<b>84</b>
<b>Figura 28: Grafica Resultado encuesta como es el acompañamiento de la familia</b>	<b>84</b>
<b>Figura 29: Grafica Resultado encuesta hace cuanto se graduó del bachillerato</b>	<b>85</b>
<b>Figura 30: Grafica Resultado encuesta como es el tiempo que le dedica al estudio y cree que este método funciona</b>	<b>85</b>
<b>Figura 31: Grafica Resultado encuesta frente a su carrera en estos momentos usted esta</b>	<b>86</b>
<b>Figura 32: Grafica Resultado encuesta cree usted que está en el programa adecuado</b>	<b>86</b>
<b>Figura 33: Grafica Resultado encuesta tenia conocimientos previos de la carrera</b>	<b>86</b>
<b>Figura 34: Grafica Resultado encuesta por qué eligió esta carrera</b>	<b>87</b>
<b>Figura 35: Grafica Resultado encuesta se ha cambiado de carrera</b>	<b>87</b>
<b>Figura 36: Grafica Resultado encuesta si la respuesta fue positiva elija un porque</b>	<b>88</b>
<b>Figura 37: Grafica Resultado encuesta la relación con los profesores es</b>	<b>89</b>
<b>Figura 38: Grafica Resultado encuesta la relación con los compañeros es</b>	<b>89</b>
<b>Figura 39: Grafica Resultado encuesta realizo usted el examen de orientación profesional</b>	<b>90</b>
<b>Figura 40: Grafica Resultado encuesta al presentar el examen</b>	<b>90</b>



<b>Figura 41: Grafica Resultado encuesta como califica el programa de estudios</b>	90
<b>Figura 42: Grafica Resultado encuesta como califica el pensum académico</b>	91
<b>Figura 43: Grafica Resultado encuesta se identifica con el programa que está estudiando</b>	91
<b>Figura 44: Grafica Resultado encuesta ha tenido dificultad para cursar el semestre</b>	91
<b>Figura 45: Grafica Resultado encuesta la respuesta fue si ¿Cuál fue la dificultad?</b>	92
<b>Figura 46: Grafica Resultado encuesta a que programa pertenece</b>	92
<b>Figura 47: Grafica Resultado encuesta categoría de respuesta</b>	93
<b>Figura 48: Grafica pantallazo 1 resultados Muestra extraída del programa SPSS</b>	103
<b>Figura 49: Grafica pantallazo 2 base de datos nueva del programa SPSS</b>	104
<b>Figura 51: pantallazo 3 observación grupo de predicciones probabilísticas pasó 6.</b>	147
<b>Figura 52: pantallazo 4 seleccionar el 70 % de los datos estos fueron los resultados</b>	152
<b>Figura 51: pantallazo 5 seleccionar el 30 % de los datos</b>	153

## LISTADO DE TABLAS

<b>Tabla 1: Ingeniera Eléctrica Matemáticas 1</b>	60
<b>Tabla 2: Ingeniería Física Matemáticas 1</b>	62
<b>Tabla 3: Ingeniería Física Matemáticas 2</b>	63
<b>Tabla 4: Ingeniería en Sistemas y Computación Matemáticas 1</b>	65
<b>Tabla 5: Ingeniería en Sistemas y Computación, Informática 1</b>	66
<b>Tabla 6: Ingeniería Mecánica, Matemática 1</b>	68
<b>Tabla 7: Ingeniería Mecánica, Matemática 2</b>	69
<b>Tabla 8: Facultad de Ingenierías, Tabla de mortalidad</b>	70
<b>Tabla 9: muestra el ajuste de bondad del modelo y su significancia la cual es muy buena</b>	110
<b>Tabla 10: contraste de la razón de probabilidad</b>	110
<b>Tabla 11: Estimación de los parámetros</b>	113
<b>Tabla 12: Resumen de la encuesta propuesta a cada uno de los participantes</b>	128
<b>Tabla 13: Observaciones ordenados de mayor a menor y formar</b>	138
<b>Tabla 14: Tabla De Clasificación.</b>	139
<b>Tabla 15: Resumen De Los Modelos</b>	140
<b>Tabla 16: Prueba De Hosmer Y Lemeshow</b>	140
<b>Tabla 17: Tabla De Clasificación.</b>	141
<b>Tabla 18: Variables En La Ecuación</b>	142

## **1. INTRODUCCION**

El Proyecto pretende clasificar las principales variables que son causantes de mortalidad académica en materias de facultades objeto de estudio (facultad de ingenierías) que a través del tiempo, de registros históricos y de estudios previos, para las variables que han sido catalogadas y estudiadas para los años anteriores y los altos niveles de mortalidad académica estudiadas hasta el año 2013-1.

Los estudios anteriores se realizaron basados en la información suministrada por la división de sistemas y fundamentalmente por el observatorio académico adjudicado a la facultad de ingeniería industrial.

## 1.1. PLANTEAMIENTO DEL PROBLEMA:

La mayor importancia para la Universidad es aumentar la calidad, disminuir la deserción con el fin de incrementar su competitividad con la relación a las otras Universidades Públicas de Colombia. Una de las principales causas de deserción es la mortalidad académica para lo cual se deben clasificar las causas de mortalidad estudiantil para la facultad de Ingenierías en los periodos anuales del 2011 hasta el 2013-1. Las variables estudiadas anteriormente son rendimiento académico, ingreso familiar, sexo, tipo de colegio, resultados pruebas icfes y la edad.

Por ejemplo el rendimiento académico depende de las materias claves dentro de los programas de la facultad de ingeniería de la Universidad, las consecuencias del problema de las materias críticas se reflejan en los índices de la mortalidad académica generando deserción estudiantil y rezago del programa.

La importancia del actual proyecto que permitirá a las entidades encargadas de velar por el normal desempeño académico de los estudiantes, tomar medidas no solo correctivas sino también preventivas, para en un futuro ojala no lejano disminuir los índices negativos de mortalidad. Para ayudar a prevenirlas generamos un modelo estadístico binomial, el cual selecciona según las variables estudiadas quienes tendrían mayor probabilidad de perder las materias de mortalidad académica para crear métodos de ayuda desde antes de que se produzca la pérdida de las materias.

Cada programa en la Universidad presenta dentro de sus currículos materias que normalmente son de difícil aprobación para la mayoría de los estudiantes de dicho programa, específicamente dentro de la facultad de ingenierías los problemas tienen que ver con las ciencias básicas que incluyen las materias Matemáticas I, II, III, IV y Álgebra lineal.

Los índices de mortalidad y repitencia en estas materias si bien se han disminuido siguen siendo altos, este problema finalmente se refleja en anomalía en el desarrollo de la carrera universitaria para los estudiantes y altos índices de deserción lo que es preocupante para la Universidad y sus facultades.

Es importante mencionar que estas asignaturas críticas en las facultades antes relacionadas, fueron identificadas por investigaciones realizadas por el Observatorio Académico de la Universidad y anteriores trabajos de grado, adscrito a la Facultad de Ingeniería Industrial.

Con el fin de tomar medidas preventivas más que correctivas por parte del observatorio y de los distintos departamentos encargados y relacionados con la problemática antes mencionada, es fundamental el desarrollo de la presente

investigación y la puesta a punto de las distintas estrategias aquí planteadas con el fin de disminuir la mortalidad académica.

### **FORMULACION DEL PROBLEMA**

¿Cómo identificar los factores que influyen en la mortalidad académica en la facultad de ingeniería entre los periodos anuales 2011-1 y 2013-1 en la Universidad Tecnológica de Pereira y crear un modelo de pronósticos estadísticos para ayudar a su prevención?

#### **1.2. DELIMITACIÓN**

La Investigación es realizada en la Universidad Tecnológica de Pereira con la información y los estudios anteriores realizados por el observatorio estudiantil y trabajos de grado de la facultad de ingeniería Industrial desde el primer semestre del 2011 hasta el primer semestre del 2013.

**TEMPORAL:** La elaboración de la evaluación se lleva a cabo un tiempo aproximado de (3) cuatro meses, entre Octubre y Diciembre del 2013.

**GEOGRAFICA:** La Investigación es realizada en la Universidad Tecnológica de Pereira con la información y los estudios anteriores realizados por el observatorio estudiantil de la facultad de ingeniería Industrial desde el primer semestre del 20011 hasta el primer semestre del 2013.

**DEMOGRAFICA:** La información para la clasificación de las causas de deserción estudiantil de la facultad de ingenierías entre los años 2011 y 2013 en la universidad tecnológica de Pereira se obtiene del observatorio de la facultad de Ingeniería Industrial, los archivos Universitarios y estudios anteriores.

### 1.3. JUSTIFICACIÓN

Se define la deserción estudiantil como el hecho de que un número de estudiantes matriculados no siga la trayectoria normal del programa académico, bien sea por retirarse de ella o por demorar más tiempo del previsto en finalizarla, por repetir cursos o por retiros temporales. El abandono o la interrupción pueden ser voluntarios o forzados. También puede presentarse el cambio de carrera dentro de la misma institución o cambio de institución donde puede continuar con la misma carrera o con otra.

La mortalidad académica es una de las principales causas de deserción estudiantil, se entiende por mortalidad académica cuando los estudiantes pierden mucho las materias y estas se consideran materias de gran dificultad como son las de ciencias básicas.

Desde el punto de vista el fenómeno comprende a quienes no siguen el curso normal de la carrera bien sea por cancelar su matrícula o renovar su matrícula. Los factores que influyen más en el fenómeno de la deserción académica son familiares, individuales, propios del sector educativo, económico social y culturales.

En los estudios anteriores se utilizaron filtros y pruebas para determinar los factores que influyen en el fenómeno de la mortalidad académica, otro proceso que puede afectar este fenómeno es la adaptación a la Universidad, a la carrera e institución, y el paso del colegio a la universidad, como también las perspectivas y expectativas que tiene el estudiante de la universidad. También podemos ver que influyen mucho en el fenómeno de la deserción, los resultados de las pruebas icfes que son solicitados para cada ingreso a una respectiva carrera.

Los estudios sobre factores asociados a la mortalidad académica estudiantil, mencionados en el plan sectorial (Gobierno Nacional, MEN) han encontrado que en la educación superior existen factores comunes con los demás niveles educativos, tales como los socioeconómicos, pérdida del sentido de educación, dificultad para combinar trabajo y estudio, embarazos, conflictos con los docentes, autoritarismo, falta de afecto y agresividad entre otros.

Por lo cual el desarrollo del proyecto genera importancia para la Universidad Tecnológica para determinar los altos índices de mortalidad académica, y clasificar las causas para realizar planes de prevención y contingencia para el presente y futuros de la Universidad mejorando su calidad y generando mayor competitividad con las demás Universidades Públicas. Generando estrategias para atacar el problema de raíz y contrarrestar la problemática ya que la mortalidad genera dificultades y problemas adicionales en la carrera de cada estudiante como lo es por ejemplo el rezago en su plan de estudio, problema que en el largo plazo generalmente se traduce en deserción académica.

Se pueden diseñar diferentes estrategias que minimicen este fenómeno, con esto se lograría un gran beneficio tanto para los estudiantes como para la Universidad y analizar si los procesos que está realizando la Universidad en el presente para disminuir el efecto de la mortalidad académica generan realmente resultados y disminuye la deserción.

Generamos mejor calidad académica para la universidad al solucionar el problema de mortalidad permitiendo tener personas graduadas con mejor nivel de conocimiento, pasión y entrega, que sean competitivos en el ámbito profesional y que la Universidad suba su estatus dentro del ámbito académico nacional.

## **2. OBJETIVOS DE LA INVESTIGACION**

### **2.1. OBJETO DE ESTUDIO**

El fenómeno de la mortalidad académica que presentan los estudiantes en materias denominadas críticas en las facultades de Ingenierías en jornada diurna.

### **2.2. OBJETIVO GENERAL**

Identificar las variables y los factores que influyen en la mortalidad académica en materias críticas de la Facultad de Ingenierías entre los periodos anuales 2009 y 2013-1 en la Universidad Tecnológica de Pereira. Y crear un modelo que permita clasificar los estudiantes a futuro en cada uno de los diferentes desempeños que pone la universidad.

### **2.3. OBJETIVOS ESPECÍFICOS**

- Analizar y revisar el término de mortalidad académica.
- Depurar la bases de datos, los estudios anteriores y los archivos del observatorio e históricos para analizar las causas de deserción estudiantil que generar la mortalidad académica.
- Conceptualizar, operacionalizar y clasificar las causas de la Mortalidad Académica obtenidas en los estudios anteriores.
- Realizar un análisis descriptivo de la información asociada a la base de datos sobre los aspectos a estudiar y que pongan en evidencia el mayor efecto causado por las variables que generan el fenómeno para solucionar la problemática.
- Realizar un análisis por medio de algunas técnicas multivariadas de clasificación entre conjunto de variables.
- Identificar y clasificar las variables de mayor peso en la mortalidad académica en la facultad de ingeniería y ciencias de la salud.
- Crear un modelo que permita clasificar los estudiantes a futuro en cada uno de los diferentes desempeños que pone la universidad.



### 3. MARCO CONCEPTUAL

#### 3.1.1 ANALISIS PREVIO DE LOS DATOS

Antes de aplicar cualquier técnica de análisis multivariante es preciso realizar un análisis previo de los datos de los que se dispone. Es necesario examinar las variables individuales y las relaciones entre ellas, así como evaluar y solucionar problemas en el diseño de la investigación y en la recolección de datos tales como el tratamiento de la información faltante (o datos ausentes) y la presencia de datos anómalos (o atípicos).

La primera tarea que suele abordarse es el análisis exploratorio y gráfico de datos. Estas técnicas permiten el examen de las características de la distribución de las variables implicadas en el análisis, las relaciones bivariantes (y multivariantes) entre ellas y el análisis de las diferencias entre grupos.

La segunda tarea que suele llevarse a cabo antes de aplicar cualquier técnica multivariante es el análisis de los datos ausentes. Cualquier recogida y procesamiento de datos presenta problemas que van a impedir obtener información de algunos de los elementos de la población en estudio.

La tercera tarea para aplicar cualquier técnica multivariante es la detección de valores atípicos. Se trata de detectar la existencia de observaciones que no siguen el mismo comportamiento que el resto. Los casos atípicos suelen deberse a errores en el procedimiento a la hora de introducir los datos o de codificarlos. También pueden ser consecuencia de acontecimientos anormales que hacen destacar determinadas observaciones.

Una última tarea previa a la aplicación de las técnicas multivariantes es la comprobación de los supuestos subyacentes en los métodos multivariantes. Estos supuestos suelen ser el contraste de la normalidad de todas y cada una de las variables que forman parte del estudio, el testeo de la linealidad de las relaciones entre variables que intervienen en el estudio, la comprobación de la homocedasticidad de los datos que consiste en ver que la variación de la variable dependiente que se intenta explicar a través de las variables independientes no se concentra en un pequeño grupo de valores independientes (se tratará por lo tanto de ver la igualdad de las varianzas para los datos agrupados según valores similares de la variable dependiente), y la comprobación de la multicolinealidad o existencia de relaciones entre las variables independientes. A veces también es necesario contrastar la ausencia de correlación serial de los residuos o auto correlación, que consiste en asegurar que cualquiera de los errores de predicción no está correlacionado con el resto.

### **3.1.2. ANALISIS EXPLORATORIO Y GRAFICO DE LOS DATOS**

Actualmente se utilizan las novedosas técnicas de análisis exploratorio de datos, mediante las cuales se pueden analizar exhaustivamente y detectar las posibles anomalías que presentan las observaciones. J.W. Tukey ha sido uno de los pioneros en la introducción de este tipo de análisis. Los estadísticos descriptivos más habitualmente utilizados han sido la media y la desviación típica. Sin embargo, estos tienen sentido cuando la distribución es aproximadamente normal o, al menos, simétrica y unimodal. Pero las variables objeto de estudio no siempre cumplen estos requisitos y se hace necesario un análisis más riguroso de los datos. Se recomienda iniciar un análisis exploratorio de datos con gráficos que permitan visualizar su estructura, por ejemplo para datos cuantitativos es aconsejable comenzar con el gráfico de tallos y hojas o histograma digital. El siguiente paso es examinar la posible presencia de normalidad, simetría y valores atípicos (outliers) en el conjunto de datos. Para ello suele utilizarse los gráficos de caja y bigotes. No obstante los diagramas de bigotes siempre deben ir acompañados de los histogramas digitales (o gráficos de hojas y tallos), ya que los primeros no detectan la presencia de distribuciones multimodales.

#### **3.1.2.1 Histograma.**

Histograma es una representación gráfica de una variable en forma de barras, el eje vertical se representan las frecuencias, y en el eje horizontal los valores de las variables, señalando las marcas de clase, es decir, la mitad del intervalo en el que están agrupados los datos.

Se utiliza para el estudio de variables continuas, como franjas de edades o alturas de la muestra, y, sus valores se agrupan en clases, de una forma continua. En los casos de datos cuantitativos (no numéricos), es preferible un diagrama de sectores.

Tipos de histograma

#### **Diagramas De Barras Simples**

Representa la frecuencia simple (absoluta o relativa) mediante la altura de la barra la cual es proporcional a la frecuencia simple de la categoría que representa.

#### **Diagramas De Barras Compuestas**

Se usa para representar la información de una tabla de doble entrada o sea a partir de dos variables, las cuales se representan así; la altura de la barra representa la frecuencia simple de las modalidades o categorías de la variable y esta altura es proporcional a la frecuencia simple de cada modalidad.

## Diagramas De Barras Agrupadas

Se usa para representar la información de una tabla de doble entrada o sea a partir de dos variables, el cual es representado mediante un conjunto de barras como se clasifican respecto a las diferentes modalidades.

## Polígono De Frecuencias

Es un gráfico de líneas que de las frecuencias absolutas de los valores de una distribución en el cual la altura del punto asociado a un valor de las variables es proporcional a la frecuencia de dicho valor.

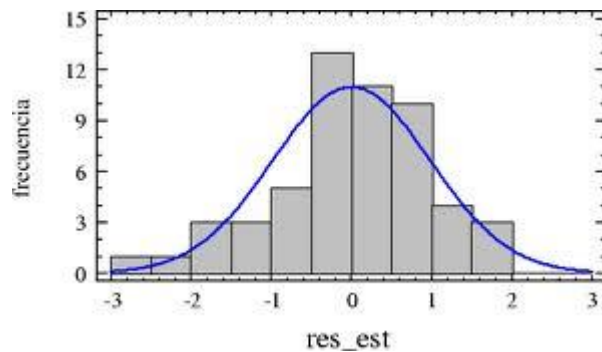
## Ojiva Porcentual

Es un gráfico acumulativo, el cual es muy útil cuando se quiere representar el rango porcentual de cada valor en una distribución de frecuencias.

En los gráficos las barras se encuentran juntas y en la tabla los números poseen en el primer miembro un corchete y en el segundo un paréntesis.

**Figura 1.** Histograma con curva de probabilidad.

NUM\_DIRECTIVOS

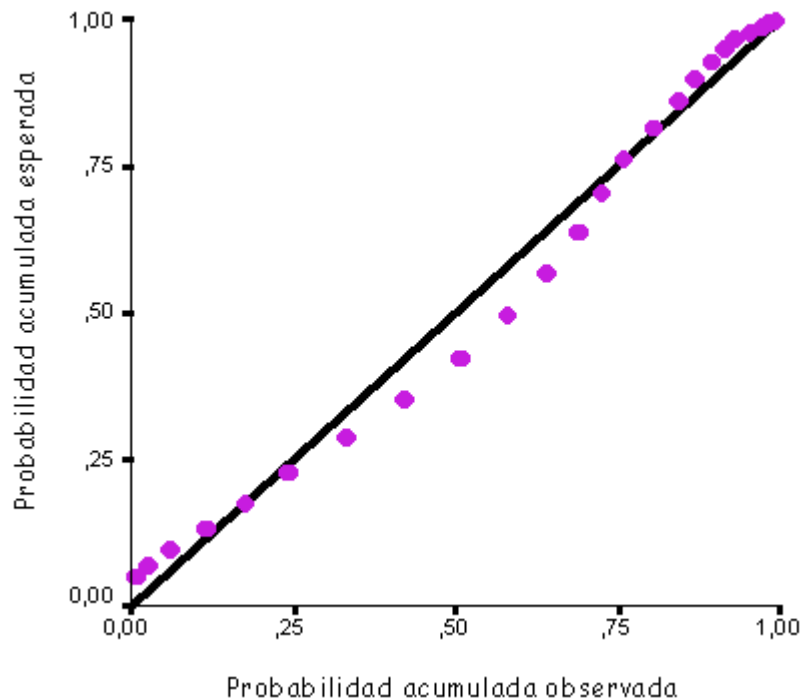


### 3.1.2.2 Gráficos de Normalidad

Los gráficos de probabilidad **P-P** suelen utilizarse para determinar si la distribución de una variable coincide con una distribución especificada. Si la variable seleccionada coincide con la distribución de estudio, los puntos se agruparán en torno a una línea recta.

**Figura 2.** Gráfico normal de probabilidad

Prob acum observada



### 3.1.2.3. Gráfico de Caja y Bigotes

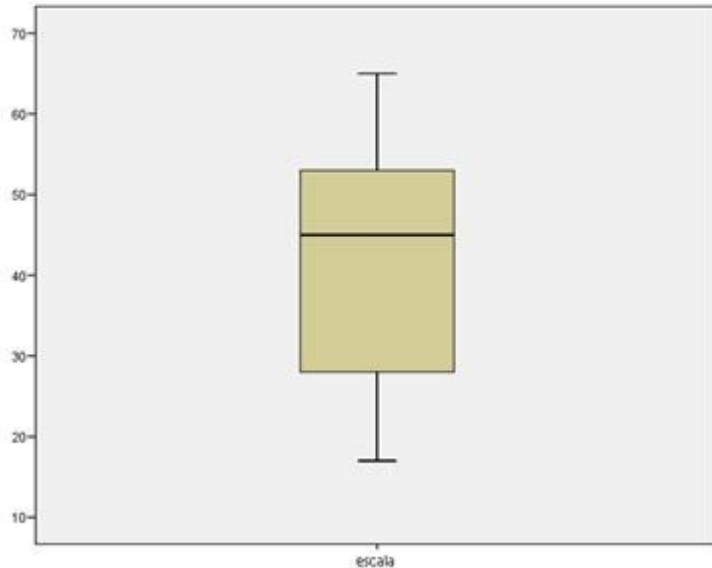
Este tipo de gráfico permite analizar y resumir un conjunto de datos univariante dado. Esta herramienta de análisis exploratorio de datos permite estudiar la simetría de los datos, detectar valores atípicos y vislumbrar un ajuste de los datos a una distribución de frecuencia determinada.

El gráfico de caja y bigotes divide los datos en cuatro áreas de igual frecuencia, una caja central dividida en dos áreas por una línea vertical y otras dos áreas representadas por segmentos horizontales (bigotes) que parten del centro de cada lado vertical de la caja. La caja central encierra el 50% de los datos. El sistema dibuja la mediana como una línea vertical en el interior de la caja. Si esta línea está en el centro de la caja no hay asimetría en la variable. El bigote 24

de la izquierda tiene un extremo en el cuartil Q1, y el otro en el valor dado por el primer cuartil menos 1.5 veces el rango intercuartílico, esto es  $Q1 - 1.5 \cdot (Q3 - Q1)$ . El bigote de la derecha tiene un extremo en el tercer cuartil y el otro en el valor dado por el tercer cuartil más 1.5 veces el rango intercuartílico, esto es,  $Q3 + 1.5 \cdot (Q3 - Q1)$ . El sistema considera valores atípicos (outliers) los que se encuentran a la izquierda del bigote izquierdo y a la derecha del bigote derecho.

**Figura 3.** Gráfico de cajas y bigotes

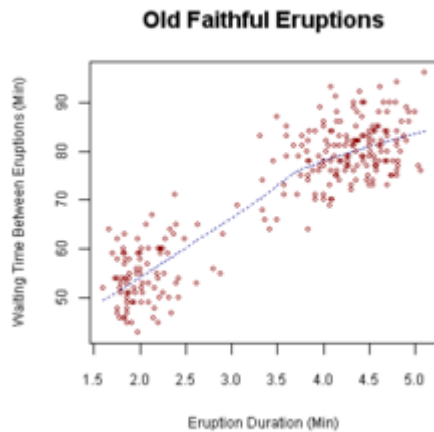
CAP\_INSTALADA



#### **3.1.2.4. Gráfico de dispersión**

Permite ver la relación entre dos o más variables. Está formado por puntos cuyas coordenadas cartesianas son los pares de valores de dos variables cuya relación se quiere estudiar, representada una en el eje vertical y la otra en el eje horizontal. El posicionamiento de los puntos del gráfico de dispersión define la relación entre las variables. Si se sitúan alrededor de una recta, existe correlación lineal entre las variables. Si los puntos siguen una pauta no lineal, la relación entre las variables no puede definirse como lineal. Si la nube de puntos es aleatoria y dispersa, no existe relación alguna entre las variables, por ejemplo, el gráfico 4 muestra un diagrama de dispersión para un conjunto de variables.

**Figura 4.** Gráfico de dispersión



### **3.1.3. ANALISIS DE DATOS AUSENTES**

Cuando se aplica un método de análisis Multivariado sobre los datos disponibles puede ser que no exista información para determinadas observaciones y variables. Estamos entonces ante valores ausentes o valores missing. La presencia de esta información faltante puede deberse a un registro defectuoso de la información, a la ausencia natural de la información buscada o a una falta de respuesta (total o parcial).

#### **3.1.3.1. Detección y Diagnóstico de los Datos Ausentes**

Tras observar la presencia de datos ausentes en una distribución, será necesario detectar si estos se distribuyen aleatoriamente. La simple presencia de datos ausentes no implica que su falta sea crítica para el análisis estadístico. Será necesario detectar que el efecto de los datos ausentes es importante mediante pruebas formales de aleatoriedad.

*Una primera prueba para valorar los datos ausentes para una única variable Y consiste en formar dos grupos para Y, los que tienen datos ausentes y los que no los tienen. A continuación, para cada variable X distinta de Y, se realiza un test para determinar si existen diferencias significativas entre los dos grupos de valores determinados por la variable Y (ausentes y no ausentes) sobre la 26*

variable X. Si se considera como Y cada una de las variables del análisis y repitiendo el proceso anterior, se encuentra que todas las diferencias son no significativas, se puede concluir que los valores missing obedecen a un *proceso completamente aleatorio* y por lo tanto se pueden realizar análisis estadísticos fiables con nuestras variables *imputando los datos ausentes* por métodos formales para tal caso.

Si un porcentaje bastante alto de las diferencias son no significativas, puede considerarse que los datos missing obedecen a un *proceso aleatorio* (no completamente aleatorio) que también permitirá realizar análisis estadísticos fiables con las variables, previa *imputación de la información faltante*, aunque con menos fiabilidad que en el caso anterior.

*Una segunda prueba para valorar los datos ausentes es la prueba de las correlaciones dicotomizadas.* Para realizar esta prueba, para cada variable del análisis se construye una variable dicotomizada asignando el *valor cero* a los valores ausentes y *uno* a los valores presentes, posteriormente se haya la matriz de correlaciones acompañada de los contrastes de significancia de cada coeficiente de correlación de la matriz. Las correlaciones indican el grado de asociación entre los valores perdidos sobre cada par de variables (bajas correlaciones indican aleatoriedad en el par de variables), con lo que se puede concluir que si los elementos de la matriz de correlaciones son no significativos, los datos ausentes son completamente aleatorios. Si existe alguna correlación significativa y la mayor parte son no significativas, los datos ausentes pueden considerarse aleatorios. En ambos casos puede realizarse análisis estadísticos, previa imputación de la información faltante.

*Una tercera prueba para valorar los datos ausentes es el test conjunto de aleatoriedad Little*, contraste formal basado en la Ji-cuadrado, cuyo p-value indica si los valores perdidos constituyen o no un conjunto de números aleatorios.

### **3.1.4. ANALISIS Y DETECCION DE VALORES ATÍPICOS**

Los casos atípicos son observaciones aisladas cuyo comportamiento se diferencia claramente del comportamiento medio del resto de las observaciones.

Existe una primera categoría de casos atípicos formado por aquellas observaciones que provienen de un error de procedimiento, como por ejemplo un error de codificación, error en la entrada de los datos, etc. Estos datos atípicos, si no se detectan mediante filtrado, deben eliminarse o recodificarse como datos ausentes.

Una segunda categoría, contempla aquellas observaciones que ocurren como consecuencia de un acontecimiento extraordinario existiendo una explicación para su presencia en la muestra.

Una tercera categoría comprende las observaciones extraordinarias para las que el investigador no tiene explicación. Normalmente estos datos se eliminan en el análisis.

Una cuarta categoría, la forman las observaciones que se sitúan fuera del rango ordinario de valores de la variable. Suelen denominarse valores extremos y se eliminan del análisis si se observa que no son elementos significativos para la población.

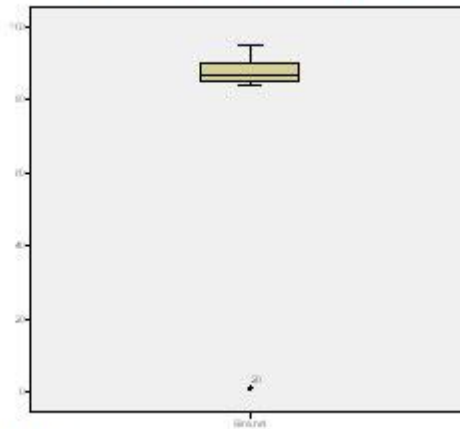
### 3.1.4.1. Detección Univariante de Valores Atípicos

Para este tipo de detección de valores atípicos en un contexto univariante, suele utilizarse los diagramas de caja y bigotes, en este gráfico los valores atípicos se presentan como puntos aislados en los extremos de los bigotes.

A continuación la figura 5 brinda una ilustración de cómo se representa un valor atípico en un gráfico de caja y bigotes.

**Figura 5.** Gráfica de valores atípicos

**Gráfico 3. Prueba de caja y bigote sin túnel**



Fuente: elaboración propia.

### VALOR ATIPICO

#### 3.1.4.2. Detección Bivariante de Valores Atípicos

Cuando se trata de detectar casos atípicos en un contexto bivariante, puede utilizarse el gráfico de caja y bigotes múltiple, que representa distintos gráficos de una variable, para diferentes niveles de otra.

#### 3.1.4.3. Detección Multivariante de Valores Atípicos

Cuando se trata de detectar casos atípicos en un contexto multivariante, pueden utilizarse estadísticos basados en distancias, para detectar los puntos influyentes. La *distancia  $D^2$  de Mahalanobis* es una medida de la distancia de cada observación en un espacio multidimensional respecto del centro medio de las observaciones. El *estadístico DFITS* mide la influencia de cada observación en caso de ser eliminada del análisis. La *influencia (leverage)* mide la influencia de cada observación.



## **3.2. MARCO TEORICO**

Durante la carrera universitaria los estudiantes se ven afectados por numerosas circunstancias que terminan por entorpecer de manera directa o indirecta su nivel académico, estas circunstancias se reflejan por ejemplo en la mortalidad de determinadas materias y por ende en su repitencia, estos dos factores a su vez hacen que el estudiante inicie un proceso de rezago que mientras más aumenta más lo desmotiva haciéndolo finalmente desertar y dejar sus estudios. Es objeto de esta investigación determinar esos aspectos, circunstancias o factores que infieren en el normal desarrollo de los estudiantes.

Para esto es necesario conocer y definir los principales aspectos objeto de este estudio, entenderlos y determinar cuáles son las variables que aumentan la probabilidad de que se presenten y afecten a los estudiantes.

Además de definir dichos aspectos, se analizarán también otros conceptos como lo son la deserción académica que definitivamente guarda una gran relación con la mortalidad, la repitencia y el rezago y de la cual se podrán determinar variables de peso de las que se hará uso en el análisis de la presente investigación.

### **3.2.1. DESERCIÓN ESTUDIANTIL**

Si bien no es el objeto de estudio de este trabajo, la deserción está demasiado ligada a los aspectos de mortalidad, repitencia y rezago ya que es casi una consecuencia de los mismos.

Es por eso que es necesario entenderla y de hecho utilizar variables que normalmente se estudian en su caracterización y que posiblemente ayudaran a analizar de una mejor manera el fenómeno de mortalidad en la Universidad.

“El retiro estudiantil puede presentarse en dos modalidades, por egreso o graduación, cuando el estudiante termina sus estudios o finalmente recibe su título académico y por deserción que es el abandono de los estudios sin haberlos terminado” . Dado este concepto puede definirse entonces la deserción académica como el abandono de los estudios por influencias positivas o negativas de aspectos internos o externos y la cual puede ser forzosa o involuntaria y voluntaria.

“La deserción forzosa o involuntaria tiene que ver cuando el estudiante es retirado de su carrera por decisión institucional fundamentada en los reglamentos vigentes, por otra parte la deserción voluntaria puede adoptar la forma de renuncia a la carrera por parte del estudiante o del abandono no informado a la institución de educación superior” .

Dentro de la deserción se pueden además diferenciar dos aspectos, estos tienen que ver con que la deserción se dé en la carrera que se está cursando o en general para la universidad. La primera no necesariamente tiene que ver con la

segunda ya que un estudiante puede abandonar una carrera y entrar en otra dentro de la misma universidad, es lo que se conoce como transferencia interna o traspaso de carrera. A la vez el abandono de una carrera en una Universidad puede concluir en la continuación de la misma en otra institución.

En la siguiente figura se ilustra la clasificación de la deserción y los posibles caminos que puede tener.

Figura 7. Clasificación de la deserción Tomada de “Modelos de análisis de la deserción estudiantil en la educación superior”.

El análisis de la deserción puede ser agrupado entonces en cinco grandes grupos que son:

- Psicológicos
- Económicos
- Sociológicos
- Organizacionales
- De interacciones

Otros textos resaltan como principales en cuanto a la influencia en la deserción estudiantil los siguientes aspectos:

- Personales: Características personales de los estudiantes, por ejemplo la falta de actitud de logro en el crecimiento profesional. La incompatibilidad del tiempo dedicado al trabajo y a los estudios. El poco interés por el estudio en general, por la carrera y por la institución en la que estudia. La poca acogida que le brinda la Universidad. Las expectativas del estudiante respecto de la importancia de la carrera que estudia.

- Institucionales y pedagógicos: La deficiente orientación vocacional recibida antes de ingresar a la carrera que provoca que los estudiantes se inscriban en las carreras profesionales sin sustentar su decisión en una sólida información sobre las mismas. Las características académicas previas del estudiante. La carencia de una práctica temprana y la ausencia de asignaturas que aproximan el ejercicio profesional desde los primeros años. Los contenidos y las formas de estudiar en que varias asignaturas se aproximan a los de las ramas de la enseñanza media. La falta de preocupación institucional frente a los resultados negativos de la organización universitaria en su conjunto, tales como el nivel de formación y profesionalización de los profesores, la manera en la que se organiza el trabajo académico, la pertinencia y la actualización del currículo, los apoyos materiales y administrativos. El cambio de carrera (el alumno o alumna continua en la misma institución, pero se incorpora a otra carrera) o de institución.

- Socioeconómicos y laborales: Las condiciones económicas desfavorables del estudiante y la carencia de financiamiento. Las bajas expectativas de encontrar trabajo estable y con una remuneración adecuada. La obligación de estar titulado para ejercer la desarticulación familiar.

La figura 8 nos muestra otra manera de clasificar los aspectos influyentes en la deserción estudiantil.

Figura 8. Aspectos influyentes en la deserción estudiantil Tomada de “Modelos de análisis de la deserción estudiantil en la educación superior” .

Las variables relacionadas a estos grandes grupos son las que ayudan a caracterizar la deserción en las universidades, estas variables a su vez pueden ser utilizadas para la caracterización y análisis de mortalidad.

### **3.2.2. REPITENCIA**

La repitencia es la acción de matricular una materia por segunda vez, esto ocurre cuando la persona cursó la misma materia con anterioridad obteniendo como resultado la reprobación de ella.

Para el caso especial de nuestra Universidad, la repitencia no es posible sino una vez, puesto que cuando se pierde una materia por segunda ocasión automáticamente se le convertirá en una prueba de suficiencia.

Como se explicó al inicio, la mortalidad académica en una materia hace que la persona entre en un periodo de repitencia, esto finalmente se refleja en el atraso o rezago universitario, es decir, en la prolongación de los estudios por sobre lo establecido formalmente para cada carrera o programa.

“La repitencia se entiende como la acción de cursar reiterativamente una actividad docente, sea por mal rendimiento del estudiante o por causas ajenas al ámbito académico. La repitencia en la educación superior puede presentarse de varias formas de acuerdo al régimen curricular. Puede estar referida a todas las actividades académicas de un período determinado (año, semestre o trimestre), o bien, a cada asignatura para el caso de currículo flexible”.

Dada la relación existente entre mortalidad y repitencia, las variables utilizadas para la aproximación de las causas de mayor peso en el segundo fenómeno, pueden extenderse y ser utilizadas para el análisis de la mortalidad.

Encontraremos entonces:

- Información sobre el colegio de procedencia
- Estudios post-colegio y pre-universidad
- Resultados en las pruebas de estado

- Información de tipo familiar
- Información de tipo económica
- Información de la actualidad laboral
- Gusto por la carrera que está haciendo

Estas son macro variables de las cuales se derivan otras tantas que servirán para el modelamiento de los aspectos a estudiar.

### **3.2.3. REZAGO**

Entiéndase el rezago como el hecho de estar cursando un semestre inferior al que por tiempo de matrículas en la Universidad debería estarse cursando, este rezago puede ser parcial (una materia) o total (todo un semestre). El rezago es ocasionado por la mortalidad en las materias, esto conlleva a repetirla haciendo que el estudiante empiece a presentar atraso en su currículo.

Esto normalmente es causa de frustración y desmotivación de los estudiantes terminando por hacerlos desfallecer y desertar de toda aspiración universitaria.

El rezago en las universidades puede verse potencializado por materias que se consideran críticas y que son el requisito para cursar otras materias en los semestres siguientes, es por eso que se deben identificar estas materias y realizar toda clase de actividades y estrategias para disminuir la mortalidad en las mismas, esto con el fin no solo de mejorar la calidad académica, sino además de disminuir el rezago académico que finalmente conlleva a la deserción universitaria.

### **3.2.4 MUESTREO.**

Dado que se debe realizar un tipo de muestreo en el momento del desarrollo del trabajo con el fin de determinar quiénes serán las personas a analizar para el periodo 20093, es necesario realizar un bosquejo sobre los principales tipos de muestreo existentes.

En ocasiones en que no es posible o conveniente realizar un censo (analizar a todos los elementos de una población), se selecciona una muestra, entendiendo por tal una parte representativa de la población. El muestreo es por lo tanto una herramienta de la investigación científica, cuya función básica es determinar que parte de una población debe examinarse, con la finalidad de hacer inferencias sobre dicha población. La muestra debe lograr una representación adecuada de la población, en la que se reproduzca de la mejor manera los rasgos esenciales de dicha población que son importantes para la investigación. Para que una muestra sea representativa, y por lo tanto útil, debe de reflejar las similitudes y diferencias encontradas en la población, es decir ejemplificar las características de ésta. Los errores más comunes que se pueden cometer son:

1.- Hacer conclusiones muy generales a partir de la observación de sólo una parte de la Población, se denomina error de muestreo.

2.- Hacer conclusiones hacia una Población mucho más grande de la que originalmente se tomó la muestra. Error de Inferencia.

### **3.2.4.1.TIPOS DE MUESTREO.**

Existen diferentes criterios de clasificación de los tipos de muestreo, aunque en general pueden dividirse en dos grandes grupos: métodos de muestreo probabilísticos y métodos de muestreo no probabilísticos.

I. Muestreo probabilístico. Los métodos de muestreo probabilísticos son aquellos que se basan en el principio de equiprobabilidad. Es decir, aquellos en los que todos los individuos tienen la misma probabilidad de ser elegidos para formar parte de una muestra y, consiguientemente, todas las posibles muestras de tamaño  $n$  tienen la misma probabilidad de ser seleccionadas. Sólo estos métodos de muestreo probabilísticos aseguran la representatividad de la muestra extraída y son, por tanto, los más recomendables. Dentro de los métodos de muestreo probabilísticos se encuentran los siguientes tipos:

1.- Muestreo aleatorio simple: El procedimiento empleado es el siguiente: 1) se asigna un número a cada individuo de la población y 2) a través de algún medio mecánico (bolas dentro de una bolsa, tablas de números aleatorios, números aleatorios generados con una calculadora u ordenador, etc.) se eligen tantos sujetos como sea necesario para completar el tamaño de muestra requerido.

Este procedimiento, atractivo por su simpleza, tiene poca o nula utilidad práctica cuando la población que se está manejando es muy grande.

2.- Muestreo aleatorio sistemático: Este procedimiento exige, como el anterior, numerar todos los elementos de la población, pero en lugar de extraer  $n$  números aleatorios sólo se extrae uno. Se parte de ese número aleatorio y, que es un número elegido al azar, y los elementos que integran la muestra son los que ocupan los lugares  $i, i+k, i+2k, i+3k, \dots, i+(n-1)k$ , es decir se toman los individuos de  $k$  en  $k$ , siendo  $k$  el resultado de dividir el tamaño de la población entre el tamaño de la muestra:  $k= N/n$ . El número  $i$  que se emplea como punto de partida será un número al azar entre 1 y  $k$ .

El riesgo de este tipo de muestreo está en los casos en que se dan periodicidades en la población ya que al elegir a los miembros de la muestra con una periodicidad constante ( $k$ ) se puede introducir una homogeneidad que no se da en la población. Por ejemplo si se está seleccionando una muestra sobre listas de 10 individuos en los que los 5 primeros son varones y los 5 últimos mujeres, empleando un muestreo aleatorio sistemático con  $k=10$  siempre se seleccionarán o sólo hombres o sólo mujeres, no podría haber una representación de los dos sexos.

3.- Muestreo aleatorio estratificado: Trata de obviar las dificultades que presentan los anteriores ya que simplifican los procesos y suelen reducir el error muestral para un tamaño dado de la muestra. Consiste en considerar categorías típicas diferentes entre sí (estratos) que poseen gran homogeneidad respecto a alguna característica (se puede estratificar, por ejemplo, según la profesión, el municipio de residencia, el sexo, el estado civil, etc.). Lo que se pretende con este tipo de muestreo es asegurarse de que todos los estratos de interés estarán representados adecuadamente en la muestra. Cada estrato funciona independientemente, pudiendo aplicarse dentro de ellos el muestreo aleatorio simple o el estratificado para elegir los elementos concretos que formarán parte de la muestra. En ocasiones las dificultades que plantean son demasiado grandes, pues exige un conocimiento detallado de la población. (Tamaño geográfico, sexos, edades, etc.).

La distribución de la muestra en función de los diferentes estratos se denomina afijación, y puede ser de diferentes tipos:

**Afijación Simple:** A cada estrato le corresponde igual número de elementos muestrales.

**Afijación Proporcional:** La distribución se hace de acuerdo con el peso (tamaño) de la población en cada estrato.

**Afijación Óptima:** Se tiene en cuenta la previsible dispersión de los resultados, de modo que se considera la proporción y la desviación típica. Tiene poca aplicación ya que no se suele conocer la desviación.

4.- Muestreo aleatorio por conglomerados: Los métodos presentados hasta ahora están pensados para seleccionar directamente los elementos de la población, es decir, que las unidades muestrales son los elementos de la población.

En el muestreo por conglomerados la unidad muestral es un grupo de elementos de la población que forman una unidad, a la que se le conoce como conglomerado. Las unidades hospitalarias, los departamentos universitarios, una caja de determinado producto, etc, son conglomerados naturales.

En otras ocasiones se pueden utilizar conglomerados no naturales como, por ejemplo, las urnas electorales. Cuando los conglomerados son áreas geográficas suele hablarse de "muestreo por áreas".

El muestreo por conglomerados consiste en seleccionar aleatoriamente un cierto número de conglomerados (el necesario para alcanzar el tamaño muestral establecido) y en investigar después todos los elementos pertenecientes a los conglomerados elegidos.

## II. Métodos de muestreo no probabilísticos

A veces, para estudios exploratorios, el muestreo probabilístico resulta excesivamente costoso y se acude a métodos no probabilísticos, aun siendo conscientes de que no sirven para realizar generalizaciones (estimaciones inferenciales sobre la población), pues no se tiene certeza de que la muestra extraída sea representativa, ya que no todos los sujetos de la población tienen la misma probabilidad de ser elegidos. En general se seleccionan a los sujetos siguiendo determinados criterios procurando, en la medida de lo posible, que la muestra sea representativa.

En algunas circunstancias los métodos estadísticos y epidemiológicos permiten resolver los problemas de representatividad aun en situaciones de muestreo no probabilístico, por ejemplo los estudios de caso-control, donde los casos no son seleccionados aleatoriamente de la población.

Entre los métodos de muestreo no probabilísticos más utilizados en investigación se encuentran:

1.- Muestreo por cuotas: También denominado en ocasiones "accidental". Se asienta generalmente sobre la base de un buen conocimiento de los estratos de la población y/o de los individuos más "representativos" o "adecuados" para los fines de la investigación. Mantiene, por tanto, semejanzas con el muestreo aleatorio estratificado, pero no tiene el carácter de aleatoriedad de aquél.

En este tipo de muestreo se fijan unas "cuotas" que consisten en un número de individuos que reúnen unas determinadas condiciones, por ejemplo: 20 individuos de 25 a 40 años, de sexo femenino y residentes en Gijón. Una vez determinada la cuota se eligen los primeros que se encuentren que cumplan esas características. Este método se utiliza mucho en las encuestas de opinión.

2.- Muestreo intencional o de conveniencia: Este tipo de muestreo se caracteriza por un esfuerzo deliberado de obtener muestras "representativas" mediante la inclusión en la muestra de grupos supuestamente típicos. Es muy frecuente su utilización en sondeos preelectorales de zonas que en anteriores votaciones han marcado tendencias de voto.

También puede ser que el investigador seleccione directa e intencionadamente los individuos de la población. El caso más frecuente de este procedimiento es el utilizar como muestra los individuos a los que se tiene fácil acceso (los profesores de universidad emplean con mucha frecuencia a sus propios alumnos).

3.- Bola de nieve: Se localiza a algunos individuos, los cuales conducen a otros, y estos a otros, y así hasta conseguir una muestra suficiente. Este tipo de muestreo se emplea muy frecuentemente cuando se hacen estudios con poblaciones "marginales", delincuentes, sectas, determinados tipos de enfermos, etc.

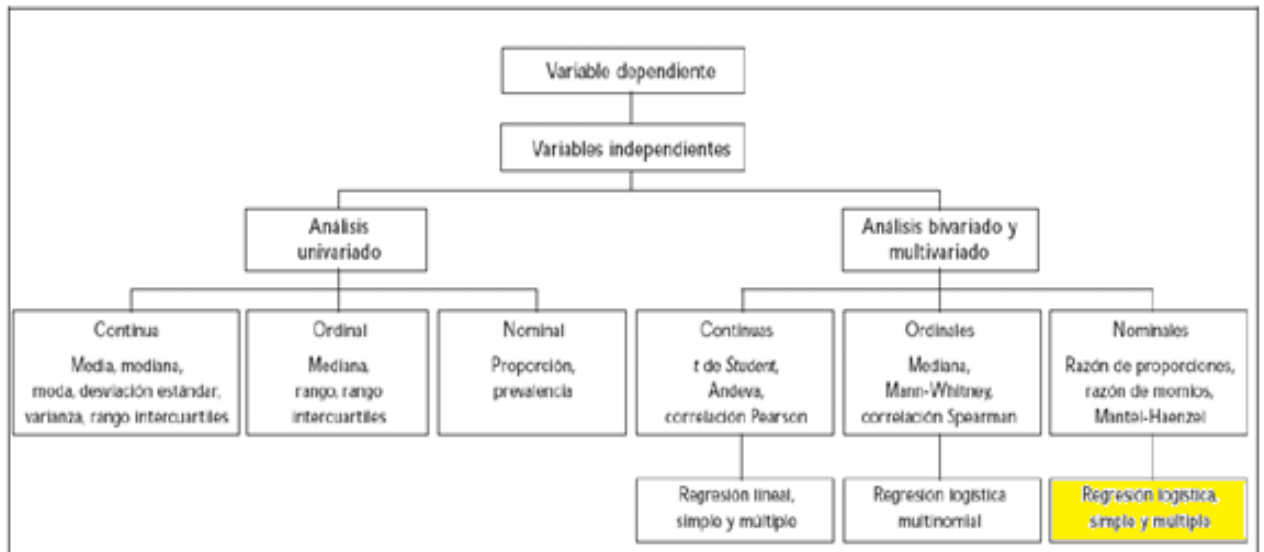
4.- Muestreo Discrecional: A criterio del investigador los elementos son elegidos sobre lo que él cree que pueden aportar al estudio.

### 3.2.5. ANALISIS MULTIVARIADOS.

Los métodos estadísticos multivariantes y el análisis multivariante son herramientas estadísticas que estudian el comportamiento de tres o más variables al mismo tiempo. Se usan principalmente para buscar las variables menos representativas para poder eliminarlas, simplificando así modelos estadísticos en los que el número de variables sea un problema y además para comprender la relación entre dichos grupos de variables.

Se pueden sintetizar dos objetivos claros:

1. Proporcionar métodos cuya finalidad es el estudio conjunto de datos multivariantes que el análisis estadístico uní y bidimensional es incapaz de conseguir.
2. Ayudar al analista o investigador a tomar decisiones óptimas en el contexto en el que se encuentre teniendo en cuenta la información disponible por el conjunto de datos analizado.



Tomada de Hernández y Velazco-Mondragón (2000)

Existen diferentes modelos y métodos, cada uno con su tipo de análisis:

#### 1. Métodos de Dependencia:

1. Un estudio de la regresión nos permite averiguar hasta qué punto una variable puede ser prevista conociendo otra. Se utiliza para intentar predecir el comportamiento de ciertas variables a partir de otras.



2. El análisis de la correlación canónica intenta analizar la posible existencia de relación entre dos grupos de variables.

3. Un análisis discriminante nos puede dar una función discriminante que puede ser utilizada para distinguir entre dos o más grupos, y de este modo tomar decisiones.

4. Un análisis multivariante de la varianza (MANOVA), extendiendo el análisis de la varianza (ANOVA), cubre los casos en los que se conozca la existencia de más de una variable dependiente sin poderse simplificar más el modelo.

5. La regresión logística permite la elaboración de un análisis de regresión para estimar y probar la influencia de una variable sobre otra, cuando la variable dependiente o de respuesta es de tipo dicotómico.

2. Métodos de Interdependencia:

1. El análisis de los componentes principales procura determinar un sistema más pequeño de variables que sinteticen el sistema original.

2. El análisis clúster clasifica una muestra de entidades (individuos o variables) en un número pequeño de grupos de forma que las observaciones pertenecientes a un grupo sean muy similares entre sí y muy disimilares del resto. A diferencia del Análisis discriminante se desconoce el número y la composición de dichos grupos.

### **3.2.5.1. MODELOS DE CLASIFICACION. MODELOS LOGIT Y PROBIT.**

#### **ELECCION DISCRETA.**

Los modelos de selección discreta predicen directamente la probabilidad de ocurrencia de un suceso que viene definido por los valores de las variables independientes. Como los valores de una probabilidad están entre cero y uno, las predicciones realizadas con los modelos de regresión discreta deben estar acotadas para que caigan en los rangos entre cero y uno. El modelo general que cumple estas condiciones es un caso particular del modelo de regresión múltiple que se denomina modelo de elección discreta, y tiene la forma funcional

$$P_i = F(X_i, \beta) + u_i$$

Se observa que si  $F$  es lineal tenemos el modelo lineal de probabilidad, pero si  $F$  es una función de distribución de una variable aleatoria, entonces  $P$  varía entre -cero y uno de modo seguro.

En el caso particular en que la función  $F$  es la función logística estaremos ante el modelo logit o regresión lógica, cuya forma funcional será la siguiente.

$$P_i = F(X_i, \beta) + u_i = \left( \frac{e^{x_i\beta}}{1 + e^{x_i\beta}} \right) + u_i$$

En el caso particular en el que la función F es la función de distribución de una normal unitaria estaremos ante el modelo probit, cuya forma funcional será la siguiente.

$$P_i = F(X_i, \beta) + u_i = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_i\beta} e^{-t^2/2} dt + u_i$$

### 3.2.6. REGRESION LOGISTICA

Un modelo de regresión con variable de pendiente binomial será un modelo que permita estudiar si dicha variable discreta depende o no de otra u otras variables. Si una variable binomial de parámetro p es independiente de otra variable X. Por consiguiente, un modelo de regresión de pendiente binomial y una única variable independiente X se materializa en una función en la que p aparece dependiendo de X y de unos coeficientes de investigación que permiten abordar la relación de dependencia.

Para una única variable independiente X, el modelo de regresión logística toma la forma que se presenta en la ecuación 1.

$$\text{Ecuación 1. } \ln\left(\frac{p}{q}\right) / X = x = \alpha_0 + \alpha_1 X$$

O de forma simplificada como se muestra en la ecuación 2.

$$\text{Ecuación 2. } \ln(p/q) = \alpha_0 + \alpha_1 X$$

Donde ln significa logaritmo neperiano, y son constantes y X una variable que puede ser aleatoria o no, continua o discreta. Este modelo se puede generalizar fácilmente para k variables independientes, dando lugar al modelo logístico múltiple que se expresa en la ecuación 3.

$$\text{Ecuación 3. } \ln(p/q) = \alpha_0 + \alpha_1 x + \alpha_2 X + \dots + \alpha_k X_k$$

El objetivo primordial que resuelve esta técnica es el de modelar cómo influye en la probabilidad de aparición de un suceso, habitualmente dicotómico, la presencia o no de diversos factores y el valor o nivel de los mismos. También puede ser usada para estimar la probabilidad de aparición de cada una de las posibilidades de un suceso con más de dos categorías (politómico).

De todos es sabido que este tipo de situaciones se aborda mediante técnicas de regresión. Sin embargo, la metodología de la regresión lineal no es aplicable ya que ahora la variable de respuesta sólo presenta dos valores (caso dicotómico), como puede ser presencia/ausencia de hipertensión.

Si se clasifica el valor de la variable respuesta como 0 cuando no se presenta el suceso (ausencia de hipertensión) y con el valor 1 cuando sí está presente (paciente hipertenso), y se busca cuantificar la posible relación entre la presencia de hipertensión y, por ejemplo, la cantidad media de sal consumida al día como posible factor de riesgo, se puede caer en la tentación de utilizar una regresión lineal como se muestra en la ecuación 4.

Ecuación 4.  $Hipertensi3n = a + b \cdot [Consumo\_sal]$

Así estimar a partir de los datos, por el procedimiento habitual de m3nimos cuadrados, los coeficientes a y b de la ecuaci3n. Sin embargo, y aunque esto es posible matem3ticamente, conduce a la obtenci3n de resultados absurdos,

ya que cuando se calcule la funci3n obtenida para diferentes valores de consumo de sal se obtendr3n resultados que, en general, ser3n diferentes de 0 y 1, los 3nicos realmente posibles en este caso, ya que esa restricci3n no se impone en la regresi3n lineal, en la que la respuesta puede en principio tomar cualquier valor.

Si se utiliza c3mo variable dependiente la probabilidad p de que un paciente padezca hipertensi3n y se construye la siguiente ecuaci3n 5.

Ecuaci3n 5.  $\ln \frac{p}{1-p}$

Ahora s3 se tiene una variable que puede tomar cualquier valor, por lo que se puede plantear el buscar para ella, una ecuaci3n de regresi3n tradicional como se plantea en la ecuaci3n 6.

Ecuaci3n 6.  $\ln \frac{p}{1-p} = a + b \cdot [consumo\_sal]$

Esta se puede convertir con una peque1a manipulaci3n algebraica en la ecuaci3n 7 que se presenta a continuaci3n.

Ecuaci3n 7.  $Pr. HTA = \frac{1}{1 + e^{(-a - b \cdot [consumo\_sal])}}$

Y este es precisamente el tipo de ecuaci3n que se conoce como modelo log3stico, donde el n3mero de factores puede ser m3s de uno, as3 en el exponente que figura en el denominador de la ecuaci3n 7 se podr3a tener la siguiente ecuaci3n 8.

Ecuaci3n 8.  $b1.consumo\_sal + b2.edad + b3.sexo + b4.fumador$

## LOS COEFICIENTES DEL MODELO LOGISTICO COMO CUANTIFICADORES DE RIESGO

Una de las características que hacen tan interesante la regresión logística es la relación que este guarda con un parámetro de cuantificación de riesgo conocido en la literatura como "odds ratio" (aunque puede tener traducción al castellano, no se tendrá en cuenta para evitar confusión ya que siempre se utiliza la terminología inglesa).

El odds asociado a un suceso es el cociente entre la probabilidad de que ocurra frente a la probabilidad de que no ocurra como se muestra en la ecuación 9.

$$\text{Ecuación 9. } odds = \frac{p}{1-p}$$

Siendo  $p$  la probabilidad del suceso. Así, por ejemplo, se puede calcular el odds de presencia de hipertensión cuando el consumo diario de sal es igual o superior a una cierta cantidad, que en realidad determina cuántas veces es más probable que haya hipertensión a que no la haya en esa situación.

Igualmente se puede calcular el odds de presencia de hipertensión cuando el consumo de sal es inferior a esa cantidad. Si se divide el primer odds entre el segundo, se abra calculado un cociente de odds, esto es un odds ratio, que de alguna manera cuantifica cuánto más probable es la aparición de hipertensión cuando se consume mucha sal (primer odds) respecto a cuándo se consume poca. La noción que se está midiendo es parecida a la que se encuentra en lo que se denomina riesgo relativo que corresponde al cociente de la probabilidad de que aparezca un suceso (hipertensión) cuando está presente el factor (consumo elevado de sal) respecto a cuándo no lo está. De hecho cuando la prevalencia del suceso es baja (< 20 %) el valor del odds ratio y el riesgo relativo es muy parecido, pero no es así cuando el suceso es bastante común, hecho que a menudo se ignora.

Si en la ecuación de regresión se tiene un factor dicotómico, como puede ser por ejemplo si el sujeto es no fumador, el coeficiente  $b$  de la ecuación para ese factor está directamente relacionado con el odds ratio OR de ser fumador respecto a no serlo, ecuación 10.

$$\text{Ecuación 10. } OR = \exp(b)$$

Es decir que  $\exp(b)$  es una medida que cuantifica el riesgo que representa poseer el factor correspondiente respecto a no poseerlo, suponiendo que el resto de variables del modelo permanecen constantes.

Cuando la variable es numérica, como puede ser por ejemplo la edad, o el índice de masa corporal, es una medida que cuantifica el cambio en el riesgo cuando se

pasa de un valor del factor a otro, permaneciendo constantes el resto de variables. Así el odds ratio que supone pasar de la edad X1 a la edad X2, siendo b el coeficiente correspondiente a la edad en el modelo logístico es el que se presenta en la ecuación 11.

Ecuación 11.  $OR = \exp[b \cdot (X2 - X1)]$

Nótese que se trata de un modelo en el que el aumento o disminución del riesgo al pasar de un valor a otro del factor, es proporcional al cambio, es decir a la diferencia entre los dos valores, pero no al punto de partida, quiere esto decir que el cambio en el riesgo, con el modelo logístico, es el mismo cuando se pasa de 40 a 50 años que cuando se pasas de 80 a 90.

Cuando el coeficiente b de la variable es positivo se obtiene un odds ratio mayor que 1 y corresponde por tanto a un factor de riesgo. Por el contrario, si b es negativo el odds ratio será menor que 1 y se trata de un factor de protección.

## LAS VARIABLES CUALITATIVAS EN EL MODELO LOGISTICO

Puesto que la metodología empleada para la estimación del modelo logístico se basa en la utilización de variables cuantitativas, al igual que en cualquier otro procedimiento de regresión, es incorrecto que en él intervengan variables cualitativas, ya sean nominales u ordinales.

La asignación de un número a cada categoría no resuelve el problema ya que si tenemos, por ejemplo, la variable ejercicio físico con tres posibles respuestas: sedentario, realiza ejercicio esporádicamente, realiza ejercicio frecuentemente, y le asignamos los valores 0, 1, 2, significa a efectos del modelo, que efectuar ejercicio físico frecuentemente es dos veces mayor que solo hacerlo esporádicamente, lo cual no tienen ningún sentido. Más absurdo sería si se trata, a diferencia de ésta, de una variable nominal, sin ninguna relación de orden entre las respuestas, como puede ser el estado civil.

La solución a este problema es crear tantas variables dicotómicas como número de respuestas - 1. Estas nuevas variables, artificialmente creadas, reciben en la literatura anglosajona el nombre de "dummy", traducándose en español con diferentes denominaciones como pueden ser variables internas, indicadoras, o variables diseño.

## REGRESION LOGISTICA BINARIA

La regresión logística resulta útil para los casos en los que se desea predecir la presencia o ausencia de una característica o resultado según los valores de un conjunto de variables predictoras. Es similar a un modelo de regresión lineal pero se encuentra adaptado para los modelos en los que la variable dependiente es dicotómica. Los coeficientes de regresión logística pueden utilizarse para estimar la razón de las ventajas de cada variable independiente del modelo. La regresión

logística se puede aplicar a un rango más amplio de situaciones de investigación que el análisis discriminante.

### 3.2.7. REGRESION LOGISTICA MULTINOMIAL

La regresión logística multinomial resulta útil en aquellas situaciones en las que se desee poder clasificar a los sujetos según los valores de un conjunto de variables predictoras. Este tipo de regresión es similar a la regresión logística, pero más general, ya que la variable dependiente no está restringida a dos categorías.

“De igual forma se conoce como análisis de regresión multivariante al método estadístico que permite establecer una relación matemática entre un conjunto de variables  $X_1, X_2 \dots X_k$  (covariantes o factores) y una variable dependiente  $Y$ . Se utiliza fundamentalmente en estudios en los que no se puede controlar por diseño los valores de las variables independientes, como suele ocurrir en los estudios epidemiológicos y observacionales.

Los objetivos de un modelo de regresión puede ser dos:

- ✓ Obtener una ecuación que nos permita "predecir" el valor de  $Y$  una vez conocidos los valores de  $X_1, X_2 \dots X_k$ . Se conocen como modelos predictivos.
- ✓ Cuantificar la relación entre  $X_1, X_2 \dots X_k$  y la variable  $Y$  con el fin de conocer o explicar mejor los mecanismos de esa relación. Se trata de modelos explicativos, muy utilizados cuando se busca encontrar qué variables afectan a los valores de un parámetro fisiológico por ejemplo, o cuáles son los posibles factores de riesgo que pueden influir en la probabilidad de que se desarrolle una patología.

La disponibilidad y facilidad de uso del software que permite la construcción de modelos de regresión, ha hecho olvidar que se trata de técnicas complejas, que requieren un cierto conocimiento de la metodología estadística subyacente, por lo que se cae con excesiva frecuencia en una pobre utilización de las técnicas de regresión y una peor descripción de cómo se emplearon en cada caso concreto, e incluso una ausencia total de esa explicación, y se comunican los resultados como si la propia ecuación de regresión fuera sin más un "artículo de fe" que no necesitara de una cuidadosa validación.

Un problema fundamental que se plantea a la hora de construir un modelo multivariante es qué los factores  $X_1, X_2 \dots X_k$ , se deben incluir en la ecuación, de tal manera que se estime el mejor modelo posible a partir de los datos del estudio. Para ello lo primero que habría que definir es qué se entiende por "mejor modelo". Si se busca un modelo predictivo, será aquél que proporcione predicciones más fiables, más acertadas; mientras que si el objetivo es construir un modelo explicativo, se busca que las estimaciones de los coeficientes de la ecuación sean

precisas, ya que a partir de ellas se van a efectuar las deducciones. Cumplidos esos objetivos es claro que otra característica deseable del modelo es que sea lo más sencillo posible”.

#### MODELO PROBIT:

Si se elige la función F como la función de distribución de una Normal (0,1), el modelo lineal general:

$$P_i = F(x_i \beta) + u_i$$

Se denomina modelo probit, cuyos parámetros  $\beta$  admiten estimación por máxima verosimilitud.

No olvidemos que:

$$P_i = F(X_i, \beta) + u_i = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_i \beta} e^{-t^2/2} dt$$

En el caso de que  $F(X_i, \beta) = \frac{e^{x_i \beta}}{1 + e^{x_i \beta}}$  estaríamos ante el modelo logit.

#### 4. ESTADO DEL ARTE TRABAJO DE GRADO

##### 1. IDENTIFICACIÓN Y ANÁLISIS DE FACTORES INCIDENTES EN EL RETIRO VOLUNTARIO DE ESTUDIANTES DE PREGRADO EN EL POLITÉCNICO COLOMBIANO JAIME ISAZA CADAVID EN EL PERÍODO 02-2004 A 01-2006

Ing. José Ignacio Maya Guerra - Universidad Nacional De Colombia -Sede Medellín -Facultad De Minas – 2008

La investigación se dedico a identificar y analizar los factores que influyen en la mortalidad académica para crear un modelo de predicción y prevención que ayude al instituto para intervenir, la investigación tuvo en cuenta los factores internos del instituto y los externos a los que se enfrentan los estudiantes de pregrado en los periodos 2004 y 2006 abarcando los 14.000 estudiantes que se encontraban matriculados en estos periodos. Para el estudio se tomo una muestra de 200 estudiantes y con la información primaria recogida se construyó un modelo de regresión logística que muestra la significancia de las variables seleccionadas en el estudio y dando la posibilidad de determinar la probabilidad de retiro de los estudiantes.

Conclusiones los resultados obtenidos permitieron determinar la significación de las variables mediante icoeficientes, la obtención de un modelo predictivo; confirman el impacto conjunto que tienen los factores financieros, laborales, académicos e institucionales sobre la deserción, pero igualmente importante al identificar la relevancia de las variables, con estos resultados se pueden crear modelos y soluciones para enfrentar la problemática educativa.

Los objetivos que motivaron a la investigación fue saber e identificar las causad de la mortalidad académica, crear un modelo de predicción y prevención para enfrentar esta problemática.

Las variables son: retiro –momento de retiro, tipo de retiro y causa de retiro; demografía – sexo, municipio de origen y edad; Condiciones Académicas – nivel académico y rendimiento académico; Ámbito – programa académico, facultad, sede, institución y modalidad; Condiciones familiares – salud, servicio militar, estado civil y personas dependientes.

Los resultados obtenidos son la identificación de las variables y la creación del modelo por regresión logística.

Las técnicas fueron regresión logística, Prueba de Hosmer – Lemeshow,

Link [http://www.bdigital.unal.edu.co/885/1/71670474\\_2009.pdf](http://www.bdigital.unal.edu.co/885/1/71670474_2009.pdf)



## 2. ESTUDIO DE LA DESERCIÓN ESTUDIANTIL DE LOS PROGRAMAS DE PREGRADO DE LA UNIVERSIDAD DE LOS LLANOS

Luz Miryam Malagón Escobar (Directora proyecto social) – Cesar Augusto Calderón Cañón (Egresado programa de Economía) Edwin Leonardo Soto Hernández (Egresado programa de Economía)

Los resultados obtenidos al estudiar el comportamiento de la deserción estudiantil en la Universidad de los Llanos a través del enfoque institucional que evidencia que la variable cuantitativa y cualitativa, a través de indicadores que describen el comportamiento de la deserción en las diferentes cohortes del periodo del estudio, los resultados generales para los doce programas de pregrado en la Universidad permiten concluir que la más alta deserción se presenta en los primeros semestres de la carrera con mayor incidencia en el primer semestre, definida como Deserción inicial. En cuanto a las facultades, la facultad de Ciencias Agropecuarias y Recursos Naturales es la que presenta mayor índice de deserción, y el programa que está siendo más afectado con mayor relevancia es Ingeniería de Sistemas con un índice de 56,0%.

Respecto a la eficiencia en la retención, el 56.0% finaliza su Plan de Estudios, y la principal causa por la que se presenta la deserción es por “motivos académicos” con un 48.0% dentro de los episodios de deserción, con el mayor determinante de la deserción en esta Universidad.

No existe correlación entre el puntaje de la Prueba de estado y el desempeño académico de los estudiantes desertores. De otra parte un alto puntaje ICFES no garantiza la permanencia del estudiante en la Universidad, pero si le permite acceder nuevamente al mismo programa o a otro en la Institución. En cuanto a las causas socioeconómicas, la incidencia de los problemas económicos y el bajo nivel de ingreso es un factor que más se presenta, pero ésta es mínima al compararla con las demás causas. Los programas de Medicina Veterinaria y Zootecnia e Ingeniería Agronómica son los que generan más episodios de deserción por motivos económicos.

La deserción por causas individuales afecta más a mujeres que a hombres y dentro de estas causas, los factores más determinantes son el problema en el entorno familiar del estudiante y el embarazo.

Conclusión los gastos por abandono, los años con mayor impacto económico, son el 2002 y el 2003 con un 16.0% y 18.0% respectivamente del total del presupuesto de la Universidad. Evidenciándose el costo económico de la deserción para el periodo del 2000 al 2004 en \$12.928'292.779 siendo la deserción aproximadamente el 50% del presupuesto de la Universidad para el año 2005

Link [http://web.unillanos.edu.co/centro-de-documentacion/doc\\_view/33-estudio-de-la-desercion-estudiantil-de-los-programas-de-pregrado-de-la-universidad-de-los-llanos.html](http://web.unillanos.edu.co/centro-de-documentacion/doc_view/33-estudio-de-la-desercion-estudiantil-de-los-programas-de-pregrado-de-la-universidad-de-los-llanos.html)

### 3. CARACTERIZACIÓN DE LA DESERCIÓN ESTUDIANTIL EN LA UNIVERSIDAD NACIONAL DE COLOMBIA SEDE MEDELLÍN, Medellín, Mayo de 2006.

El trabajo realizado como se menciona en el comprende el primer esfuerzo sistemático para estudiar el fenómeno en esta sede, recopilado, depurando y consolidando las bases de datos y los históricos de los estudios realizados anteriormente de deserción estudiantil para los periodos 2001 al 2005. En el estudio se analizaron los factores académicos y socioeconómicos que pueden ser causa de la deserción estudiantil y también se tiene en cuenta las transferencias y movilidad estudiantil que son tomadas como una deserción interna. Y por último con los resultados obtenidos se generaron conclusiones y recomendaciones para enfrentar el problema y ayudas a disminuir el índice de deserción estudiantil.

El objetivo principal que motivo a realizar este proyecto fue determinar los factores causantes de la deserción y crear soluciones para enfrentar el problema y disminuir el índice de deserción estudiantil.

Las conclusiones son: La deserción no académica contribuye en promedio con el 40% de la deserción general. Las causas socioeconómicas son la determinante principal de este tipo de deserción. La deserción estudiantil en la Sede Medellín de la Universidad Nacional es alta, aunque su tasa promedio acumulada es inferior a la del país, que es del 52% en la educación superior, por término medio está entre 45% y 50% en la mayoría de sus cohortes.

Las variables manejadas son las socioeconómicas y las académicas manejadas por notas, promedios, transferencias entre otras.

Link [http://www.alfaguia.org/alfaguia/files/1320765488\\_2242.pdf](http://www.alfaguia.org/alfaguia/files/1320765488_2242.pdf)

### 4. ESTUDIO SOBRE DESERCIÓN Y PERMANENCIA ACADÉMICA EN LA FACULTAD DE INGENIERÍA DE LA UNIVERSIDAD DE LA GUAJIRA DESDE EL II PA 2005 HASTA EL II PA 2007

Investigador Principal: RAFAEL MELENDEZ SURMAY

UNIVERSIDAD DE LA GUAJIRA FACULTAD DE INGENIERÍA PROGRAMA DE INGENIERÍA INDUSTRIAL RIOHACHA 2008

El estudio realizado pretende analizar más a fondo la deserción estudiantil, identificando las variables que más afectan esta problemática teniendo presente que son muchas las variables que influyen en dicho efecto, poder pasar de las conclusiones y resultados comunes. Tener resultados que hagan la diferencia creando un modelo que ayude realmente a disminuir el índice de la deserción estudiantil. Es por esto que se hace necesario organizar la información para crear documentos que permitan realizar proyecciones de la gestión administrativa, y a la vez servir como herramienta para la evaluación institucional, generando posibles políticas que permitan mejorar la eficacia y eficiencia institucional de la facultad. En el trabajo se tienen en cuenta los estudios anteriores no solo realizados en la universidad si no también en otras universidades ya que este problema afecta a todas las universidades e instituciones educativas de nivel superior. También pretende analizar los diferentes tipos de deserción posible que se presentan como también sus índices. Se tiene muy presente que la deserción estudiantil es un gran problema y la universidad quiere disminuir su índice y prevenir que aumenten nuevamente.

Conclusiones: Se concluyeron porcentajes de deserción. Así mismo, la deserción total está motivada principalmente porque los estudiantes no presentan el perfil adecuado para los programas a los cuales se matriculan.

La Facultad de Ingenierías presenta una tasa de deserción acumulada del 16,12%, es decir que de 5796 estudiantes matriculados durante los cinco periodos académicos objeto de este estudio, 930 en total presentaron algún tipo de deserción, lo que equivale a un promedio de 186 estudiantes retirados en cada periodo.

Los resultados obtenidos fueron las causas de la deserción, los porcentajes de deserción de las facultades analizadas y cómo influyen las causas en estos porcentajes.

Las variables estudiadas son las socioeconómicas como la edad, la población de la que viene, los ingresos familiares, entre otras. Las variables académicas como la deserción por programas, como influyen las materias entre otras.

Link [http://www.colombiaaprende.edu.co/html/directivos/1598/articles-250629\\_recurso\\_6.pdf](http://www.colombiaaprende.edu.co/html/directivos/1598/articles-250629_recurso_6.pdf)

## 5. DESERCIÓN ESTUDIANTIL EN LA EDUCACIÓN SUPERIOR.

Autores Carolina Guzmán Ruiz - Diana Durán Muriel - Jorge Franco Gallego

Ministerio de Educación Nacional

En el sistema de educación superior colombiano la deserción académica es un tema preocupante, pese a que los últimos años se han caracterizado por el aumento de cobertura e ingreso de estudiantes nuevos, los primeros tres semestres de la educación superior es cuando se presenta mayor deserción cerca de un 60% son desertores del total de ingresados, las causales son académicas y de orientación profesional y vocacional.

En conclusión se puede decir que se arrojan muchas sugerencias para que sean aplicadas en las entidades superiores en Colombia, como crear programas de apoyo estudiantil, de adaptación social y estrategia académica de enseñanza y aprendizaje.

Finalmente, las autoridades educativas nacionales deberían garantizar la existencia en todas las instituciones de educación superior del proceso del seguimiento y evaluación de la deserción estudiantil y graduación a través del establecimiento de un marco legal que las incentiven y que muestre resultados en términos de retención estudiantil, así como su incorporación en los procesos de evaluación de la calidad. Esto garantizaría que los esfuerzos sean realizados por todos los entes participantes en el sistema de educación superior y no que se concentre en el esfuerzo de unos pocos.

Link [http://www.mineducacion.gov.co/sistemasdeinformacion/1735/articles-254702\\_libro\\_desercion.pdf](http://www.mineducacion.gov.co/sistemasdeinformacion/1735/articles-254702_libro_desercion.pdf)

#### 6. ESTUDIO DE LA DESERCIÓN ESTUDIANTIL DE LOS PROGRAMAS DE PREGRADO DE LA UNIVERSIDAD DE LOS LLANOS (1988 – 2004). UNIVERSIDAD DEL LOS LLANOS, ENERO DE 2006.

El propósito del estudio es determinar el comportamiento de la deserción estudiantil en la Universidad de los llanos, bajo la premisa de generar un seguimiento permanente que se deberá asumir desde la administración académica a partir de los resultados obtenidos. Se sigue un reglamento de analizar la información de todos los estudiantes matriculados en estos periodos determinando los índices posibles de deserción, luego se realizaron los análisis longitudinal y transversal, a continuación se precisaron los factores que generaban los índices y cuantificaron los impactos de estos factores. En la investigación se tuvieron en cuenta los posibles tipos de deserción. Como la deserción precoz, temprana y la deserción tardía. En el estudio intentan caracterizar los estudiantes de primer semestre para identificar sus condiciones y los factores que inciden en la mortalidad académica generando conclusiones y recomendaciones para enfrentar la problemática e intentar prevenirla generando mejoramientos en la calidad de la universidad.

Las variables e índices manejados en la investigación son los siguientes:

Índice de deserción semestral, índice de deserción por cohorte, índice de deserción promedio por nivel, tasa ponderada de deserción por niveles. Y las variables son la información personal, información académica.

Los resultados obtenidos en este estudio fue evidenciar las variables que inciden en el fenómeno de mortalidad académica, a través de los indicadores lograron describir el comportamiento de la deserción en los diferentes periodos de estudio y generar conclusiones y recomendaciones para prevenir este fenómeno en la universidad del llano mejorando su calidad.

Las conclusiones

Que para los programas la deserción más alta se presenta en los primeros cinco semestres definida como deserción inicial. En la investigación se tuvieron en cuenta los posibles tipos de deserción. Como la deserción precoz, temprana y la deserción tardía. Otra de las causas que influyen en la deserción para no finalizar su plan de estudios son causas académicas. El resultado de las pruebas icfes y el desempeño académico no existe correlación entre ellos. Y el hecho de no recibir orientación vocacional es una de las causas académicas que influyen de manera directa en la alta deserción inicial y temprana.

Link [http://web.unillanos.edu.co/centro-de-documentacion/doc\\_view/33-estudio-de-la-desercion-estudiantil-de-los-programas-de-pregrado-de-la-universidad-de-los-llanos.html](http://web.unillanos.edu.co/centro-de-documentacion/doc_view/33-estudio-de-la-desercion-estudiantil-de-los-programas-de-pregrado-de-la-universidad-de-los-llanos.html)

## 7. FACTORES DETERMINANTES EN LA MORTALIDAD ACADÉMICA PARA LA FACULTAD DE INGENIERÍAS EN LA UNIVERSIDAD TECNOLÓGICA DE PEREIRA ENTRE LOS AÑOS 2004 Y 2009.

La investigación pretendía identificar las variables que tienen mayor influencia en la mortalidad estudiantil para la facultad de ingeniería. Se utilizó como base la información con la que cuenta el observatorio académico adscrito a la facultad de ingeniería industrial, otras bases de datos de la universidad y información primaria para realizar el estudio, con esta información se realizó el estudio estadístico descriptivo donde se evidencia el fenómeno objeto de la investigación y se utilizaron las técnicas estadísticas de regresión logística multivariadas para identificar los factores objeto del estudio.

Las variables utilizadas en el estudio fueron de tipo académico como: resultados pruebas de estado – número de créditos vistos durante el último semestre – número de créditos aprobados durante el último semestre - Promedio obtenido en el último semestre - Estado de la matrícula - Materias perdidas durante la carrera - Materias aprobadas durante la carrera - Profesor Clasificación- Profesor Original.

Variables de tipo Socioeconómico: Género - Estrato socio económico - Departamento de procedencia - Municipio de procedencia - Género del colegio en el que estudio - Tipo de colegio en el que estudio - Estado civil – Edad. Variables de tipo Personal: Aspectos que considera importantes para el buen desempeño académico - Motivación para cursar las materias del semestre - Materias con dificultad al inicio - Satisfacción con el programa - Conocimiento previo del programa - Esfuerzo para cursar el semestre - Apoyo familiar en los estudios - Actividades extra clase.

El objetivo de dicho análisis era el de evidenciar los altos índices de mortalidad que se presentan en las materias críticas de estos programas, relacionando el Índice de Mortalidad (IM), obtenido de la relación entre el número de estudiantes que inscribieron la materia (EI) y el número de estudiantes que la perdieron o reprobaron (EP).

Los resultados obtenidos se pudieron criticar las materias que generan gran impacto en la deserción estudiantil desde las variables académicas, el estudio se realizó en cada una de las materias consideradas como críticas y en cada uno de los semestres relacionados (I-2004 hasta II-2009).

Las conclusiones obtenidas Los niveles de mortalidad más altos se dieron en el primer semestre del estudio (I 2004) y el más bajo se dio en el último (II 2009), esto para la facultad de Ingenierías, Los índices de mortalidad más altos se presentan en las materias que se dictan en los primeros semestres de cada programa, esto puede estar ligado a cierto grado de irresponsabilidad por parte de los estudiantes en el momento de ingresar a la Universidad, así como a deficiencias de carácter académico en la formación de los estudiantes por parte de los colegios de los que provienen, Se concluye que son las variables de tipo académico las que más influyen en el desempeño que los estudiantes tienen en la Universidad, es por eso que es allí donde deben concentrarse las medidas preventivas, Se concluye además que variables de tipo socioeconómico y de tipo personal no presentaron mucha importancia en el desarrollo del presente trabajo

Link

[http://www.google.com.co/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=3&sqj=2&ved=0CCkQFjAC&url=http%3A%2F%2Frevistas.utp.edu.co%2Findex.php%2Frevistaciencia%2Farticle%2Fdownload%2F1237%2F809&ei=qYBvUKudB4Lu9ASt\\_oH4DQ&usq=AFQjCNEFIVJ6UyqYjFOKP5qQxqnhshLg](http://www.google.com.co/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=3&sqj=2&ved=0CCkQFjAC&url=http%3A%2F%2Frevistas.utp.edu.co%2Findex.php%2Frevistaciencia%2Farticle%2Fdownload%2F1237%2F809&ei=qYBvUKudB4Lu9ASt_oH4DQ&usq=AFQjCNEFIVJ6UyqYjFOKP5qQxqnhshLg)

## 8. ESTUDIO SOBRE DESERCIÓN Y PERMANENCIA ACADÉMICA EN LA FACULTAD DE INGENIERÍA DE LA UNIVERSIDAD DE LA GUAJIRA DESDE EL II PA 2005 HASTA EL II PA 2007

Rafael Melendez Surmay (Investigador Principal) - Diber Meriño Mendoza (Coinvestigador) - Ana Maria Londoño Colorado - Faner Johjan Pana Van-Grieken (Auxiliares de Investigación)

Este estudio apunta a un indicador que corresponde al tiempo de permanencia de los estudiantes, teniendo en cuenta el abandono voluntario o forzado de la carrera en la que el estudiante se matriculo, donde el estudiante se retira sin haber culminado sus estudios.

Una de los posibles variables que se analizan son las condiciones como las entidades privadas o públicas, la calidad de los programas académicos ofrecidos, las políticas y estrategias de la institución que garanticen la permanencia del estudiante, el factor geográfico, demográfico, económico y político del medio en el cual se desarrollan las actividades del ente educativo entre otras, estas características hacen que los estudios se direcciones según las características

Generalmente se ven una constante en los estudiantes, el bajo rendimiento académico por parte de los estudiantes, una edad entre los 20 y 24 años, la procedencia de grupos minoritarios, la dedicación parcial al estudio y como característica relevante la variable financiera que tiene una gran influencia en la decisión de persistencia. El factor económico si se tiene en cuenta que la mayoría de los alumnos proceden de hogares de bajos recursos, el peso real que posee el factor económico en el problema de deserción al interior de la Universidad de le Guajira y la facultad de Ingeniería es lo que se busca como resultado por medio del desarrollo de este estudio.

Los estudiantes que ingresan a la educación superior no saben estudiar. Elemento que junto con la voluntad personal determina el éxito en la universidad. Este estudio revelo, que a la hora de programar las asignaturas incide en el aprovechamiento, o la preparación de los docentes no solo en el conocimiento de la materia, sino también en cuanto a las herramientas pedagógicas utilizadas y su formación pedagógica como tal son determinantes entre los factores de riesgos. También influyen las dificultades presentadas frente a ciertas asignaturas que suelen retrasar la terminación a tiempo la carrera y cuya perdida y repetición es un factor de riesgo muy alto en la deserción universitaria, las asignaturas que causan mayor dificultad que son en un su orden: matemáticas, física, estadística, bioquímica y química.

La Universidad de Cartagena ostenta el título de la universidad con menor deserción de educación superior en el país, este no supera el 4% (estudio de la deserción estudiantil en la educación superior en Colombia. Convenio 107/2002. Universidad Nacional – ICFES) , resultado que se debe al estudio y análisis de las posibles causas y factores de riesgo de deserción al interior del estudiantado de la

institución que la llevan a aplicar estrategias efectivas para el control de dicho fenómeno.

Conclusión el Estudio de Deserción y Permanencia Académica en la Facultad de Ingenierías de la Universidad de La Guajira desde el II PA 2005 hasta el II PA 2007. Y basados en los resultados arrojados por el estudio realizado se concluye que: la facultad de Ingeniería tiene una tasa de deserción acumulada del 16,12% , es decir 5796 estudiantes matriculados durante los cinco periodos académicos objeto de este estudio, 930 en total presentaron algún tipo de deserción, lo que equivale a un promedio de 186 estudiantes retirados en cada periodo. El periodo con mayor índice de deserción fue el II semestre del 2006 con un porcentaje total del 17,84%, mientras que el periodo inmediatamente anterior presento el nivel más bajo correspondiente al 13,82%. Según el Ministerio de Educación el promedio de deserción para el año 2005 la deserción anual es del 13,4% en las instituciones de educación superior lo que indica que en Facultad de Ingeniería de la Universidad de La Guajira se encuentra por encima del promedio, es un dato preocupante para la Universidad de la Guajira el estudio ayudo también que ayudo a encontrar el motivo de la deserción de los estudiantes y es por que los estudiantes no tiene el perfil adecuado para los programas que se matriculan (40%), terminan transfiriéndose a otras carreras de las misma universidad en la mayoría ciencias económicas y administrativas, la facultad que tiene mayor índice de recepción de traslados es Ingeniería Industrial, con esto se debe tener en cuenta que la Universidad debe ser hacer un cuestionamiento a la hora de la admisión del estudiantado. Sin embargo, otros motivos que cuentan con un peso significativo dentro de la deserción definitiva están el bajo rendimiento académico (18%) y el factor económico (15%).

Link [http://www.colombiaaprende.edu.co/html/directivos/1598/articles-250629\\_recurso\\_6.pdf](http://www.colombiaaprende.edu.co/html/directivos/1598/articles-250629_recurso_6.pdf)

## 9. ENCUESTA SOBRE LAS TASAS Y RAZONES DE DESERCIÓN DE LOS ESTUDIANTES EN EL CURSO DE INFORMÁTICA DE "HELLENIC OPEN UNIVERSITY"

Es un hecho que la deserción escolar ocurre con bastante frecuencia en instituciones que poseen métodos de educación abierta o a distancia. (Eisenberg & Dowsett, 1990). Las tasas de deserción en el aprendizaje a distancia en el nivel universitario son definitivamente más altas que aquellas en las universidades convencionales y varían dependiendo del sistema educativo adoptado por el instituto que provee educación a distancia y de las materias de estudio seleccionadas (Narasimharao,1990). En Europa, las tasas de deserción van desde el 20% al 30% (Rumble,1992) mientras que en los países asiáticos estos



porcentajes pueden ser tan altos como el 50% (Shin & Kim, 1999), (Narasimharao, 1999). Estudios relevantes han mostrado que uno de los mayores factores que influyen en la deserción es el número de módulos completados por un estudiante; los estudiantes desertores son en la mayoría de los casos, estudiantes que dejaron de registrarse después de su primer o segundo módulo (Chyung et. al, 1998).

Los factores históricos que influyen en las deserciones a nivel universitario de educación a distancia pueden ser categorizados como: a) factores internos relacionados a la percepción y al lugar de control de los estudiantes, b) factores relacionados con el curso y los tutores y c) factores relacionados con ciertas características demográficas de los estudiantes.

La deserción, éste estudio ha mostrado que 7.9% de los estudiantes ni siquiera empezaron a seguir el curso de Informática, debe decirse que de este porcentaje 12.0% son mujeres y el 6.2% hombres. La edad de la mayoría de los estudiantes desertores oscila entre los 29 y 35 años, una edad que se caracteriza por una intensa actividad de trabajo. El porcentaje mencionado previamente (7.9%) incrementa si el porcentaje de (20.5%) correspondiente a los estudiantes que empezaron pero por alguna razón desertaron se adiciona a la cifra; además el porcentaje total de deserciones alcanza el 28.4% mientras que el de aquellos que continuaron es de 71.6%. Es importante mencionar que la gran mayoría de los estudiantes que desertaron lo hicieron sin completar el primer o segundo trabajo escrito. Como consecuencia es acertado concluir que la gran mayoría de los estudiantes que entregan los primeros dos trabajos de un módulo completarán el módulo y serán intitulados para participar de los exámenes finales.

Otro hecho es la existente pero no estadísticamente significativa correlación de los estudiantes desertores; edad: a mayor edad mayor es la posibilidad de desertar, llegando a la conclusión que los estudiantes mayores necesitan más apoyo por parte del tutor. Por el contrario la correlación entre las deserciones y el género no es tan importante pero existe; parece ser más difícil para las mujeres decidir empezar un curso pero ellas no desertan tan fácil como lo hacen los hombres. De todos modos en total las mujeres y los hombres tienen casi las mismas probabilidades de terminar sus estudios (el porcentaje de deserción en las mujeres es de 29.6% por el contrario el de los hombres es de 27.8%, este número resulta después de sumar las cifras de aquellos que desertaron y aquellos que ni siquiera empezaron el curso).

De acuerdo con éste estudio el estatus de la familia del estudiante no está relacionado con la tasa de deserción, sin embargo existe una fuerte correlación entre las deserciones y la existencia de educación previa en el campo de la informática y el desempeño de actividades profesionales que demandan el uso del computador. En cuanto a éste último también se indicó que aquellos que trabajan en el campo de la informática tienen las mismas posibilidades de desertar de aquellos que sólo usan un computador en el trabajo.

En cuanto a las razones dadas por los estudiantes para desertar, más de la mitad afirmó que no fueron capaces de estimar el tiempo preciso que tendrían que dedicar a su actividad profesional y como resultado el tiempo dedicado a su educación decreció inesperadamente. La segunda razón dada por aproximadamente uno de cada cuatro estudiantes fue la sensación de no tener el conocimiento suficiente para estudios de nivel universitario. Otras razones dadas por menos cantidad de estudiantes fueron el nacimiento de un hijo, un cambio grande en el trabajo, problemas de salud, fallar en la realización de los trabajos y el cumplimiento de plazos, asistencia insuficiente por parte del tutor y otras razones personales.

En resumen, éste estudio presentó una encuesta de dos años hecha a estudiantes desertores que se llevó a cabo en el campo de la educación universitaria en el área de Informática. La correlación entre las deserciones de los estudiantes y sus perfiles fueron examinadas y las razones de dichas deserciones fueron investigadas y presentadas. Los resultados podrían ser útiles como referencia en casos similares en programas de educación abierta que usan métodos de aprendizaje a distancia en el campo de la informática.

## **5. TIPO DE INVESTIGACIÓN**

Los métodos de investigación utilizados son: analítico, abstracción, observación científica, experimental, cuantitativo y dialectico.

Se eligen estos métodos de investigación por que permiten a través de la observación de fenómenos particulares la contemplación de alternativas que conduzcan a la evaluación anual o semestral y hasta la implementación de nuevos sistemas que beneficien la empresa, se pueden llegar a conclusiones y de igual manera cuantificar las variables analizadas, llevándolas a un comparativo entre costo -beneficio para la toma de decisiones.

### **5.1 METODO DESCRIPTIVO.**

El objetivo de la investigación descriptiva consiste en llegar a conocer las situaciones, costumbres y actitudes predominantes a través de la descripción exacta de las actividades, objetos, procesos y personas. Su meta no se limita a la recolección de datos, sino a la predicción e identificación de las relaciones que existen entre dos o más variables. Los investigadores no son meros tabuladores, sino que recogen los datos sobre la base de una hipótesis o teoría, exponen y resumen la información de manera cuidadosa y luego analizan minuciosamente los resultados, a fin de extraer generalizaciones significativas que contribuyan al conocimiento.

### **5.2 MÉTODO ANALÍTICO.**

Se distinguen los elementos de un fenómeno y se procede a revisar ordenadamente cada uno de ellos por separado. La física, la química y la biología utilizan este método; a partir de la experimentación y el análisis de gran número de casos se establecen leyes universales. Consiste en la extracción de las partes de un todo, con el objeto de estudiarlas y examinarlas por separado, para ver, por ejemplo las relaciones entre las mismas. Estas operaciones no existen independientes una de la otra; el análisis de un objeto se realiza a partir de la relación que existe entre los elementos que conforman dicho objeto como un todo; y a su vez, la síntesis se produce sobre la base de los resultados previos del análisis.

### **5.3 MÉTODO DE LA ABSTRACCIÓN.**

Es un proceso importantísimo para la comprensión del objeto, mediante ella se destaca la propiedad o relación de las cosas y fenómenos. No se limita a destacar

y aislar alguna propiedad y relación del objeto asequible a los sentidos, sino que trata de descubrir el nexo esencial oculto e inasequible al conocimiento empírico.

#### **5.4 OBSERVACIÓN CIENTÍFICA.**

El investigador conoce el problema y el objeto de investigación, estudiando su curso natural, sin alteración de las condiciones naturales, es decir que la observación tiene un aspecto contemplativo. La observación configura la base de conocimiento de toda ciencia y, a la vez, es el procedimiento empírico mas generalizado de conocimiento. Mario Bunge reconoce en el proceso de observación cinco elementos:

El objeto de la observación

El sujeto u observador

Las circunstancias o el ambiente que rodean la observación

Los medios de observación

El cuerpo de conocimientos de que forma parte la observación

#### **5.5 EL MÉTODO EXPERIMENTAL.**

Porque se comparan las mediciones del comportamiento en la UNIVERSIDAD TECNOLÓGICA DE PEREIRA con las mediciones y evaluaciones realizadas anteriormente.

#### **5.6 EL MÉTODO CUANTITATIVO.**

Porque se explican situaciones a través de la recolección de datos que permitan dar soluciones al problema planteados.

#### **5.7 MÉTODO DIALÉCTICO.**

La característica esencial del método dialéctico es que considera los fenómenos históricos y sociales en continuo movimiento. Dio origen al materialismo histórico, el cual explica las leyes que rigen las estructuras económicas y sociales, sus correspondientes superestructuras y el desarrollo histórico de la humanidad. Aplicado a la investigación, afirma que todos los fenómenos se rigen por las leyes de la dialéctica, es decir que la realidad no es algo inmutable, sino que está sujeta a contradicciones y a una evolución y desarrollo perpetuo. Por lo tanto propone que todos los fenómenos sean estudiados en sus relaciones con otros y en su

estado de continuo cambio, ya que nada existe como un objeto aislado. Este método describe la historia de lo que nos rodea, de la sociedad y del pensamiento, a través de una concepción de lucha de contrarios y no puramente contemplativa, más bien de transformación. Estas concepciones por su carácter dinámico exponen no solamente los cambios cuantitativos, sino los radicales o cualitativos.

## **5.8. RECURSOS DE LA INVESTIGACIÓN**

### **HUMANOS**

Docentes de la Universidad Tecnológica de Pereira:

✓ M.Sc. SANTIAGO VASQUEZ ARTUNDUAGA

### **DOCUMENTALES**

- IDENTIFICACION DE ASPECTOS QUE INFLUYEN EN LA MORTALIDAD ACADEMICA EN MATERIAS CRITICAS DE LOS PROGRAMAS DIURNOS DE LAS FACULTADES DE INGENIERIAS Y CIENCIAS DE LA SALUD ENTRE LOS AÑOS 2004 Y 2009 EN LA UNIVERSIDAD TECNOLOGICA DE PEREIRA
- ESTUDIOS DE DESERCIÓN ESTUDIANTIL  
<http://www.javeriana.edu.co/biblos/tesis/educacion/tesis81.pdf>
- TRABAJO DE DESERCIÓN ESTUDIANTIL  
[http://www.mineduacion.gov.co/1621/articles-85600\\_Archivo\\_pdf3.pdf](http://www.mineduacion.gov.co/1621/articles-85600_Archivo_pdf3.pdf)
- OBSERVATORIO DE LA FACULTAD DE INGENIERIA INDUSTRIAL UNIVERSIDAD TECNOLOGICA DE PEREIRA.
- SANTIAGO VASQUEZ ARTUNDUAGA

### **EQUIPOS E INSTRUMENTOS**

- Evaluaciones y Certificados anteriores
- Historiales y estudios anteriores de la empresa
- Computador.
- SPSS

- Impresora.
- Software Office e Internet.
- Calculadora.
- USB.
- Instrumentos de recolección de datos.
- Papelería.
- Utensilios de Oficina (Lápiz, Lapicero, Borrador, Cosedora, Block)

## **LOCALES**

- Instalaciones de la UNIVERSIDAD TECNOLÓGICA
- Oficina de la Facultad de Ingeniería Industrial.
- Biblioteca Jorge Roa Martínez, Universidad Tecnológica de Pereira.
- Salas virtuales Facultad de Ingeniería Industrial y el Centro de Recursos.
- Centro de Documentación de la Facultad de Ingeniería Industrial.
- Observatorio de la Facultad de Ingeniería Industrial.

## PRESUPUESTO Y FINANCIACION

CONCEPTO	VALOR
Papelería, Fotocopias e Internet	\$ 450.000
Investigación Exploratoria	\$ 80.000
Transporte	\$ 350.000
Manejo de Información (Software Office, Asesoría)	\$ 3.000.000
Documento Final (Software, Planos, Fotografías, Impresión)	\$ 500.000
<b>TOTAL</b>	<b>\$ 4.380.000</b>





## 6. MORTALIDAD ACADÉMICA

Las universidades orientan la mortalidad en dos aspectos, la mortalidad por semestre y la mortalidad por materias.

- La mortalidad por materias surge cuando se tiene nota definitiva al terminar el semestre por debajo de 3,0 o al final del semestre se cancela dicha materia para repetirla al próximo semestre. Esta es la clase de mortalidad objeto de estudio de este proyecto.
- La mortalidad por semestre se genera cuando el promedio al finalizar el semestre está por debajo de 3,0 generando los estados de clasificación utilizados por el observatorio de la universidad tecnológica. Estos estados de bajo rendimiento son:  
**Estado periodo de prueba:** promedio (0,0-2.5)  
**Estado de transición:** promedio (2,6-2,9)  
**Estado de cancelación de matrícula:** promedio (0)

Encontramos factores o variables que generan y aumentan la mortalidad académica. Es necesario identificar estos factores y variables para poder pronosticarlo y combatirlo desde un principio. Generando soluciones desde el inicio previniendo que aumente o se genere mortalidad académica.

Algunos factores pueden ser internos de la Universidad y asociados al rendimiento académico, a interactuar con los compañeros y los profesores, o al grado de dificultad de las materias asociados al nivel académico. También hay factores externos que influyen en la mortalidad académica como son los socios económicos o la vida familiar.

Los factores y las variables pueden clasificarse en académicos, interpretación personal y socioeconómica. En los académicos están las exigencias mínimas que los debe estudiantes deben cumplir y los prerrequisitos para cursar una materia. Los factores que afecten más probablemente la mortalidad son los académicos porque son los que generan las exigencias y los prerrequisitos para el estudiante cursar y aprobar las materias afectando directamente la capacidad de los estudiantes y demostrando como responden a cumplir dichas exigencias.

Estas exigencias tienen que ver por ejemplo con el número de materias por semestre dentro de la carrera, número de horas a la semana y composición del plan de estudios por áreas del conocimiento.

Los factores externos son mas con la formación preuniversitaria de los estudiantes, para esto se analiza las procedencias y los niveles educativos de los estudiantes, como la proveniencia del colegio entre otros.

## 6.1. ANALISIS DESCRIPTIVO

Según los estudios anteriores (información secundaria) las materias que generan deserción en la Universidad Tecnológica de Pereira según el observatorio académico que mantiene pendiente de el rendimiento académico de los estudiantes.

Las materias de mortalidad son:

Primera materia por carrera que genera mayor mortalidad

- ING. SISTEMAS=Matemáticas I
- ING ELECTRICA= Matemáticas I
- ING. FISICA= Matemáticas I
- ING. MECANICA= Matemáticas I

Segunda materia por carrera que genera mortalidad académica

- ING. SISTEMAS=Introducción a la Informática
- ING ELECTRICA= Física I
- ING. FISICA= Matemáticas II
- ING. MECANICA= Matemáticas II

Tercera materia por carrera que genera mortalidad académica

- ING. SISTEMAS=Introducción a la Informática
- ING ELECTRICA= Física I
- ING. FISICA= Matemáticas II
- ING. MECANICA= Matemáticas II

Estos resultados fueron obtenidos del observatorio académico.

Este análisis descriptivo de la situación de mortalidad trata no solo de evidenciar la tendencia y el comportamiento de los estudiantes de las facultades y programas objeto de estudio dentro de dichas materias, sino además de justificar y dejar en claro que el problema es real, que se presenta desde hace un tiempo y que es necesario por parte de los directivos universitarios tomar cartas en el asunto con distintas estrategias tanto académicas como extra académicas con el fin de minimizarlo.

El análisis se realizo por programa estudiando algunas de las materias que generan deserción y se realizo para facultadde ingenierías como un conglomerado para matemáticas 1 que es la materia que mas mortalidad académica genera desde los periodos 2011-1 al 2013-1. El estudio se realizo con la poca información que la Universidad nos brindo en la base de datos.

Vale la pena aclarar que para la cuantificación de personas que matricularon las materias y por ende el posterior cálculo del índice de mortalidad no se tuvieron en cuenta:

- Personas que matricularon y aprobaron o reprobaron en intersemestral
- Personas que matricularon y aprobaron o reprobaron en pruebas de suficiencia
- Personas a las que se les fue acreditada u homologada la materia
- Personas que cancelaron
- Personas que abandonaron

La razón es que se desea calcular índices de mortalidad en periodos normales y por medio de la asistencia completa al curso durante un periodo académico.

Analizaremos alguna de las materias que generan mortalidad académica que se nos fue entregada en la base de datos que se solicitó a la universidad tecnológica, la base de datos que recibimos no cumplió con todos los requisitos o exigencias que se pidió en la base, por ende solo miraremos algunas materias de las cuales tenemos información.

El año 2010 fue pronosticado por qué no se tiene información de estos dos periodos.

En cada tabla se presenta el número de estudiantes que se matricularon en cada semestre, cuantos perdieron por semestre y el índice de mortalidad por semestre.

## INGENIERÍA ELÉCTRICA

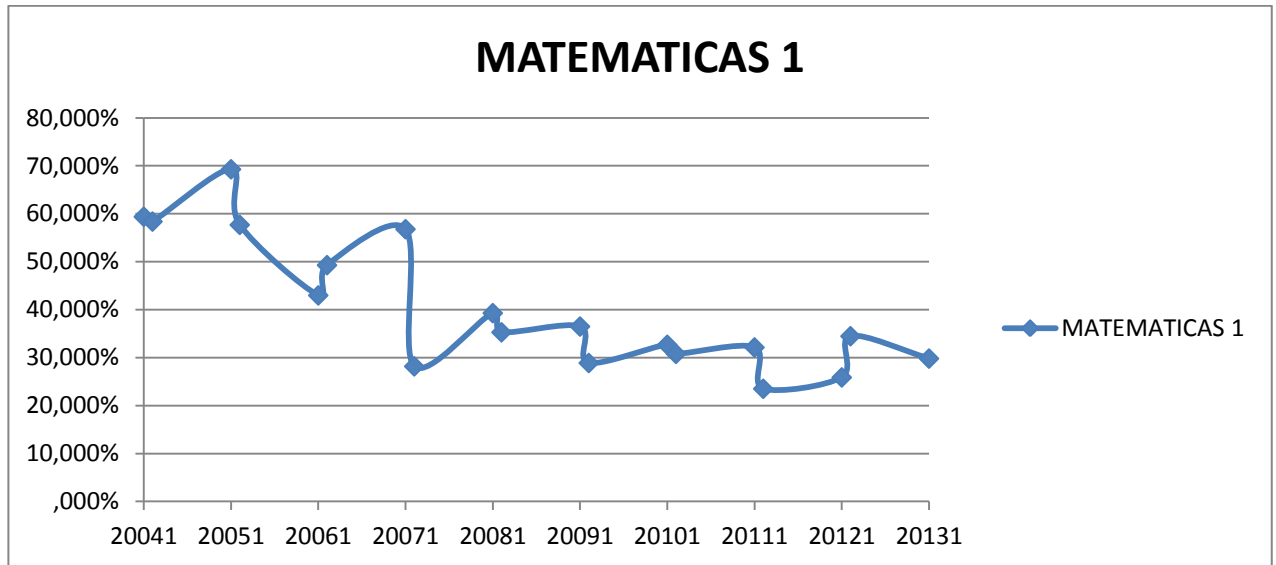
La siguiente tabla contiene la información de matemáticas 1 en los diferentes periodos analizados.

**Tabla 1: Ingeniera Eléctrica Matemáticas 1**

Semestre	mortalidad	Total estudiantes perdieron	total estudiantes inscritos
20041	59,40%	63	106
20042	58,40%	52	89
20051	69,30%	70	101
20052	57,70%	45	78
20061	43,00%	43	100
20062	49,30%	34	69
20071	56,80%	46	81
20072	28,20%	22	78
20081	39,30%	33	84
20082	35,30%	24	68
20091	36,50%	38	104
20092	28,90%	22	76
20101	32,70%	28	86
20102	30,80%	25	81
20111	32,14%	27	84
20112	23,53%	16	68
20121	25,89%	29	112
20122	34,48%	30	87
20131	29,82%	34	114

Podemos observar que en los últimos semestres han bajado los niveles de mortalidad. En la siguiente grafica se ilustran los índices.

**Figura 6: Grafica Matemáticas 1 Ingeniería Eléctrica**



Podemos observar que el comportamiento tiene una tendencia decreciente después de haber presentado su máximo valor en el 2005-1 del 69.3 % disminuyendo en el último periodo al 29.82% en el primer periodo del 2013-1.

El promedio de mortalidad en esta materia para Ingeniería Eléctrica es de 40,60%, valor que es realmente alto ya que es cercano al 50% lo que indicaría que casi la mitad de personas que matriculan Matemáticas I en Ingeniería Eléctrica la pierden.

## INGENIERÍA FÍSICA

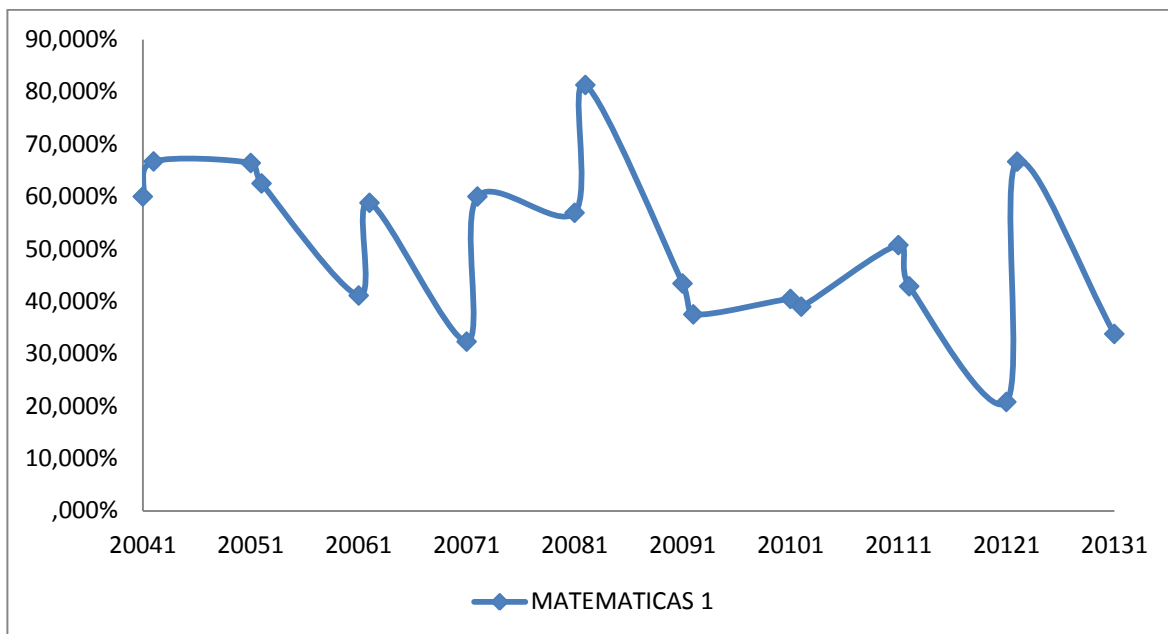
La siguiente tabla contiene la información de matemáticas 1 en los diferentes periodos analizados.

**Tabla 2: Ingeniería Física Matemáticas 1**

Semestre	mortalidad	Total estudiantes perdieron	total estudiantes inscritos
20041	60,00%	51	85
20042	66,70%	10	15
20051	66,40%	71	107
20052	62,50%	15	24
20061	41,10%	30	73
20062	58,80%	10	17
20071	32,30%	20	62
20072	60,00%	9	15
20081	56,90%	37	65
20082	81,30%	13	16
20091	43,40%	23	53
20092	37,50%	6	16
20101	40,45%	15	35
20102	38,98%	10	25
20111	50,72%	35	69
20112	42,86%	6	14
20121	20,78%	16	77
20122	66,67%	8	12
20131	33,77%	26	77

Se evidencian los altos índices de mortalidad de matemáticas 1 en la carrera de ingeniería física solo en el periodo 2012-1 tuvo el índice más bajo del 20,78 y el índice más alto fue en el 2008-2 del 81,30 %.

**Figura 7: Grafica Matemáticas 1 Ingeniería Física**



En la grafica estos índices son muy variables entre ellos no poseen una tendencia estable ni de crecimiento o decrecimiento. En la siguiente grafica podemos evidenciar bien los índices de mortalidad de matemáticas 1. El promedio es de 50.59 % el cual es el más alto presentado hasta ahora en las distintas materias y programas

A continuación la tabla presenta información relacionada con el programa de ingeniería física en lo que compete con la materia Matemáticas II.

**Tabla 3: Ingeniería Física Matemáticas 2**

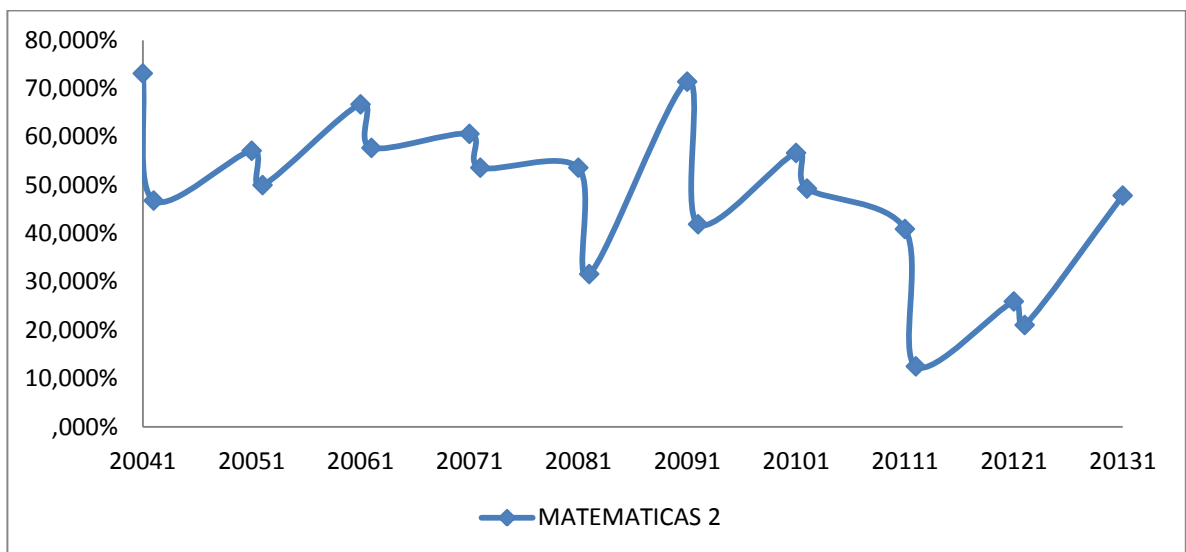
Semestre	mortalidad	Total estudiantes perdieron	total estudiantes inscritos
20041	73,10%	19	26
20042	46,80%	22	47
20051	57,10%	12	21
20052	50,00%	22	44
20061	66,70%	20	30
20062	57,70%	30	52
20071	60,60%	20	33
20072	53,60%	15	28
20081	53,60%	15	28
20082	31,60%	6	19
20091	71,40%	15	21
20092	41,90%	13	31
20101	56,65%	14	26

20102	49,28%	13,5	28,5
20111	40,91%	9	22
20112	12,50%	3	24
20121	25,93%	7	27
20122	21,05%	8	38
20131	47,83%	11	23

De la tabla anterior se puede observar que los principales índices de mortalidad se dan en los primeros periodos de la cohorte.

A continuación se grafican los datos para tener una idea más clara sobre este comportamiento de los índices de mortalidad en la materia Matemáticas II que se dan en el programa de Ingeniería Física.

**Figura 8: Grafica Matemáticas 2 Ingeniería Física**



Como se evidencia en la gráfica no sigue una tendencia, ni es creciente ni decreciente, es muy variable según los periodos. El mayor índice se presentó en el periodo en el 2004-1 del 73,10% y el índice más bajo en el 2011-2 de 12,50%.

El promedio total de mortalidad en Matemáticas II es de 48,33 % lo cual sigue siendo muy alto, este promedio está cerca del 50%, esto indica que casi mitad de las personas que matriculan esta materia en este programa la pierden normalmente.



## INGENIERÍA EN SISTEMAS Y COMPUTACIÓN.

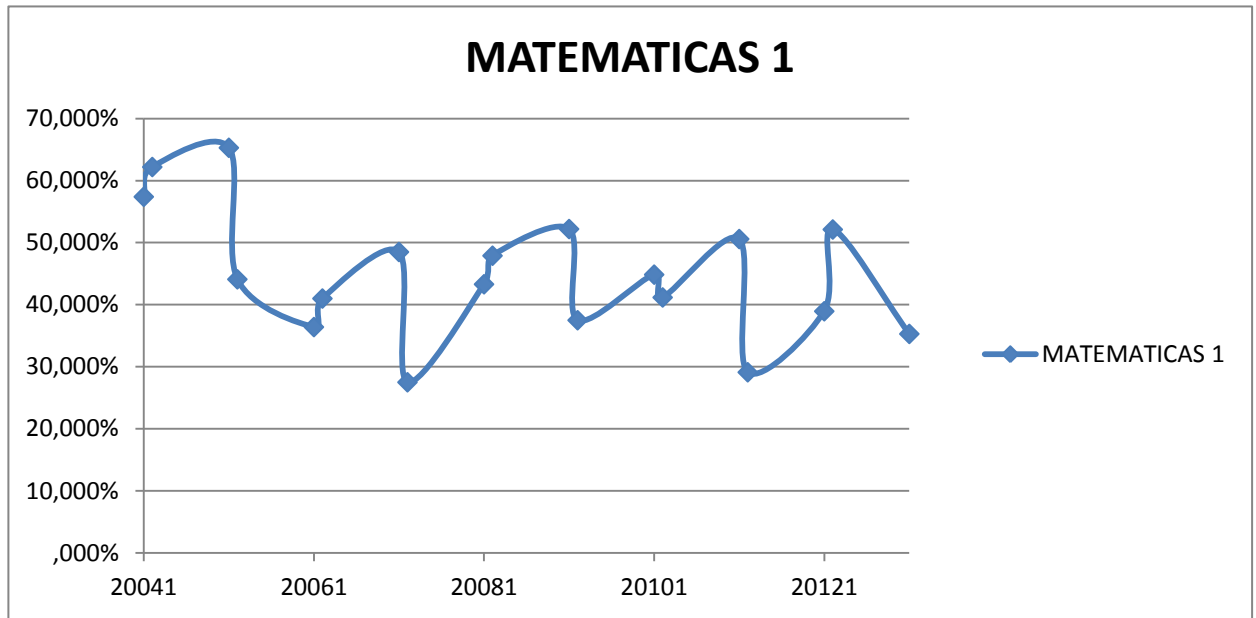
En la tabla siguiente se presentan los índices de mortalidad académica de matemáticas 1 para ingeniería en sistemas.

**Tabla 4: Ingeniería en Sistemas y Computación Matemáticas 1**

Semestre	mortalidad	Total estudiantes perdieron	total estudiantes inscritos
20041	57,40%	66	115
20042	62,20%	56	90
20051	65,30%	62	95
20052	44,10%	30	68
20061	36,40%	28	77
20062	41,00%	25	61
20071	48,50%	33	68
20072	27,50%	14	51
20081	43,30%	42	97
20082	47,90%	23	48
20091	52,20%	48	92
20092	37,50%	30	80
20101	44,85%	39	86
20102	41,18%	35	83
20111	50,59%	43	85
20112	29,11%	23	79
20121	38,95%	37	95
20122	52,13%	49	94
20131	35,29%	36	102

Se evidencian los índices de mortalidad más altos en los periodos cohorte. En la gráfica se muestra más claro los índices y su comportamiento a través de los periodos.

**Figura 9: Grafica Matemáticas 1 Ingeniería en Sistemas y Computación Matemáticas** **1**



Como se evidencia en la gráfica no sigue una tendencia, ni es creciente ni decreciente, es muy variable según los periodos. El mayor índice se presentó en el periodo en el 2005-1 del 65,30% y el índice más bajo en el 2007-2 de 27,50%.

El promedio total de mortalidad en Matemáticas II es de 45,02 % lo cual sigue siendo muy alto, ya que casi la mitad de las personas que matriculan esta materia en este programa la pierden normalmente.

**INFORMATICA 1**

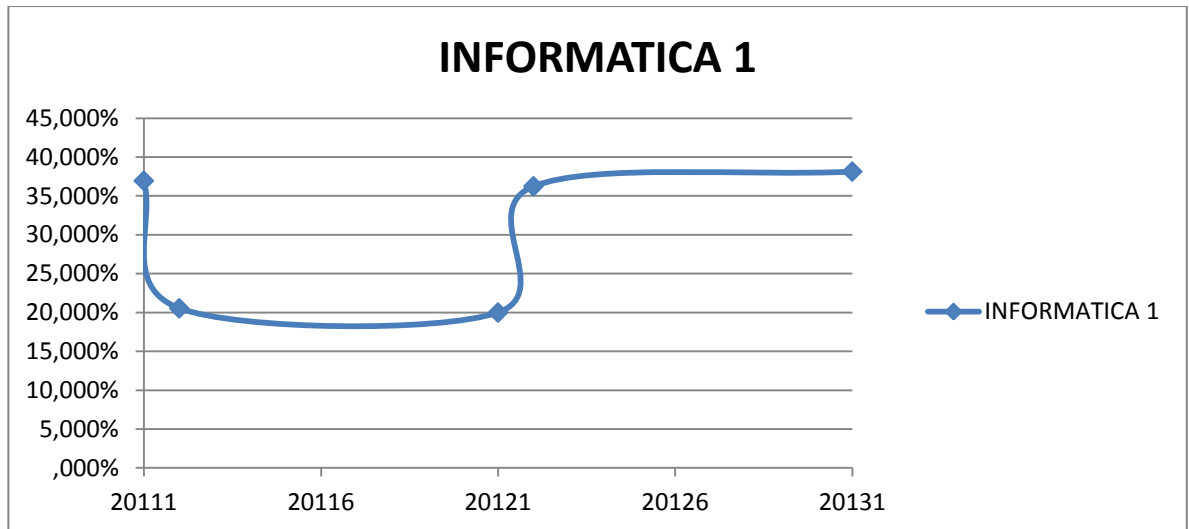
En la tabla siguiente se presentan los índices de mortalidad académica de informática 1 para ingeniería en sistemas en los periodos del 2011-1 hasta el 2013-1.

**Tabla 5: Ingeniería en Sistemas y Computación, Informática 1**

Semestre	mortalidad	Total estudiantes perdieron	total estudiantes inscritos
20111	36,96%	34	92
20112	20,55%	15	73
20121	20,00%	17	85
20122	36,25%	29	80
20131	38,14%	37	97

En los últimos periodos aumento el índice de mortalidad académica. En la grafica se pueden observar mejor las tendencias de los resultados.

**Figura 10: Grafica Informática 1 Ingeniería en Sistemas y Computación**



Como se evidencia en la gráfica sigue una tendencia creciente en los últimos periodos. El mayor índice se presentó en el periodo en el 2013-1 del 38,14% y el índice más bajo en el 2012-1 de 20,00%.

El promedio total de mortalidad en Matemáticas II es de 30,38 % lo cual es uno de los índices más bajo en el estudio, igualmente podemos observar que tiene una tendencia creciente.

## INGENIERÍA MECÁNICA.

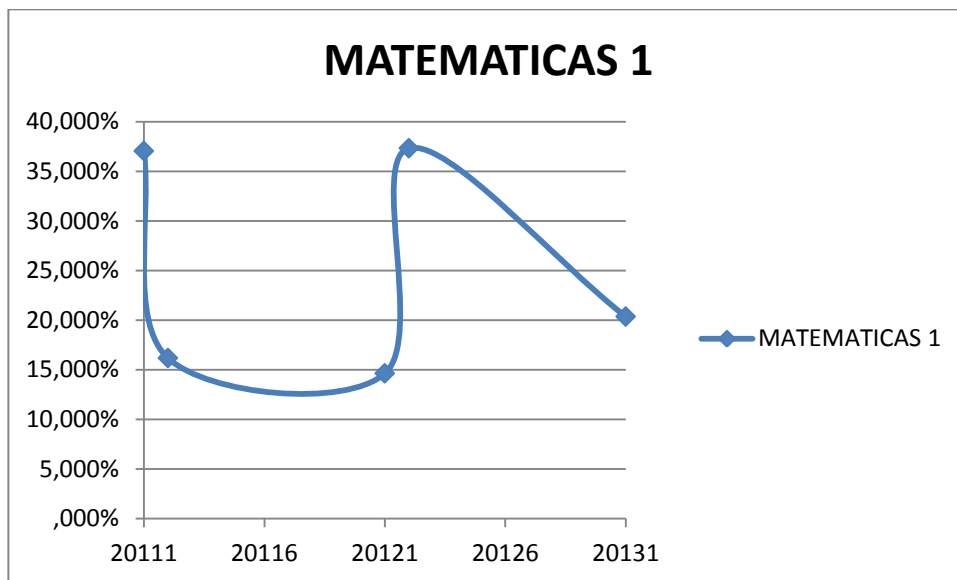
En la tabla siguiente se presentan los índices de mortalidad académica de matemáticas 1 para ingeniería en mecánica.

**Tabla 6: Ingeniería Mecánica, Matemática 1**

Semestre	mortalidad	Total estudiantes perdieron	total estudiantes inscritos
20111	37,04%	40	108
20112	16,18%	11	68
20121	14,63%	12	82
20122	37,31%	25	67
20131	20,35%	23	113

Se evidencian que los índices de mortalidad no siguen una tendencia, los índices son variables. En la grafica se muestra más claro los índices y su comportamiento a través de los periodos.

**Figura 11: Grafica Matemáticas 1 Ingeniería Mecánica**



Como se evidencia en la grafica sus índices son muy variables. El mayor índice se presentó en el periodo en el 2012-2 del 37,31% y el índice más bajo en el 2012-1 de 14,63%.

El promedio total de mortalidad en Matemáticas II es de 25,10 % lo cual es el índice más bajo en el estudio, igualmente podemos observar que tiene no tiene una tendencia a seguir y que bajo mucho en el último periodo.

## MATEMATICAS 2

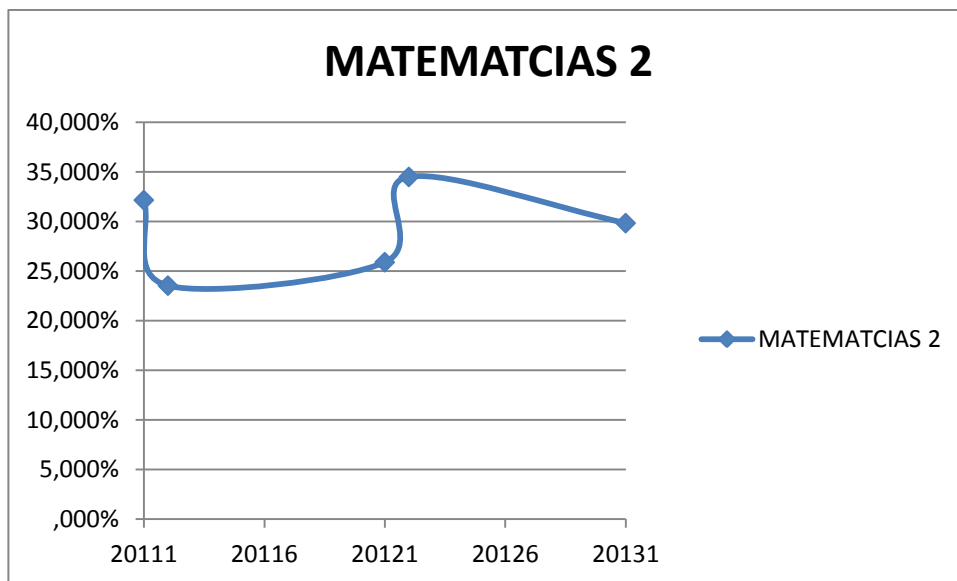
En la tabla siguiente se presentan los índices de mortalidad académica de matemáticas 2 para ingeniería en mecánica.

**Tabla 7: Ingeniería Mecánica, Matemática 2**

Semestre	mortalidad	Total estudiantes perdieron	total estudiantes inscritos
20111	32,14%	27	84
20112	23,53%	16	68
20121	25,89%	29	112
20122	34,48%	30	87
20131	29,82%	34	114

Se evidencian que los índices de mortalidad no siguen una tendencia, los índices son variables. En la grafica se muestra más claro los índices y su comportamiento a través de los periodos.

**Figura 12: Grafica Matemáticas 2 Ingeniería Mecánica**



Como se evidencia en la grafica sus índices son muy variables. El mayor índice se presento en el periodo en el 2012-2 del 34,48% y el índice más bajo en el 2011-2 de 23,53%.

El promedio total de mortalidad en Matemáticas II es de 29,17 % lo cual es uno de los índices más bajo en el estudio, igualmente podemos observar que tiene no tiene una tendencia a seguir y que bajo mucho en el último periodo.

## **FACULTAD DE INGENIERÍAS**

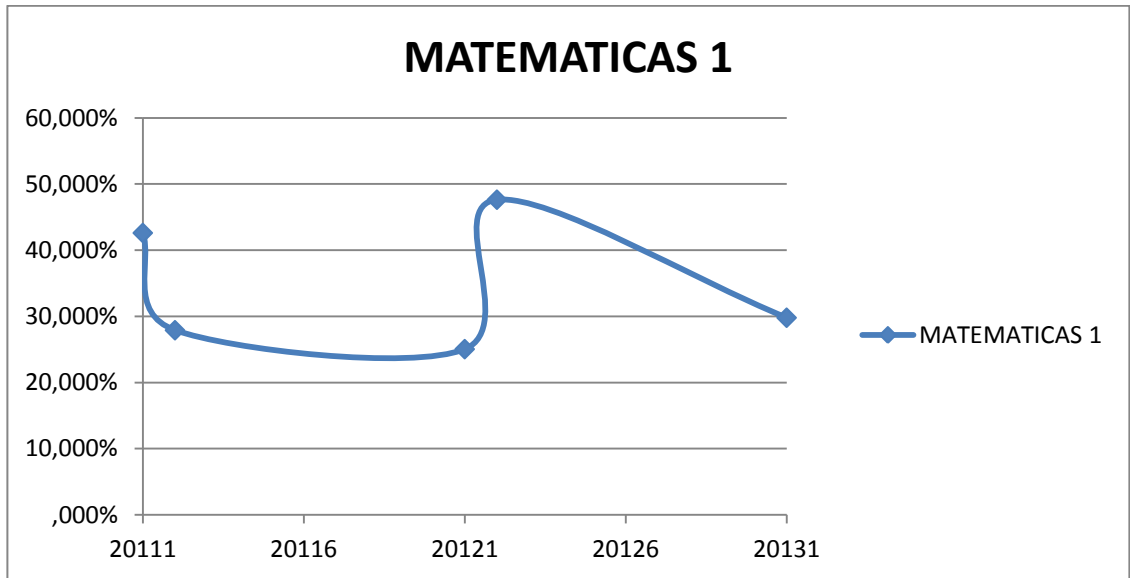
En esta parte se realiza un estudio conglomerado de los programas estudiados en los periodos de 2011-1 hasta el 2013-1. En la tabla se encuentran los índices de mortalidad para la materia Matemáticas I en la Facultad de Ingenierías.

**Tabla 8: Facultad de Ingenierías, Tabla de mortalidad**

Semestre	mortalidad	Total estudiantes perdieron	total estudiantes inscritos
20111	42,62%	36	87
20112	27,92%	14	57
20121	25,06%	24	92
20122	47,65%	28	65
20131	29,81%	30	102

Se evidencia que los índices no sigue un comportamiento, no poseen una tendencia ni creciente ni decreciente. En la grafica se puede ver más claro el comportamiento de los índices.

**Figura 13: Grafica Matemáticas 1 Facultad de Ingeniería**



Como se evidencia en la gráfica sus índices son muy variables. El mayor índice se presentó en el periodo en el 2012-2 del 47,65% y el índice más bajo en el 2012-1 de 25,06%.

El promedio total de mortalidad en Matemáticas I es de 34,61 % para los periodos 2011-1 al 2013-1 lo cual es uno de los índices más bajo en el estudio, igualmente podemos observar que tiene no tiene una tendencia a seguir y que bajo mucho en el último periodo.

En general podemos concluir para la facultad de ingenierías lo siguiente:

- Los índices de mortalidad de la materia matemáticas 1 en la Facultad de Ingenierías son altos, presenta su menor promedio en matemáticas IV con el 25,06% aproximadamente en los periodos 2011-1 al 2013-1.
- El alto índice de mortalidad en matemáticas 1 en la facultad de ingenierías como en las otras materias que generan mortalidad que se pudieron analizar por cada programa sus índices son muy altos, siendo matemáticas 1 el mayor de ellos.
- El alto índice de mortalidad en matemáticas 1 puede ser atribuido por la inmadurez de los estudiantes al momento de ingresar a la Universidad y evidenciando las altas deficiencias en las materias de ciencias básicas que poseen los colegios de procedencia de los estudiantes.

- En casi todas las gráficas es común que los índices de mortalidad bajen, exceptuando a matemáticas 2 en ingeniería física y a informática 1 en ingeniería en sistemas, la razón de que la mayoría de los índices bajen puede ser la aplicación de las estrategias que se han aplicado en la Universidad como soluciones que genera el observatorio estudiantil para combatir esta problemática, ejemplo los tutoriales, el programa de introductorio a matemáticas entre otros.
- Igualmente se evidencia que todavía son altos los índices de mortalidad académica por lo cual es necesario la intervención con nuevos programas y soluciones para disminuir la problemática en estas materias.
- Por eso se hace necesario saber las causas directas e indirectas que generan esta problemática en las materias que generan mortalidad y así poder plantear soluciones para disminuir los índices.

## **7. APLICACIÓN DE MODELOS MULTIVARIADOS EN EL PLANTEAMIENTO DE UN MODELO DE PRONÓSTICO PARA LA CLASIFICACION DE ESTUDIANTES CON RIESGO DE DESERCIÓN ESTUDIANTIL CRÍTICAS DE LAS FACULTADES DE INGENIERIAS**

El objetivo de la presente encuesta es recolectar información que permita caracterizar las condiciones socioeconómicas, personales e internas de la Universidad Tecnológica de Pereira para realizar el debido estudio de deserción estudiantil.

### **GENERO**

### **3. En que rango se encuentra su**



**edad?**

- a) De 15 a 18 años
- b) De los 18 a los 20 años
- c) De los 20 a los 25 años
- d) Mayor de 25 años

**\*4. Su estrato social es**

- a) 1
- b) 2
- c) 3
- d) 4
- e) 5
- f) 6

**\*5. De qué departamento procede usted?**

- a) Risaralda
- b) Quindío
- c) Choco
- d) Caldas
- e) Valle
- f) Tolima
- Otro (Por favor especifique)

**\*6. El colegio del que proviene es**

- a) Femenino
- b) Masculino
- c) MIXTO

**\*7. El colegio del que proviene es:**

- a) Privado
- b) Semiprivado\*
- c) Publico

**\*8. Estado civil:**

- a) Soltero
- b) Casado
- c) Unión libre
- d) Viudo

**\*9. Tiene casa:**

- a) Propia
- b) Arrendada
- c) Usufructo

**\*10. Vive usted en estos momentos:**

- a) Solo
- b) Con sus padres
- c) Con algún familiar
- d) Con su pareja
- e) Con compañeros de la u

**\*11. Su situación económica en la universidad es:**

- a) Depende de sus padres
- b) Se encuentra laborando
- c) Tiene una beca estudiantil
- d) Está financiando por algún programa
- e) Tiene apoyo económico de un familiar

**\*12. Presenta alguna dificultad familiar?**

- a) Si
- b) No

**\*13. Se le presentan cruces del horario laboral con el horario de clases**

- a) Si
- b) No

**\*14.Cuál es el nivel educativo alcanzado por su padre**

- a) Básica primaria
- b) Básica secundaria
- c) Cursos básico
- d) Técnicos
- e) Universitarios

**\*15.Cuál es el nivel educativo alcanzado por su madre**

- a) Básica primaria
- b) Básica secundaria
- c) Cursos básico
- d) Técnicos
- e) Universitarios

**\*16. En que se desempeña su padre**

- a) Trabajador independiente
- b) Empleado público
- c) Empresario
- d) Empleo formal en compañía privada
- e) Jubilado

**\*17. En que se desempeña su madre**

- a) Trabajador independiente
- b) Empleado público
- c) Empresario
- d) Empleo formal en compañía privada
- e) Jubilado
- f) Ama de casa

**\*18. Presenta alguna discapacidad?**

- a) Si
- b) No

**\*19. Tiene usted personas a cargo:**

- a) Si

- b) No

**20. Como es el acompañamiento de la familia con su vida universitaria**

- a) Siempre existe total apoyo económico, intelectual y afectivo
- b) Recibe algún respaldo económico y a veces se interesan por su vida universitaria
- c) Tiene apoyo económico limitado o condicionado; y rara vez preguntan por su vida universitaria
- d) No recibe ningún apoyo , ni reciben ninguna oposición
- e) Su familia se opone y obstaculiza su vida universitaria

**21. Hace cuanto se graduó del bachillerato**

- a) El año pasado
- b) 1 a 2 años
- c) 3 a 5 años
- d) Más de 5 años

**\*22. El tiempo que le dedica a sus estudios es**

- a) Muy bueno
- b) Bueno
- c) Regular
- d) Malo
- e) Muy malo

**\*23. Cree usted que sus métodos de estudio si funcionan**

- a) Si
- b) No

**\*24. En estos momentos usted se encuentra**

- a) Muy contento en su carrera
- b) Motivado para cursas este semestre
- c) Cansado en el semestre
- d) No está feliz en su carrera

**\*25. Cree que esta en el programa adecuado**

- a) Si
- b) No

**\*26. Tenía conocimientos previos del programa académico?**

- a) Si
- b) No

**\*27. Por qué eligió su carrera**

- a) Por gusto
- b) Por sus conocimientos previos
- c) Por orientación profesional
- d) Por seguir las recomendaciones familiares
- e) El puntaje del icfes no le alcanzo para lo que quería estudiar
- f) Porque tiene buen campo de acción laboral

**\*28. Se ha cambiado en el transcurso de la universidad de carrera?**

- a) Si
- b) No

**\*29. Si a la pregunta anterior respondió si por qué?**

- a) Situación económica
- b) Orientación profesional
- c) Decisión propia
- d) Por dificultades de aprendizaje
- e) Por la dificultad de la carrera

**\*30. En el transcurso de la carrera se ha encontrado docente con**

- a) Mala metodología para transmitir sus conocimientos
- b) Sin conocimiento sobre la materia
- c) El docente dictaba una asignatura que no se correspondía con el perfil para el cual estaba formado.
- d) El docente manifestaba serias deficiencias en pedagogía

**\*31. Cual ha sido la relación con sus profesores?**

- a) Muy buena
- b) Buena
- c) Regular
- d) Mala
- e) Muy mala

**\*32. Las relaciones con sus compañeros han sido:**

- a) Una buena relación
- b) Tiene problemas con sus compañeros
- c) No interactúa con sus compañeros
- d) La relación con sus compañeros le ha traído dificultades en la carrera

**\*33. Realizo usted en primer semestre el examen de orientación profesional?**

- a) Si
- b) No

**\*34. Al presentar el examen de orientación profesional:**

- a) Le ayudo a seleccionar una buena carrera
- b) Se encuentra usted en la carrera correcta

**\*35. Como califica la calidad de su programa de estudio?**

- a) Muy buena
- b) Buena
- c) Regular
- d) Mala
- e) Muy mala

**\*36. Como califica su pensum académico?**

- a) Muy buena
- b) Buena
- c) Regular
- d) Mala

e) Muy mala

**\*37. Se identifica con el programa de estudio de su carrera?**

- a) Si
- b) No

**\*38. Ha tenido dificultades para cursar este semestre**

- a) Si
- b) No

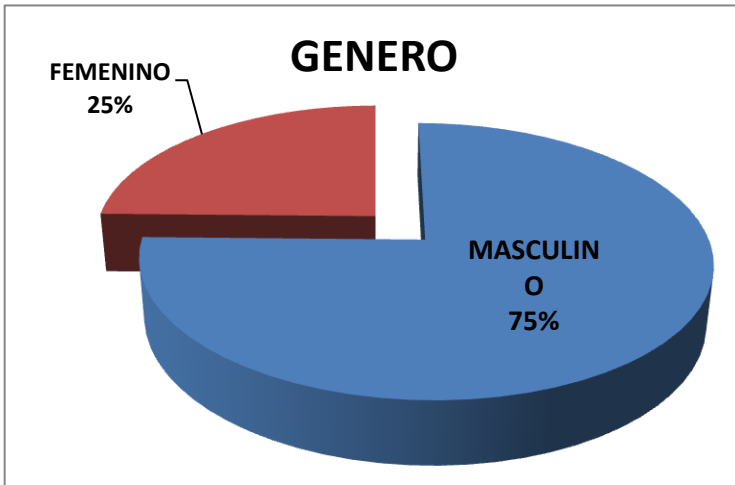
**\*39. Si a la pregunta anterior contesto si, qué tipo de dificultad**

- a) Económicas
- b) Familiares
- c) Laborales
- d) Con sus compañeros
- e) Con sus profesores
- f) O personales

**40. A que programa pertenece?**

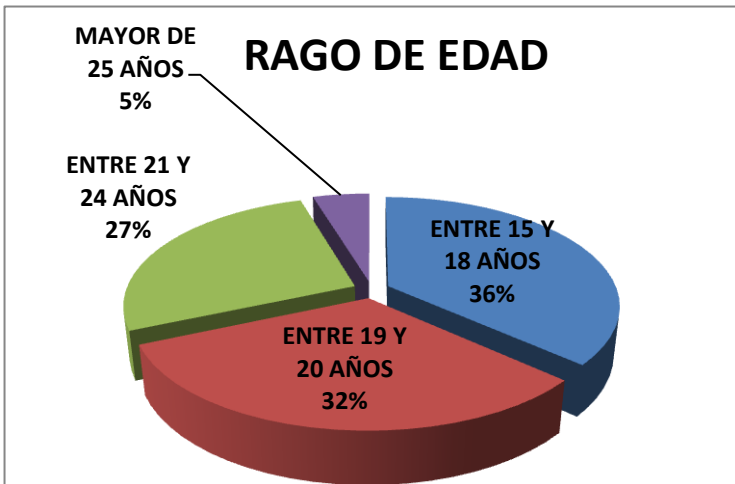
## 8. GRAFICAS DE CADA UNA DE LAS PREGUNTAS

Figura 14: Grafica Resultado encuesta Género



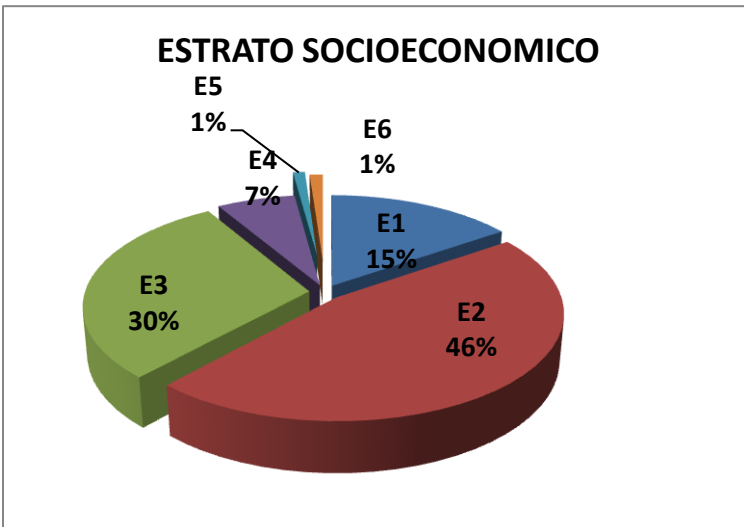
**GENERO.** Se observa que en la población encuestada tenemos un 25% femenino y un 75% es masculino, teniendo en cuenta que las Ingenierías que se encuestaron son las que se evidencia menor número de mujeres matriculadas tienen

Figura 15: Grafica Resultado encuesta Rango de edad



**RANGO DE EDAD.** Aquí evidenciamos que la población pertenece a la gente de la jornada diurna, es una población muy joven, el 85% de la población no supera los 25 años y la población más grande esta entre los 15 y 18 años

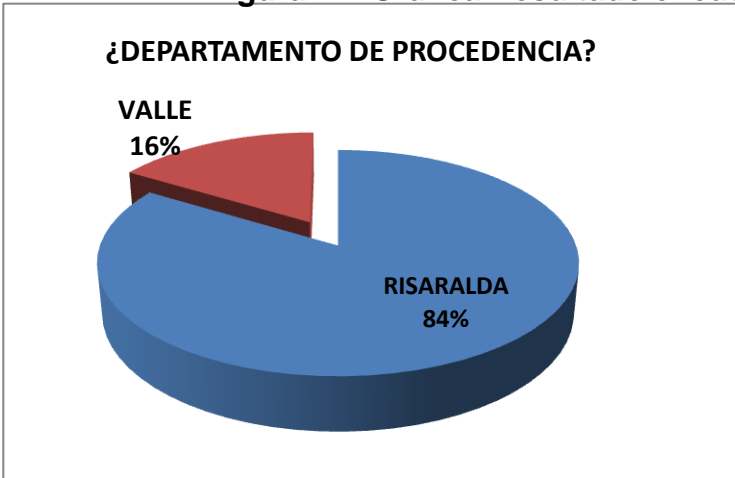
Figura 16: Grafica Resultado encuesta Estrato socioeconómico



**ESTARTO SOCIOECONOMICO.** Se evidencia en esta grafica que el 61% de la población pertenecen a los estratos más vulnerables, quienes hacen los esfuerzos para salir adelante y mejorar el estilo de vida propio y de su familia

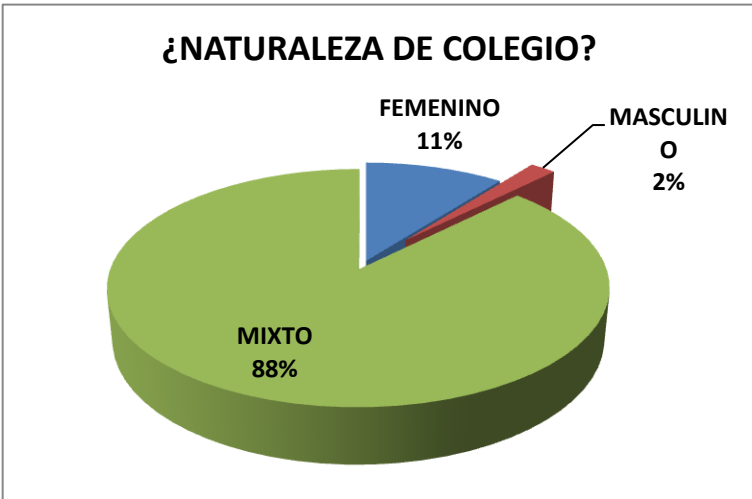


Figura 17: Grafica Resultado encuesta Departamento procedencia



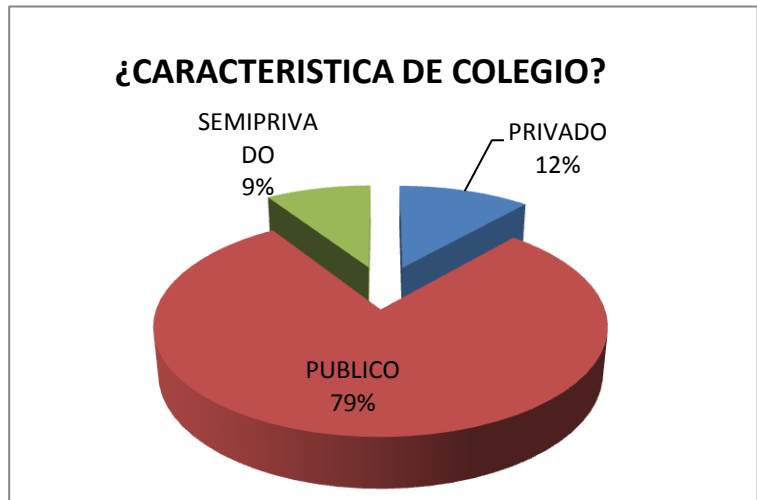
**DEPARTAMENTO DE PROCEDENCIA.** La población dominante es la de la región con un 84%

Figura 18: Grafica Resultado encuesta Naturaleza colegio



**NATURALEZA DEL COLEGIO.** Como vemos la gran mayoría de los estudiantes vienen de colegios mixtos con un 88% de la población

Figura 19: Grafica Resultado encuesta característica de colegio



**CARACTERISTICAS DEL COLEGIO.** Como se ve en la gráfica **estrato socioeconómico** y **naturaleza de colegio** esta grafica coincide con los resultados de las anteriores, pues los colegios mixtos son donde estudian las poblaciones de estratos 1 y 2 , además los colegios públicos en su mayoría son mixtos, vemos que los públicos tienen un 79% de la población encuestada

**Figura 20: Grafica Resultado encuesta ha tenido dificultad para pagar la universidad**



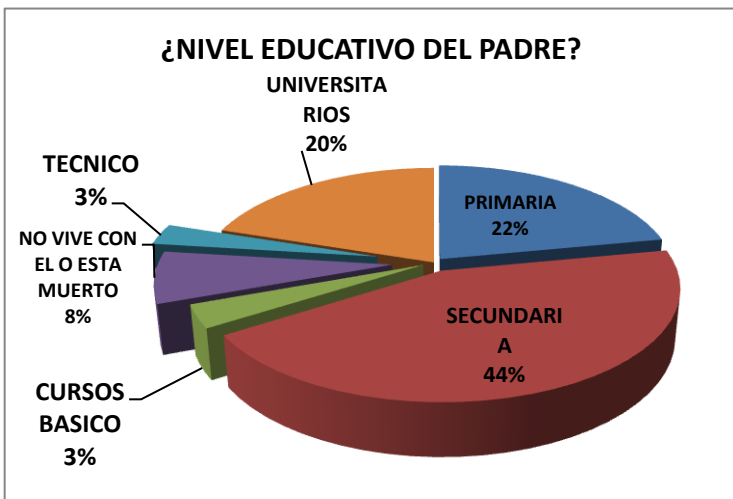
**DIFICULTAD PARA PAGAR LA UNIVERSIDAD.** A la respuesta de esta pregunta se analiza que tiene relación con las anteriores, pues así el semestre sea más económico que cualquier otra universidad, los padres de estos estudiantes se ven en dificultades, a la hora de pagar el semestre ya que sus salarios no superan los 2 salarios mínimos legales vigentes y los núcleos familiares son algo numerosos

**Figura 21: Grafica Resultado encuesta se le presenta cruces con el horario de trabajo**



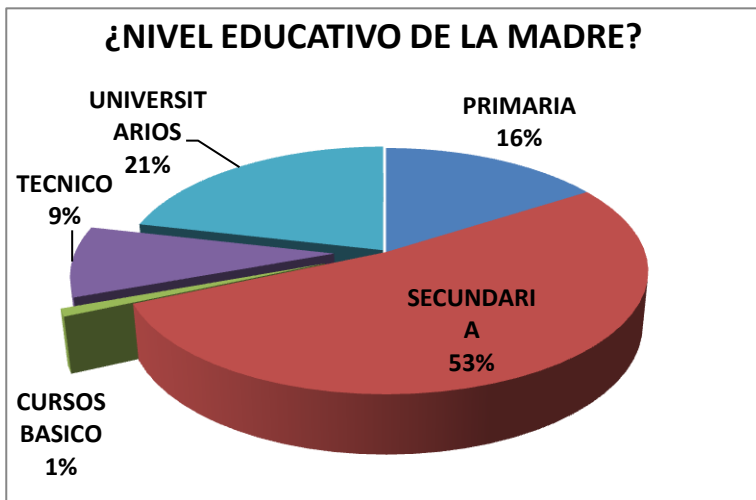
**SE LE PRESENTAN CRUCES CON EL HORARIO DE TRABAJO.** Es evidente que el 97% de las respuesta sea **NO** la población encuestada pertenece a la jornada diurna y no tienen tiempo para conseguir un trabajo y los que trabajan, lo hacen con la Universidad para conseguir auxilios o beneficios por parte de la Universidad

**Figura 22: Grafica Resultado encuesta nivel educativo del padre**



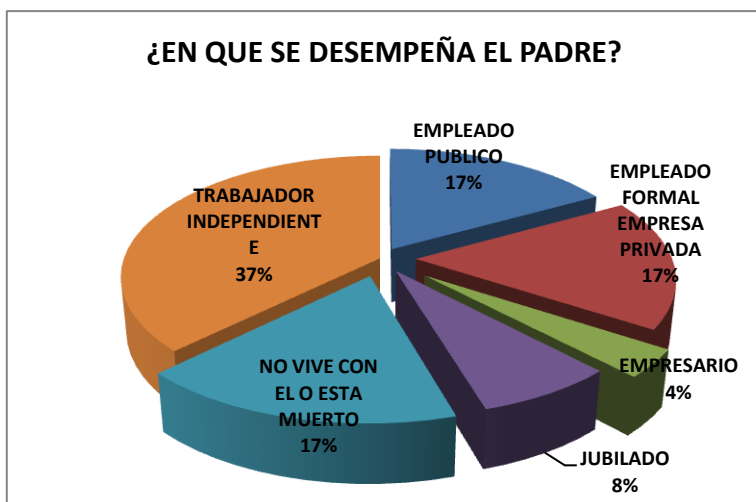
**NIVEL EDUCATIVO DEL PADRE.** Es notorio que los padres de la población busquen lo mejor para sus hijos y quieren que sean mejores que ellos, dándoles una educación superior vemos que el 92% de la población estudio algo

Figura 23: Grafica Resultado encuesta nivel educativo de la madre



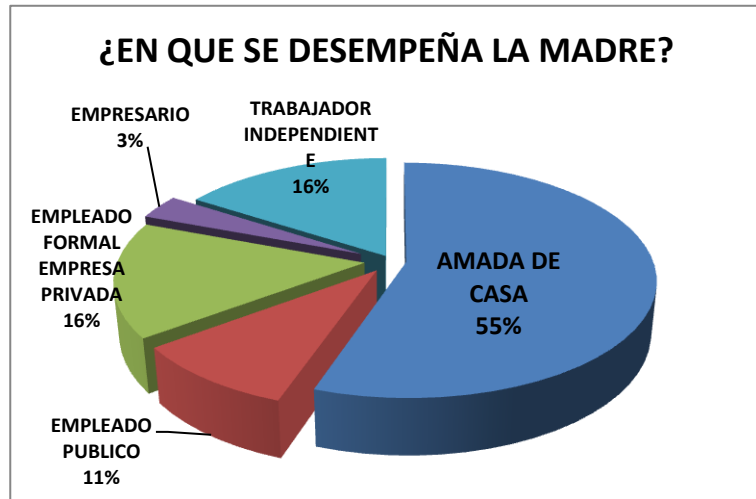
**NIVEL EDUCATIVO DE LA MADRE.** Este dato es interesante a comparación con el padre vemos que las madres tienen datos más altos en educación, con los padres vemos datos más bajos, con solo ver los de primaria y secundaria, los universitarios y técnicos no tienen diferencias tan grandes, pero se ve que los padres buscan más superarse que los padres

Figura 24: Grafica Resultado encuesta en que se desempeña el padre



**EN QUE SE DESEMPEÑA EL PADRE.** En esta grafica se observa que solo el 46% tienen un salario estable y el 37% tienen un salario variable o inestable, tiene relación también con la gráfica de **dificultad para pagar la universidad**

Figura 25: Grafica Resultado encuesta en que se desempeña la madre



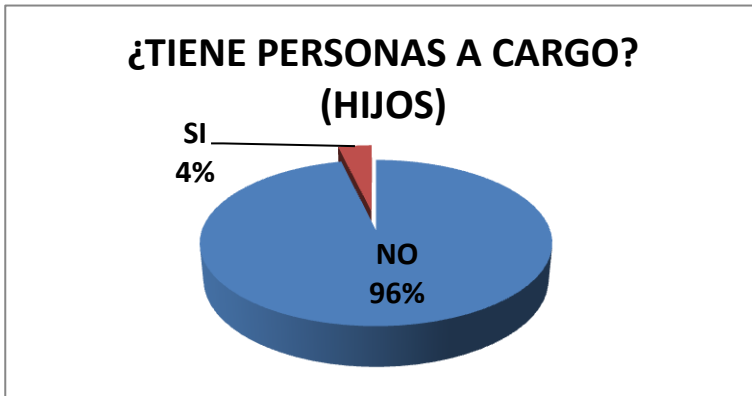
**EN QUE SE DESEMPEÑA LA MADRE.** A pesar que las madres lideran la gráfica de la educación evidenciamos que las mujeres se están quedando en la casa y las empresarias independientes en la mayoría son mujeres que trabajan con catálogos y los ingresos son variables, y las que trabajan en empleos formales los ingresos son muy bajos (manifestado por los encuestados)

Figura 26: Grafica Resultado encuesta presenta alguna discapacidad



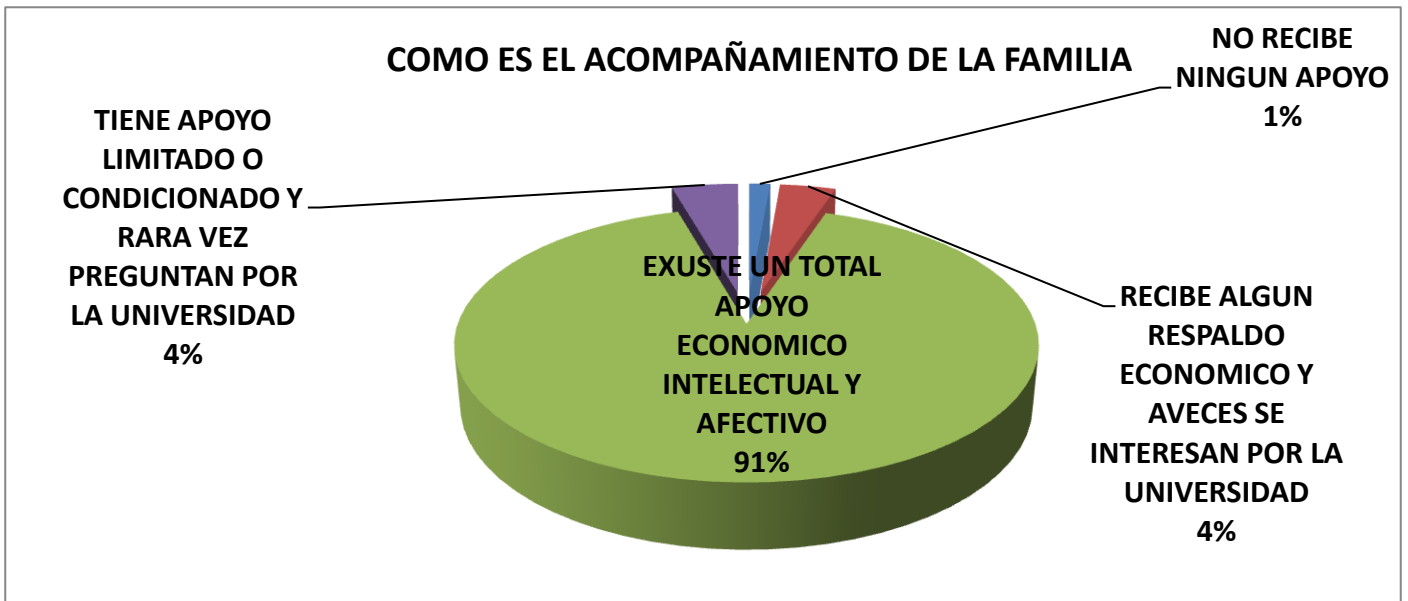
**PRESENTA ALGUNA DISCAPACIDAD.** La población encuestada ninguno presenta discapacidad alguna

Figura 27: Grafica Resultado encuesta tiene personas a cargo (Hijos)



**TIENE PERSONAS A CARGO.** Es evidente que los jóvenes están tomando conciencia y solo el 4% de la población tienen un hijo y siguen con los estudios, algunos un hijo puede ser un tropiezo, postergando sus estudios o abandonándolo

Figura 28: Grafica Resultado encuesta como es el acompañamiento de la familia



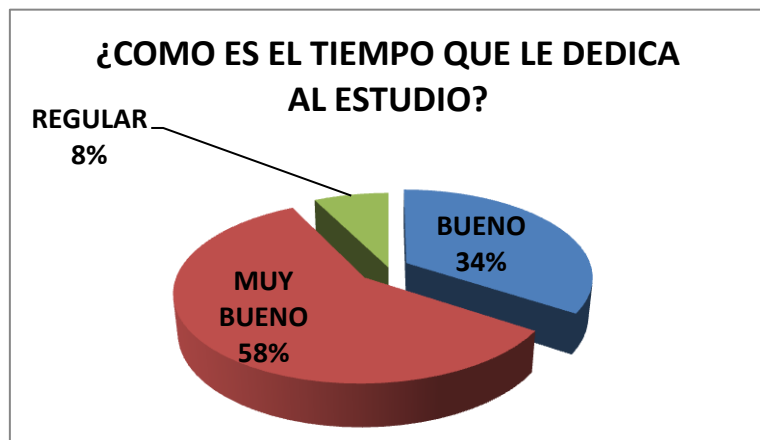
**COMO ES EL ACOMPAÑAMIENTO DE LA FAMILIA.** Se evidencia que la familia juega un papel importante en la vida universitaria, tanto económico como emocional, los estudiantes de los estratos 1 y 2 saben que no se pueden dar el lujo de estar perdiendo dinero, para estar perdiendo materias y cambiando de carrera

**Figura 29: Grafica Resultado encuesta hace cuanto se graduó del bachillerato**



**HACE CUANTO SE GARDUO DEL BACHILLERATO.** Esta grafica se evidencia que la población es relativamente joven que pertenecen a la jornada diurna, pues el 65% se graduó hace 2 años máximos

**Figura 30: Grafica Resultado encuesta como es el tiempo que le dedica al estudio y cree que este método funciona**

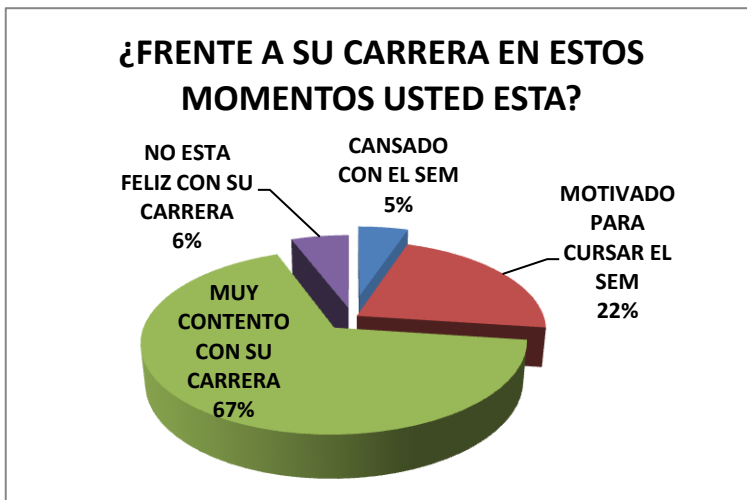


**COMO ES EL TIEMPO QUE LE DEDICA AL ESTUDIO y CREE QUE ESTE METODO FUNCIONA.**

Son dos preguntas que van de la mano que nos asegura que en promedio la población tiene un buen rendimiento académico



**Figura 31: Grafica Resultado encuesta frente a su carrera en estos momentos usted esta**



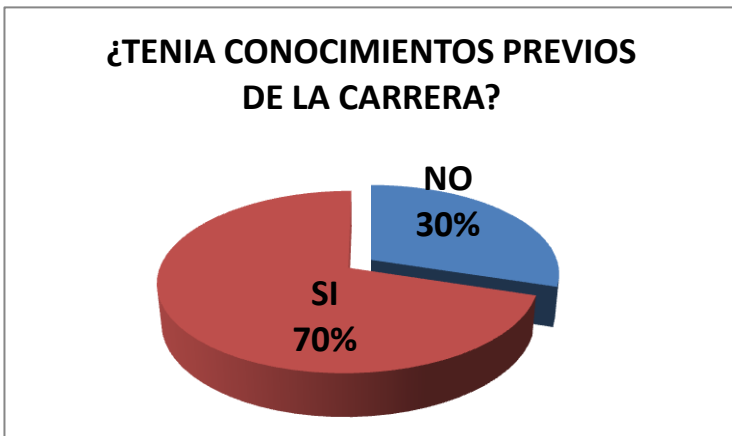
**FRENTE A SU CARRERA EN ESTOS MOMENTOS USTED ESTA.** Los estudiantes manifiestan una conformidad con un 89 % y una inconformidad de un 11% que pueden estar en la carrera equivocada

**Figura 32: Grafica Resultado encuesta cree usted que está en el programa adecuado**



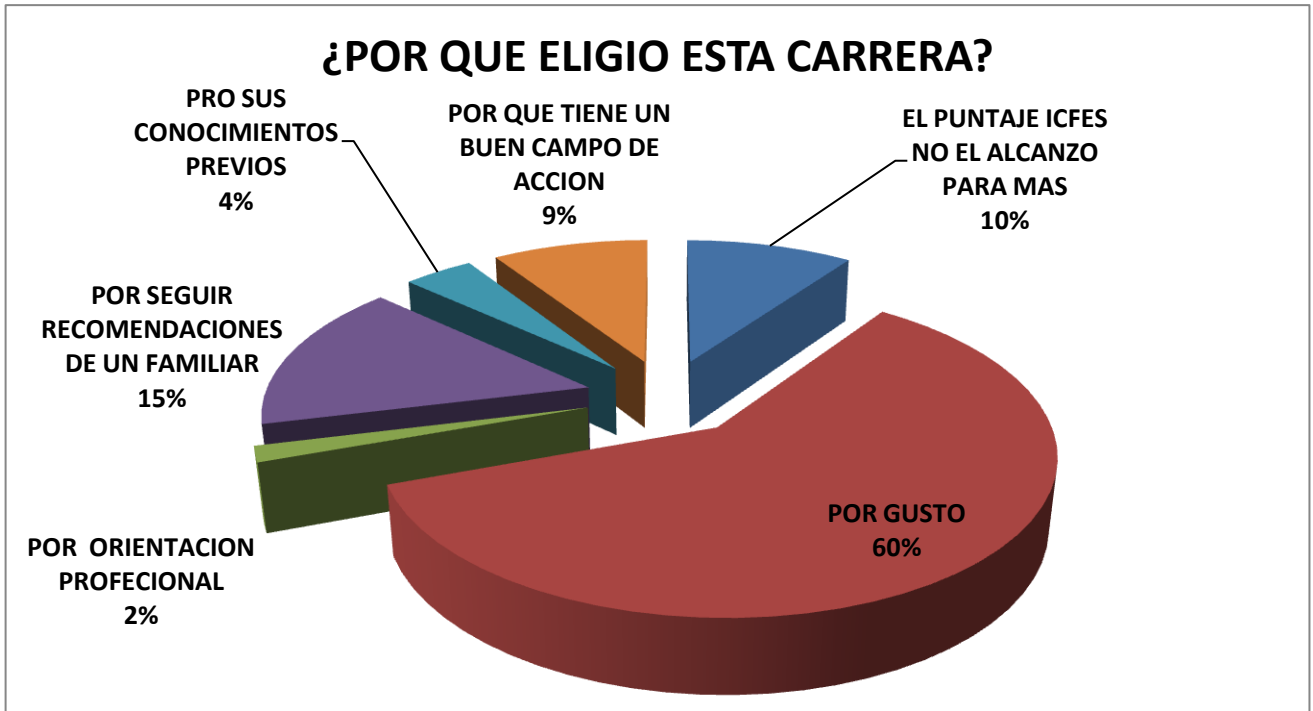
**CREE USTED QUE ESTA EN EL PROGRAMA ADECUADO.** Es bueno saber que el 85% de la población han elegido, pero es preocupante ver que un 15 % manifiesta inconformidad con la carrera elegida

**Figura 33: Grafica Resultado encuesta tenia conocimientos previos de la carrera**



**TENÍA CONOCIMIENTOS PREVIOS DE LA CARRERA.** Es preocupante ver un 30% de la población encuestada manifestar que no tenía conocimiento de lo que comenzaría a estudiar

**Figura 34: Grafica Resultado encuesta por qué eligió esta carrera**



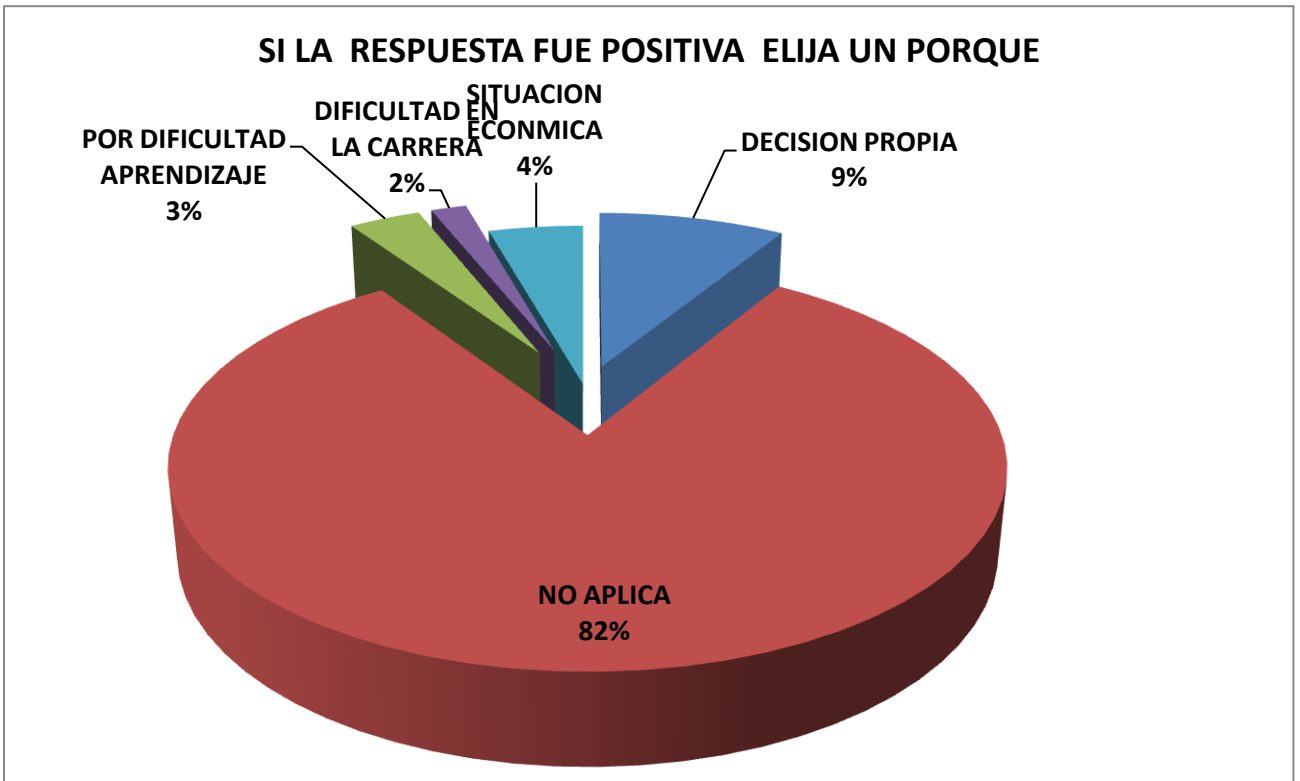
**POR QUE ELIJO ESTA CARRERA.** Al mirar este grafico se puede analizar que el 25% está estudiando algo que no le gusta o por que le toco, el 11 % buscaron carreras que tengan una buena acogida en el mercado, solo un 4% sabían de que trataba la carrera y un 60% manifiesta por gusto, donde estos gusto pueden ser un gusto temporal

**Figura 35: Grafica Resultado encuesta se ha cambiado de carrera**



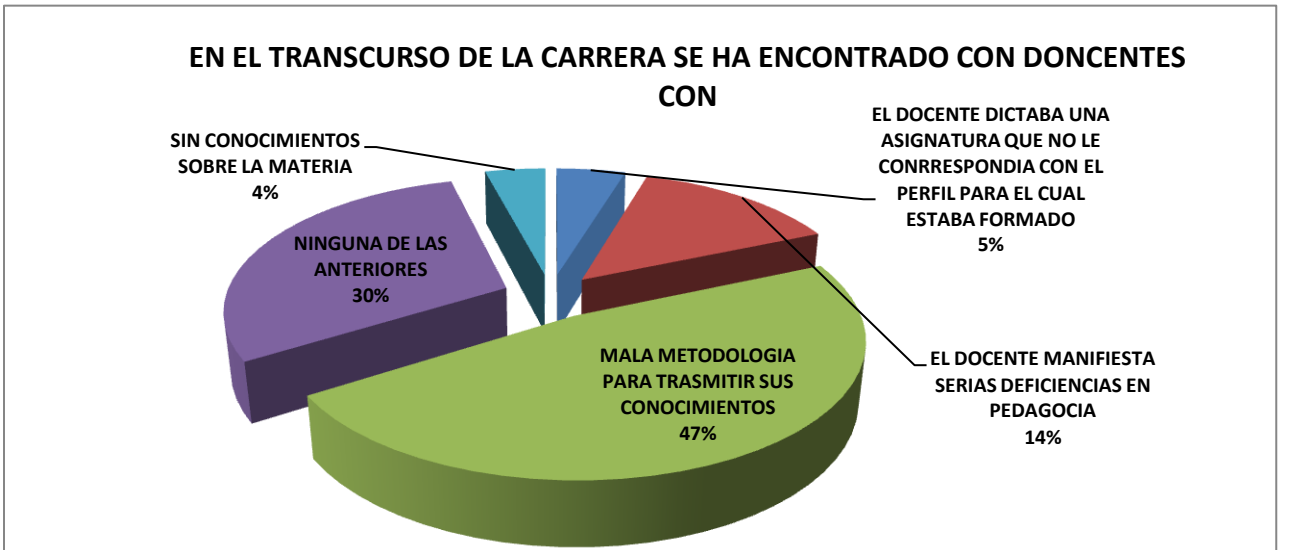
**SE HA CAMBIADO DE CARRERA.** Se observa que un 80% de la población no ha cambiado la carrera que está estudiando cuando el índice de inconformidad es de un 25% según lo que se manifiesta en la gráfica anterior

**Figura 36: Grafica Resultado encuesta si la respuesta fue positiva elija un porque**



**SI LA RESPUESTA FUE POSITIVA ELIJA UN PORQUE.** A la respuesta anterior la población manifestó solo el 5% justifico su cambio por motivos que le son de índole educativo el 4% por la situación económica, y el 9% manifiesta que por decisión propia que puede ser que no tuvo una buera orientación previa al ingreso a la universidad

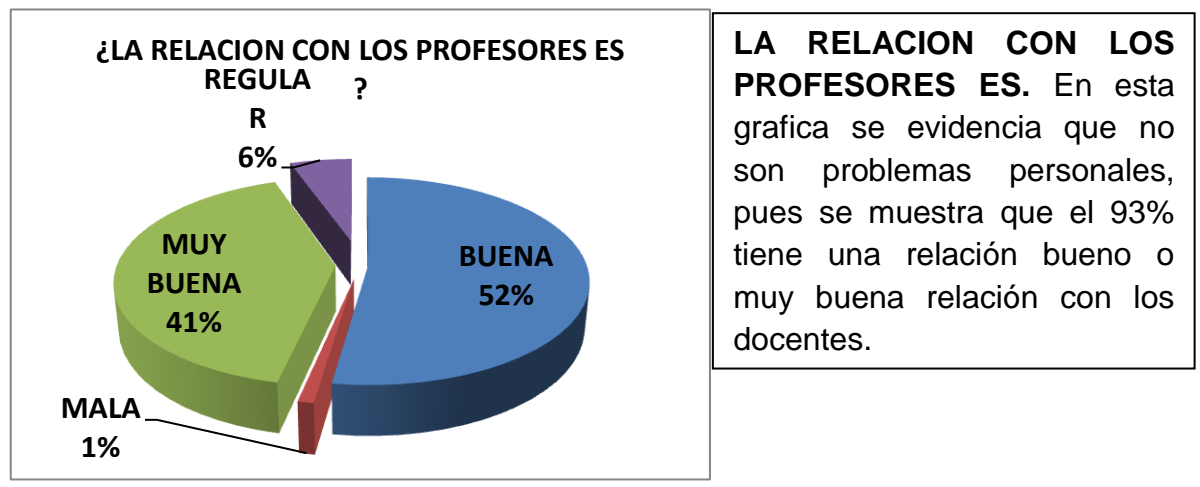
**Figura 36: Grafica Resultado encuesta en el transcurso de la carrera se ha encontrado con docentes con**



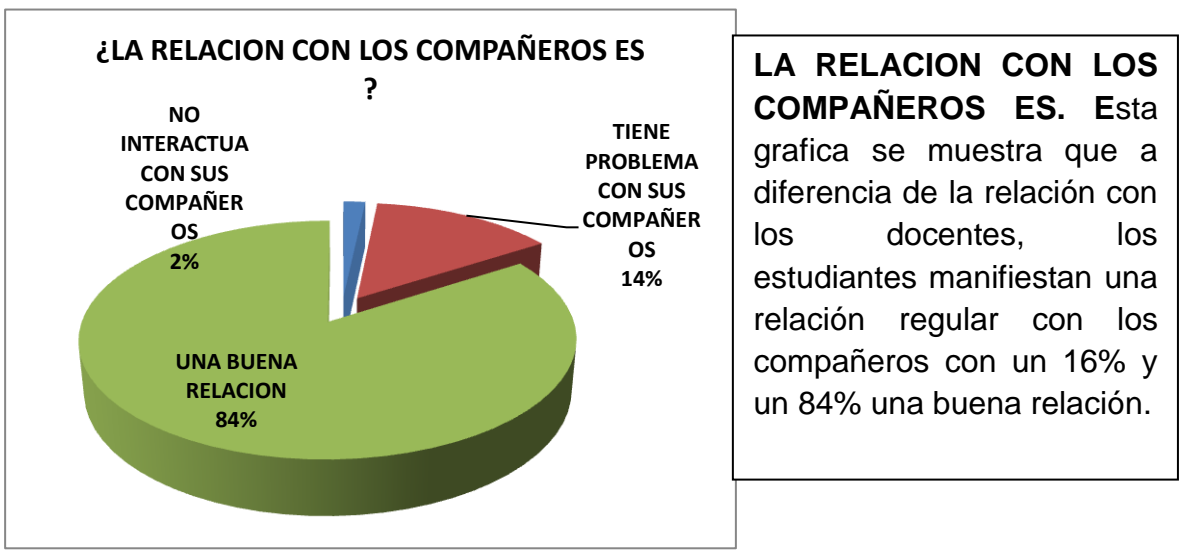


**EN EL TRANCURSO DE LA CARRERA SE HA ENCONTRADO CON DONCENTES CON.** El 30% de la población manifiesta no tener inconvenientes con los docentes, mientras un 70% declaran alguna molestia con algún

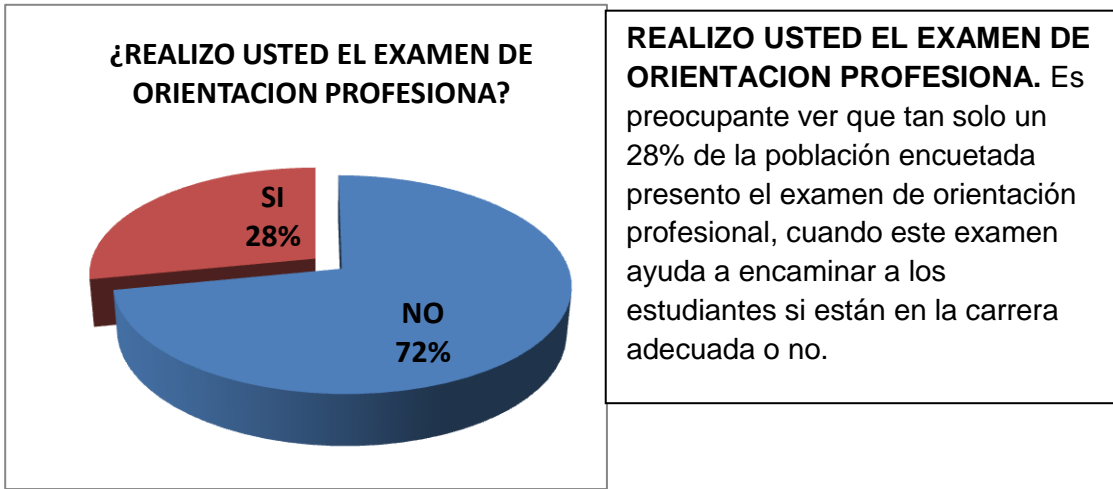
**Figura 37: Grafica Resultado encuesta la relación con los profesores es**



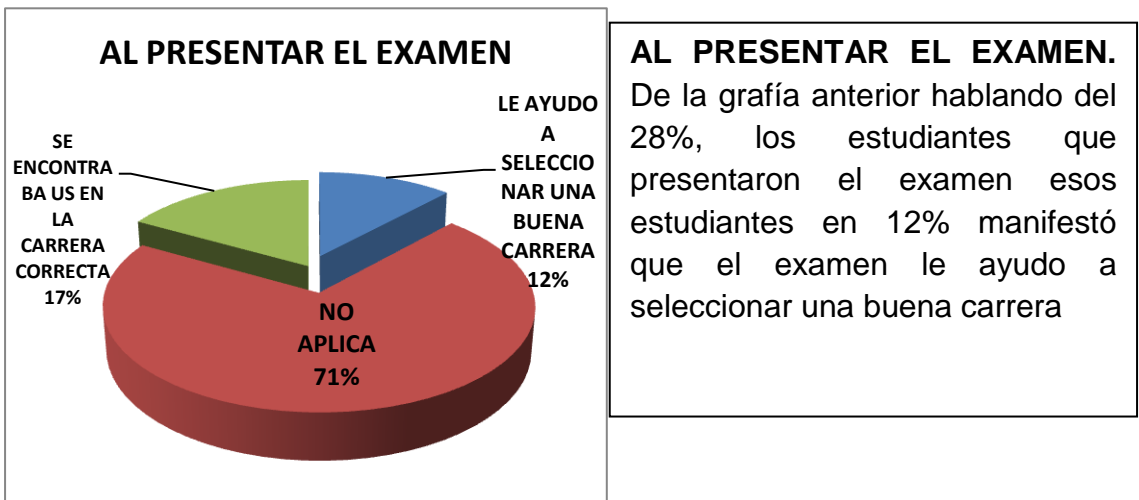
**Figura 38: Grafica Resultado encuesta la relación con los compañeros es**



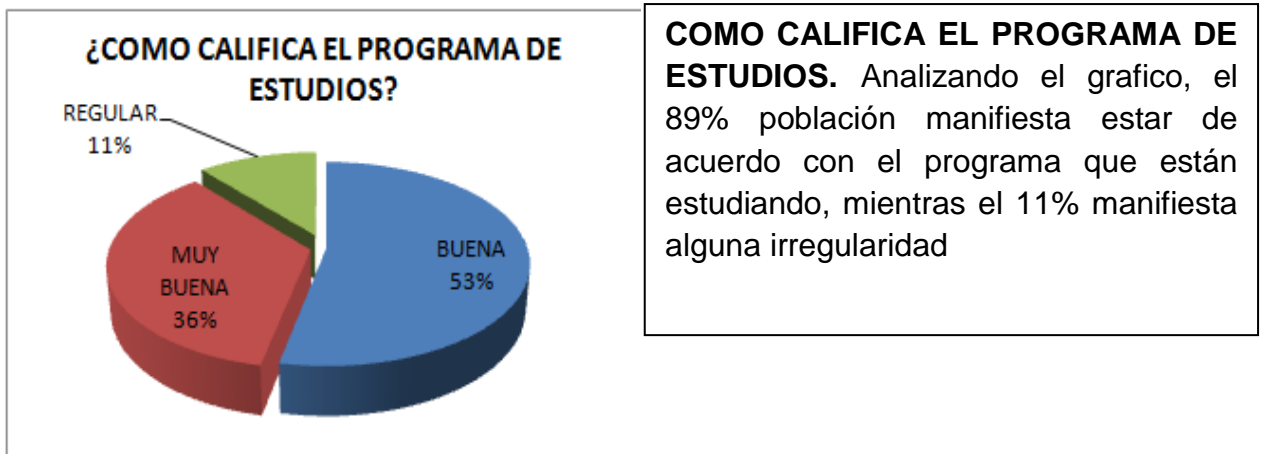
**Figura 39: Grafica Resultado encuesta realizo usted el examen de orientación profesional**



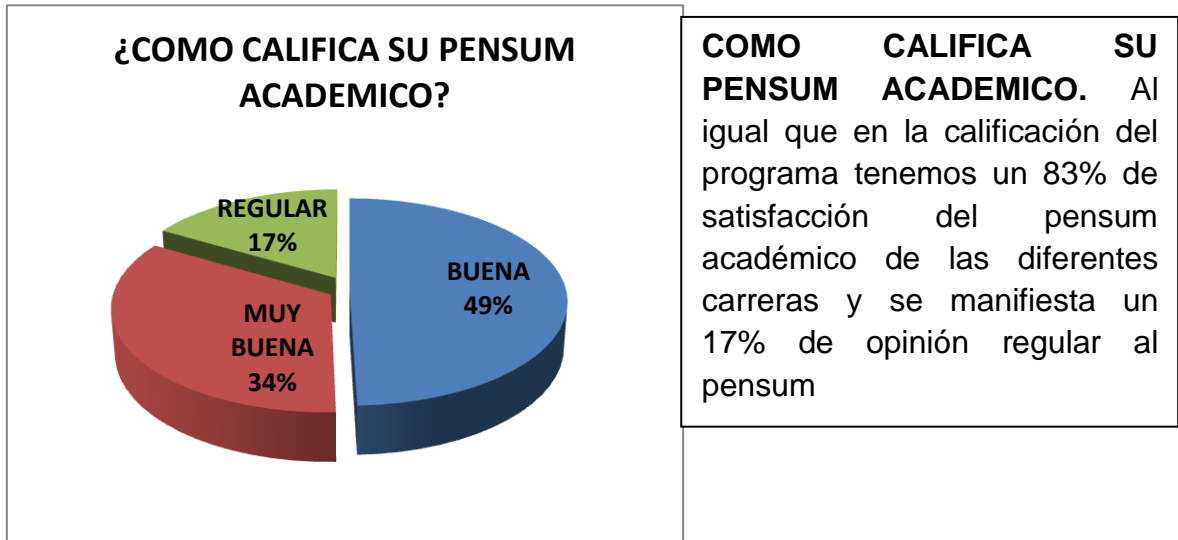
**Figura 40: Grafica Resultado encuesta al presentar el examen**



**Figura 41: Grafica Resultado encuesta como califica el programa de estudios**



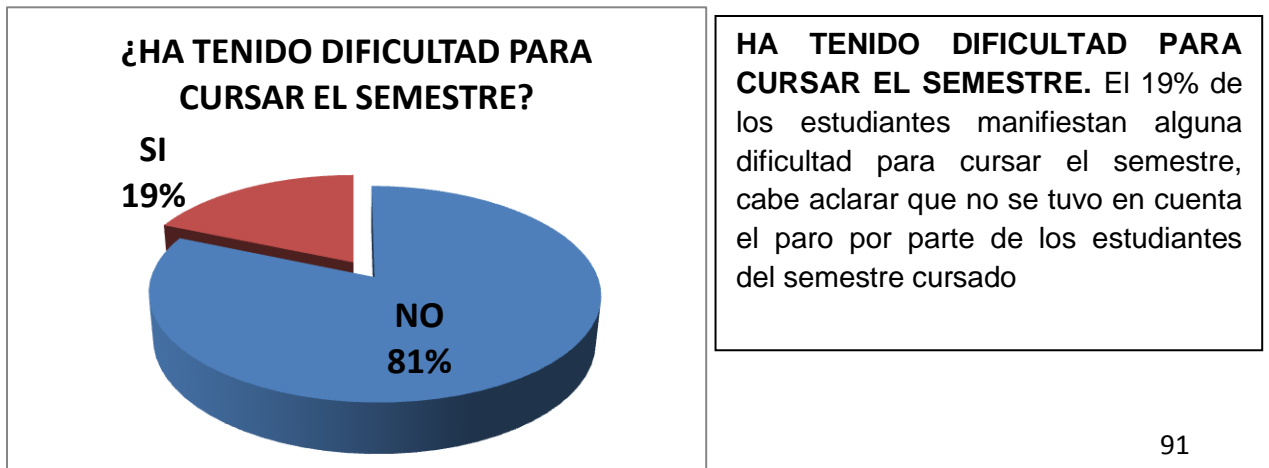
**Figura 42: Grafica Resultado encuesta como califica el pensum académico**



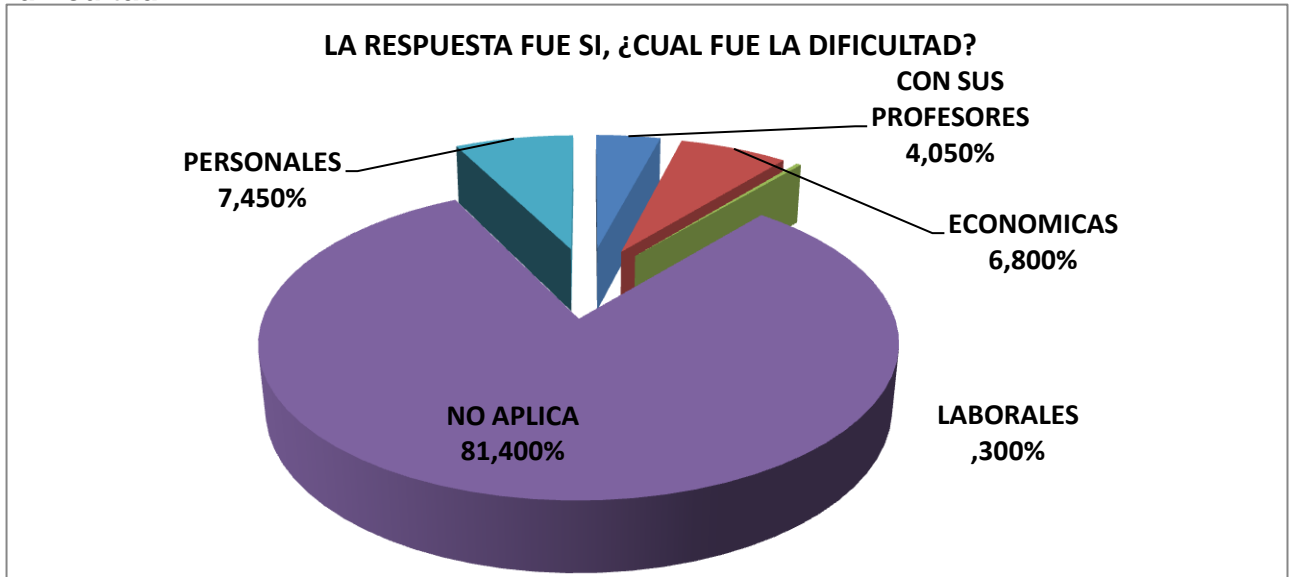
**Figura 43: Grafica Resultado encuesta se identifica con el programa que esta estudiando**



**Figura 44: Grafica Resultado encuesta ha tenido dificultad para cursar el semestre**

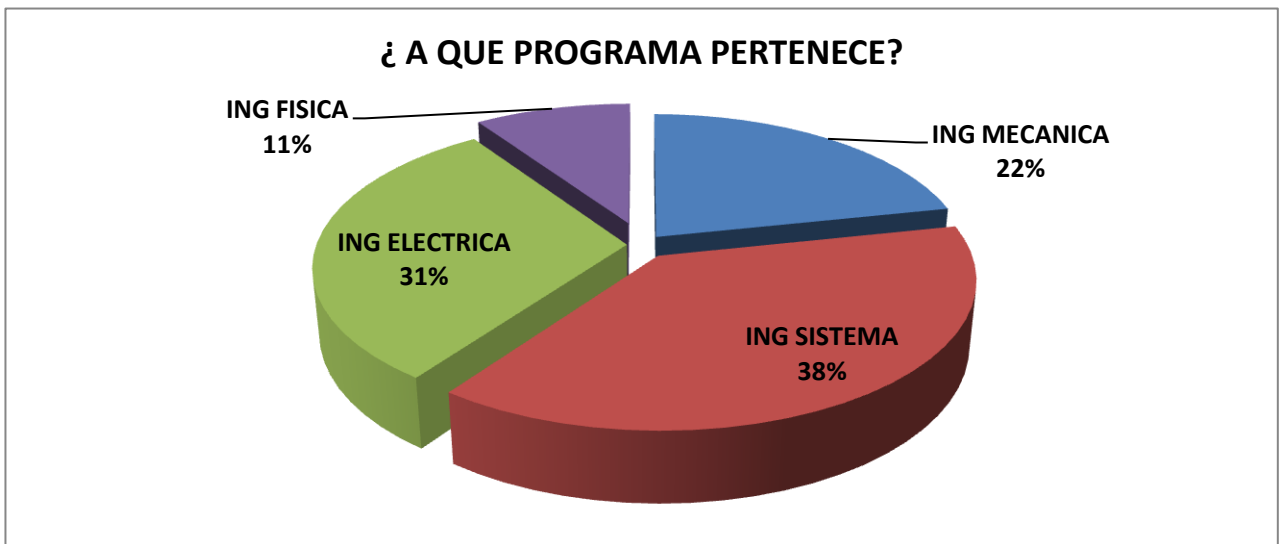


**Figura 45: Grafica Resultado encuesta la respuesta fue si ¿Cuál fue la dificultad?**



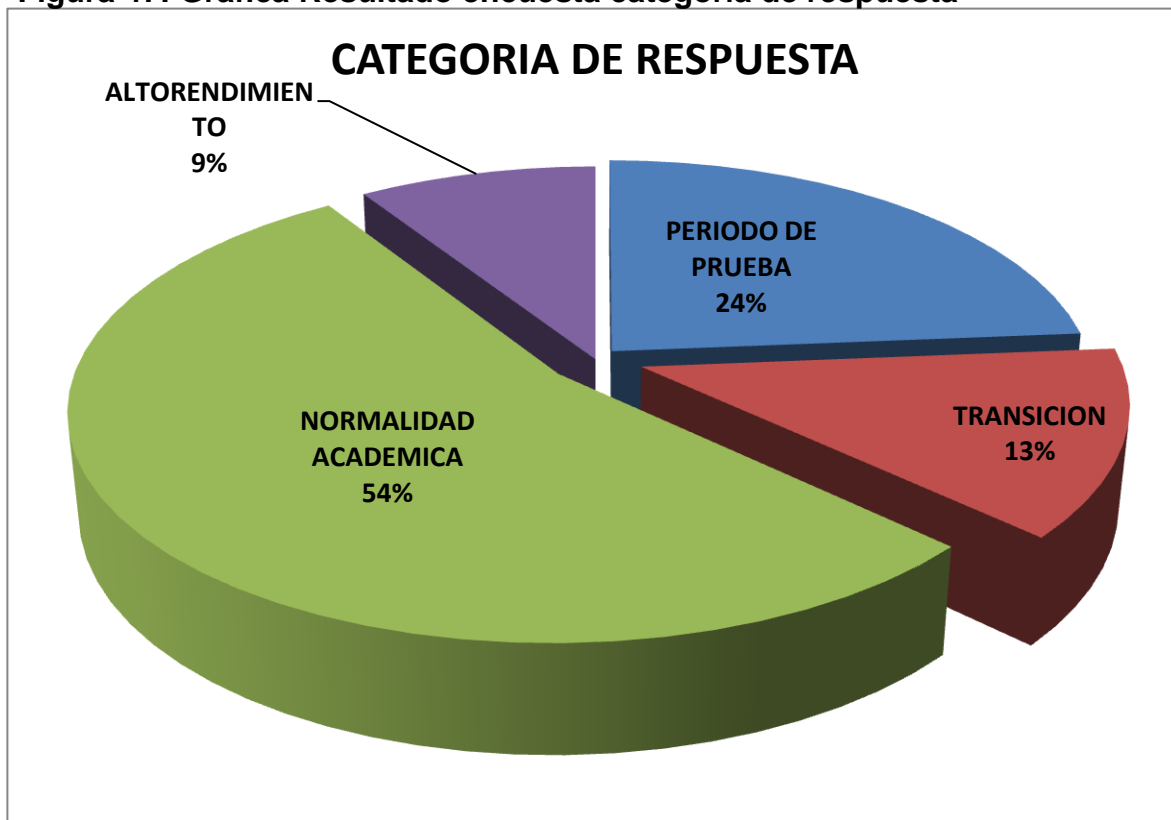
**LA RESPUESTA FUE SI, ¿CUAL FUE LA DIFICULTAD?** El 18,6% de la población manifiesta tener dificultad de alguna índole, donde las más representativas son motivos personales 7.5% y económicos 6.8%

**Figura 46: Grafica Resultado encuesta a que programa pertenece**



**A QUE PROGRAMA PERTENECE.** En la población que realizó la encuesta encontramos que la de menos participación fue la ingeniería Física con un 9%, las demás carreras tuvieron una buena participación, se aclara que la Ingeniería Industrial no se tuvo en cuenta, por carecer de información en lo entregado por registro y control

Figura 47: Grafica Resultado encuesta categoría de respuesta



**CATEGORIA DE RESPUESTA.** Con respecto a las materias vemos que el 37% de la población presenta dificultad en la carrera y en la población de normalidad académica con el 54% tenemos que en ella se encuentran estudiantes que están repitiendo materias, solo el 9% de la población tiene un alto rendimiento.

## **9. VARIABLES DE IMPORTANCIA EN LA MORTALIDAD ACADEMICA Y EN EL DESEMPEÑO DE LOS ESTUDIANTES EN LA UNIVERSIDAD**

La gran mayoría de la información que corresponde a las variables que se mencionan a continuación fue brindada por la división de sistemas la cual fue tomada de la base de datos de la Universidad. Esta información secundaria es obtenida mediante el diligenciamiento de algunos formatos en el momento en el que los estudiantes ingresan a los distintos programas que ofrece la Universidad y de su historial académico para ver como se encuentran en sus diferentes programas académicos.

Por otra parte existen variables de las que no se obtuvo información por parte de la Universidad, lo cual hizo necesario la aplicación de un instrumento de muestreo como es la encuesta para la recolección de la misma información primaria, así mismo existieron otras variables que definitivamente no se trabajaron por la imposibilidad de levantar la información.

A continuación se presentan las variables que fueron el objeto de estudio de la investigación al considerarlas necesarias para la realización de el importante proyecto y de resultados confiables por parte de los investigadores, pero será el programa y el modelo quienes finalmente dirán cuáles son las variables realmente importantes, de peso y que verdaderamente influyen en la problemática, es importante aclarar que no se pudo obtener información de todas ellas, es por eso que además se realizó la recolección de cierta parte de la información. En el capítulo de Análisis Multivariado se relacionan las variables que entraron y que hicieron parte de dicho análisis.

### **Variables de tipo Académico:**

- 1. Resultados pruebas de estado**
- 2. Número de Créditos vistos durante el último semestre.**
- 3. Profesor Clasificación**
- 4. Profesor Original**
- 5. Metodología del Profesor**
- 6. Orientación Profesional**
- 7. Calidad del Programa**
- 8. Relaciones con los profesores u otros estudiantes**

### **Variables de tipo Socioeconómico:**

- 1. Género**
- 2. Estrato socio económico**
- 3. Departamento de procedencia**
- 4. Municipio de procedencia**
- 5. Género del colegio en el que estudio**
- 6. Tipo de colegio en el que estudio**
- 7. Estado civil**
- 8. Edad**
- 9. Entorno Familiar**
- 10. Integración Social.**
- 11. Becas y Formas de financiamiento**
- 12. Dependencia Económica**
- 13. Nivel Educativo de los padres**
- 14. situación laboral de los padres**

**Variables de tipo Personal:**

- 1. Motivación para cursar las materias del semestre**
- 2. Satisfacción con el programa**
- 3. Conocimiento previo del programa**
- 4. Esfuerzo para cursar el semestre**
- 5. Apoyo familiar en los estudios**
- 6. Actividades extra clase.**
- 7. Calamidad y problemas de salud**
- 8. Incompatibilidad horario con actividades extra académicas (trabajo del estudiante)**
- 9. Método de estudio**
- 10. Personas a cargo**

**9.1 Variables de tipo socioeconómico**

**Género:** Informativa, arroja información en un futuro cruce de variables, permite medir y comparar el desempeño académico dependiendo del género de la muestra analizada.

**Edad:** Informativa, arroja información en un futuro cruce de variables, sirve para determinar la influencia de la edad en el desempeño académico.

**Estado civil:** Informativa, arroja información en un futuro cruce de variables, sirve para medir además como afecta una relación de pareja el desempeño académico.

**Teléfono:** Informativa, en caso de necesitar alguna información extra del estudiante.

**Programa:** Para hacer el análisis estadístico por programas y realizar comparativos entre ellos.

**Facultad:** Para hacer el análisis estadístico por facultades y después realizar comparativos entre ellas.

**Estrato social:** Sirve para medir la influencia de la situación económica en el desempeño académico.

**Tipo de Colegio (Público o Privado):** Permite analizar la influencia de la procedencia de los colegios (Públicos o Privados) en los estudiantes, así también permite evaluar la calidad de la educación en los distintos tipos de colegio.

**Genero del Colegio (Mixto, Femenino, Masculino):** Permite medir la incidencia en los estudiantes dado el tipo de colegio del cual provengan, ya sean masculinos, femeninos o mixtos.

**Lugar de procedencia al ingresar a la Universidad:** Permite medir la influencia en el desempeño académico del lugar de procedencia.

**Horas de trabajo:** El hecho de trabajar mientras se estudia puede influir de manera drástica en el desempeño académico, es por eso de la necesidad de esta variable.

**Número de personas a cargo:** La responsabilidad de tener familia a cargo no solo hace que la persona necesariamente tenga que trabajar, sino que además tenga que invertir tiempo en otras actividades que lo alejen de un buen ritmo de estudio repercutiendo en su situación académica.

**Paga su estudio:** Esta variable indica si la persona corre con los gastos de su carrera o si se la financian, esto entrega un indicio de que tanto valora el estudiante su carrera y la relación con su comportamiento académico.



## 9.2 Variables de tipo académico

**Población estudiantil matriculada por semestre:** Número de personas que matriculan por facultad cada semestre.

**Numero de semestres del programa:** Numero de semestres que conforman cada programa.

**Fecha de ingreso a la Universidad:** Fecha exacta en la que el estudiante se matriculo y aprobó el ingreso a la Universidad por primera vez.

**Promedio acumulado de carrera:** Promedio actual del estudiante.

**Puntaje en la prueba de estado:** Puntaje ICFES.

**Semestre en el cual está clasificado:** Según la Universidad en que semestre académico se encuentra clasificado el estudiante.

**Número de créditos aprobados:** Numero de créditos que a la fecha ha aprobado el estudiante.

**Profesor:** Es importante analizar hasta qué punto ciertos profesores empiezan a ser la consecuencia principal para el buen o mal desempeño de los estudiantes, parte de la información para esta variable es brindada por la Universidad la otra se debe levantar. Esta variable se dividió en dos partes, la primera en Profesor Clasificación, que según unas variables que se pensaron importantes en el perfil del docente, obtuvieron cierta ponderación una vez recolectada dicha información, se crearon entonces cuatro grupos de categorías de profesores de acuerdo a las calificaciones que ellos obtenían. La otra división consiste simplemente en alimentar el programa con la ponderación original que tuvieron los profesores basados en la información o variables que moldean su perfil. Estas variables que se tuvieron en cuenta para la clasificación y calificación de los docentes fueron el tipo de contratación, la cantidad de años que lleva ejerciendo docencia en la Universidad, el título o los títulos que posea y el número de horas que este dedica a la investigación.

## 9.3 Variables de Interpretación personal

**Aspectos relevantes para el buen desempeño académico:** Que factores consideran los estudiantes que son determinantes para un buen rendimiento académico.

**Motivación con las asignaturas cursadas:** Dado que la motivación hacia la carrera es importante para un buen desempeño, es necesario saber que tan motivados se encuentran los estudiantes.

**Satisfacción con la carrera:** La satisfacción que se tenga con el programa que se cursa es fundamental para un desempeño alto en los estudios.

**Conocimiento previo del programa:** En ocasiones no se tiene un conocimiento del programa al cual se accede y puede haber inconformidad una vez empiezan las clases, esto se refleja en el desempeño académico.

**Esfuerzo para aprobar las materias:** Es importante saber la dedicación que los estudiantes ponen en el momento de cursarlas.

**Apoyo familiar:** Es importante determinar que tanto influye el apoyo familiar en el buen desempeño académico.

**Actividades extra clase:** Las actividades que se llevan a cabo aparte del estudio, como trabajo, deportes u otras, pueden influir de manera directa en los resultados que se obtienen en el semestre.

## 10. MUESTREO

Definidas las variables para realizar el análisis en el año 2011, se hizo necesario la recolección de más información sobre todo de tipo personal que no era suministrada por parte de la Universidad, pero que de una u otra forma es importante y podría entregar resultados muy interesantes en el análisis estadístico, es por eso que se llevó a cabo la aplicación de un instrumento (encuesta) de recolección de información primaria para anexar a las variables a analizar. El instrumento se puede observar en la encuesta, este no es más que una entrevista de tipo telefónica que se realiza para la muestra seleccionada entre los estudiantes que matricularon las materias definidas como críticas durante el año 2011.

Vale la pena aclarar que la entrevista se realiza a las personas seleccionadas en la muestra, dado que de la base de datos original se puede contar con el número telefónico y de celular de cada uno de los estudiantes, se decidió entonces realizar la recolección de la información por este medio, puesto que ahorraría tiempo, se interactúa con el estudiante y agilizaría dicha toma de la información con confiabilidad.

De este trabajo de campo se obtienen entonces ocho variables más de tipo cualitativo y de carácter personal que amplían la posibilidad de obtener más resultados y más conclusiones acercándonos así de una manera más precisa a lo que realmente afecta los estudiantes en la Universidad.

### 10.1 TIPO DE MUESTREO Y CLASIFICACIÓN

Dada la problemática, se decidió aplicar para la recolección de información restante y el posterior análisis de variables, muestreo estratificado, es así como se pretende diferenciar entre cuatro distintos tipos de desempeños académicos dentro de los cuales se clasificarán todos los estudiantes relacionados en el estudio. Estos cuatro tipos de desempeños serán los estratos de nuestro muestreo y se basa su definición en la clasificación que les ha asignado el observatorio estudiantil el cual realiza periódicamente el seguimiento académico de la Universidad Tecnológica de Pereira, realizado una vez culminado el semestre y basados en el promedio obtenido por ellos. Esta clasificación se explica en el siguiente texto.

“El observatorio académico de la UTP se ha interesado en los últimos años en identificar el estado en el que ingresan a la universidad los estudiantes de primer semestre. Para tales efectos ha desarrollado una serie de instrumentos o pruebas de valoración que buscan dar cuenta de las variables más significativas que puedan afectar o estar en relación con el desempeño académico futuro del estudiante.

Tradicionalmente se ha venido midiendo el desempeño académico de los alumnos mediante el promedio semestral. Con base a los valores arrojados por el promedio en el semestre, la universidad tiene políticas para clasificar a los estudiantes en diferentes estados a saber”.

1. **Estado normal:** el estudiante obtiene un promedio al cierre del primer semestre entre [3.0-4.2)
2. **Estado periodo de prueba:** el estudiante obtiene un promedio al cierre del primer semestre entre [0,0-2.5)
3. **Estado de transición:** el estudiante obtiene un promedio al cierre del primer semestre entre [2,6-2,9)
4. **Estado de alto rendimiento:** el estudiante obtiene un promedio superior o igual a 4,3
5. **Estado de cancelación de matrícula:** el estudiante cancela su primer semestre en la universidad.”

Basados en esta clasificación, y con la necesidad de identificar la influencia de las variables estudiadas en el desempeño de los estudiantes, se hace necesaria una clasificación más cerrada y es allí donde surgen tres segmentos o estratos de desempeño estudiantil para cada facultad.

La nueva clasificación que se realizó es la siguiente:

1. Estudiante en cancelación (Nota entre 0)
2. Estudiante en periodo de prueba (Nota entre 0 y 2,5)
3. Estudiante en transición (Nota entre 2,6 y 2,9)
4. Estudiantes en normalidad académica (Nota entre 3 y 4,2)
5. Estudiante en alto rendimiento (Nota entre 4,3 y 5)

Tomando como Base la clasificación de la Universidad y según el criterio de los investigadores, se estableció la anterior clasificación, esta destaca que los estudiantes con alto desempeño necesariamente deben tener notas superiores a 4,3.

Cabe aclarar que la clasificación que se hace en la Universidad por el respectivo observatorio académico y está dada por el promedio semestral, para efectos de la presente investigación la clasificación se hará según las notas obtenidas en las materias críticas definidas con anterioridad.

También se realiza un muestro con regresión logística binomial para general el modelo estadístico de pronóstico. Para este modelo se genero otra clasificación en la variable nota con la necesidad de identificar y analizar cómo afectan las variables directamente al modelo. La clasificación para la regresión logística binomial es la siguiente:

1. No aprobado
2. Aprobado

Con esta clasificación pudimos general el modelo de pronóstico como una solución para la prevención de la mortalidad académica en la Universidad Tecnológica de Pereira.

## **10.2 METODOLOGIA DEL MUESTREO**

La primera parte se realizó la identificación de los estados de estudios asignados por el observatorio de la Universidad Tecnológica de Pereira. Estos estados son los que diferencian los tipos de desempeño de los estudiantes y su debida clasificación.

Al obtener estos estados de clasificación que ayudan a determinar la mortalidad de las materias se realizó la depuración de la base de datos para tener la información estrictamente necesaria para realizar el estudio según las variables definidas que son objeto principal de dicho estudio. La base de datos que nos entregó la universidad no cumplía con todos los requisitos solicitados pero por el tiempo que se demoró la universidad para entregarnos dicha información se decido depurarla y trabajar con ella.

Esta base de datos esta desde los periodos 2011-1 al 2013-1, la información obtenida del 2004-1 al 2009-2 son de estudios anteriores realizados en la faculta. Y los datos del 2010-1 y 2010-2 fueron pronosticados para darle una continuidad al estudio.

En el estudio es muy posible tener miembros de las poblaciones repetidos en el año analizado por las opciones que le brinda la Universidad a los estudiantes para aprobar una materia como son los intersemestrales o los exámenes de suficiencia, por ejemplo, una persona de Ingeniería Eléctrica que pierda Matemáticas I con una nota de 2,5 en el primer periodo del año 2011, estaría clasificada dentro del sub estrato bajo desempeño de Ingeniería Eléctrica, muy posiblemente esta persona realizó intersemestral de esta materia y lo aprobó, esto haría que con seguridad esta persona curse Matemáticas II el siguiente periodo, y es posible también que repruebe la materia con una nota de 2,5 por ejemplo. Este estudiante estaría clasificado o sumaria dos veces dentro del sub estrato.

Por ende se realizó una verificación por cada semestre, materia y programa para mirar los datos duplicados y promediarlos en una sola nota por periodo. Y se asignó dicho promedio para efectos de estudio y de desviación estándar entre otros de dicha población. Es del resultado de este largo análisis y filtro, que se obtienen los estratos de donde se extraerán las muestras que indicaran a que personas aplicar el instrumento.

Por consiguiente se realizó la aplicación de los instrumentos para el cálculo de la debida muestra para la información primaria que se obtiene por medio de la encuesta aplica telefónicamente a dicha población.

$$n = \frac{n'}{1 + \frac{n'}{N}}$$

$$n' = \frac{s^2}{\sigma^2}$$

Siendo:  $\sigma^2$ , de donde:

$\sigma^2$  es la varianza de la población respecto a determinadas variables.

$s^2$  es la varianza de la muestra, la cual podrá determinarse en términos de probabilidad como  $s^2 = p(1 - p)$

$se$  es error estándar que está dado por la diferencia entre  $(\mu - \bar{x})$  la media poblacional y la media muestral.

$(se)^2$  es el error estándar al cuadrado, que nos servirá para determinar  $\sigma^2$ , por lo que  $\sigma^2 = (se)^2$  es la varianza poblacional.

$$N = 384 / 1 + (384/2133)$$

$$N = 326$$

Se aplico la encuesta a 350 estudiantes para tener mayor seguridad.

Teniendo el tamaño de muestra para realizar el estudio fue necesario llevar las poblaciones al programa SPSS para que el software entregue con precisión las personas que serán entrevistadas en cada caso.

Abriendo los datos de Excel de la población en spss, se selecciona la opción de extraer una muestra en el programa SPSS y se obtiene en una nueva hoja el listado de personas a muestrear.

Muestra extraída del programa SPSS.

**Figura 48: Grafica pantallazo 1 resultados Muestra extraída del programa SPSS**

CONTEO	PERIODOACADEMICO	DOCUMENTO	NOMBRECOMPLETO	TELEFONO
1	20121:Pregrado:Pereira:Pregrado	1088296697	ABRIL PEREZ JUAN SEBASTIAN	3454168
2	20131:Pregrado:Pereira	1088322756	ACEVEDO FRANCO VANESSA ALEJANDRA	3346891
3	20111:Pregrado:Pereira	1087488837	ACEVEDO GALVIZ FABIO ANDRÉS	000
4	20131:Pregrado:Pereira	1004789708	ACEVEDO GARZON OSCAR EDUARDO	3284874
5	20121:Pregrado:Pereira:Pregrado	96021914549	ACEVEDO MÚNERA JUAN ALBERTO	2323240
6	20131:Pregrado:Pereira	1143949253	ACEVEDO ROJAS JHON EDINSON	4139028
7	20131:Pregrado:Pereira	1088018117	ACOSTA DELGADO JOHAN STIVEN	3286749 , 3286749
8	20112:Pregrado:Pereira	1030620169	ACOSTA GUTIERREZ ERICK DAMIAN	3229228
9	20131:Pregrado:Pereira	95111720585	ACOSTA HENAO OMAR	3286509
10	20131:Pregrado:Pereira	95011915052	ACOSTA MARTINEZ VIVIANA	3280385
11	20122:Pregrado:Pereira:Pregrado	95090122348	ACOSTA VELEZ MIGUEL ANGEL	3326974
12	20122:Pregrado:Pereira:Pregrado	1088327736	AGUALIMPIA COPETE KAREN TATIANA	3213283
13	20122:Pregrado:Pereira:Pregrado	2474134	AGUDELO AGUIRRE EDGARDO	3437306
14	20121:Pregrado:Pereira:Pregrado	1088019500	AGUDELO ÁLVAREZ JORGE ALEXÁNDER	3428270 , 3326996 , 3326996
15	20111:Pregrado:Pereira	1088319191	AGUDELO ESTRADA SANTIAGO ALBERTO	3378913
16	20131:Pregrado:Pereira	96042219352	AGUDELO MARÍN KELLY YOJANA	3326631
17	20111:Pregrado:Pereira	1088255523	AGUDELO MARTINEZ NESTOR WILLIAM	3210117 , 3210117 , 3210117 , 3210117
18	20122:Pregrado:Pereira:Pregrado	1088315453	AGUDELO MUÑOZ MAYRA ALEJANDRA	3311718 , 3311718 , 3311718
19	20131:Pregrado:Pereira	95122626220	AGUDELO OROZCO NICOLAS	3455489
20	20122:Pregrado:Pereira:Pregrado	1112766226	AGUDELO RAMIREZ JUAN DAVID	3136490248
21	20111:Pregrado:Pereira	1114787496	AGUDELO RENDÓN PAOLA ANDREA	3122904754
22	20121:Pregrado:Pereira	96042406288	AGUDELO SALAZAR SEBASTIAN	3457444

Los estudiantes que aparecen en la casi numero con una diagonal son aquellos que no se tendrán en cuenta para aplicar las encuesta.

Para que fuera más fácil se le solicito al software crear una base de datos nueva con los estudiantes seleccionados para aplicar la encuesta como se muestra en la siguiente imagen.

**Figura 49: Grafica pantallazo 2 base de datos nueva del programa SPSS**

Visible: 28 de 28 variables

	CONTEO	PERIODOACADEMICO	DOCUMENTO	NOMBRECOMPLETO	TELEFONO
1	3	20111:Pregrado Pereira	1087488837	ACEVEDO GALVIZ FABIO ANDRÉS	000
2	19	20131:Pregrado Pereira	95122626220	AGUDELO OROZCO NICOLAS	3455489
3	33	20111:Pregrado Pereira	1087990618	AGUIRRE CUERVO CAMILO ERNESTO	3325643 , 3325643
4	38	20122:Pregrado Pereira:Pregrado	1088249641	AGUIRRE ROMÁN DIEGO ALEJANDRO	3324453
5	47	20121:Pregrado Pereira:Pregrado	1152436849	ALVAREZ ALZATE DIEGO ANDRES	2144024
6	50	20112: Pregrado Pereira	1115421597	ALVAREZ CORRERA JULIAN ANDRES	3212667 , 3118854617
7	53	20112: Pregrado Pereira	1089483340	ALVAREZ MUÑOZ ANDRES MAURICIO	7264906
8	55	20111:Pregrado Pereira	1093219252	ALVAREZ OCHOA JUAN FERNANDO	3659280 , 3659280 , 3113509780 , 3113509780
9	56	20112: Pregrado Pereira	94060829289	ALVAREZ VELEZ CRISTIAN CAMILO	3104936721
10	68	20122:Pregrado Pereira:Pregrado	1088247663	AMEZQUITA SANTA OSCAR ANDRÉS	3269112 , 3269112
11	69	20111:Pregrado Pereira	1123200028	ANDINO MORA RODRIGO ANTONIO	4260433 , 4290735
12	84	20112: Pregrado Pereira	1093223189	ARANGO GRANADA DIEGO ALEJANDRO	3640643
13	107	20112: Pregrado Pereira	1004686327	ARDILA GRANADA SANTIAGO	3404107 , 3162280
14	122	20131:Pregrado Pereira	95021918467	ARIAS MEJIA SEBASTIAN	3105477383
15	126	20131:Pregrado Pereira	96012105468	ARIAS PÉREZ FERNANDO	3152235444
16	132	20131:Pregrado Pereira	1112880865	ARIAS VALENCIA JHON SEBASTIÁN	2533846
17	134	20121:Pregrado Pereira:Pregrado	1088323562	ARIAS VILLADA YENIFFER	3259520
18	136	20131:Pregrado Pereira	95111818755	ARICAPA MUÑOZ DANIELA	3292054
19	138	20131:Pregrado Pereira	1053835124	ARISTIZABAL CASTRILLON JACOBO	8851843
20	146	20111:Pregrado Pereira	1088295445	ARREDONDO ARTEAGA DEBBIE JOHAN	3290876
21	154	20111:Pregrado Pereira	1088300834	ARROYAVE PÉREZ ALEJANDRO	3210794
22	155	20111:Pregrado Pereira	1088006611	ASDILLA RAMÍREZ JHON ALEXANDER	2201006

Vista de datos Vista de variables

IBM SPSS Statistics Processor está listo

ES 04:15 p.m. 16/12/2013



## **11. ANÁLISIS MULTIVARIADO**

Una vez realizado el muestreo y teniendo la información de las personas a analizar en los periodos 2011-1 al 2013-1, se procede a realizar el análisis multivariado para determinar la importancia de las variables a estudiar que según estudios anteriores se relacionan directamente como posibles causas de mortalidad académica.

Para el estudio de mortalidad académica se tuvieron en cuenta las siguientes variables debido a que no fue posible tener información de todas las variables mencionadas anteriormente.

### **Variables de tipo Académico:**

- 1. Resultados pruebas de estado**
- 2. Número de Créditos vistos durante el último semestre.**
- 3. Profesor Clasificación**
- 4. Metodología del Profesor**
- 5. Orientación Profesional**
- 6. Calidad del Programa**
- 7. Relaciones con los profesores u otros estudiantes**

### **Variables de tipo Socioeconómico:**

- 1. Género**
- 2. Estrato socio económico**
- 3. Departamento de procedencia**
- 4. Municipio de procedencia**
- 5. Género del colegio en el que estudio**
- 6. Tipo de colegio en el que estudio**
- 7. Estado civil**
- 8. Edad**
- 9. Entorno Familiar**
- 10. Becas y Formas de financiamiento**
- 11. Dependencia Económica**

**12. Nivel Educativo de los padres**

**13. situación laboral de los padres**

**Variables de tipo Personal:**

**1. Motivación para cursar las materias del semestre**

**2. Satisfacción con el programa**

**3. Conocimiento previo del programa**

**4. Esfuerzo para cursar el semestre**

**5. Apoyo familiar en los estudios**

**6. Actividades extra clase.**

**7. Calamidad y problemas de salud**

**8. Incompatibilidad horario con actividades extra académicas (trabajo del estudiante)**

**9. Método de estudio**

**10. Personas a cargo.**

### **11.1 TÉCNICA DE ANÁLISIS MULTIVARIADO**

De acuerdo a las necesidades y tipo de variables que tenemos disponibles, se seleccionó La Regresión Logística Multinomial como la técnica a utilizar en el análisis multivariado teniendo en cuenta los aspectos fundamentales del proyecto y lo mencionado en los puntos anteriores.

Uno de ellos es la variable dependiente que contiene 5 categorías de respuesta:

1. Estudiante en cancelación (Nota entre 0)
2. Estudiante en periodo de prueba (Nota entre 0 y 2,5)
3. Estudiante en transición (Nota entre 2,6 y 2,9)
4. Estudiantes en normalidad académica (Nota entre 3 y 4,2)
5. Estudiante en alto rendimiento (Nota entre 4,3 y 5)

Esto hace que se imposibilite trabajar la regresión logística binaria, la cual solo admite dos categorías en la variable dependiente.

Así mismo las variables a trabajar puede que no sigan una distribución normal, es allí donde este modelo es fundamental ya que es flexible y poco riguroso con estas variables y permite operacionalizarlas sin problemas ni rigurosas pruebas de normalidad por ejemplo.

De igual forma el análisis por medio de la regresión logística multinomial es un método fácil de ejecutar el cual brinda unos resultados confiables y de calidad.

## **11.2 METODOLOGIA DEL ANALISIS MULTIVARIADO**

El procedimiento que se siguió es el de correr el programa primero con todas las variables que se consideraron, tomando como referencia la categoría:

1. Estudiante en cancelación (Nota entre 0)
2. Estudiante en periodo de prueba (Nota entre 0 y 2,5)
3. Estudiante en transición (Nota entre 2,6 y 2,9)
4. Estudiantes en normalidad académica (Nota entre 3 y 4,2)
5. Estudiante en alto rendimiento (Nota entre 4,3 y 5)

para cualquiera que sea el programa, una vez se tienen los resultados de esta primera corrida se analiza el ajuste de bondad del modelo el cual deberá de ir aumentando y tendiendo a uno con cada una de las filtraciones, esto indicara que las variables que se seleccionan se ajustan al modelo y que este cada vez es mejor. Se debe analizar además la significancia del modelo, la cual debe ir disminuyendo y acercarse cada vez a cero. Estos dos parámetros son indicadores de que el modelo está mejorando y que las variables son buenas y brindan información precisa y de calidad. De igual manera se analiza en la tabla de estimación de los parámetros, el error típico y nuevamente la significancia de cada variable ahora para cada una de las categorías.

Esta significancia debe tener valores cercanos a cero y pequeños para que la variable se considere como importante para el modelo. El error típico debe ser inferior a uno, cuando esto no pasa se analiza la significancia, pero de ser

necesario se seleccionan como importantes aquellas que tengan los errores más bajos. Como se mencionó en estudios anteriores al utilizar este modelo se menciona en este estudio.

## 12.RESULTADOS

A continuación veremos los resultados obtenidos después de correr toda la información en IBB SPSS Statistic 20

GET

```
FILE='C:\Users\VICA\Desktop\TESIS\BASE PROCESAMIENTO.sav'.
DATASET NAME Conjunto_de_datos1 WINDOW=FRONT.
DATASET ACTIVATE Conjunto_de_datos1.
SAVE OUTFILE='C:\Users\VICA\Desktop\TESIS\BASE PROCESAMIENTO.sav'
/COMPRESSED.
DATASET ACTIVATE Conjunto_de_datos1.
SAVE OUTFILE='C:\Users\VICA\Desktop\TESIS\BASE PROCESAMIENTO.sav'
/COMPRESSED.
DATASET ACTIVATE Conjunto_de_datos1.
SAVE OUTFILE='C:\Users\VICA\Desktop\TESIS\BASE PROCESAMIENTO.sav'
/COMPRESSED.
NOMREG GEN (BASE=LAST ORDER=ASCENDING) BY ED EST DEP_PROC
COLEGIO TIPO_COL ESTADO_CIVIL VIVE_UD_CON LA_CASA_ES
QUIEN_PAGA_LA_U DIFICULTAD_PAGAR_SEM
SE_PRESENTAN_CRUCES_EN_EL_HORARIO NIVEL_ED_PADRE
NIVEL_ED_MADRE EN_QUE_SE_DESEMPEÑA_PADRE
EN_QUE_SE_DESEMPEÑA_MADRE PRESENTA_ALGUNA_DISCAPACIDAD
TIENE_HIJOS ACOMPAÑAMIENTO_DE_LA_FAMILIA_EN_LA_U
HACE_CUANTO_SE_GRADUO EL_TIEMPO_Q_LE_DEDICA_AL_ESTUDIO
TIENE_BUENOS_RESULTADOS SE_ENCUNETRA_UD_CON_SU_CARRERA
CREE_Q_ESTA_EN_EL_PROGRAMA_ADECUADO
TENIA_CONOCIMINETOS_PREVIOS_DE_LA_CARRERA
POR_Q_ELIGIO_SU_CARRERA HA_CAMBIADO_DE_CARRERA
SI_LA_RESPUESTA_ES_SI_POR_QUE
EN_LA_CARRERA_SE_HA_ENCONTRADO_DOCENTES_CON
RELACION_CON_LOS_DOCENTES RELACION_LOS_COMPAÑEROS
REALIZO_EL_EXAMEN_DE_ORIENTACION_PROFECIONAL
AL_PRESENTAR_EL_EXAMEN
COMO_CALIFICA_LA_CALIDAD_DE_SU_PROGRAMA_DE_ESTUDIOS
COMO_CALIFICA_SU_PENS
```

Se encuentran singularidades inesperadas en la matriz hessiana. Esto indica que algunas variables predictoras deben excluirse o algunas categorías deberían fusionarse.

El procedimiento NOMBRE continúa a pesar de la advertencia anterior (s). Resultados posteriores muestran están basadas en la última iteración.

**Tabla 9: muestra el ajuste de bondad del modelo y su significancia la cual es muy buena**

**Modelo de Información de montaje**

Modelo	Modelo Criterios de montaje	Probabilidad ratio de pruebas		
	-2 Log Likelihood	Chi-Square	df	Sig.
Sólo Intercepción final	330,120 ,000	330,120	95	,000

Como se muestra en la tabla tenemos un ajuste de bondad de 330,120 y una significancia de cero

**Tabla 10: contraste de la razón de probabilidad**

**Probabilidad ratio de pruebas**

Efecto	Modelo Criterios de montaje	Probabilidad ratio de pruebas		
	-2 Log verosimilitud	Chi-Cuadrado	df	Sig.
Intercept	,000 <sup>a</sup>	0.000	0	
RESULTADO_S_M	,000 <sup>b</sup>	.000	1	1.000
RESULTADO_S_F	,000 <sup>b</sup>	.000	1	1.000
RESULTADO_S_L	,000 <sup>b</sup>	.000	1	1.000

VAR_SAL	,000 <sup>a</sup>	0.000	0	00
ED	,000 <sup>b</sup>	.000	3	1.0 00
EST	,000 <sup>b</sup>	.000	4	1.0 00
DEP_PROC	,000 <sup>b</sup>	.000	1	1.0 00
COLEGIO	,000 <sup>b</sup>	.000	2	1.0 00
TIPO_COL	,000 <sup>b</sup>	.000	2	1.0 00
ESTADO_CIVIL	,000 <sup>a</sup>	0.000	0	
VIVE_UD_CON	,000 <sup>b</sup>	.000	2	1.0 00
LA_CASA_ES	,000 <sup>b</sup>	.000	2	1.0 00
QUIEN_PAGA_LA_U	,000 <sup>b</sup>	.000	4	1.0 00
DIFICULTAD_PAGAR_SEM	,000 <sup>b</sup>	.000	1	1.0 00
SE_PRESENTAN_CRUCES_EN_EL_HORARIO	,000 <sup>b</sup>	.000	1	1.0 00
NIVEL_ED_PADRE	,000 <sup>b</sup>	.000	5	1.0 00
NIVEL_ED_MADRE	,000 <sup>b</sup>	.000	3	1.0 00
EN_QUE_SE_DESEMPEÑA_PADRE	,000 <sup>b</sup>	.000	5	1.0 00
EN_QUE_SE_DESEMPEÑA_MADRE	,000 <sup>b</sup>	.000	4	1.0 00
PRESENTA_ALGUNA_DISCAPACIDAD	,000 <sup>a</sup>	0.000	0	
TIENE_HIJOS	,000 <sup>b</sup>	.000	1	1.0 00
ACOMPANAMIENTO_DE_LA_FAMILIA_EN_LA_U	,000 <sup>b</sup>	.000	3	1.0 00
HACE_CUANTO_SE_GRADUO	,000 <sup>b</sup>	.000	3	1.0 00
EL_TIEMPO_Q_LE_DEDICA_AL_ESTUDIO	,000 <sup>b</sup>	.000	2	1.0 00
TIENE_BUENOS_RESULTADOS	,000 <sup>b</sup>	.000	1	1.0 00
SE_ENCUNETRA_UD_CON_SU_CARRERA	,000 <sup>b</sup>	.000	3	1.0

				00
CREE_Q_ESTA_EN_EL_PROGRAMA_ADECUADO	,000 <sup>b</sup>	.000	1	1.0 00
TENIA_CONOCIMIENTOS_PREVIOS_DE_LA_CARRERA	,000 <sup>b</sup>	.000	1	1.0 00
POR_Q_ELIGIO_SU_CARRERA	,000 <sup>b</sup>	.000	5	1.0 00
HA_CAMBIADO_DE_CARRERA	,000 <sup>b</sup>	.000	1	1.0 00
SI_LA_RESPUESTA_ES_SI_POR_QUE	,000 <sup>b</sup>	.000	4	1.0 00
EN_LA_CARRERA_SE_HA_ENCONTRADO_DOCENTES_CON	,000 <sup>b</sup>	.000	4	1.0 00
RELACION_CON_LOS_DOCENTES	,000 <sup>b</sup>	.000	2	1.0 00
RELACION_LOS_COMPAÑEROS	,000 <sup>b</sup>	.000	2	1.0 00
REALIZO_EL_EXAMEN_DE_ORIENTACION_PROFESIONAL	,000 <sup>b</sup>	.000	1	1.0 00
AL_PRESENTAR_EL_EXAMEN	,000 <sup>b</sup>	.000	2	1.0 00
COMO_CALIFICA_LA_CALIDAD_DE_SU_PROGRAMA_DE_ESTUDIOS	,000 <sup>b</sup>	.000	2	1.0 00
COMO_CALIFICA_SU_PENSUM_ACADEMICO	,000 <sup>b</sup>	.000	2	1.0 00
SE_IDENTIFICA_CON_EL_PROGRAMA_QUE_ESTA_ESTUDIANDO	,000 <sup>b</sup>	.000	1	1.0 00
HA_TENIDO_DIFICULTADES_PARA_CURSAR_EL_SEMESTRE	,000 <sup>a</sup>	0.000	0	
SI_LA_RESPUESTA_FUE_SI_CUAL_DIFICULTAD	,000 <sup>b</sup>	.000	2	1.0 00
PROG_ACA	,000 <sup>b</sup>	0.000	3	1.0 00
filter_\$	,000 <sup>a</sup>	0.000	0	
PRE_1	,000 <sup>b</sup>	.000	2	1.0 00

La estadística de chi-cuadrado es la diferencia de  $-2 \log$ -verosimilitudes entre el modelo final y un modelo reducido. El modelo reducido se forma omitiendo un efecto del modelo final. La hipótesis nula es que todos los parámetros de este efecto son 0.



a. Este modelo reducido es equivalente al modelo final porque omitiendo el efecto no aumenta los grados de libertad.

b. Se encuentran singularidades inesperadas en la matriz hessiana. Esto indica que algunas variables predictoras deben excluirse o algunas categorías deberían fusionarse.

**Tabla 11: Estimación de los parámetros**

GENERO <sup>a</sup>	B	Std. Error	Wald	df	Sig.	Exp(B)	95% intervalo de confianza para Exp (B)	
							Baja	límite superior
Intercept	-2103.664	1181791.054	.000	1	.999			
RESULTADO_S_M	-4.714	2481.685	.000	1	.998	0.01	0.00	.b
RESULTADO_S_F	2.515	1492.790	.000	1	.999	12.37	0.00	.b
RESULTADO_S_L	1.409	3042.019	.000	1	1.000	4.09	0.00	.b
VAR_SAL	-261.936	123803.382	.000	1	.998	0.00	0.00	.b
[ED=1]	186.193	111031.732	.000	1	.999	7.29E+80	0.00	.b
[ED=2]	-110.515	63226.510	.000	1	.999	0.00	0.00	.b
[ED=3]	-280.733	145657.914	.000	1	.998	0.00	0.00	.b
[ED=4]	0 <sup>c</sup>			0				
[EST=1]	636.978	327470.270	.000	1	.998	4.32E+276	0.00	.b
[EST=2]	241.499	130318.867	.000	1	.999	7.61E+104	0.00	.b
[EST=3]	514.092	264481.946	.000	1	.998	1.85E+223	0.00	.b
[EST=4]	144.444	94492.	.000	1	.999	5.39E+62	0.00	.b

		163					00	
[EST=5]	131.285	0.000		1		1.04E+57	1.0 4E +57	1.0 4E +57
[EST=6]	0 <sup>c</sup>			0				
[DEP_PROC=1]	132.918	70234. 982	.000	1	.998	5.32E+57	0.0 00	. b
[DEP_PROC=2]	0 <sup>c</sup>			0				
[COLEGIO=FEMEN INO]	-282.405	133788 .065	.000	1	.998	0.00	0.0 00	. b
[COLEGIO=MASC ULINO]	-211.900	100126 .242	.000	1	.998	0.00	0.0 00	. b
[COLEGIO=MIXTO]	0 <sup>c</sup>			0				
[TIPO_COL=PRIVA DO]	228.502	99620. 038	.000	1	.998	1.73E+99	0.0 00	. b
[TIPO_COL=PUBLI CO]	-426.840	222087 .811	.000	1	.998	0.00	0.0 00	. b
[TIPO_COL=SEMIP RIVADO]	0 <sup>c</sup>			0				
[ESTADO_CIVIL=S OLTERO]	439.553	208985 .118	.000	1	.998	7.86E+190	0.0 00	. b
[ESTADO_CIVIL=U NION LIBRE]	0 <sup>c</sup>			0				
[VIVE_UD_CON=C ON ALGUN FAMILIAR]	-54.608	60692. 809	.000	1	.999	0.00	0.0 00	. b
[VIVE_UD_CON=C ON SU PAREJA]	0 <sup>c</sup>			0				
[VIVE_UD_CON=C ON SUS PADRES]	-13.229	23177. 465	.000	1	1.000	0.00	0.0 00	. b
[VIVE_UD_CON=S OLO]	0 <sup>c</sup>			0				
[LA_CASA_ES=AR RENDADA]	77.926	58214. 430	.000	1	.999	6.96E+33	0.0 00	. b
[LA_CASA_ES=PR OPIA]	-20.083	40517. 401	.000	1	1.000	0.00	0.0 00	. b
[LA_CASA_ES=US UFRUCTO]	0 <sup>c</sup>			0				

[QUIEN_PAGA_LA _U=DEPENDE DE SUS PADRES]	314.972	208385 .055	.000	1	.999	6.18E+136	0.0 00	. b
[QUIEN_PAGA_LA _U=ESTA FINANCIADO POR ALGUN PROGRAMA]	-89.000	94336. 212	.000	1	.999	0.00	0.0 00	. b
[QUIEN_PAGA_LA _U=SE ENCUENTRA LABORANDO]	411.827	230497 .905	.000	1	.999	7.15E+178	0.0 00	. b
[QUIEN_PAGA_LA _U=TIENE APOYO ECONOMICO DE UN FAMILIAR]	1350.451	752741 .017	.000	1	.999	. b	0.0 00	. b
[QUIEN_PAGA_LA _U=TIENE UNA BECA ESTUDIANTIL]	0 <sup>c</sup>			0				
[DIFICULTAD_PAG AR_SEM=NO]	100.296	53452. 836	.000	1	.999	3.61E+43	0.0 00	. b
[DIFICULTAD_PAG AR_SEM=SI]	0 <sup>c</sup>			0				
[SE_PRESENTAN_ CRUCES_EN_EL_ HORARIO=NO]	4.504	37754. 558	.000	1	1.000	90.34	0.0 00	. b
[SE_PRESENTAN_ CRUCES_EN_EL_ HORARIO=SI]	0 <sup>c</sup>			0				
[NIVEL_ED_PADR E=BASICA PRIMARIA]	-115.882	48290. 441	.000	1	.998	0.00	0.0 00	. b
[NIVEL_ED_PADR E=BASICA SECUNDARIA]	-185.296	84653. 354	.000	1	.998	0.00	0.0 00	. b

[NIVEL_ED_PADR E=CURSOS BASICOS]	113.379	79029. 904	.000	1	.999	1.74E+49	0.0 00	. b
[NIVEL_ED_PADR E=NO VIVE CON ELLOS O ESTA MUERTO]	-198.763	101652 .185	.000	1	.998	0.00	0.0 00	. b
[NIVEL_ED_PADR E=TECNICOS]	347.969	222927 .027	.000	1	.999	1.32E+151	0.0 00	. b
[NIVEL_ED_PADR E=UNIVERSITARI OS]	0 <sup>c</sup>			0				
[NIVEL_ED_MADR E=BASICA PRIMARIA]	346.791	157068 .952	.000	1	.998	4.07E+150	0.0 00	. b
[NIVEL_ED_MADR E=BASICA SECUNDARIA]	197.074	97676. 268	.000	1	.998	3.87E+85	0.0 00	. b
[NIVEL_ED_MADR E=CURSOS BASICOS]	0 <sup>c</sup>			0				
[NIVEL_ED_MADR E=TECNICOS]	-35.885	48564. 454	.000	1	.999	0.00	0.0 00	. b
[NIVEL_ED_MADR E=UNIVERSITARI OS]	0 <sup>c</sup>			0				
[EN_QUE_SE_DES EMPEÑA_PADRE= EMPLEADO PUBLICO]	-178.429	95195. 898	.000	1	.999	0.00	0.0 00	. b
[EN_QUE_SE_DES EMPEÑA_PADRE= EMPLEO FORMAL EN COMPAÑIA PRIVADA]	48.403	31744. 479	.000	1	.999	1.05E+21	0.0 00	. b
[EN_QUE_SE_DES EMPEÑA_PADRE= EMPRESARIO]	421.874	185789 .559	.000	1	.998	1.65E+183	0.0 00	. b

[EN_QUE_SE_DES EMPEÑA_PADRE= JUBILADO]	244.799	134385 .692	.000	1	.999	2.07E+106	0.0 00	. b
[EN_QUE_SE_DES EMPEÑA_PADRE= NO VIVE CON ELLOS O ESTA MUERTO]	87.338	107722 .379	.000	1	.999	8.52E+37	0.0 00	. b
[EN_QUE_SE_DES EMPEÑA_PADRE= TRABAJADOR INDEPENDIENTE]	0 <sup>c</sup>			0				
[EN_QUE_SE_DES EMPEÑA_MADRE= AMA DE CASA]	-377.370	172160 .361	.000	1	.998	0.00	0.0 00	. b
[EN_QUE_SE_DES EMPEÑA_MADRE= EMPLEADO PUBLICO]	-604.466	277728 .707	.000	1	.998	0.00	0.0 00	. b
[EN_QUE_SE_DES EMPEÑA_MADRE= EMPLEO FORMAL EN COMPAÑIA PRIVADA]	-158.839	78648. 268	.000	1	.998	0.00	0.0 00	. b
[EN_QUE_SE_DES EMPEÑA_MADRE= EMPRESARIO]	-9.597	52439. 827	.000	1	1.000	0.00	0.0 00	. b
[EN_QUE_SE_DES EMPEÑA_MADRE= TRABAJADOR INDEPENDIENTE]	0 <sup>c</sup>			0				
[PRESENTA_ALGU NA_DISCAPACIDA D=NO]	1420.967	688677 .554	.000	1	.998	. b	0.0 00	. b
[PRESENTA_ALGU NA_DISCAPACIDA D=SI]	0 <sup>c</sup>			0				

[TIENE_HIJOS=NO ]	706.379	435935 .176	.000	1	.999	5.98E+306	0.0 00	. b
[TIENE_HIJOS=SI]	0 <sup>c</sup>			0				
[ACOMPÑAMIEN TO_DE_LA_FAMILI A_EN_LA_U=NO RECIBE NINGUN APOYO, NI RECIBE NINGUNA OPOSICION]	422.100	211871 .358	.000	1	.998	2.07E+183	0.0 00	. b
[ACOMPÑAMIEN TO_DE_LA_FAMILI A_EN_LA_U=RECI BE ALGUN RESPALDO ECONOMICO Y A VECES SE INTERESAN POR SU VIDA UNIVERSITARIA]	-226.529	169225 .988	.000	1	.999	0.00	0.0 00	. b
[ACOMPÑAMIEN TO_DE_LA_FAMILI A_EN_LA_U=SIEM PRE EXISTE TOTAL APOYO ECONOMICO, INTELECTUAL Y AFECTIVO]	207.695	109548 .862	.000	1	.998	1.59E+90	0.0 00	. b

[ACOMPAÑAMIENTO DE LA FAMILIA EN LA UNIVERSIDAD APOYO ECONOMICO LIMITADO O CONDICIONAL, Y RARA VEZ PREGUNTAN POR SU VIDA UNIVERSITARIA]	0 <sup>c</sup>			0				
[HACE CUANTO SE GRADUO=1 A 2 AÑOS]	-120.953	92168.850	.000	1	.999	0.00	0.000	.b
[HACE CUANTO SE GRADUO=3 A 5 AÑOS]	158.787	98331.724	.000	1	.999	9.12E+68	0.000	.b
[HACE CUANTO SE GRADUO=EL AÑO PASADO]	-192.938	125360.277	.000	1	.999	0.00	0.000	.b
[HACE CUANTO SE GRADUO=MAS DE 5 AÑOS]	0 <sup>c</sup>			0				
[EL TIEMPO QUE LE DEDICA AL ESTUDIO=BUENO]	-40.926	37400.991	.000	1	.999	0.00	0.000	.b
[EL TIEMPO QUE LE DEDICA AL ESTUDIO=MUY BUENO]	-58.886	38140.243	.000	1	.999	0.00	0.000	.b
[EL TIEMPO QUE LE DEDICA AL ESTUDIO=REGULAR]	0 <sup>c</sup>			0				
[TIENE BUENOS RESULTADOS=NO]	-333.074	196678.102	.000	1	.999	0.00	0.000	.b
[TIENE BUENOS RESULTADOS=SI]	0 <sup>c</sup>			0				

[SE_ENCUNETRA_UD_CON_SU_CARRERA=CANSADO EN EL SEMESTRE]	-260.465	187905.709	.000	1	.999	0.00	0.000	.b
[SE_ENCUNETRA_UD_CON_SU_CARRERA=MOTIVADO PARA CURSAR ESTE SEMESTRE]	-356.409	225087.478	.000	1	.999	0.00	0.000	.b
[SE_ENCUNETRA_UD_CON_SU_CARRERA=MUY CONTENTO EN SU CARRERA]	-597.610	329624.090	.000	1	.999	0.00	0.000	.b
[SE_ENCUNETRA_UD_CON_SU_CARRERA=NO ESTA FELIZ EN SU CARRERA]	0 <sup>c</sup>			0				
[CREE_Q_ESTA_EN_EL_PROGRAMA_ADECUADO=NO]	-31.620	43751.333	.000	1	.999	0.00	0.000	.b
[CREE_Q_ESTA_EN_EL_PROGRAMA_ADECUADO=SI]	0 <sup>c</sup>			0				
[TENIA_CONOCIMIENTOS_PREVIOS_DE_LA_CARRERA=NO]	49.377	30826.902	.000	1	.999	2.78E+21	0.000	.b
[TENIA_CONOCIMIENTOS_PREVIOS_DE_LA_CARRERA=SI]	0 <sup>c</sup>			0				



[POR_Q_ELIGIO_S U_CARRERA=EL PUNTAJE DEL ICFES NO LE ALCANZO PARA LO QUE QUERIA ESTUDIAR]	-29.935	29185. 730	.000	1	.999	0.00	0.0 00	. b
[POR_Q_ELIGIO_S U_CARRERA=POR GUSTO]	-168.114	110846 .378	.000	1	.999	0.00	0.0 00	. b
[POR_Q_ELIGIO_S U_CARRERA=POR ORIENTACION PROFESIONAL]	-616.620	357293 .448	.000	1	.999	0.00	0.0 00	. b
[POR_Q_ELIGIO_S U_CARRERA=POR SEGUIR LAS RECOMENDACIONES FAMILIARES]	116.057	59369. 544	.000	1	.998	2.53E+50	0.0 00	. b
[POR_Q_ELIGIO_S U_CARRERA=POR SUS CONOCIMIENTOS PREVIOS]	-113.589	81368. 835	.000	1	.999	0.00	0.0 00	. b
[POR_Q_ELIGIO_S U_CARRERA=POR QUE TIENE BUEN CAMPO DE ACCION LABORAL]	0 <sup>c</sup>			0				
[HA_CAMBIADO_D E_CARRERA=NO]	314.661	184201 .269	.000	1	.999	4.53E+136	0.0 00	. b
[HA_CAMBIADO_D E_CARRERA=SI]	0 <sup>c</sup>			0				
[SI_LA_RESPUESTA ES_SI_POR_QUE E=DECISION PROPIA]	280.294	179453 .604	.000	1	.999	5.37E+121	0.0 00	. b

[SI_LA_RESPUESTA_ES_SI_POR_QU E=NO APLICA]	336.644	187586 .384	.000	1	.999	1.59E+146	0.0 00	. b
[SI_LA_RESPUESTA_ES_SI_POR_QU E=POR DIFICULTADES DE APRENDIZAJE]	-505.713	232508 .458	.000	1	.998	0.00	0.0 00	. b
[SI_LA_RESPUESTA_ES_SI_POR_QU E=POR LA DIFICULTAD DE LA CARRERA]	-596.510	269002 .422	.000	1	.998	0.00	0.0 00	. b
[SI_LA_RESPUESTA_ES_SI_POR_QU E=SITUACION ECONOMICA]	0 <sup>c</sup>			0				
[EN_LA_CARRERA _SE_HA_ENCONT RADO_DOCENTE S_CON=EL DOCENTE DICTABA UNA ASIGNATURA QUE NO LE CORRESPONDIA CON EL PERFIL PARA EL CUAL ESTABA FORMADO]	-263.962	187411 .716	.000	1	.999	0.00	0.0 00	. b

[EN_LA_CARRERA _SE_HA_ENCONT RADO_DOCENTE S_CON=EL DOCENTE MANIFESTABA SERIAS DEFICIENCIAS EN PEDAGOGIA]	236.333	116554 .981	.000	1	.998	4.35E+102	0.0 00	. <sup>b</sup>
[EN_LA_CARRERA _SE_HA_ENCONT RADO_DOCENTE S_CON=MALA METODOLOGIA PARA TRANSMITIR SUS CONOCIMIENTOS]	167.133	90857. 118	.000	1	.999	4.35E+102	0.0 00	. <sup>b</sup>
[EN_LA_CARRERA _SE_HA_ENCONT RADO_DOCENTE S_CON=NINGUNA DE LAS ANTERIORES]	-17.918	96584. 794	.000	1	1.000	0.00	0.0 00	. <sup>b</sup>
[EN_LA_CARRERA _SE_HA_ENCONT RADO_DOCENTE S_CON=SIN CONOCIMIENTO SOBRE LA MATERIA]	0 <sup>c</sup>			0				
[RELACION_CON_ LOS_DOCENTES= BUENA]	-292.303	182317 .690	.000	1	.999	0.00	0.0 00	. <sup>b</sup>
[RELACION_CON_ LOS_DOCENTES= MALA]	0 <sup>c</sup>			0				

[RELACION_CON_LOS_DOCENTES=MUY BUENA]	-492.970	277657.304	.000	1	.999	0.00	0.000	.b
[RELACION_CON_LOS_DOCENTES=REGULAR]	0 <sup>c</sup>			0				
[RELACION_LOS_COMPAÑEROS=NO INTERACTUA CON SUS COMPAÑEROS]	-1407.120	786808.159	.000	1	.999	0.00	0.000	.b
[RELACION_LOS_COMPAÑEROS=TIENE PROBLEMAS CON SUS COMPAÑEROS]	-255.191	117633.182	.000	1	.998	0.00	0.000	.b
[RELACION_LOS_COMPAÑEROS=UNA BUENA RELACION]	0 <sup>c</sup>			0				
[REALIZO_EL_EXAMEN_DE_ORIENTACION_PROFECIONAL=NO]	235.328	108140.319	.000	1	.998	1.59E+102	0.000	.b
[REALIZO_EL_EXAMEN_DE_ORIENTACION_PROFECIONAL=SI]	0 <sup>c</sup>			0				
[AL_PRESENTAR_EL_EXAMEN=LE AYUDO A SELECCIONAR UNA BUENA CARRERA]	-5.955	25837.181	.000	1	1.000	0.00	0.000	.b
[AL_PRESENTAR_EL_EXAMEN=NO APLICA]	60.880	64838.423	.000	1	.999	2.75E+26	0.000	.b

[AL_PRESENTAR_EL_EXAMEN=SE_ENCONTRABA_USTED_EN_LA_CARRERA_CORRECTA]	0 <sup>c</sup>			0				
[COMO_CALIFICA_LA_CALIDAD_DE_SU_PROGRAMA_DE_ESTUDIOS=BUENA]	280.659	170816.337	.000	1	.999	7.74E+121	0.000	.b
[COMO_CALIFICA_LA_CALIDAD_DE_SU_PROGRAMA_DE_ESTUDIOS=MUY BUENA]	481.300	284076.989	.000	1	.999	1.06E+209	0.000	.b
[COMO_CALIFICA_LA_CALIDAD_DE_SU_PROGRAMA_DE_ESTUDIOS=REGULAR]	0 <sup>c</sup>			0				
[COMO_CALIFICA_SU_PENSUM_ACADEMICO=BUENA]	18.267	27889.561	.000	1	.999	85771438.27	0.000	.b
[COMO_CALIFICA_SU_PENSUM_ACADEMICO=MUY BUENA]	-80.131	61078.550	.000	1	.999	0.00	0.000	.b
[COMO_CALIFICA_SU_PENSUM_ACADEMICO=REGULAR]	0 <sup>c</sup>			0				
[SE_IDENTIFICA_CON_EL_PROGRAMA_QUE_ESTA_ESTUDIANDO=NO]	39.873	55660.645	.000	1	.999	2.07241E+17	0.000	.b

[SE_IDENTIFICA CON_EL_PROGRA MA_QUE_ESTA_E STUDIANDO=SI]	0 <sup>c</sup>			0				
[HA_TENIDO_DIFI CULTADES_PARA _CURSAR_EL_SE MESTRE=NO]	-88.791	72003. 327	.000	1	.999	0.00	0.0 00	. b
[HA_TENIDO_DIFI CULTADES_PARA _CURSAR_EL_SE MESTRE=SI]	0 <sup>c</sup>			0				
[SI_LA_RESPUEST A_FUE_SI_CUAL_ DIFICULTAD=CON SUS PROFESORES]	-115.304	76155. 830	.000	1	.999	0.00	0.0 00	. b
[SI_LA_RESPUEST A_FUE_SI_CUAL_ DIFICULTAD=ECO NOMICAS]	-426.740	246583 .649	.000	1	.999	0.00	0.0 00	. b
[SI_LA_RESPUEST A_FUE_SI_CUAL_ DIFICULTAD=LAB ORALES]	0 <sup>c</sup>			0				
[SI_LA_RESPUEST A_FUE_SI_CUAL_ DIFICULTAD=NO APLICA]	0 <sup>c</sup>			0				
[SI_LA_RESPUEST A_FUE_SI_CUAL_ DIFICULTAD=O PERSONALES]	0 <sup>c</sup>			0				
[PROG_ACA=1]	.000	9434.7 75	.000	1	1.000	1.00	0.0 00	. b
[PROG_ACA=2]	.000	7066.0 15	.000	1	1.000	1.00	0.0 00	. b

[PROG_ACA=3]	.000	8242.5 49	.000	1	1.000	1.00	0.0 00	. <sup>b</sup>
[PROG_ACA=4]	0 <sup>c</sup>			0				
[filter_\$=1]	0 <sup>c</sup>			0				
[PRE_1=1]	-484.966	220108 .252	.000	1	.998	0.00	0.0 00	. <sup>b</sup>
[PRE_1=2]	-85.021	58214. 863	.000	1	.999	0.00	0.0 00	. <sup>b</sup>
[PRE_1=3]	0 <sup>c</sup>			0				
[PRE_1=4]	0 <sup>c</sup>			0				

Con respecto a la Tabla 3 vemos que cada ítem tiene igual comparación, por lo cual no ninguna influye más que las otras.

**Tabla 12: Resumen de la encuesta propuesta a cada uno de los participantes**

Resumen de procesamiento de casos			
		N	Porcentaje Marginal
GENERO	MASCULINO	222	75,3%
	FEMENINO	73	24,7%
RANGO DE EDAD	ENTRE 15 Y 18 AÑOS	108	36,6%
	ENTRE 18 Y 20 AÑOS	94	31,9%
	ENTRE 20 Y 25 AÑOS	79	26,8%
	MAYOR DE 25 AÑOS	14	4,7%
ESTRATO SOCIOECONOMICO	E1	45	15,3%
	E2	136	46,1%
	E3	87	29,5%
	E4	19	6,4%
	E5	3	1,0%
	E6	5	1,7%
DEPARTAMENTO DE PROCEDENCIA	RISARALDA	247	83,7%
	VALLE	48	16,3%
NATURALEZA COLEGIO	FEMENINO	31	10,5%
	MASCULINO	6	2,0%
	MIXTO	258	87,5%
SI ES PUBLI O	PRIVADO	34	11,5%

PRIVAD	PUBLICO	234	79,3%
	SEMIPRIVADO	27	9,2%
ESTADO CIVIL	SOLTERO	291	98,6%
	UNION LIBRE	4	1,4%
CON QUIEN VIVE UD	CON ALGUN FAMILIAR	27	9,2%
	CON SU PAREJA	4	1,4%
	CON SUS PADRES	178	60,3%
	SOLO	86	29,2%
LA CASA ES	ARRENDADA	65	22,0%
	PROPIA	192	65,1%
	USUFRUCTO	38	12,9%
QUIEN PAGA LA UNIVERSIDAD	DEPENDE DE SUS PADRES	229	77,6%
	ESTA FINANCIADO POR ALGUN PROGRAMA	12	4,1%
	SE ENCUENTRA LABORANDO	16	5,4%
	TIENE APOYO ECONOMICO DE UN FAMILIAR	20	6,8%
	TIENE UNA BECA ESTUDIANTIL	18	6,1%
TIENE DIFICULTAD PARA PAGAR SEMESTRE	NO	235	79,7%
	SI	60	20,3%
SE LE PRESENTAN CRUCES, CON EL HORARIO DE TRABAJO	NO	285	96,6%
	SI	10	3,4%
NIVEL EDUCATIVO DEL PADRE	BASICA PRIMARIA	65	22,0%
	BASICA SECUNDARIA	130	44,1%
	CURSOS BASICOS	10	3,4%
	NO VIVE CON ELLOS O ESTA MUERTO	22	7,5%
	TECNICOS	9	3,1%
	UNIVERSITARIOS	59	20,0%
NIVEL EDUCATIVO DE LA MADRE	BASICA PRIMARIA	47	15,9%
	BASICA SECUNDARIA	155	52,5%



	CURSOS BASICOS	3	1,0%
	TECNICOS	27	9,2%
	UNIVERSITARIOS	63	21,4%
EN QUE LABORA EL PADRE	EMPLEADO PUBLICO	50	16,9%
	EMPLEO FORMAL EN COMPAÑIA PRIVADA	51	17,3%
	EMPRESARIO	11	3,7%
	JUBILADO	22	7,5%
	NO VIVE CON ELLOS O ESTA MUERTO	51	17,3%
	TRABAJADOR INDEPENDIENTE	110	37,3%
	EN QUE LABORA LA MADRE	AMA DE CASA	163
EMPLEADO PUBLICO		28	9,5%
EMPLEO FORMAL EN COMPAÑIA PRIVADA		47	15,9%
EMPRESARIO		10	3,4%
TRABAJADOR INDEPENDIENTE		47	15,9%
PRESENTA ALGUNA DISCAPACIDAD	NO	294	99,7%
	SI	1	,3%
TIENE HIJOS	NO	284	96,3%
	SI	11	3,7%
COMO ES EL ACOMPAÑAMIENTO DE LA FAMILIA EN LA U	NO RECIBE NINGUN APOYO, NI RECIBE NINGUNA OPOSICION	4	1,4%
	RECIBE ALGUN RESPALDO ECONOMICO Y A VECES SE INTERESAN POR SU VIDA UNIVERSITARIA	11	3,7%
	SIEMPRE EXISTE TOTAL APOYO ECONOMICO, INTELLECTUAL Y AFECTIVO	267	90,5%
	TIENE APOYO ECONOMICO LIMITADO O CONDICIONAL, Y	13	4,4%

	RARA VEZ PREGUNTAN POR SU VIDA UNIVERSITARIA		
HACE CUANTO SE GRADUO	1 A 2 AÑOS	98	33,2%
	3 A 5 AÑOS	104	35,3%
	EL AÑO PASADO	38	12,9%
	MAS DE 5 AÑOS	55	18,6%
COMO ES EL TIEMPO QUE LE DEDICA AL ESTUDIO	BUENO	101	34,2%
	MUY BUENO	172	58,3%
	REGULAR	22	7,5%
CREE QUE EL METODO DE ESTUDIO FUNCIONA	NO	41	13,9%
	SI	254	86,1%
EN ESTOS MOMENTOS UD ESTA	CANSADO EN EL SEMESTRE	15	5,1%
	MOTIVADO PARA CURSAR ESTE SEMESTRE	65	22,0%
	MUY CONTENTO EN SU CARRERA	198	67,1%
	NO ESTA FELIZ EN SU CARRERA	17	5,8%
CREE QUE ESTA EN EL PROGRAMA ADECUADO	NO	44	14,9%
	SI	251	85,1%
TENIA CONOCIMIENTOS PREVIOS DE LA CARRERA	NO	88	29,8%
	SI	207	70,2%
POR QUE ELIGIO ESTA CARRERA	EL PUNTAJE DEL ICFES NO LE ALCANZO PARA LO QUE QUERIA ESTUDIAR	29	9,8%
	POR GUSTO	176	59,7%
	POR ORIENTACION PROFESIONAL	5	1,7%
	POR SEGUIR LAS RECOMENDACIONES FAMILIARES	46	15,6%

	POR SUS CONOCIMIENTOS PREVIOS	12	4,1%
	PORQUE TIENE BUEN CAMPO DE ACCION LABORAL	27	9,2%
SE HA CAMBIADO DE CARRERA	NO	237	80,3%
	SI	58	19,7%
SI LA RESPUESTA ES SI POR QUE	DECISION PROPIA	26	8,8%
	NO APLICA	241	81,7%
	POR DIFICULTADES DE APRENDIZAJE	10	3,4%
	POR LA DIFICULTAD DE LA CARRERA	5	1,7%
	SITUACION ECONOMICA	13	4,4%
EN EL TRANCURSO DE LA CARRERA SE HA ENCONTRADO CON DOCENCTES CON	EL DOCENTE DICTABA UNA ASIGNATURA QUE NO LE CORRESPONDIA CON EL PERFIL PARA EL CUAL ESTABA FORMADO	14	4,7%
	EL DOCENTE MANIFESTABA SERIAS DEFICIENCIAS EN PEDAGOGIA	41	13,9%
	MALA METODOLOGIA PARA TRANSMITIR SUS CONOCIMIENTOS	140	47,5%
	NINGUNA DE LAS ANTERIORES	88	29,8%
	SIN CONOCIMIENTO SOBRE LA MATERIA	12	4,1%
LA RELACION CON LOS PROFESORES	BUENA	154	52,2%
	MALA	3	1,0%
	MUY BUENA	122	41,4%
	REGULAR	16	5,4%
LA RELACION CON	NO INTERACTUA	5	1,7%

LOS COMPAÑEROS	CON SUS COMPAÑEROS		
	TIENE PROBLEMAS CON SUS COMPAÑEROS	42	14,2%
	UNA BUENA RELACION	248	84,1%
REALIZO UD EL EXAMEN DE ORIENTACION PROFECIONAL	NO	212	71,9%
	SI	83	28,1%
AL PRESENTAR EL EXAMEN	LE AYUDO A SELECCIONAR UNA BUENA CARRERA	35	11,9%
	NO APLICA	210	71,2%
	SE ENCONTRABA USTED EN LA CARRERA CORRECTA	50	16,9%
COMO CALIFICA EL PROGRAMA DE ESTUDIOS	BUENA	157	53,2%
	MUY BUENA	105	35,6%
	REGULAR	33	11,2%
COMO CALIFICA SU PENSUM ACADEMICO	BUENA	146	49,5%
	MUY BUENA	100	33,9%
	REGULAR	49	16,6%
SE IDENTIFICA CON EL PROGRAMA QUE ESTA ESTUDIANDO	NO	53	18,0%
	SI	242	82,0%
HA TENIDO DIFICULTADES PARA CURSAR EL SEMESTE	NO	240	81,4%
	SI	55	18,6%
SI LA RESPUESTA FUE SI CUAL FUE LA DIFICULTAD	CON SUS PROFESORES	12	4,1%
	ECONOMICAS	20	6,8%
	LABORALES	1	,3%
	NO APLICA	240	81,4%
	O PERSONALES	22	7,5%
PROGRAMA ACADEMICO	INGENIERIA MECANICA	64	21,7%
	INGENIERIA D ESITEMAS	113	38,3%

	INGENIERIA ELECTRICA	90	30,5%
	INGENIERIA FISICA	28	9,5%
Aproximadamente 80% de los casos (SAMPLE)	1	295	100,0%
Categoría de respuesta pronosticada	PERIODO DE PRUEBA	70	23,7%
	TRANSICION	39	13,2%
	NORMALIDAD ACADEMICA	159	53,9%
	ALTO RENDIMIENTO	27	9,2%
valido		295	100,0%
Error		55	
Total		350	
subpoblación		112 <sup>a</sup>	
a. La variable dependiente sólo tiene un valor observado en 112 (100,0%) subpoblaciones.			

### Pseudo R-Square

Cox and Snell	,673
Nagelkerke	1,000
McFadden	1,000

### 13. ANALISIS REGRESION LOGISITICA BINARIA

Primero se realizó el análisis multivariado. Luego se clasificó la variable salida correspondiente a la nota en 1 no aprobado y 2 aprobado y se corrió el modelo en regresión logística binaria para crear el modelo.

La regresión logística binaria se utilizó para predecir el resultado de una variable categórica (una variable que puede adoptar un número limitado de categorías) en función de las variables independientes o predictivas. Es útil para modelar la probabilidad de un evento ocurriendo como función de otros factores.

La regresión logística analiza datos distribuidos binomialmente de la forma

$$Y_i \sim B(p_i, n_i), \text{ para } i = 1, \dots, m,$$

Donde los números de ensayos Bernoulli  $n_i$  son conocidos y las probabilidades de éxito  $p_i$  son desconocidas. Un ejemplo de esta distribución es el porcentaje de semillas ( $p_i$ ) que germinan después de que  $n_i$  son plantadas.

El modelo es entonces obtenido a base de lo que cada ensayo (valor de  $i$ ) y el conjunto de variables explicativas/independientes puedan informar acerca de la probabilidad final. Estas variables explicativas pueden pensarse como un vector  $X_i$   $k$ -dimensional y el modelo toma entonces la forma

$$p_i = E \left( \frac{Y_i}{n_i} \middle| X_i \right).$$

Los logits de las probabilidades binomiales desconocidas (*i.e.*, los logaritmos de la razón de momios) son modeladas como una función lineal de los  $X_i$ .

$$\text{logit}(p_i) = \ln \left( \frac{p_i}{1 - p_i} \right) = \beta_0 + \beta_1 x_{1,i} + \dots + \beta_k x_{k,i}.$$

Note que un elemento particular de  $X_i$  puede ser ajustado a 1 para todo  $i$  obteniéndose una constante independiente en el modelo. Los parámetros desconocidos  $\beta_j$  son usualmente estimados a través de máxima verosimilitud.

La interpretación de los estimados del parámetro  $\beta_j$  es como los efectos aditivos en el logaritmo de la razón de momios para una unidad de cambio en la  $j$ ésima variable explicativa. En el caso de una variable explicativa dicotómica, por ejemplo género,  $e^{\beta}$  es la estimación del odds ratio de tener el resultado para, por decir algo, hombres comparados con mujeres.

El modelo tiene una formulación equivalente dada por

$$p_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_{1,i} + \dots + \beta_k x_{k,i})}}$$

Esta forma funcional es comúnmente identificada como un "perceptrón" de una capa simple or red neuronal artificial de una sola capa. Una red neuronal de una sola capa calcula una salida continua en lugar de una función definida a trozos. La derivada de  $p_i$  con respecto a  $X = x_1 \dots x_k$  es calculada de la forma general:

$$y = \frac{1}{1 + e^{-f(X)}}$$

donde  $f(X)$  es una función analítica en  $X$ . Con esta escogencia, la red de capa simple es idéntica al modelo de regresión logística. Esta función tiene una derivada continua, la cual permite ser usada en propagación hacia atrás. Esta función también es preferida pues su derivada es fácilmente calculable:

$$y' = y(1 - y) \frac{df}{dX}$$

La regresión logística unidimensional puede usarse para tratar de correlacionar la probabilidad de una variable cualitativa binaria (asumiremos que puede tomar los valores reales "0" y "1") con una variable escalar  $x$ . La idea es que la regresión

logística aproxime la probabilidad de obtener "0" (no ocurre cierto suceso) o "1" (ocurre el suceso) con el valor de la variable explicativa  $x$ . En esas condiciones, la probabilidad aproximada del suceso se aproximará mediante una función logística del tipo:

$$\pi(x) = \frac{e^{(\beta_0 + \beta_1 x)}}{e^{(\beta_0 + \beta_1 x)} + 1} = \frac{1}{e^{-(\beta_0 + \beta_1 x)} + 1},$$

que puede reducirse al cálculo de una regresión lineal para la función [logit](#) de la probabilidad:

$$g(x) = \ln \frac{\pi(x)}{1 - \pi(x)} = \beta_0 + \beta_1 x,$$

o una regresión exponencial:

$$\frac{\pi(x)}{1 - \pi(x)} = e^{(\beta_0 + \beta_1 x)}.$$

El grafo de la función logística se muestra en la figura que encabeza esta sección, la variable independiente es la combinación lineal  $\beta_0 + \beta_1 x$  y la variable dependiente es la probabilidad estimada  $\pi(x)$ . Si se realiza la regresión lineal, la forma de la probabilidad estimada puede ser fácilmente recuperada a partir de los coeficientes calculados:<sup>1</sup>

Para hacer la regresión deben tomarse los valores  $X_i$  de las observaciones ordenados de mayor a menor y formar la siguiente tabla:



**Tabla 13: observaciones ordenados de mayor a menor y formar**

Valores ordenados	Valor de categoría	Probabilidad estimada	Logit
$X_1$	$\varepsilon_1$	$\pi(X_1)$	$g(X_1)$
$X_2$	$\varepsilon_2$	$\pi(X_2)$	$g(X_2)$
...	...	...	...
$X_n$	$\varepsilon_n$	$\pi(X_n)$	$g(X_n)$

Donde  $\varepsilon_i$  es "0" o "1" según el caso y además:

$$0 \leq \pi(X_i) = \frac{\sum_{k=1}^i \varepsilon_k}{i} \leq 1, \quad g(X_i) = \ln \left( \frac{\pi(X_i)}{1 - \pi(X_i)} \right) = \beta_0 + \beta_1 X_i$$

El el cálculo de  $g$  pueden aparecer problemas al principio del intervalo si  $\pi(X_j) = 0$  para algunos valores de  $j$ .

## 14.RESULTADOS

A continuación veremos los resultados obtenidos después de correr toda la información en IBB SPSS Statistic 20

Bloque 0:

**Tabla 14: Tabla De Clasificación**

**Tabla de clasificación<sup>a,b</sup>**

	Observed		Predicted		
			VARIABLE DE RESPUESTA		Porcentaje Correcto
			NO APROBATORIO	APROBATORIO	
Step 0	VARIABLE DE RESPUESTA	NO APROBATORIO	0	131	,0
		APROBATORIO	0	219	100,0
	Overall Percentage				62,6

a. Constante está incluido en el modelo.

a. El valor de corte es, 500

Para el paso cero el modelo no clasifica los no aprobados, clasificando solo los datos aprobados en la variable salida y es solo clasifica el 62.6 % de datos. Por eso se generan el método de wald paso a paso hasta obtener el mejor modelo en el sexto paso.

Bloque 1: Método = Por pasos hacia adelante (Wald)

**Tabla 15: Resumen De Los Modelos**

**Reaumen de los modelos**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	428,089 <sup>a</sup>	,095	,129
2	398,838 <sup>a</sup>	,167	,228
3	367,060 <sup>a</sup>	,239	,326
4	320,860 <sup>b</sup>	,333	,455
5	300,067 <sup>b</sup>	,372	,507
6	342,035 <sup>c</sup>	,292	,398

a. Estimación termina en la iteración número 5 porque las estimaciones de los parámetros modificados por menos de, 001.

b. Estimación por terminado en número de iteración 20 porque se ha alcanzado el máximo de iteraciones. La solución final no puede ser encontrado.

c. Estimación termina en la iteración número 6 porque las estimaciones de los parámetros modificados por menos de, 001.

**Tabla 16: Prueba De Hosmer Y Lemeshow**

**Prueba de Hosmer y Lemeshow**

Step	Chi-square	df	Sig.
1	,000	2	1,000
2	9,169	4	,057
3	12,887	8	,116

4	26,528	8	,001
5	19,603	8	,012
6	23,488	8	,003

En la prueba de Hosmer y Lemeshow podemos observar que en el sexto paso el valor es de .003 indicando que no se ajusta el modelo a los datos, pero en el paso sexto de Wald el modelo se ajusta a un 76% en la clasificación de los datos como lo muestra la tabla de clasificación.

**Tabla 17: Tabla De Clasificación.**

**Tabla de Clasificación<sup>a</sup>**

	Observed	Predicted			
		VARIABLE DE RESPUESTA		Porcentaje Correcto	
		NO APROBATORIO	APROBATORIO		
Step 1	VARIABLE DE RESPUESTA	NO APROBATORIO	21	110	16,0
		APROBATORIO	2	217	99,1
	Overall Percentage				68,0
Step 2	VARIABLE DE RESPUESTA	NO APROBATORIO	76	55	58,0
		APROBATORIO	47	172	78,5
	Overall Percentage				70,9
Step 3	VARIABLE DE RESPUESTA	NO APROBATORIO	60	71	45,8
		APROBATORIO	19	200	91,3
	Overall Percentage				74,3

Step 4	VARIABLE DE RESPUESTA	NO APROBATORIO	69	62	52,7
		O APROBATORIO	17	202	92,2
	Overall Percentage				77,4
Step 5	VARIABLE DE RESPUESTA	NO APROBATORIO	83	48	63,4
		O APROBATORIO	30	189	86,3
	Overall Percentage				77,7
Step 6	VARIABLE DE RESPUESTA	NO APROBATORIO	70	61	53,4
		O APROBATORIO	23	196	89,5
	Overall Percentage				76,0

a. El valor de corte es, 500

En la tabla de clasificación podemos observar que el modelo clasifica el 76% de los datos correctamente.

**Tabla 18: Variables En La Ecuación**

		Variables En la ecuación							
		B	E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1a	SE_ENCUNETRA_UD_CON_SU_CARRE			19,492	3	,000			
	SE_ENCUNETRA_UD_CON_SU_CARRE RA(1)	2,862	,902	10,061	1	,002	17,500	2,985	102,599

S te p 2 b	SE_ENCUNETRA_U D_CON_SU_CARRE RA(2)	2,679	,777	11,89 6	1	,001	14,565	3,17 9	66,7 35
	SE_ENCUNETRA_U D_CON_SU_CARRE RA(3)	3,161	,753	17,60 7	1	,000	23,589	5,38 9	103, 251
	Constant	-2,351	,740	10,09 6	1	,001	,095		
	SE_ENCUNETRA_U D_CON_SU_CARRE RA			21,36 4	3	,000			
	SE_ENCUNETRA_U D_CON_SU_CARRE RA(1)	2,630	,907	8,402	1	,004	13,869	2,34 3	82,0 92
	SE_ENCUNETRA_U D_CON_SU_CARRE RA(2)	3,054	,792	14,88 1	1	,000	21,205	4,49 3	100, 084
	SE_ENCUNETRA_U D_CON_SU_CARRE RA(3)	3,393	,763	19,75 7	1	,000	29,767	6,66 6	132, 919
	TENIA_CONOCIMIN ETOS_PREVIOS_DE _LA_CARRERA(1)	-1,379	,259	28,32 8	1	,000	,252	,152	,419
	Constant	-2,119	,746	8,069	1	,005	,120		
	EN_QUE_SE_DESE MPEÑA_PADRE			28,12 5	5	,000			
S te p 3 c	EN_QUE_SE_DESE MPEÑA_PADRE(1)	1,366	,400	11,65 6	1	,001	3,922	1,79 0	8,59 3
	EN_QUE_SE_DESE MPEÑA_PADRE(2)	1,234	,405	9,296	1	,002	3,437	1,55 4	7,59 9
	EN_QUE_SE_DESE MPEÑA_PADRE(3)	-,514	,669	,589	1	,443	,598	,161	2,22 1
	EN_QUE_SE_DESE MPEÑA_PADRE(4)	-,578	,550	1,105	1	,293	,561	,191	1,64 8
	EN_QUE_SE_DESE MPEÑA_PADRE(5)	1,409	,401	12,36 8	1	,000	4,091	1,86 6	8,97 1

	SE_ENCUNETRA_U D_CON_SU_CARRE RA			21,22 6	3	,000			
	SE_ENCUNETRA_U D_CON_SU_CARRE RA(1)	3,101	,959	10,44 2	1	,001	22,210	3,38 7	145, 631
	SE_ENCUNETRA_U D_CON_SU_CARRE RA(2)	3,071	,846	13,17 9	1	,000	21,559	4,10 8	113, 150
	SE_ENCUNETRA_U D_CON_SU_CARRE RA(3)	3,616	,820	19,46 6	1	,000	37,177	7,46 0	185, 286
	TENIA_CONOCIMIN ETOS_PREVIOS_DE _LA_CARRERA(1)	-1,711	,300	32,58 2	1	,000	,181	,100	,325
	Constant	-2,803	,813	11,87 2	1	,001	,061		
	EN_QUE_SE_DESE MPEÑA_PADRE			25,87 6	5	,000			
	EN_QUE_SE_DESE MPEÑA_PADRE(1)	1,460	,445	10,73 8	1	,001	4,305	1,79 8	10,3 06
	EN_QUE_SE_DESE MPEÑA_PADRE(2)	1,307	,439	8,859	1	,003	3,696	1,56 3	8,74 0
	EN_QUE_SE_DESE MPEÑA_PADRE(3)	- 20,19 3	8261,5 41	,000	1	,998	,000	,000	.
S te p 4 d	EN_QUE_SE_DESE MPEÑA_PADRE(4)	-,933	,610	2,341	1	,126	,393	,119	1,30 0
	EN_QUE_SE_DESE MPEÑA_PADRE(5)	1,572	,443	12,62 1	1	,000	4,818	2,02 4	11,4 69
	EN_QUE_SE_DESE MPEÑA_MADRE			10,32 1	4	,035			
	EN_QUE_SE_DESE MPEÑA_MADRE(1)	,448	,414	1,170	1	,279	1,566	,695	3,52 7
	EN_QUE_SE_DESE MPEÑA_MADRE(2)	1,621	,632	6,588	1	,010	5,060	1,46 7	17,4 51
	EN_QUE_SE_DESE MPEÑA_MADRE(3)	-,472	,512	,852	1	,356	,623	,229	1,70 0

EN_QUE_SE_DESE MPEÑA_MADRE(4)	41,60 5	12010, 574	,000	1	,997	117127 391420 059238 0,000	,000	.
SE_ENCUNETRA_U D_CON_SU_CARRE RA			23,87 8	3	,000			
SE_ENCUNETRA_U D_CON_SU_CARRE RA(1)	3,244	,973	11,10 9	1	,001	25,645	3,80 6	172, 812
SE_ENCUNETRA_U D_CON_SU_CARRE RA(2)	3,438	,870	15,62 0	1	,000	31,137	5,65 9	171, 324
SE_ENCUNETRA_U D_CON_SU_CARRE RA(3)	4,119	,867	22,56 8	1	,000	61,476	11,2 39	336, 271
TENIA_CONOCIMIN ETOS_PREVIOS_DE _LA_CARRERA(1)	-2,337	,362	41,68 5	1	,000	,097	,048	,196
Constant	-3,438	,949	13,13 1	1	,000	,032		
TIPO_COL			17,64 6	2	,000			
TIPO_COL(1)	3,271	,797	16,82 4	1	,000	26,337	5,51 8	125, 706
TIPO_COL(2)	,632	,493	1,641	1	,200	1,881	,715	4,94 4
S EN_QUE_SE_DESE te MPEÑA_PADRE			31,56 2	5	,000			
p EN_QUE_SE_DESE 5 MPEÑA_PADRE(1)	2,171	,508	18,29 0	1	,000	8,770	3,24 2	23,7 22
e EN_QUE_SE_DESE MPEÑA_PADRE(2)	1,731	,475	13,25 7	1	,000	5,645	2,22 4	14,3 32
EN_QUE_SE_DESE MPEÑA_PADRE(3)	- 19,91 7	8119,6 46	,000	1	,998	,000	,000	.
EN_QUE_SE_DESE MPEÑA_PADRE(4)	-,922	,655	1,982	1	,159	,398	,110	1,43 6



EN_QUE_SE_DESE MPEÑA_PADRE(5)	1,824	,483	14,26 4	1	,000	6,200	2,40 5	15,9 79
EN_QUE_SE_DESE MPEÑA_MADRE			4,097	4	,393			
EN_QUE_SE_DESE MPEÑA_MADRE(1)	,255	,424	,363	1	,547	1,291	,562	2,96 2
EN_QUE_SE_DESE MPEÑA_MADRE(2)	,265	,752	,124	1	,724	1,304	,299	5,68 8
EN_QUE_SE_DESE MPEÑA_MADRE(3)	-,757	,548	1,907	1	,167	,469	,160	1,37 4
EN_QUE_SE_DESE MPEÑA_MADRE(4)	41,73 4	11806, 743	,000	1	,997	133251 800622 182349 0,000	,000	.
SE_ENCUNETRA_U D_CON_SU_CARRE RA			25,06 3	3	,000			
SE_ENCUNETRA_U D_CON_SU_CARRE RA(1)	3,939	1,035	14,48 1	1	,000	51,350	6,75 4	390, 432
SE_ENCUNETRA_U D_CON_SU_CARRE RA(2)	3,670	,911	16,23 6	1	,000	39,245	6,58 5	233, 904
SE_ENCUNETRA_U D_CON_SU_CARRE RA(3)	4,432	,904	24,02 9	1	,000	84,092	14,2 95	494, 693
TENIA_CONOCIMIN ETOS_PREVIOS_DE _LA_CARRERA(1)	-2,819	,420	45,04 4	1	,000	,060	,026	,136
Constant	-4,340	1,128	14,79 2	1	,000	,013		
S te p 6 e			17,74 0	2	,000			
TIPO_COL(1)	2,854	,731	15,24 7	1	,000	17,352	4,14 3	72,6 81
TIPO_COL(2)	,416	,444	,876	1	,349	1,515	,635	3,61 9

EN_QUE_SE_DESE MPEÑA_PADRE			32,41 2	5	,000			
EN_QUE_SE_DESE MPEÑA_PADRE(1)	1,876	,435	18,63 5	1	,000	6,525	2,78 4	15,2 91
EN_QUE_SE_DESE MPEÑA_PADRE(2)	1,494	,438	11,63 7	1	,001	4,457	1,88 9	10,5 19
EN_QUE_SE_DESE MPEÑA_PADRE(3)	-1,180	,691	,068	1	,794	,835	,216	3,23 4
EN_QUE_SE_DESE MPEÑA_PADRE(4)	-627	,608	1,064	1	,302	,534	,162	1,75 8
EN_QUE_SE_DESE MPEÑA_PADRE(5)	1,629	,429	14,42 1	1	,000	5,101	2,20 0	11,8 27
SE_ENCUNETRA_U D_CON_SU_CARRE RA			22,82 0	3	,000			
SE_ENCUNETRA_U D_CON_SU_CARRE RA(1)	3,724	1,001	13,83 4	1	,000	41,444	5,82 3	294, 972
SE_ENCUNETRA_U D_CON_SU_CARRE RA(2)	3,261	,874	13,91 0	1	,000	26,071	4,69 8	144, 664
SE_ENCUNETRA_U D_CON_SU_CARRE RA(3)	3,868	,842	21,11 2	1	,000	47,854	9,19 0	249, 180
TENIA_CONOCIMIN ETOS_PREVIOS_DE _LA_CARRERA(1)	-1,997	,331	36,39 0	1	,000	,136	,071	,260
Constant	-3,717	,954	15,19 8	1	,000	,024		

a. Variable (s) introducida en el paso 1:

SE\_ENCUNETRA\_UD\_CON\_SU\_CARRERA.

b. Variable (s) introducida en el paso 2:

TENIA\_CONOCIMINETOS\_PREVIOS\_DE\_LA\_CARRERA.

El sexto paso es el que nos genera un modelo más ajustado a los datos.



## 15. CREACION DE UN MODELO DE PRONOSTICO

Los pronósticos son una de las herramientas fundamentales para la toma dentro de las organizaciones tanto productivas como sin fines de lucro. Algunas de las áreas en donde se utilizan pronósticos en la industria son la planeación y control de inventarios, producción, finanzas, ventas, comercialización, entre muchas otras.

### Objetivo de un Pronóstico

Reducir la incertidumbre acerca de lo que puede acontecer en el futuro proporcionando información cercana a la realidad que permita tomar decisiones sobre los cursos de acción a tomar tanto en el presente como en el futuro. Modelo Matemático de Pronóstico

Es una expresión matemática que representa en forma simplificada el fenómeno por medio del cual se obtienen los valores idealizados que toma una variable aleatoria en un periodo de tiempo determinado.

### Técnicas Estadísticas para la Obtención de Pronósticos

Estas técnicas se basan en la existencia de patrones, en el estudio de los mismos, las transformaciones que sufren, y la influencia del ruido o perturbación causado por factores de naturaleza aleatoria.

Dentro de estas técnicas se utilizan dos enfoques.

En el primero se obtiene el pronóstico basado en el razonamiento de que los datos de la serie de tiempo se pueden dividir o descomponer en componentes identificables que pueden presentarse o no en una determinada serie, estos componentes pueden ser la tendencia, la estacionalidad, la ciclisidad y la

aleatoriedad de los datos. El pronóstico se realiza combinando la proyección de los componentes que se presentan dentro de la serie de tiempo.

En el segundo el pronóstico se obtiene a partir del análisis estadístico de los datos que integran la serie de tiempo.

## 15.1 MODELO LOGÍSTICO

En este caso se supone que la variable dependiente  $Y$  tiene más de dos categorías y utiliza como distribución subyacente la distribución binomial.

Sea  $g$  el número de categorías de  $Y$ . El modelo logístico binomial supone que:

$$P[Y=i/x] = p_i = \frac{e^{x'\beta_i}}{1 + \sum_{j=1}^{g-1} e^{x'\beta_j}} \quad i=1, \dots, g-1$$

$$P[Y=g/x] = p_g = \frac{1}{1 + \sum_{j=1}^{g-1} e^{x'\beta_j}}$$

Observar que:

$$\frac{p_r}{p_s} = e^{x'(\beta_r - \beta_s)} \quad \text{si } r, s \neq g \text{ y } \frac{p_r}{p_g} = e^{x'\beta_r}$$

$$\pi(x) = \frac{e^{(\beta_0 + \beta_1 x)}}{e^{(\beta_0 + \beta_1 x)} + 1} = \frac{1}{e^{-(\beta_0 + \beta_1 x)} + 1},$$

Por lo que el ratio de dos probabilidades de cada categoría es independiente de los parámetros del resto de las categorías. Esta propiedad se conoce con el nombre de **independencia de alternativas irrelevantes**

## 16.RESULTADOS DE NUESTRO MODELO

$$\pi(x) = \frac{e^{(\beta_0 + \beta_1 x)}}{e^{(\beta_0 + \beta_1 x)} + 1} = \frac{1}{e^{-(\beta_0 + \beta_1 x)} + 1},$$

P step6=  $\frac{1}{\exp(-3,717 - 1,997 \text{conocimiento previo carrera} + 3,868 \text{se encuentra con su carrera} + 3,261 \text{ se encuentra con su carrera} + 3,724 \text{se encuentra con su carrera} + 1 + 1,629 \text{desempeño padre} + 5 - 0,627 \text{desempeño padre} + 0,180 \text{desempeño padre} + 3 + 1,494 \text{desempeño padre} + 2 + 1,876 \text{desempeño padre} + 1 + 0,416 \text{colegio} + 2 + 2,854 \text{colegio} + 1)}$

Un método matemático que valora el rendimiento académico y sus posibles causas. Es el proceso de estimación en situaciones de incertidumbre. Una solución para la problemática, generando soluciones desde un principio antes de que se genere la mortalidad académica estudiando a esos posibles casos de mortalidad.

Estas son las variables que actúan directamente en el modelo de pronóstico:

- Conocimiento previo carrera
- Se encuentra con su carrera 3.
- Se encuentra con su carrera 2.
- Se encuentra con su carrera 1.
- Desempeño padre 5.
- Desempeño padre 4.
- Desempeño padre 3.
- Desempeño padre 2.
- Desempeño padre 1.
- Colegio 2.
- Colegio 1.

Estas son las variables que nos ayudaran a pronosticar los casos que generaran mortalidad académica y crear nuevas ayudas atacando la problemática desde un comienzo antes de que se generen los índices de mortalidad.

## VALIDACION DEL MODELO

Para realizar la validación del modelo se volvió a correr en análisis de regresión logística binaria el 70% de los datos y los datos pertenecientes al 30% se les aplicó la fórmula del modelo para realizar el análisis y compara los datos generados por spss y los por el modelo.

Al seleccionar el 70 % de los datos estos fueron los resultados obtenidos

**Figura 50: pantallazo 4 seleccionar el 70 % de los datos estos fueron los resultados**

	RESULTADO_S_F	RESULTADO_S_L	VAR_SAL	PGR_1	PRE_1	var	var	var	var	var	var	var
2	58,00	55,00	APROBATORIO	APROBATORIO	,74719							
<del>3</del>	43,00	77,00	APROBATORIO		.							
4	43,00	63,00	NO APROBATORIO	NO APROBATORIO	,00000							
5	52,00	52,00	APROBATORIO	NO APROBATORIO	,43243							
6	47,00	47,00	APROBATORIO	APROBATORIO	,74719							
7	49,00	49,00	NO APROBATORIO	APROBATORIO	,74719							
8	54,00	43,00	NO APROBATORIO	NO APROBATORIO	,43243							
9	59,00	78,00	APROBATORIO	APROBATORIO	,74719							
10	56,00	69,00	NO APROBATORIO	NO APROBATORIO	,43243							
11	47,00	62,00	APROBATORIO	APROBATORIO	,74719							
12	52,00	51,00	APROBATORIO	APROBATORIO	,74719							
13	70,00	54,00	APROBATORIO	APROBATORIO	,74719							
14	46,00	55,00	NO APROBATORIO	NO APROBATORIO	,00000							
<del>15</del>	55,00	48,00	APROBATORIO		.							
16	47,00	42,00	APROBATORIO	APROBATORIO	,74719							
<del>17</del>	51,00	54,00	APROBATORIO		.							
18	52,00	62,00	APROBATORIO	APROBATORIO	,74719							
19	49,00	63,00	NO APROBATORIO	APROBATORIO	,74719							
<del>20</del>	47,00	62,00	APROBATORIO		.							
21	57,00	52,00	APROBATORIO	APROBATORIO	,74719							
<del>22</del>	50,00	41,00	NO APROBATORIO		.							
23	52,00	40,00	APROBATORIO	NO APROBATORIO	,43243							



**Figura 51: pantallazo 5 seleccionar el 30 % de los datos**

	RESULTADO_S_F	RESULTADO_S_L	VAR_SAL	PGR_1	PRE_1	var	var	var	var	var	var
111	59,00	78,00	APROBATORIO	APROBATORIO	,74719						
112	56,00	69,00	NO APROBATORIO		.						
113	47,00	62,00	APROBATORIO		.						
114	52,00	51,00	APROBATORIO	APROBATORIO	,74719						
115	70,00	54,00	APROBATORIO	APROBATORIO	,74719						
116	46,00	55,00	NO APROBATORIO		.						
117	55,00	48,00	APROBATORIO	APROBATORIO	,74719						
118	47,00	42,00	APROBATORIO	APROBATORIO	,74719						
119	51,00	54,00	APROBATORIO		.						
120	52,00	62,00	APROBATORIO	APROBATORIO	,74719						
121	49,00	63,00	NO APROBATORIO		.						
122	49,00	63,00	APROBATORIO	NO APROBATORIO	43243						
123	41,00	57,00	APROBATORIO		.						
124	56,00	46,00	NO APROBATORIO	NO APROBATORIO	43243						
125	51,00	40,00	APROBATORIO		.						
126	41,00	52,00	NO APROBATORIO	APROBATORIO	,74719						
127	42,00	59,00	NO APROBATORIO		.						
128	39,00	48,00	NO APROBATORIO		.						
129	49,00	69,00	APROBATORIO	NO APROBATORIO	43243						
130	49,00	63,00	APROBATORIO	NO APROBATORIO	43243						
131	41,00	57,00	APROBATORIO	APROBATORIO	,74719						
132	49,00	46,00	NO APROBATORIO	NO APROBATORIO	43243						

Con el 30% se probó el modelo inicialmente obtenido en los resultados.

P step6=

1

$$\exp(-3,717-1,997\text{conocimientopreviocarrera}+3,868\text{se encuentra con su carrera}+3,261\text{ se encuentra con su carrera}+3,724\text{se encuentra con su carrera}+1,629\text{desempeñopadre}+5,0,627\text{desempeñopadre}+0,180\text{desempeñopadre}+3+1,494\text{desempeñopadre}+2+1,876\text{desempeñopadre}+1+0,416\text{colegio}+2,854\text{colegio}+1)$$

- ✓ Caso 3= 0,68293. El resultado indica que la predicción es aprobado y en la variable original salida su resultado es aprobado.
- ✓ Caso 15 = 0.62745. El resultado indica que la predicción es aprobado y en la variable original salida su resultado es aprobado.
- ✓ Caso 36 = 0.00000. El resultado indica que la predicción es no aprobado y en la variable original salida su resultado es no aprobado.
- ✓ Caso 68 = 0.68293. El resultado indica que la predicción es aprobado y en la variable original salida su resultado es aprobado.

- ✓ Caso 75 = 0.72727. El resultado indica que la predicción es aprobado y en la variable original salida su resultado es aprobado.
- ✓ Caso 112 = 0.62745. El resultado indica que la predicción es aprobado pero en la variable original salida su resultado es no aprobado.
- ✓ Caso 116 = 0.00000. El resultado indica que la predicción es no aprobado y en la variable original salida su resultado es no aprobado.
- ✓ Caso 121 = 0.62745. El resultado indica que la predicción es aprobado pero en la variable original de salida su resultado es no aprobado.
- ✓ Caso 127 = 0.68293. El resultado indica que la predicción es aprobado y en la variable original salida su resultado es aprobado.
- ✓ Caso 141 = 0.62745. El resultado indica que la predicción es aprobado y en la variable original salida su resultado es aprobado.

Se determina que de los 10 casos, 8 de ellos fueron pronosticados correctamente de acuerdo a la variable de salida. Indicando que el modelo está ajustado a un 80% de los datos. Se puede observar los casos que no fueron pronosticados correctamente por el modelo en rojo.

## 17. CONCLUSIONES

- Según la información suministrada por el grupo de investigación y los datos de registro y control, vemos que en la actualidad se sigue viendo dificultades con las mismas materias y se presentan con gran frecuencia en las ingenierías (Física, Mecánica, eléctrica y sistemas) dichas materias son: matemáticas I y II , física III algebra lineal, este dato lo vemos reflejado en el 37% en el grafico categoría de respuesta que lo componen los estudiantes que se encuentran en periodo de prueba y transición, solo se pudo estudiar matemáticas I y II, por que estas se encuentran en la base de datos que nos fue suministrada por la Universidad Tecnológica de Pereira división de sistemas, de las otras materias no se nos suministró ningún tipo de información.
- La mortalidad académica puede presentarse con mucha frecuencia en ciertas materias y programas, esto se puede presentar por en el entorno diario de los estudiantes en la vida universitaria, socioeconómica, familiar y dificultad con algunos docentes, pero la que se caracteriza con mayor influencia es la parte económica donde los resultados nos muestran un 80% de la población encuestada manifestó una dificultad para pagar la universidad.
- Al realizar en el análisis descriptivo de los resultados, se logra evidenciar algunos aspectos sobresalientes que son : opinión de los docentes, el examen de orientación profesional, categoría de respuestas (información que tiene que ver con el rendimiento académico de la población estudiada) y la dificultad para pagar el semestre, siendo esta ultima un tema que es difícil de manejar. También se evidencia que muchos estudiantes ingresan por orientación familiar (obligados por los padres o por la economía que presenta) a estudiar carreras que no son de su lección, vocación o de sus habilidades.
- Partiendo de los antecedentes históricos se llega a la conclusión que la mortalidad académica no solo depende del tipo de carrera que se esté haciendo, sino que tiene algunos aspectos ajenos a la universidad que se debe tener en cuenta como es su entorno familiar y económico, se nota que poseen más dificultades aquellas familias que solo cuentan con madres cabezas de hogar en su composición.

- Se evidencia que la mortalidad académica es una de las principales causas que generan la deserción estudiantil en las Universidades, por esto se debe encontrar soluciones para controlar la mortalidad académica creando modelos de predicción a futuro y atendiendo estos casos desde su inicio.
- Es importante promover más el examen de orientación profesional, para que a futuro no se presente deserción estudiantil y a futuro una mortalidad académica, pues en ocasiones los estudiantes se encuentran que no era lo que ellos esperaban, se evidencia que un 12% de la población que presento el examen les ayudo a seleccionar o corroborar la carrera que querían estudiar, y un dato preocupante de un 71% de personas que no lo presentaron o no tenían conocimiento de este examen.
- Debido a los altos índices de mortalidad académica que generan más deserción se recomienda estar estudiando cada semestre su comportamiento para generar soluciones y disminuir dichos índices y mitigar el problema.
- En comparación con los estudios anteriores tenidos en cuenta para realizar este estudio y con los resultados obtenidos se evidencia que los índices de mortalidad académica han disminuido pero sigue siendo altos, se debe generar medidas de solución y prevención a esta problemática que ayuda a aumentar la deserción.
- La Universidad debe enfocarse en trabajar en las materias críticas que generan mortalidad planteando soluciones correctivas y preventivas.
- Es necesario para la realización de cualquier investigación de tipo descriptivo con análisis de datos multivariado, conocer las técnicas que se aplican para dichos análisis y saber cuál de ellas utilizar dadas ciertas circunstancias. De esta manera concluimos que para una variable multicategorica de salida y con variables independientes tanto cuantitativas como cualitativas, es la técnica de análisis de regresión logística multinomial una técnica ideal para llevar a cabo el estudio.
- Se evidencia que el nivel educativo de los padres es una de las variables que tiene peso para generar mortalidad académica.

## 18.RECOMENDACIONES:

- La información es la parte más importante de una investigación, si la base de datos no se encuentra en un buen estado, no posee toda la información solicitada, no se le da un buen uso y manejo, no se generan buenas depuraciones y filtros, dejara una matriz de números o información sin sentido. Es por eso que se solicita un mejor apoyo a este tipo de estudios por la Universidad Tecnológica de Pereira en acompañamiento y en la información que brindan a los estudiantes cumpliendo todos los requisitos solicitados en las bases de datos para realizar este estudio, así se generan resultados más confiables y de alta calidad. Realizando un correcto trabajo de campo e investigación para la recolección de la información necesaria. Ya que el trabajo es para mejoras a futuro de la misma Universidad.
- Es necesario la constante actualización de las bases de datos de los estudiantes por parte de la Universidad, esto para tener información veras que permita el desarrollo coherente de los estudios que se realicen.
- Se recomienda realizar periódicamente por cada semestre investigaciones de mortalidad y deserción no solo realizadas por el observatorio si no por estudiantes que pueden generar nuevos puntos de vista y posibles soluciones para esta problemática.
- Es necesario monitorias las materias críticas que generan mortalidad para analizar sus comportamientos y generar soluciones.
- Se recomienda la interacción de la Universidad con los estudiantes para informales las herramientas y programas que tiene la Universidad para mejorar su rendimiento académico y que no poseen ningún costo.
- Es necesario el acompañamiento total y pleno de los docentes con el estudiante en las materias críticas brindándoles más herramientas o diferentes metodologías para disminuir la problemática que generan dichas materias.
- Se recomienda un seguimiento a los docentes, sus conocimientos y metodologías para evitar que esta sea una causa de mortalidad, teniendo un control sobre el comportamiento docente y materia críticas.
- Se recomienda que el introductorio a matemáticas sea evaluado con notas reales, se pueda perder y se tenga que repetir para combatir las falencias

que traen los estudiantes desde el colegio y así se puede disminuir el índice de mortalidad en estas materia crítica.

- Se necesita un análisis y control total de los resultados icfes en matemáticas y física para ingresar a las ingenierías así se segmenta la información desde que entra el estudiante y se pueden analizar sus falencias en estas materias críticas que generan mortalidad.

## 19. BIBLIOGRAFIA Y WEBGRAFIA

[1] Estudio sobre La deserción estudiantil Universidad Nacional de Medellín

[2] Estudio deserción estudiantil Universidad de los Llanos

En la página [www.google.com](http://www.google.com) “Estudios sobre deserción en las Universidades”

[http://web.unillanos.edu.co/centro-de-documentacion/doc\\_download/33-estudio-de-la-desercion-estudiantil-de-los-programas](http://web.unillanos.edu.co/centro-de-documentacion/doc_download/33-estudio-de-la-desercion-estudiantil-de-los-programas)

[3] Supervivencia, graduación, deserción y rezago en La Universidad Nacional de Colombia

En la página [www.google.com](http://www.google.com) “El rezago en las Universidades”

<http://www.google.com.cosearchhl=es&q=+el+rezago+en+las+universidades&bt>

[4] Propuesta para disminuir la deserción el rezago y la repitencia en I Universidad Tecnológica de Pereira

En la página [www.google.com](http://www.google.com) “Rezago en las Universidades”

<http://www.utp.edu.co/php/revistas/ScientiaEtTechnica/docsFTP/163153145-150.pdf>

[5] Estudio sobre la repitencia y la deserción en la educación superior Chilena

En la página [www.google.com](http://www.google.com) “Variables que inciden en la repitencia en las Universidades”

[http://www.cedus.cl/files/REPITENCIA\\_DESERCION.pdf](http://www.cedus.cl/files/REPITENCIA_DESERCION.pdf)

[6] Estudio sobre deserción Universidad Pedagógica Nacional

En la página [www.google.com](http://www.google.com) “Estudios sobre deserción en universidades”  
Universidad Pedagógica Nacional

[www.mineducacion.gov.co/1621/articles-85600\\_Archivo\\_pdf3.pdf](http://www.mineducacion.gov.co/1621/articles-85600_Archivo_pdf3.pdf)

[7] <http://www.estadistica.mat.uson.mx/Material/elmuestreo.pdf>

[8] PEREZ L Cesar, Técnicas de análisis de datos con SPSS 15. Pearson Educación, S.A. España. Pág. 488. 2009

- [9] PEREZ L Cesar, Técnicas de análisis de datos con SPSS 15. Pearson Educación, S.A. España. Pág. 127. 2009
- [10] [http://es.wikipedia.org/wiki/Muestreo\\_en\\_estad%C3%ADstica](http://es.wikipedia.org/wiki/Muestreo_en_estad%C3%ADstica)
- [11] <http://www.seh-lelha.org/rlogis1.htm>
- [12] <http://www.seh-lelha.org/regresion1.htm>
- [13] [www.paginasprodigy.com/sylsr/ingenierias/pronosticos/conceptos%20b%C3%A1sicos%20de%20pron%C3%B3sticos.html](http://www.paginasprodigy.com/sylsr/ingenierias/pronosticos/conceptos%20b%C3%A1sicos%20de%20pron%C3%B3sticos.html)
- [14] [http://es.wikipedia.org/wiki/Regresi%C3%B3n\\_log%C3%ADstica](http://es.wikipedia.org/wiki/Regresi%C3%B3n_log%C3%ADstica)