



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Instituto de Ciencias Matemáticas

“Uso de Regresión Multinomial Logística para determinar el perfil del estudiante de la ESPOL, según preferencias alimenticias.”

TESIS DE GRADO

Previo a la obtención del Título de:

INGENIERO EN ESTADÍSTICA E INFORMÁTICA

Presentada por:

Betzabeth Reina Yonson Sangurima

Guayaquil – Ecuador

AÑO 2005

AGRADECIMIENTO

Al Dios todo poderoso quien me ha acompañado y ayudado en todo momento, quien me ha levantado en mis caídas y guiado hasta el momento, a mi padre por apoyarme y confiar en mi, a mi madre por estar conmigo siendo mi amiga , a ellos por ser motivo de mi inspiración , a mi enamorado Dalton Martillo por apoyarme moralmente aunque no este presente he sentido su amor y paciencia incondicional, a mí Director de Tesis por su colaboración y a todas las demás personas que de una u otra manera me apoyaron para la realización de esta Tesis.

DEDICATORIA

A mis padres: Fernando Yonson y Edita Sangurima por ser las personas más importantes en mi vida, por ser el pilar fundamental en mi educación, a ellos por el infinito amor que les tengo al igual que mis hermanos y mi familia por creer y confiar en mi, a ellos y a mi Dios todopoderoso que sin ellos no hubiera podido realizar mis metas.

TRIBUNAL DE GRADUACIÓN

MAT. JORGE MEDINA
PRESIDENTE DEL TRIBUNAL

MAT. FERNANDO SANDOYA
DIRECTOR DE TESIS

MAT. RIVADENEIRA
VOCAL

ING. PABLO ALVAREZ
VOCAL

DECLARACIÓN EXPRESA

“La responsabilidad del contenido de esta Tesis de Grado, me corresponde exclusivamente; y el patrimonio intelectual de la misma a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL”

(Reglamento de Graduación de la ESPOL)

Betzabeth Reina Yonson Sangurima

RESUMEN

El presente estudio trata sobre el “Uso de Regresión Multinomial Logística para determinar el perfil del estudiante politécnico según preferencias alimenticias” enfocados a los estudiantes registrados en el I termino del presente año, con ello se permitirá obtener una visión clara de cómo los estudiantes están alimentándose y si ello influye en el rendimiento académico.

Este estudio se desarrolla en seis capítulos, donde se presentan los resultados de esta investigación. En la primera parte se menciona una serie de conceptos que nos facilitan una mejor comprensión sobre el tema de la alimentación, en relación a la educación.

En la segunda parte se profundizan los conceptos de regresión logística y sus aplicaciones e interpretaciones.

En la tercera parte se revisan los fundamentos teóricos para el análisis de las variables, se desarrolla la descripción de cómo se realizó el diseño muestral, la determinación del tamaño de muestra y el plan de muestreo y la descripción de las variables seleccionadas en el estudio.

En el cuarto capítulo se presenta el análisis univariado de todas las variables que intervienen en este estudio.

En el quinto capítulo se realizarán inferencias sobre la independencia de las variables (tablas de contingencia) y el uso de regresión multinomial logística para el total de variables en el estudio.

En el sexto capítulo se ponen a consideración las conclusiones y recomendaciones a las que nos llevó los capítulos anteriores mencionados.

BIBLIOGRAFÍA

1. William Mendenhall, 1994, Estadística Matemática con Aplicaciones, México, Grupo Editorial Iberoamérica
2. César Pérez, 2000, Técnicas de Muestreo, México, RA-MA Editorial.
3. Luis Carlos Silva, Isabel María Barroso Utra, 2004, Regresión Logística, Madrid, La Muralla S.A., Editorial.
4. César Pérez, 2001, Técnicas Estadísticas con SPSS , Madrid, Prentice Hall.
5. Kerlly Bermúdez , El perfil del estudiante de la ESPOL al iniciarse el siglo 21: Un análisis estadístico para las carreras tradicionales, Tesis de Grado, Instituto de Ciencias Matemáticas, Ingeniería en Estadística Informática, Escuela Superior Politécnica del Litoral, 2002.
6. <http://www.fisterra.com/mbe/investiga.htm>
7. <http://ciberconta.unizar.es/LECCION/logis/INICIO.HTML>
8. <http://Tuotromedico Control del colesterol.htm>
9. <http://www.texasheartinstitute.org/cholspan.html>

ÍNDICE GENERAL

	Pág.
RESUMEN.....	II
ÍNDICE GENERAL.....	III
ABREVIATURAS.....	IV
SIMBOLOGÍA.....	V
ÍNDICE DE CUADROS.....	VI
ÍNDICE DE GRÁFICOS.....	VII
ÍNDICE DE TABLAS.....	VIII
INTRODUCCIÓN.....	1

CAPÍTULO I

1. PLANTEAMIENTO DEL PROBLEMA DE ESTUDIO

1.1 Introducción.....	3
1.2 Definición de Alimentación	3
1.3 Alimentación Saludable	4
1.3.1. Nutrientes.....	4
1.3.1.1. Funciones de los Nutrientes.....	5
1.3.1.2. Vitaminas	6

1.3.1.3. Minerales	8
1.4 Mala Alimentación	9
1.4.1. Causas de Mala Alimentación.....	9
1.4.1.1 Gastritis y Úlcera.....	10
1.4.1.2 Colesterol.....	11
1.4.1.3 Diabetes.....	12
1.4.1.4 Cálculos.....	13
1.4.1.5 Problemas de garganta.....	14

CAPÍTULO II

2. REGRESIÓN LOGÍSTICA

2.1 Introducción.....	15
2.2 Tipos de Variables	17
2.3 Modelo Logit	18
2.4 Modelo Probit	19
2.5 Odds Ratio	19
2.6 Función de Verosimilitud.....	21
2.7 Modelo de Regresión Logística Binaria Simple.....	22

2.8 Modelo de Regresión Logística Múltiple.....	22
2.8.1 Ajuste del Modelo de Regresión Logística Múltiple.....	23
2.8.2 Método de Newton – Raphson	25
2.8.3 Pruebas de Hipótesis Sobre los Parámetros	27
2.8.4 Pruebas de Significancia para Regresión Logística Múltiple	28
2.8.5 La Prueba de Wald	29
2.8.6 Prueba de Hosmer-Lemeshow	30
2.9 Regresión Logística Multinomial.....	30
2.9.1 Regresión Logística con respuesta Nominal Politómica.....	31

CAPÍTULO III

3. DISEÑO MUESTRAL

3.1 Introducción.....	33
3.2 Población Objetivo	34
3.3 Marco Muestral.....	34
3.4 Diseño del Cuestionario.....	34
3.5 Muestra.....	35
3.6 Tipos de Muestreo.....	35
3.6.1. Muestreo Probabilísticos.....	35
3.6.1.1 Muestreo Aleatorio Simple.....	36
3.6.1.2 Muestreo Estratificado.....	36
3.7 Tamaño de la Muestra.....	38
3.7.1. Determinación del tamaño de la Muestra	38
3.8 Técnicas de Muestreo	40
3.9 Descripción de las variables de nuestro estudio.....	42

CAPÍTULO IV

4. ANÁLISIS ESTADÍSTICO DE LOS DATOS OBSERVADOS

4.1 Análisis Descriptivo.....	57
-------------------------------	----

CAPÍTULO V

5. ANÁLISIS MULTIVARIADO

5.1	
Introducción.....	100
5.2 Tabla de Contingencia.....	101
5.2.1. Análisis de Contingencia Para Las Variables	103
5.3 Análisis de Regresión Multinomial Logística.....	118

CAPÍTULO VI

5. CONCLUSIONES Y RECOMENDACIONES

5.1 Conclusiones	123
5.2 Recomendaciones.....	125

ANEXOS

BIBLIOGRAFÍA

ABREVIATURAS

CRECE Centro de Registros, Calificaciones y Estadísticas

RL Regresión Logística

Mg Miligramos

Cmts Centímetros

SIMBOLOGÍA

N	Tamaño de la Población
n	Tamaño Muestral
χ^2	Estadístico Chi Cuadrado
gl.	Grados de Libertad
P	Pearson
H ₀	Hipótesis Nula
H ₁	Hipótesis Alternativa

ÍNDICE DE GRÁFICOS

	Pág.
Gráfico 4.1 Histograma de frecuencia de género.....	58
Gráfico 4.2 Distribución de la variable edad.....	60
Gráfico 4.3 Diagrama de cajas de la variable edad.....	60
Gráfico 4.4 Distribución de la variable Factor P.....	62
Gráfico 4.5 Diagrama de cajas de la variable Factor P.....	62
Gráfico 4.6 Distribución de la variable promedio.....	64
Gráfico 4.7 Diagrama de cajas de la variable promedio.....	64
Gráfico 4.8 Distribución de Rendimiento Académico.....	66
Gráfico 4.9 Diagrama de cajas de Rendimiento Académico.....	66
Gráfico 4.10 Distribución de la variable Peso.....	68
Gráfico 4.11 Diagrama de cajas de peso.....	68
Gráfico 4.12 Distribución de la variable estatura.....	70
Gráfico 4.13 Diagrama de cajas de estatura.....	70
Gráfico 4.14 Histograma de frecuencia de la variable trabaja	71
Gráfico 4.15 Histograma de frecuencia de la variable horas que permanece en la ESPOL.....	72
Gráfico 4.16 Histograma de frecuencia de días promedio a la semana que Realiza alguna actividad deportiva	73

Gráfico 4.17	Histograma de frecuencia de cómo considera la Alimentación para desempeñar su rol de estudiantes.....	75
Gráfico 4.18	Distribución de la variable porcentaje de alimentación es preparada en su casa	76
Gráfico 4.19	Diagrama de cajas del porcentaje de alimentación Es hecha en casa.....	77
Gráfico 4.20	Histograma de frecuencia de días a la semana que desayuna.....	78
Gráfico 4.21	Histograma de frecuencia de días a la semana que come carne roja.....	79
Gráfico 4.22	Histograma de frecuencia de días a la semana que come pescado.....	80
Gráfico 4.23	Histograma de frecuencia de días a la semana que come pollo.....	81
Gráfico 4.24	Histograma de frecuencia de días a la semana que toma leche.....	82
Gráfico 4.25	Histograma de frecuencia de días a la semana que come vegetales.....	83
Gráfico 4.26	Histograma de frecuencia de días a la semana que come frutas.....	84
Gráfico 4.27	Histograma de frecuencia de comidas que realiza en el día	85

Gráfico 4.28	Histograma de frecuencia para la variable Comida chatarra.....	86
Gráfico 4.29	Histograma de frecuencia de horario fijo para la alimentación diaria del estudiante.....	87
Gráfico 4.30	Histograma de frecuencia de bebidas para Satisfacer su sed.....	88
Gráfico 4.31	Histograma de frecuencia que fuma.....	89
Gráfico 4.32	Histograma de frecuencia con que ingiere bebida Alcohólica.....	90
Gráfico 4.33	Histograma de frecuencia de la variable gastritis.....	91
Gráfico 4.34	Histograma de frecuencia de la variable úlcera.....	92
Gráfico 4.35	Histograma de frecuencia de problema de garganta.....	93
Gráfico 4.36	Histograma de frecuencia de colesterol.....	94
Gráfico 4.37	Histograma de frecuencia de diabetes.....	95
Gráfico 4.38	Histograma de frecuencia de cálculo.....	96
Gráfico 4.39	Histograma de frecuencia de otras enfermedades.....	97
Gráfico 4.40	Histograma de frecuencia de ausencia de clases Por problemas alimenticios.....	98
Gráfico 4.41	Histograma de frecuencia de estrés.....	99

ÍNDICE DE TABLAS

	Pág.
Tabla I	Distribución de frecuencias de género.....58
Tabla II	Estadística descriptiva de edad.....59
Tabla III	Estadística descriptiva de facto p.....61
Tabla IV	Estadística descriptiva de Promedio.....63
Tabla V	Estadística descriptiva de Rendimiento académico65
Tabla VI	Estadística descriptiva de Peso.....67
Tabla VII	Estadística descriptiva de estatura.....68
Tabla VIII	Distribución de frecuencia de trabaja.....71
Tabla IX	Distribución de frecuencia de horas que permanece en la ESPOL72
Tabla X	Distribución de frecuencia de días promedio a la Semana que realiza deporte.....73
Tabla XI	Distribución de frecuencia de de cómo considera la Alimentación para desempeñar su rol de estudiante74
Tabla XII	Estadística descriptiva de porcentaje de alimentación es preparada en su casa76
Tabla XIII	Distribución de frecuencia de días a la semana Que el estudiante desayuna.....77
Tabla XIV	Distribución de frecuencia de días a la semana Que come carne roja.....78

Tabla XV	Distribución de frecuencia de días a la semana Que come pescado.....	79
Tabla XVI	Distribución de frecuencia de días a la semana que come pollo.....	80
Tabla XVII	Distribución de frecuencia de días a la semana Que toma leche.....	81
Tabla XVIII	Distribución de frecuencia de días a la semana que come vegetales.....	82
Tabla XIX	Distribución de frecuencia de días a la semana que come frutas.....	83
Tabla XX	Distribución de frecuencia de comidas que realiza en el día	83
Tabla XXI	Distribución de frecuencia para la variable Comida chatarra.....	85
Tabla XXII	Distribución de frecuencia de horario fijo para la alimentación diaria del estudiante.....	86
Tabla XXIII	Distribución de frecuencia de bebidas para Satisfacer su sed.....	88
Tabla XXIV	Distribución de frecuencia que fuma.....	89
Tabla XXV	Distribución de frecuencia con que ingiere bebida Alcohólica.....	90
Tabla XXVI	Distribución de frecuencia de la variable gastritis.....	91

Tabla XXVII	Distribución de frecuencia de la variable úlcera.....	92
Tabla XXVIII	Distribución de frecuencia de problema de garganta.....	93
Tabla XXIX	Distribución de frecuencia de colesterol.....	94
Tabla XXX	Distribución de frecuencia de diabetes.....	95
Tabla XXXI	Distribución de frecuencia de cálculo.....	96
Tabla XXXII	Distribución de frecuencia de otras enfermedades.....	97
Tabla XXXIII	Distribución de frecuencia de ausencia de clases Por problemas alimenticios.....	98
Tabla XXXIV	Distribución de frecuencia de estrés.....	99
Tabla XXXV	Tabla de Contingencia: Rendimiento Académico vs. Edad	103
Tabla XXXVI	Tabla de Contingencia: Rendimiento Académico vs. Género	104
Tabla XXXVII	Tabla de Contingencia: Días a la semana que desayuna vs. Género	105
Tabla XXXVIII	Tabla de Contingencia: Peso vs. Género.....	106
Tabla XXXIX	Tabla de Contingencia: Peso vs. Colesterol en niveles anormales.....	107
Tabla XL	Tabla de Contingencia: Peso vs. Estatura.....	108
Tabla XLI	Tabla de Contingencia: Rendimiento académico vs. Promedio.....	109
Tabla XLII	Tabla de Contingencia: Porcentaje de su alimentación	

	es a través de comida preparada en casa vs. Género110
Tabla XLIII	Tabla de Contingencia: Rendimiento académico vs. Estrés.....112
Tabla XLIV	Tabla de Contingencia: Rendimiento académico vs. Trabaja.....113
Tabla XLV	Tabla de Contingencia: Rendimiento académico vs. Comida chatarra115
Tabla XLVI	Tabla de Contingencia: Realiza Actividad Deportiva vs. Estrés.....117
Tabla XLVII	Variables utilizadas en el análisis para RL.....119
Tabla XLVIII	Modelo ajustado.....120
Tabla XLIX	Bondad de ajuste.....121
Tabla L	Modelo reajustado121
Tabla LI	Bondad de ajuste.....122

ÍNDICE DE CUADROS

	Pág.
Cuadro 1.1	Vitaminas Esenciales.....7
Cuadro 1.2	Minerales Esenciales.....8
Cuadro 3.1	Distribución de la Población.....40
Cuadro 3.2	Distribución de la Población, Subestratos a cada Unidad Académica41
Cuadro 3.3	Distribución de la Muestra, Subestratos a cada Unidad Académica..... 42
Cuadro 3.4	Variable X_1 : Género.....43
Cuadro 3.5	Variable X_3 : Carrera.....44
Cuadro 3.6	Variable X_9 : Actualmente Trabaja.....46
Cuadro 3.7	Variable X_{10} : Horas que permanece en la ESPOL.....46
Cuadro 3.8	Variable X_{11} :Días promedio que realiza actividad deportiva.....46
Cuadro 3.9	Variable X_{12} : Desempeño para su rol de estudiante.....47
Cuadro 3.10	Variable X_{14} : Días a la semana que desayuna.....47
Cuadro 3.11	Variable X_{15} : Días a la semana que come carne roja.....48
Cuadro 3.12	Variable X_{16} :Días a la semana que come pescado.....48
Cuadro 3.13	Variable X_{17} : Días a la semana que come pollo.....49
Cuadro 3.14	Variable X_{18} :Días a la semana que toma leche.....49
Cuadro 3.15	Variable X_{19} : Días a la semana que consume vegetales.....49

Cuadro 3.16	Variable X ₂₀ :Días a la semana que come frutas.....	50
Cuadro 3.17	Variable X ₂₁ : Comidas que realiza en el día.....	50
Cuadro 3.18	Variable X ₂₂ : En la hora de alimentación que tan frecuente opta por comida rápida.....	51
Cuadro 3.19	Variable X ₂₃ : Horario fijo para su alimentación diaria.....	51
Cuadro 3.20	Variable X ₂₄ : Tipo de bebida de preferencia para satisfacer su sed.....	51
Cuadro 3.21	Variable X ₂₅ : Frecuencia que fuma.....	52
Cuadro 3.22	Variable X ₂₆ : Frecuencia con que ingiere bebida Alcohólica.....	52
Cuadro 3.23	Variable X ₂₇ : Problema de Gastritis.....	53
Cuadro 3.24	Variable X ₂₈ : Problema de Úlcera.....	53
Cuadro 3.25	Variable X ₂₉ : Problemas de garganta	53
Cuadro 3.26	Variable X ₃₀ : Problema de Colesterol en niveles Anormales.....	54
Cuadro 3.27	Variable X ₃₁ : Problema de Diabetes.....	54
Cuadro 3.28	Variable X ₃₂ : Problema de Cálculos.....	54
Cuadro 3.29	Variable X ₃₃ : Problema Otros.....	55
Cuadro 3.30	Variable X ₃₄ : Ausencia de clases por problemas alimenticios.....	55
Cuadro 3.31	Variable X ₃₅ : Frecuencia con la que ha pasado por el llamado estrés.....	56

INTRODUCCIÓN

Sabemos que la salud es uno de los factores más importantes que influyen en el desarrollo intelectual y rendimiento académico de los estudiantes, y que a su vez para lograr una buena salud es necesario tener una alimentación sana y equilibrada, sin embargo muchas veces por falta de tiempo o presupuesto, o simplemente por desconocimiento de los efectos nocivos de tener una mala alimentación, encontramos estudiantes que no desayunan, otros abusan de bebidas gaseosas o comida chatarra, de las cuales se conoce que tienen un efecto farmacológico, no aportan ningún nutriente y dificultan las funciones gástricas o digestivas; en otros casos por razones de tiempo no se respeta el horario de las tres comidas diarias, etc, así como estos, existen otros factores ligados a la alimentación que inciden en el desempeño académico, problemas por los que generalmente pasan la mayoría de los estudiantes.

Es por esta razón, que en el presente estudio se busca evaluar el perfil de los estudiantes de la ESPOL según preferencias alimenticias para lo cual se usarán principalmente técnicas de regresión logística.

El objetivo primordial de la regresión logística es 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).

La técnica de regresión logística resulta especialmente útil para identificar factores de riesgo y factores de prevención de enfermedades en muestras prospectivas en las que la metodología de la regresión lineal (variables continuas) no es aplicable, dado que la variable de respuesta sólo presenta dos valores (caso dicotómico) como puede ser presencia/ausencia de un suceso.

Así la regresión logística resulta de gran utilidad para su aplicación en cualquier campo de la investigación médica, biológico o de comportamiento humano cuando necesitamos precisar el efecto de un grupo de variables, consideradas potencialmente influyentes, sobre la ocurrencia de un determinado proceso.

CAPÍTULO I

1. PLANTEAMIENTO DEL PROBLEMA DE ESTUDIO

1.1. Introducción

El propósito de este capítulo es dar a conocer una serie de conceptos que nos facilitan una mejor comprensión sobre el tema de la alimentación, y conceptos básicos en relación a la educación.

Este trabajo tratará en lo posible de describir mejor las preferencias alimenticias de los estudiantes de la ESPOL y para entender cuales son los problemas de esta índole que enfrentan los estudiantes.

1.2. Definición de alimentación.

La alimentación consiste en proporcionar al cuerpo humano alimentos, los cuales tienen valores nutricionales (nutrientes) que necesita, no solo para estar en forma sino, ante todo, para dotarle de energía al cuerpo de manera eficiente para subsistir. En general consideramos como alimento a todo lo que comemos y bebemos habitualmente.

1.3 Alimentación Saludable

Para que un adulto sano obtenga todos los nutrientes y la energía que necesita se debe comer todos los días las raciones debidas de cada grupo de alimentos.

Hay que comer una diversidad de alimentos que proporcionan las cantidades indispensables de nutrientes y de energía y que a la vez permita mantener el peso corporal dentro de márgenes saludables. Se aconseja comer alimentos ricos en almidón y fibra, evitar el exceso de grasa, colesterol, azúcar y sodio, ingerir con mucha moderación las bebidas alcohólicas o suprimirlas por completo.

1.3.1. Nutrientes

Los nutrientes son sustancias presentes en los alimentos que nos proporcionan la energía que necesitamos para nuestras actividades diarias, ayudan a nuestro cuerpo a crecer y a mantenerse sano y también sirven para fortalecer las defensas contra las enfermedades.

Las tres principales clases de nutrientes son las proteínas, las grasas y los carbohidratos.

PROTEÍNAS.- Las proteínas aportan los "ladrillos" para formar y reponer tejidos. Se encuentran en las carnes rojas, leche y derivados, pescados, pollo, huevo y vegetales.

LOS CARBOHIDRATOS.- proporcionan energía al organismo, lo ayudan a regular la desintegración de las proteínas y lo protegen de las toxinas. Se hallan en las frutas, verduras y gramíneas y son muy nutritivos.

LAS GRASAS.- Las grasas son parte indispensable de la alimentación, pero no todas son saludables si se ingieren en cantidad excesiva.

La principal característica de las grasas es su grado de saturación, que se refiere a su estructura molecular. Las grasas insaturadas no propician tanta acumulación de colesterol en la sangre como las saturadas, como el exceso de colesterol en la sangre puede causar trastornos cardiacos, lo mas aconsejable es comer pocas grasas saturadas.

1.3.1.1. Funciones de los Nutrientes

Se ha establecido la siguiente clasificación de los alimentos según su función nutricional:

- *Nutrientes energéticos*: Ricos en carbohidratos o grasas. Cuanto mayor sea la actividad física realizada, mayor será la demanda de calorías a consumir y, por lo tanto, si no hay calorías, no hay energía. La cantidad de calorías que requiere el organismo depende del sexo, talla, edad y actividad individual.
- *Nutrientes constructores*: En los cuales predominan las proteínas y el calcio.
- *Nutrientes reguladores*: El organismo necesita nutrientes que regulen su funcionamiento, como las vitaminas y los minerales.

1.3.1.2. Vitaminas

Las vitaminas son sustancias que el organismo necesita para asimilar otros nutrientes, regular el sistema nervioso y ayudar a la formación de las proteínas, los glóbulos rojos y las hormonas. La carencia de vitaminas se puede producir por alimentación insuficiente o poco variada. El cuadro 1.1 muestra las vitaminas esenciales para el organismo:

Cuadro 1.1
Vitaminas Esenciales

Vitamina	Fuentes Mas Abundantes	Función
Vitamina A	Hígado, leche, huevos, mantequilla, zanahorias, espinacas, acelgas, mango y mamey.	Necesaria para las membranas del organismo, necesaria también para los dientes y los huesos.
Vitamina B1	Carne de chanco, harinas y cereales, nueces; cacahuates, chícharos, frijoles y ajonjolí.	Permite el eficiente aprovechamiento de los carbohidratos.
Vitamina B2	Leche, queso, huevos, hígado y carne de aves.	Necesaria para que las células liberen energía y se regeneren
Vitamina B6	Hígado, cereales enteros, leche y huevos.	Necesaria para el sistema nervioso y los glóbulos rojos
Vitamina B12	Huevos, carne y lácteos.	Necesaria para que la médula ósea produzca glóbulos rojos para el sistema nervioso
Ácido Fólico	Verduras frescas, carnes de aves y pescado.	Necesario para la producción de glóbulos rojos
Vitamina C	Cítricos, pimiento, papas, fresas, guayabas y zapote.	Necesaria para regenerar los huesos, dientes y tejidos.
Vitamina D	Pescados, aceite de hígado de pescado, lácteos y huevos.	Se necesita para mantener el nivel de calcio en la sangre y para el crecimiento óseo.
Vitamina E	Aceites vegetales y muchos otros comestibles	Necesaria para las membranas celulares.
Vitamina K	Verduras de hojas grandes y de color verde oscuro.	Necesaria para la coagulación normal de la sangre

Fuente: <http://www.texasheartinstitute.org>

Elaboración: Betzabeth Yonson S.

1.3.1.3. Minerales

Los minerales son sustancias inorgánicas, que el cuerpo necesita para formar los huesos, dientes, y glóbulos rojos, para regular los líquidos corporales. Los minerales esenciales (o indispensables para la vida del organismo). El cuadro 1.2 muestra los minerales esenciales para el organismo:

Cuadro 1.2
Minerales Esenciales

Minerales	Fuentes Mas Abundantes	Función
Calcio	Lácteos, verduras verdes, maíz y frijol.	Básico para la formación de huesos y dientes, necesario para el sistema nervioso.
Fósforo	Carne, lácteos, chícharos, frijoles, garbanzos, cereales	Reserva básica de energía para las células
Potasio	Aguacate, plátano, acelgas, papas y lentejas.	Esencial para el equilibrio de los líquidos corporales.
Magnesio	Frijoles, chícharos, nueces, cereales y verduras verdes de hojas grandes.	Necesario para las células e importante para la actividad eléctrica muscular y nerviosa
Yodo	Pescados y mariscos y sal yodatada.	necesario para la glándula tiroides
Hierro	Hígado, carne, cereales enriquecidos, huevos, berros y acelgas.	Necesario para la formación de hemoglobina, portadora de oxígeno en la sangre
Cromo	Muchos alimentos lo contienen en cantidades mínimas pero suficientes	desempeña funciones secundarias en la actividad química del organismo
Sodio	Casi todos los alimentos salvo las frutas.	Necesario para el equilibrio de los líquidos corporales, los músculos y los nervios

Fuente: <http://www.texasheartinstitute.org>

Elaboración: Betzabeth Yonson S.

1.4. Mala Alimentación

La alimentación es mala cuando es incompleta, insuficiente o excesiva, es decir, cuando faltan ciertos alimentos necesarios para nuestro organismo o no consumimos una cantidad adecuada de los mismos, ya sea por falta o por exceso.

1.4.1. Causas de Mala Alimentación

La mala alimentación puede ocasionar diversos trastornos, desde caries hasta cálculos biliares y quizá incluso algunos cánceres del aparato digestivo, sobrepeso. El estrés y los trastornos afectivos también perjudican la función digestiva y pueden ocasionar úlceras gástricas.

En la práctica se observa que tendemos a comer demasiadas grasas, sodio, colesterol y azúcar y que en cambio comemos pocas verduras. En cuanto a bebidas, preferimos los refrescos o gaseosas a los jugos de frutas y por lo que se refiere a las fuentes de proteínas, tendemos a comer poco pescado. Es evidente que muchas veces comemos no para alimentarnos sino para mitigar el hambre incluso con adversas consecuencias como lo son los alimentos chatarra. A continuación algunas enfermedades y problemas relacionados con las deficiencias y los excesos de alimentos:

1.4.1.1 Gastritis y Úlcera

Gastritis es la inflamación de la mucosa gástrica. Existen varias clasificaciones, las más conocidas es la gastritis aguda y la gastritis crónica

Gastritis aguda: cuando la mucosa gástrica está enrojecida presentándose en diversas formas de imágenes rojizas, hemorragias; puede ser que sólo una parte del estómago esté afectada o toda la esfera gástrica, sus causas pueden ser: abuso de bebidas alcohólicas, de analgésicos aspirina, el cigarro, estrés, etc.

Gastritis crónica: Estas gastritis son enfermedades más complejas, por ello y por su larga evolución, exigen un estudio y un tratamiento especializado. Entre ellas se encuentran:

- Gastritis de origen inmunológico: es la llamada gastritis tipo A, que afecta al cuerpo y al fondo gástrico. Es más frecuente en mujeres y ocurre por mecanismos inmunes (anticuerpos) que atacan la pared del estómago
- Gastritis asociada a la bacteria *Helicobacter pylori*: es la llamada gastritis tipo B, que afecta a otra zona del estómago, es la más frecuente, no predominando en ningún sexo. Produce gastritis crónica activa y se puede asociar con úlcera gastroduodenal.

1.4.1.2 Colesterol

El colesterol es una sustancia blanda y grasosa que proviene de dos fuentes: el cuerpo y los alimentos. El colesterol que circula por la sangre se llama colesterol total en sangre. El colesterol que proviene de la comida se llama colesterol de la dieta. Para funcionar normalmente, el cuerpo requiere de colesterol, y el hígado produce todo lo necesario.

En general el nivel de colesterol total en sangre por debajo de 200 miligramos (mg), es deseable. Un valor entre 200 mg y 239 mg se considera al límite del colesterol alto. Cuando el colesterol total se encuentra en este rango (200-239), el nivel de riesgo de un ataque al corazón comienza a elevarse. Una cifra de 240 o más se considera colesterol en sangre alto, el cual eleva aún más el riesgo de un ataque al corazón.

En los alimentos, el colesterol se encuentra en los huevos, productos lácteos, la carne de res y la carne de aves. La yema de huevo y las vísceras (el hígado, los riñones, la molleja y el cerebro) son ricos en colesterol. El pescado generalmente, contiene menos colesterol que otras carnes, pero algunos mariscos son ricos en colesterol.

Los alimentos de origen vegetal (verduras, frutas, granos, cereales, nueces y semillas) no contienen colesterol.

1.4.1.3 Diabetes

La diabetes es un desorden del metabolismo, el proceso que convierte el alimento que ingerimos en energía. Durante la digestión se descomponen los alimentos para crear glucosa, la mayor fuente de combustible para el cuerpo. Esta glucosa pasa a la sangre, donde la insulina le permite entrar en las células. (La insulina es una hormona segregada por el páncreas, una glándula grande que se encuentra detrás del estómago).

En personas con diabetes, una de dos componentes de este sistema falla:

- El páncreas no produce, o produce poca insulina (Tipo I); o
- las células del cuerpo no responden a la insulina que se produce (Tipo II).

Al tipo I, surge en los jóvenes. Como el cuerpo no produce insulina, personas con diabetes del tipo I deben inyectarse insulina para poder vivir.

En el tipo II, que surge en adultos, el cuerpo sí produce insulina, pero, o bien, no produce suficiente, o no puede aprovechar la que produce.

1.4.1.4 Cálculos

Existen diferentes tipos de cálculo, entre ellos se tiene el cálculo a la vesícula, también llamados de colélitiasis, son piedras biliares, que se forman en la vesícula. Pueden ser tan pequeños como granos de arenilla, tan grandes que ocupan toda la vesícula. La mayor parte, sin embargo, suelen ser de tamaño intermedio, es decir de menos de 20mm, o sea aproximadamente una pulgada. Así como los cálculos renales y otros, independientemente del tipo de cálculo y de la fase en la que se encuentre el enfermo, es fundamental el consumo de agua.

- Pero en todo caso todos somos candidatos en potencia para los cálculos, pero las mujeres son especialmente vulnerables a esta enfermedad. Es más, las de 20 a 60 años de edad corren el triple de riesgo de formar cálculos biliares que los varones.

1.4.1.5 Problemas de Garganta

Los problemas de garganta son causados por muchas cosas. Los virus (tales como los que causan los resfriados) pueden causar dolor de garganta. El dolor de garganta también puede ser causado por bacterias o por fumar, respirar aire contaminado, beber alcohol y muchos tipos de alergia. Entre ellos se encuentran: la faringitis, la amigdalitis, irritación a la garganta, etc.

La amigdalitis es la inflamación de las amígdalas (en la parte del fondo de su boca en cada lado de la garganta). Usualmente es causada por una infección bacteriana, pero a veces puede ser un virus.

Además de las enfermedades anteriores se pueden hablar de más, entre ellas, como las enfermedades de la nariz, el estreñimiento, etc. Y no se terminará de hablar de las enfermedades generadas por la mala alimentación, es por eso que en este estudio se ha analizado las más comunes por las que los estudiantes pasan o han pasado, en su trayecto académico.

CAPÍTULO II

2. REGRESIÓN LOGÍSTICA

2.1 INTRODUCCIÓN

La regresión Logística es una técnica estadística multivariante que nos permite estimar la relación existente entre una variable dependiente no métrica, en particular dicotómica y un conjunto de variables independientes métricas o no métricas.

Así, la regresión Logística constituye un conjunto de métodos para predecir resultados o consecuencias binarias sobre la base de una o más variables predictoras.

Es decir, si tenemos una variable dependiente que no es continua sino discreta que describe una respuesta en forma de dos posibles eventos o dicho de otra manera, dos posibles valores (Sí-No, 0-1, Verdadero-Falso, etc.), y queremos estudiar el efecto que otras variables (independientes) tienen sobre ella.

Dado que el modelo logístico no es lineal, sino exponencial, se utilizan transformaciones logarítmicas para linealizar el modelo, lo cual hace que los coeficientes no pueden interpretarse directamente.

El modelo logístico es:

$$\log\left(\frac{p(x_1, \dots, x_k; \beta)}{1 - p(x_1, \dots, x_k; \beta)}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k$$

o también

$$p_i = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}}$$

Esta técnica resulta especialmente útil para identificar factores de riesgo y factores de prevención de enfermedades en muestras prospectivas en los cuales la metodología de la regresión lineal no es aplicable, dado que la variable respuesta sólo presenta dos valores (caso dicotómico) como puede ser presencia/ausencia de un suceso.

Así, de manera general, para construir un modelo de regresión logística necesitamos:

- Una o más variables independientes o predictoras, representadas con X , en el estilo de la regresión lineal múltiple.

- Una variable de respuesta, representada con Y , que puede ser dicotómica o politómica. En esto se diferencia del modelo de regresión múltiple, en el cual la variable respuesta es numérica.

2.2 Tipos de Variables

Los valores que puede tomar cualquier variable, ya sea explicativa o de respuesta, constituyen los niveles de las variables. La naturaleza de estos niveles es lo que determina el tipo de las variables. Así se tienen los siguientes tipos de variables:

Binaria o Dicotómica : es aquella que sólo puede adquirir dos posibles valores (Sí-No, 0-1, Verdadero-Falso, etc.). Estas variables binarias constituyen un subconjunto muy importante de las llamadas variables categóricas o cualitativas

Nominal: aquella variable que puede tomar varios valores

Ordinal: aquella variable que puede tomar varios valores y en los que se puede establecer un orden entre las categorías

Cuantitativa: aquella variable que puede tomar un rango numérico de valores

Censurada: aquella variable cuantitativa que puede tener información parcial o no completa, por ejemplo el tiempo hasta que ocurre algo en el Análisis de Supervivencia

Las variables independientes (también llamadas explicativas X) pueden ser de cualquier naturaleza: cualitativas o cuantitativas. La probabilidad de que $Y=1$ se denotará por p .

En la regresión logística se emplean los métodos de Máxima Verosimilitud (MV) para llevar a cabo la estimación de los parámetros del modelo.

2.3 Modelo Logit

Cuando se realiza un estudio de regresión logística, se debe considerar la transformación logia, ya que el objetivo de este método es que cualquier otra técnica de construcción de modelos utilizada en estadística que es encontrar el modelo de mejor ajuste, que permita describir la relación entre la variable de respuesta y un conjunto de variables independientes.

Entonces, teniendo una función F con distribución Normal(0,1), el modelo logit será de la forma:

$$F(\beta_0, \beta_1 X_1, \beta_2 X_2, \dots, \beta_K X_K) = \frac{e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K)}}{1 + e^{(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K)}}$$

Donde:

$$F(\beta_0, \beta_1 X_1, \beta_2 X_2, \dots, \beta_K X_K) = \int_{-\infty}^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K} \frac{1}{\sqrt{2\pi}} \exp\left(-t^2/2\right) dt$$

2.4. Modelo Probit

Otra posible solución para explicar modelos con variable independiente dicotómica es usar un modelo Probit (también llamado modelo Normit) de la forma:

$$P_i = F(\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_K X_{iK}) + \mu_i$$

La estimación de los parámetros respectivos en los modelos Logit y Probit se realiza mediante el método de Máxima Verosimilitud (MV).

2.5. Odds Ratio:

En primer lugar se define el odds como una medida de ocurrencia de un evento, mediante el cociente entre el número de veces que ocurre el

suceso frente a cuántas veces no ocurre, así, se trata de un número que expresa cuanto más probable es que se produzca frente a que no se produzca el hecho en cuestión, entonces se tiene que el valor del odds se calcula como:

$$odds = \frac{p}{1-p} \quad (2.1)$$

Siendo p la probabilidad del suceso.

Ahora bien, el odds ratio se define como la razón del odds correspondiente a un suceso bajo cierta condición entre el odds que le corresponde al mismo suceso bajo otra condición.

$$odds - ratio = \frac{odds2}{odds1} \quad (2.2)$$

El OR (odds ratio) cuantifica cuánto más probable es la aparición por ejemplo de una enfermedad bajo una condición (odds2) respecto de cuando se está en la otra condición (odds1).

El OR resulta de gran utilidad en el modelo de regresión logística para evaluar la influencia que cada variable independiente tiene sobre la respuesta, en forma de OR. Un OR mayor que uno indica aumento en la probabilidad del evento y un OR menor que uno implica disminución en dicha probabilidad.

2.6 Función de Verosimilitud

Con el fin de estimar $\beta = (\beta_0 \beta_1 \dots \beta_k)^T$ y analizar el comportamiento del modelo considerado, observamos una muestra aleatoria simple de tamaño n dada por $\{(x_i, y_i) ; i = 1, \dots, n\}$ donde $x_i = (x_{i1}, \dots, x_{ik})'$ es el valor de la variables independientes e

$y_i \in \{0, 1\}$ es el valor observado de Y en el i -ésimo elemento de la muestra.

Por tanto se tiene que $Y/(x_1, \dots, x_k)$ tiene una distribución binomial $(1, p(y = 1 / x_1, \dots, x_k ; \beta))$, utilizando el hecho de que la variable dependiente toma sólo dos resultados (*éxito* y *fracaso*), cuando el número de éxitos en n repeticiones tiene una distribución binomial $B(n, p)$.

Luego, la función de verosimilitud es:

$$L(\beta / (x_1, y_1), \dots, (x_n, y_n)) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1-y_i} \quad (2.3)$$

donde

$$p_i = p(x_i ; \beta) = p(x_{i1}, \dots, x_{ik} ; \beta); \quad i=1, \dots, n$$

2.7. Modelo de Regresión Logística Binaria Simple

El objetivo de este análisis es encontrar el mejor ajuste, teniendo un modelo razonable que describa la relación entre la variable dependiente o de respuesta y una única variable independiente X.

Así, tenemos una variable dependiente binaria Y que toma solo dos posibles valores, consideremos 0 y 1; para ajustar el modelo de regresión logística según (2.4) a un conjunto de datos se requiere en primer lugar estimar los parámetros β_0 y β_1 los cuales son desconocidos.

$$\ln \frac{P}{1-P} = \beta_0 + \beta_1 x \quad (2.4)$$

La expresión anterior es equivalente también a:

$$P = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

Donde P, representa la probabilidad de que un individuo desarrolle la característica de interés.

2.8. Modelo de Regresión Logística Múltiple

Su objetivo es de trabajar con más de una variable (variables independientes X), aún cuando éstas puedan estar en diferentes

escalas de medición, para que nos ayude a entender por qué varía la respuesta entre unos individuos y otros.

Considerando la variable dicotómica de respuesta Y que toma solo dos posibles valores, 0 y 1, el modelo logístico múltiple es:

$$\text{Log}\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k \quad (2.5)$$

o también:

$$P(Y=1) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k}}$$

A continuación se escribirá la manera de construirlo y sus conceptos.

2.8.1 Ajuste del Modelo de Regresión Logística Múltiple

El ajuste se efectúa a través del uso de los métodos de máxima verosimilitud, para lo cual asumiremos que disponemos de una muestra de n observaciones independientes (x_i, y_i) , $i = 1, 2, \dots, n$; donde y_i toma valores 0 ó 1, para estimar $\beta^T = (\beta_0 \beta_1 \dots \beta_k)$ que es el vector de los parámetros desconocidos.

Para este modelo cuando la variable respuesta es binaria se usa el método de Máxima Verosimilitud, ya que obtendremos parámetros estimados que maximizan la probabilidad de obtener un conjunto de datos observados.

Las ecuaciones de verosimilitud son:

$$\sum_{i=1}^n [y_i - p_i] = 0 \quad (2.6)$$

y

$$\sum_{i=1}^n x_{ij} [y_i - p_i] = 0, j = 1, \dots, k; i = 1, \dots, n \quad (2.7)$$

Para encontrar la solución a este conjunto de ecuaciones se requiere del cálculo diferencial, hoy en día existen software estadísticos para estimar los parámetros de manera automática.

Sea $\hat{\beta}$ el estimador de máxima verosimilitud para el sistema de ecuaciones de tal modo \hat{p}_i es el modelo logístico múltiple de la ecuación (2.5). El método de estimación de varianzas y covarianzas de los coeficientes estimados, se realizará a través del método de máxima verosimilitud con procesos iterativos, el cual consiste en obtener la matriz de la segunda derivada parcial de la función de verosimilitud, éstas derivadas parciales tienen la siguiente forma general:

$$\frac{\partial^2 L(\beta)}{\partial \beta_j^2} = - \sum_{i=1}^n \frac{x_{ij}^2 e^{\beta_0 + \beta_1 X_i}}{(1 + e^{\beta_0 + \beta_1 X_i})^2} = - \sum_{i=1}^n x_{ij}^2 p_i (1 - p_i) \quad (2.8)$$

$$\frac{\partial^2 L(\beta)}{\partial \beta_j \partial \beta_l} = - \sum_{i=1}^n \frac{x_{ij} x_{il} e^{\beta_0 + \beta_1 X_i}}{(1 + e^{\beta_0 + \beta_1 X_i})^2} = - \sum_{i=1}^n x_{ij} x_{il} p_i (1 - p_i) \quad (2.9)$$

Sea una matriz $(k+1) \times (k+1)$ que contiene los términos negativos en las ecuaciones (2.8) y (2.9) denotado por $I(\beta)$, llamada matriz de información con varianzas y covarianzas de los coeficientes, estos términos son estimados por la inversa de la matriz de la siguiente forma $\Sigma(\beta) = I^{-1}(\beta)$, cuyos elementos de la diagonal son $\sigma^2(\beta_j)$ es decir, la varianza de $\hat{\beta}_j$ y $\sigma(\hat{\beta}_j, \hat{\beta}_k)$ son las covarianzas de $\hat{\beta}_j$ y $\hat{\beta}_k$. Para estimar la matriz de información del modelo estimado se tiene: $\hat{I}(\hat{\beta}) = X'VX$, donde $X_{n \times (k+1)}$ es la matriz de datos de los sujetos y $V_{n \times n}$ es una matriz diagonal cuyo elemento general es $\hat{P}_i(1 - \hat{P}_i)$.

Así, las matrices son:

$$X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix}_{n \times (k+1)} \quad (2.10)$$

$$V = \begin{bmatrix} \hat{P}(x_1)(1 - \hat{P}(x_1)) & 0 & \dots & 0 \\ 0 & \hat{P}(x_2)(1 - \hat{P}(x_2)) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \hat{P}(x_n)(1 - \hat{P}(x_n)) \end{bmatrix}_{n \times n} \quad (2.11)$$

2.8.2 Método de Newton – Raphson

Este método es uno de los mejores algoritmos utilizados para encontrar los estimadores de máxima verosimilitud en regresión logística ya que nos ayuda a resolver ecuaciones no lineales, que requieren una

solución mediante métodos iterativos, para hallar la estimación de los parámetros.

Para la estimación del vector β se requiere hallar el máximo de una función; para lo cual se usa entonces la primera derivada, porque se anula en el punto máximo y la segunda derivada para calcular las tangentes. En este caso es necesario hallar la segunda derivada para obtener la matriz de varianzas y covarianzas de los parámetros estimados.

Así tenemos:

$$\begin{bmatrix} \frac{\partial L(\beta)}{\partial(\beta_0)} \\ \frac{\partial L(\beta)}{\partial(\beta_1)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Y en cada paso del algoritmo de Newton Raphson, el estimador $\beta(t)$ de los parámetros se actualiza de la siguiente forma:

$$\beta^{(t+1)} = \beta^{(t)} + [-l''(\beta^{(t)})]^{-1} l'(\beta^{(t)})$$

Donde l' es el vector de primeras derivadas con respecto a cada uno de los parámetros, también conocido como vector de calificación y l'' es la matriz de segundas derivadas de la función del logaritmo de la verosimilitud, también llamada matriz Hessiana. Es decir, es una matriz de $p \times p$ con el elemento (j,k) éximo igual a $\frac{\partial^2}{\partial\beta_j\partial\beta_k}$. Este paso se

repite hasta lograr que los resultados converjan, es decir, hasta que la diferencia en el valor del estimador del parámetro entre un paso y otro sea insignificante.

2.8.3 Pruebas de Hipótesis Sobre los Parámetros.

Luego de realizar la estimación de los coeficientes, el siguiente paso, por lo general incluye una revisión de la significancia de las variables incluidas en el modelo. Esto generalmente implica formular y probar una hipótesis estadística para determinar si la(s) variable(s) independiente(s) del modelo están relacionadas en forma significativa a la variable dependiente o de respuesta. Es por ello que es necesario hacer una comparación de los valores observados de la variable respuesta respecto de los valores predichos por ambos modelos, aquél que incluye a la variable en cuestión y aquél que no lo hace. En caso de que el modelo que incluye a la variable estime de forma más precisa los valores observados que el modelo que no la incluye, es posible decir que la variable es significativa. Es importante recalcar que en esta parte del análisis no se está analizando que tan bien se ajustan los valores predichos en un sentido absoluto, sino que se está analizando en forma relativa comparando dos modelos distintos para la misma variable. Para observar el ajuste del modelo en un sentido absoluto se realizan

posteriormente las pruebas de bondad de ajuste, que responden a la pregunta en el sentido absoluto, no relativo.

Para revisar la significancia de la variable es necesario comparar los valores observados de la variable respuesta con los valores predichos obtenidos de los modelos que contiene y que no contiene la variable en cuestión. Para hacer esto, resulta muy útil pensar en los valores observados de la variable dependiente como valores predichos o estimados a partir de un modelo saturado. Un modelo saturado es aquel en el que existen tantos parámetros como observaciones. Entonces, la comparación de los valores observados respecto a los valores predichos usando la función de verosimilitud se basa en la siguiente expresión:

$$D = -2 \ln \left[\frac{\text{verosimilitud modelo ajustado}}{\text{verosimilitud modelo saturado}} \right]$$

La cantidad dentro del paréntesis en la expresión anterior se conoce como razón de verosimilitud. Se utiliza menos dos veces su logaritmo para obtener así una cantidad cuya distribución es conocida y que por lo tanto pueda ser utilizada para pruebas de hipótesis.

2.8.4. Pruebas de Significancia para Regresión Logística Múltiple.

El contraste de hipótesis para determinar si las variables independientes influyen significativamente en la probabilidad del suceso del modelo relacionado a la variable del resultado es del siguiente modo:

$$H_0 : \beta_0 = \beta_1 = \dots = \beta_k = 0$$

$$H_1 : \text{para algún } \beta_i \neq 0$$

Estadístico de prueba:

$$D \sim X^2 \text{ con } k \text{ grados de libertad} \quad (2.12)$$

Decisión: si $D > X_{\alpha,k}^2$ rechazamos H_0 , entonces al menos uno de los coeficientes es diferente de cero y la variable correspondiente influye en la probabilidad del suceso estudiado.

Una vez encontrado el mejor conjunto de variables explicativas que predicen la variable Y , luego se debe evaluar mediante Wald cada coeficiente para determinar cuál o cuáles ingresan al modelo.

2.8.5 La Prueba de Wald

Para el caso múltiple, el contraste de hipótesis es:

$$H_0 : \beta_i = 0$$

$$H_1 : \text{algún } \beta_i \neq 0$$

Estadístico de prueba:

$$W = \hat{\beta}' [\hat{I}(\hat{\beta})]^{-1} \hat{\beta}' = \hat{\beta}' (X' V X) \hat{\beta} \sim X_{\alpha, k+1}^2 \quad (2.13)$$

Donde X y V , son las matrices de las ecuaciones (2.10) y (2.11).

Decisión: si $W > X_{\alpha, k}^2$ rechazamos H_0 con un nivel de significancia fijado, concluimos que la variable independiente influye en la probabilidad del suceso.

2.8.6 Prueba de Hosmer-Lemeshow

La prueba de Hosmer-Lemeshow es otra prueba para evaluar la bondad del ajuste de un modelo de regresión logística, aunque su uso es más discutido que el anterior. La idea es que si el ajuste es bueno, se asociará un valor alto de la p predecida (con una frecuencia parecida a la p) con el resultado 1 de la variable binomial. Se trata de calcular para cada observación del conjunto de datos las probabilidades de la variable dependiente que predice el modelo, agruparlas y calcular, a partir de ellas, las frecuencias esperadas y compararlas con las observadas mediante la prueba χ^2 .

2.9. Regresión Logística Multinomial

La regresión logística multinomial es útil para las situaciones en las cuales se desea poder clasificar a los individuos según los valores de un grupo de variables predictoras. Este tipo de regresión es similar a la regresión logística, pero es más general porque la variable dependiente

no se restringe a dos categorías; posteriormente es utilizado este análisis para derivar estimaciones de la razón, para cada una de los factores indicando cuanto mas probable son las características que intervienen en el modelo.

2.9.1. Regresión Logística con Respuesta Nominal Politémica

La estrategia de este análisis es similar a la RL binaria. En este caso se toma una categoría o condición de interés que posee un individuo como referencia y se enfrenta a la complementaria. Se adoptó como “categoría de referencia” $Y=0$ y se definió el odds como $P(y=1)/ P(y=0)$. En el caso nominal politémico, como existen r categorías, se toma una de ellas como referencia y se enfrentan a ella las $r-1$ categorías restantes.

El modelo de regresión logística politémica (RLP), viene dado por la expresión:

$$P(Y = s) = \frac{e^{(\beta_s + \beta_{1s}X_1 + \dots + \beta_{ks}X_k)}}{1 + e^{(\beta_s + \beta_{1s}X_1 + \dots + \beta_{ks}X_k)}}$$

de manera que:

$$\ln \frac{P(Y = s)}{P(Y = 1)} = \beta_s + \beta_{1s}X_1 + \dots + \beta_{ks}X_k \quad s=2,3,\dots,r$$

Es decir que para cada categoría distinta de la de referencia tenemos un conjunto de parámetros $\beta_{0s}, \beta_{1s}, \dots, \beta_{ks}$, que contempla un conjunto de $c-1$ ecuaciones de regresión y, como para cada una de ellas tenemos $k-1$ parámetros, en el modelo logístico están implicados $(c-1)(k-1)$ parámetros. La definición del modelo politómico exige entonces que para cada variable independiente haya que estimar X_i efectos (coeficientes).

Lo cual se deben construir funciones igual que en el caso dicotómico tantas funciones como posibles respuestas menos 1 y es menester, en este caso, estimar no uno sino $r-1$ conjuntos de $k+1$ parámetros.

En síntesis lo que se trata es de obtener una manera de estimar la probabilidad de pertenencia a cada una de las r categorías para un sujeto que tiene un perfil específico de entrada X_1, X_2, \dots, X_k .

CAPÍTULO III

3. DISEÑO MUESTRAL

3.1. Introducción

Los problemas más frecuentes en la investigación en diferentes campos son resueltos mediante el diseño adecuado de estudios. La estadística ha contribuido el desarrollo y éxito de dichos estudios.

Por esta razón se desea determinar mediante esta investigación si la alimentación adecuada genera o refuerza en el desarrollo intelectual de los estudiantes una actitud positiva o negativa hacia la misma.

Para ello se efectuó una encuesta con sus correspondientes etapas como: determinación de la población objetivo, determinación del marco muestral, diseño del cuestionario, diseño muestral, captura de la información y digitación de los datos. Adicionalmente se obtuvo una

base de datos de los estudiantes registrados de la ESPOL con la información necesaria para dicho estudio.

En el presente capítulo se mostrará la información de las definiciones antes escritas, definiciones de la estadística inferencial y técnicas usadas para la recopilación, tabulación, análisis, interpretación y descripción de las variables utilizadas en el cuestionario y las que adicionalmente se obtuvieron mediante la base de datos.

3.2. Población Objetivo

Es el conjunto de individuos o elementos que han sido considerados para la investigación, conformada por los estudiantes registrados en el primer semestre del año 2005, el mismo que está comprendido entre mayo y septiembre del 2005 en la ESPOL en el Campus “Gustavo Galindo Velasco” y se totalizan $N=7299$ estudiantes registrados.

3.3. Marco Muestral

Es un listado de unidades o elementos pertenecientes a la población de donde se selecciona la muestra.

Para nuestro estudio el marco muestral está conformado por el listado de estudiantes registrados en las diferentes Unidades Académicas de la ESPOL, información tomada de la base de datos proporcionado por el Centro de Registros, Calificaciones y Estadísticas (CRECE), ver cuadro 3.2.

3.4 Diseño del Cuestionario

El cuestionario fue diseñado como instrumento de medición, para interrogar a los estudiantes politécnicos sobre sus datos generales, sociales y alimenticios; se contó con la ayuda parcialmente de personas especializadas en el área de la salud, alimentación y nutrición. En el anexo 1 se muestra el cuestionario utilizado.

La recopilación de esta información aplicada a los estudiantes se la obtuvo en las aulas (paralelos) seleccionados en la muestra.

3.5. Muestra

Es un subconjunto de observaciones $x_1, x_2, x_3, \dots, x_n$, de tamaño n seleccionados de una población X de tamaño N , donde $n \leq N$.

3.6 Tipos de Muestreo

Muestreo es la selección aleatoria de una muestra para hacer inferencias de una población a la cual se extrajo dicha muestra. Existen diferentes tipos de muestreo, los cuales pueden dividirse en dos grandes grupos: muestreo probabilísticos y muestreo no probabilísticos, en este estudio se trabajará con muestreo probabilístico.

3.6.1. Muestreo Probabilísticos

Es aquel en los que todos los individuos tienen la misma probabilidad de ser elegidos para formar parte de una muestra de tamaño n extraída de una población de tamaño N , y por consiguiente, que todas las posibles muestras de tamaño n tienen la misma probabilidad de ser elegidas.

3.6.1.1. Muestreo Aleatorio Simple

Consiste en elegir en forma aleatoria muestras de una población, este proceso debe permitir la misma probabilidad de selección a todas las muestras y que cada elemento de la población tenga una probabilidad igual de ser incluido en la muestra.

Cuándo se debe usar el muestreo aleatorio simple:

- Cuando se sabe que la variable de mayor interés se distribuye aleatoriamente en la población.
- Cuando se conoce o se tiene un listado de todos los elementos de la población.
- Cuando se conoce el tamaño de la población N .

3.6.1.2. Muestreo Estratificado

Es cuando a la población X , de tamaño N se la particiona en k subpoblaciones o estratos heterogéneos entre ellos y homogéneos dentro de cada estrato. Este tipo de muestreo se usa principalmente en poblaciones donde se conoce que la distribución de las variables de mayor interés es diferente entre subpoblaciones fácilmente identificables (véase el cuadro 3.1).

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

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

Afijación Proporcional: La distribución se hace de acuerdo con el tamaño de la población en cada estrato, tomándose muestras en cada estrato en forma proporcional al número total de unidades muestrales que incluye cada estrato.

Sea W_h la proporción de unidades muestrales en cada estrato

$$W_h = N_h / N$$

Donde:

N_h es el número de unidades en el estrato h

N es el número total de unidades. (véase cuadro 3.3)

Afijación Óptima: Se tiene en cuenta la dispersión de los resultados, de modo que se considera la proporción y la desviación típica.

Afijación de Mínima Varianza: También conocida afijación de Neyman consiste en determinar los valores de n_h (número unidades que se extraen del estrato k-ésimo de la muestra) de forma que para un tamaño de muestra fijo igual a n la varianza sea mínima.

3.7 Tamaño de la Muestra

El tamaño de la muestra es el número de unidades muestrales tomadas de una población para estimar la característica deseada.

El tamaño de la muestra depende de la variabilidad del parámetro a estudiar, del nivel de confianza fijado y del error.

3.7.1. Determinación del tamaño de la Muestra

Sea P la proporción de elementos en la población que tienen la característica de interés, N el tamaño de la población finita conocida es necesario estimar la varianza de la proporción.

$$\sigma_{\hat{P}}^2 = \frac{P(1-P)}{n} \left(\frac{N-n}{N-1} \right)$$

Que debido a la ecuación:

$$E = Z_{\alpha/2} \sqrt{\frac{N-n}{N-1} \left(\frac{p(1-p)}{n} \right)}$$

Y mediante cálculo algebraico se obtiene:

$$n_0 = \frac{Z_{\alpha/2}^2 p(1-p)}{E^2}$$

Para finalmente obtener el tamaño de la muestra como:

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

Por medio del Muestreo Aleatorio Simple se obtuvo una muestra piloto (prueba que se realiza para obtener el tamaño de la muestra) conformada por 25 estudiantes de las diferentes unidades académicas de la ESPOL, Campus "Gustavo Galindo Velasco".

A los cuales se les cuestionó la variable (X_{23} = Tiene un horario establecido fijo para su alimentación diaria?), donde 19 estudiantes respondieron Si y 6 respondieron No, a continuación los siguientes resultados para el tamaño de la muestra:

Se consideró una significancia (alfa) de $\alpha = 0.05$, un nivel de confianza del 95%, un error de 0.05, con $p = 0.76$, $(1-p) = 0.24$ y $N = 7299$; entonces tenemos:

$$n_0 = \frac{1.96^2 (0.76)(0.24)}{0.05^2} = 280.28$$

$$n = \frac{280.28}{1 + \frac{280.28}{7299}} = 269.9$$

Se determinó que el tamaño de la muestra es $n = 270$ estudiantes

3.8 Técnicas de Muestreo

Para el presente estudio se consideró realizar muestreo multietápico; en donde fue necesario realizar tres etapas utilizando el muestreo estratificado y por conglomerados.

Primera Etapa

Se clasificó a la población ($N = 7299$) en tres estratos, distribuyéndolos como diferentes tipos de Unidades académicas, los cuales se los identificaron como: Facultades de Ingeniería, Institutos de Ciencias y Tecnologías; como se muestra a continuación en el cuadro 3.1.

Cuadro 3.1
Distribución de la Población

Unidad Académica	No. de estudiantes registrados	Porcentaje (%)
FACULTADES DE INGENIERÍA	4229	58
INSTITUTOS DE CIENCIAS	2612	36
TECNOLOGÍAS	458	6
Total	7229	100

Fuente: CRECE, ESPOL, 2005
Elaboración: Betzabeth Yonson S.

Segunda Etapa

En esta segunda etapa se partió en subestratos a cada estrato clasificados en la primera etapa, como se verá a continuación en el cuadro 3.2:

Cuadro 3.2
Distribución de la Población
Subestratos a cada Unidad Académica

Población		No. de estudiantes registrados	Porcentaje (%)
Unidad Académica	Subestratos		
FACULTADES DE INGENIERÍA	FICT	379	9
	FIEC	2278	54
	FIMCP	1431	34
	FMAR	141	3
	Total de Facultades	4229	100
INSTITUTOS DE CIENCIAS	ICHE	1897	73
	ICM	715	27
	Total de Institutos	2612	100
TECNOLOGÍAS	INTEC	62	14
	PROTEL	184	40
	PROTEP	151	33
	PROTMEC	61	13
	Total de Tecnologías	458	100
Total Población		7229	

Fuente: CRECE, ESPOL, 2005
Elaboración: Betzabeth Yonson S.

Tercera Etapa

En esta etapa para cada subestrato de las unidades académicas se aplicó muestreo por conglomerados, los cuales representan las materias

con sus respectivos paralelos dictados en el primer semestre del año 2005. Cada paralelo fue enumerado de acuerdo como va en el orden del pensul académico de cada subestrato, para luego mediante el formato de generación de números aleatorios en Excel, se realizó la selección aleatoria de los paralelos hasta obtener el tamaño de muestra asignado para cada subestrato. A continuación en el cuadro 3.3 presentamos la distribución de cada subestrato para el tamaño de la muestra.

Cuadro 3.3
Distribución de la Muestra
Subestratos a cada Unidad Académica

Muestra		No. de estudiantes
Unidad Académica	Subestratos	registrados
FACULTADES DE INGENIERÍA	FICT	14
	FIEC	85
	FIMCP	53
	FMAR	5
	Total de Facultades	157
INSTITUTOS DE CIENCIAS	ICHE	71
	ICM	26
	Total de Institutos	97
TECNOLOGÍAS	INTEC	2
	PROTEL	7
	PROTEP	2
	PROTMEC	5
	Total de Tecnologías	16
Total Muestra		270

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

3.9 Descripción de las variables de nuestro estudio

Para el presente estudio se trabajó con treinta y cinco variables, de las cuales treinta y un variables se les interrogó al estudiante, mediante el cuestionario presentado en el anexo 1; y cuatro variables adicionales (edad, Factor P, Promedio y Rendimiento académico), tomadas de la hoja de datos de los estudiantes que nos facilitó el CRECE las cuales no se la pidió en el cuestionario por motivos de mala información o negarse a contestar las preguntas por parte de los entrevistados.

Las treinta y cinco variables seleccionadas para nuestro estudio constan en tres secciones:

- *Datos Generales.*- En esta sección se analiza al estudiante su edad, Factor P, Promedio, Rendimiento académico, Peso y Estatura.

- *Datos sociales.*- Se le pregunta al estudiante si trabaja, número de horas que permanece en la ESPOL, si realiza algún deporte y su perspectiva de cuan importante es la alimentación en su rol como estudiante.

- *Datos alimenticios.*- En esta sección se desea saber que es lo que el estudiante consume y si debido a ello ha pasado o no por problemas alimenticios.

3.9.1 Codificación de las variables de nuestro estudio

Sección I: Datos Generales

Variable X₁: Género

Variable cualitativa en escala nominal, indica si el estudiante entrevistado es del género masculino o femenino.

Cuadro 3.4
Variable X₁
Género

Código	Género
1	Masculino
2	Femenino

Variable X₂: Edad

Variable cuantitativa que describe la edad del estudiante entrevistado

Variable X₃: Carrera

Variable cualitativa en escala nominal, indica a que carrera pertenece estudiante entrevistado.

Cuadro 3.5
Variable X₃
Carrera

Código	Carreras	Significado de siglas
1	FICT	Facultad de Ingeniería de Ciencias de la Tierra
2	FIEC	Facultad de Ingeniería Eléctrica y Computación
3	FIMCP	Facultad de Ingeniería Mecánica y Ciencias de la Producción
4	FMAR	Facultad de Ingeniería Marítima y Ciencias del Mar
5	ICHE	Instituto de Ciencias Humanísticas y Económicas
6	ICM	Instituto de Ciencias Matemáticas
7	INTEC	Tecnología en Alimentos
8	PROTEL	Tecnología Eléctrica y Electrónica
9	PROTEP	Tecnología Pesquera
10	PROTMEC	Tecnología Mecánica y Mecánica Automotriz

Variable X₄: Factor P

Variable cuantitativa que describe el indicador socio-económico que la ESPOL asigna a cada estudiante al momento en que ingresa a estudiar a la Institución Académica.

Variable X₅: Promedio

Variable cuantitativa que describe el promedio de notas totales de las materias tomadas del estudiante.

Variable X₆: Rendimiento Académico

Variable cuantitativa que describe el porcentaje de rendimiento académico expresado en términos del número de materias tomadas con el número de materias aprobadas del estudiante.

Variable X₇: Peso

Variable cuantitativa describe el peso en libras del estudiante, valor real ya que fue pesado el estudiante en el momento de la entrevista.

Variable X₈: Estatura

Variable cuantitativa que describe la estatura en centímetros del estudiante entrevistado, medida que fue tomada en el momento de la entrevista con el estudiante.

Sección II: Datos Sociales

Variable X₉: Actualmente trabaja

Variable cualitativa en escala nominal, indica si el estudiante además de estudiar actualmente trabaja o no. A esta variable se le asignó la siguiente codificación:

Cuadro 3.6
Variable X₉
Actualmente Trabaja

Código	Trabaja
1	Dentro de la Espol
2	Fuera de la Espol
3	No trabaja

Variable X₁₀: Número de horas que diariamente permanece en la ESPOL

Variable categórica, indica el tiempo promedio que el estudiante permanece en la ESPOL. A esta variable se le asignó la siguiente codificación:

Cuadro 3.7
Variable X₁₀
Horas que permanece en la Espol

Código	Rango
1	Menos de 3 horas
2	De 3 a 6 horas
3	De 6 a 9 horas
4	Más de 9 horas

Variable X_{11} : Número de días a la semana que realiza algún tipo de actividad deportiva

Variable categórica, indica el número de días a la semana promedio que el estudiante realiza alguna actividad deportiva. A esta variable se le asignó la siguiente codificación:

Cuadro 3.8
Variable X_{11}
Días promedio que realiza actividad deportiva

Código	Rango
1	0 - 1 día
2	2 - 3 días
3	4 - 5 días
4	6 - 7 días

Variable X_{12} : ¿Cómo considera la alimentación para desempeñar su rol de estudiante?

Variable categórica, indica que tan importante toma el estudiante su alimentación para su desempeño académico. A esta variable se le asignó la siguiente codificación:

Cuadro 3.9
Variable X_{12}
Desempeño para su rol de estudiante

Código	Rango
1	Nada Importante
2	Poco Importante
3	Importante
4	Muy Importante
5	Extremadamente Importante

Variable X₁₃: Porcentaje de su alimentación es a través de comida preparada dentro de casa

Variable cuantitativa que describe el porcentaje promedio que el estudiante se alimenta dentro de casa.

Variable X₁₄: Días a la semana que desayuna el estudiante

Variable categórica, indica el número de días a la semana promedio que el estudiante desayuna. A esta variable se le asignó la siguiente codificación:

Cuadro 3.10
Variable X₁₄
Días a la semana que desayuna

Código	Rango
1	0 - 1día
2	2 - 3días
3	4 - 5días
4	6 - 7días

Variable X₁₅: Días a la semana que come carne roja

Variable categórica, indica el número de días a la semana promedio que el estudiante come carne roja. A esta variable se le asignó la siguiente codificación:

Cuadro 3.11
Variable X₁₅
Días a la semana que come carne roja

Código	Rango
1	0 - 1día
2	2 - 3días
3	4 - 5días
4	6 - 7días

Variable X₁₆: Días a la semana que come pescado

Variable categórica, indica el número de días a la semana promedio que el estudiante come pescado. A esta variable se le asignó la siguiente codificación:

Cuadro 3.12
Variable X₁₆
Días a la semana que come pescado

Código	Rango
1	0 - 1 día
2	2 - 3 días
3	4 - 5 días
4	6 - 7 días

Variable X₁₇: Días a la semana que come pollo

Variable categórica, indica el número de días a la semana promedio que el estudiante come pollo. A esta variable se le asignó la siguiente codificación:

Cuadro 3.13
Variable X₁₇
Días a la semana que come pollo

Código	Rango
1	0 - 1 día
2	2 - 3 días
3	4 - 5 días
4	6 - 7 días

Variable X₁₈: Días a la semana que toma leche

Variable categórica, indica el número de días a la semana promedio que el estudiante toma leche. A esta variable se le asignó la siguiente codificación:

Cuadro 3.14
Variable X₁₈
Días a la semana que toma leche

Código	Rango
1	0 - 1 día
2	2 - 3 días
3	4 - 5 días
4	6 - 7 días

Variable X₁₉: Días a la semana que consume vegetales

Variable categórica, indica el número de días a la semana promedio que el estudiante consume vegetales. A esta variable se le asignó la siguiente codificación:

Cuadro 3.15
Variable X₁₉
Días a la semana que consume vegetales

Código	Rango
1	0 - 1día
2	2 - 3días
3	4 - 5días
4	6 - 7días

Variable X₂₀: Días a la semana que come frutas

Variable categórica, indica el número de días a la semana promedio que el estudiante come frutas. A esta variable se le asignó la siguiente codificación:

Cuadro 3.16
Variable X₂₀
Días a la semana que come frutas

Código	Rango
1	0 - 1día
2	2 - 3días
3	4 - 5días
4	6 - 7días

Variable X₂₁: Comidas que realiza en el día

Variable categórica, indica cuantas comidas realiza en el día. A esta variable se le asignó la siguiente codificación:

Cuadro 3.17
Variable X₂₁
Comidas que realiza en el día

Código	Rango
1	Una vez
2	Dos veces
3	Tres veces
4	Más de tres veces

Variable X₂₂: En la hora de alimentación que tan frecuente opta por comida rápida (comida chatarra)

Variable categórica, indica que tan frecuente consume las comidas rápidas más conocidas como comida chatarra. A esta variable se le asignó la siguiente codificación:

Cuadro 3.18
Variable X₂₂
En la hora de alimentación que tan frecuente opta por comida rápida

Código	Rango
1	Nunca
2	Rara vez
3	Algunas veces
4	La mayoría de las veces
5	Siempre

Variable X₂₃: Horario fijo para su alimentación diaria

Variable cualitativa en escala nominal, indica si el estudiante entrevistado se alimenta en horas fijas.

Cuadro 3.19
Variable X₂₃
Horario fijo para su alimentación diaria

Código	Horario fijo
1	No
2	Si

Variable X₂₄: Tipos de bebida de preferencia para satisfacer su sed

Variable cualitativa en escala nominal, indica que tipo de bebida es de su preferencia para el estudiante para satisfacer su sed. A esta variable se le asignó la siguiente codificación:

Cuadro 3.20
Variable X₂₄
Tipo de bebida de preferencia para satisfacer su sed

Código	Rango
1	Agua con gas
2	Agua sin gas
3	Jugos preparados
4	Jugos envasados
5	Bebida con contenido alcohólico
6	Gaseosas
7	Otros

Variable X₂₅: Frecuencia que fuma

Variable categórica, indica que tan frecuente fuma el estudiante. A esta variable se le asignó la siguiente codificación:

Cuadro 3.21
Variable X₂₅
Frecuencia que fuma

Código	Rango
1	Nunca
2	Rara vez
3	Algunas veces
4	La mayoría de las veces
5	Siempre

Variable X₂₆: Frecuencia que ingiere bebida alcohólica

Variable categórica, indica que tan frecuente ingiere bebida alcohólica el estudiante. A esta variable se le asignó la siguiente codificación:

Cuadro 3.22
Variable X₂₆
Frecuencia con que ingiere bebida alcohólica

Código	Rango
1	Nunca
2	Rara vez
3	Algunas veces
4	La mayoría de las veces
5	Siempre

Variable X₂₇: Problema de Gastritis

Variable cualitativa en escala nominal, indica si el estudiante entrevistado en su trayecto en la ESPOL, sufre, ha sufrido o pasado por los problemas de gastritis. A esta variable se le asignó la siguiente codificación:

Cuadro 3.23
Variable X₂₇
Problema de Gastritis

Código	Gastritis
1	No
2	Si

Variable X₂₈: Problema de Úlcera

Variable cualitativa en escala nominal, indica si el estudiante entrevistado en su trayecto en la ESPOL, sufre, ha sufrido o pasado por los problemas de úlcera.

Cuadro 3.24
Variable X₂₈
Problema de Úlcera

Código	Úlcera
1	No
2	Si

Variable X₂₉: Problema de Garganta

Variable cualitativa en escala nominal, indica si el estudiante entrevistado en su trayecto en la ESPOL, sufre, ha sufrido o pasado por problemas de garganta.

Cuadro 3.25
Variable X₂₉
Problemas de garganta

Código	Problemas de garganta
1	No
2	Si

Variable X₃₀: Problema de Colesterol en Niveles Anormales

Variable cualitativa en escala nominal, indica si el estudiante entrevistado en su trayecto en la ESPOL, sufre, ha sufrido o pasado por problemas de colesterol.

Cuadro 3.26
Variable X₃₀
Problema de Colesterol en niveles anormales

Código	Problemas de Colesterol
1	No
2	Si

Variable X₃₁: Problema de Diabetes

Variable cualitativa en escala nominal, indica si el estudiante entrevistado en su trayecto en la ESPOL, sufre, ha sufrido o pasado por problemas de Diabetes.

Cuadro 3.27
Variable X₃₁
Problema de Diabetes

Código	Problemas de Diabetes
1	No
2	Si

Variable X₃₂: Problema de Cálculos

Variable cualitativa en escala nominal, indica si el estudiante entrevistado en su trayecto en la ESPOL, sufre, ha sufrido o pasado por los llamados cálculos.

Cuadro 3.28
Variable X₃₂
Problema de Cálculos

Código	Problemas de Cálculos
1	No
2	Si

Variable X₃₃: Problemas Otros

Variable cualitativa en escala nominal, indica si el estudiante entrevistado en su trayecto en la ESPOL, sufre, ha sufrido o pasado por otros problemas no mencionados anteriormente.

Cuadro 3.29
Variable X₃₃
Problema Otros

Código	Problemas Otros
1	No
2	Si

Variable X₃₄: Ausencia de clases por problemas alimenticios

Variable categórica, indica que tan frecuente el estudiante ha tenido que ausentarse de clases por problemas alimenticios. A esta variable se le asignó la siguiente codificación:

Cuadro 3.30
Variable X₃₄
Ausencia de clases por problemas alimenticios

Código	Rango
1	Nunca
2	Rara vez
3	Algunas veces
4	La mayoría de las veces
5	Siempre

Variable X₃₅: Frecuencia con la que ha pasado por el llamado “Estrés”

Variable categórica, indica que tan frecuente el estudiante ha pasado por el llamado Estrés. A esta variable se le asignó la siguiente codificación:

Cuadro 3.31
Variable X₃₅
Frecuencia con la que ha pasado por el llamado Estrés

Código	Rango
1	Nunca
2	Rara vez
3	Algunas veces
4	La mayoría de las veces
5	Siempre

CAPÍTULO IV

4. ANÁLISIS ESTADÍSTICO DE LOS DATOS OBSERVADOS

4.1 ANÁLISIS DESCRIPTIVO

Introducción

En el presente capítulo se realizará el análisis descriptivo para cada una de las variables consideradas en el capítulo anterior que fueron seleccionadas mediante la base de datos de los Registros del CRECE de los estudiantes registrados en el segundo término del año 2005 y junto a ello las variables consideradas en el cuestionario presentado para los 270 estudiantes tomados en la muestra.

De estas variables obtendremos sus respectivas distribuciones, frecuencias, diagramas de barra, medidas de tendencia central, dispersión, sesgo y curtosis; para este propósito se utilizó el paquete estadístico de SPSS 12.0

Sección I: Datos Generales

Variable X_1 : Género

En cuanto al género de los estudiantes, del total de una muestra de 270 estudiantes entrevistados en el I término del año 2005, se estima que el 63.7% son hombres y el 36.3% son mujeres.

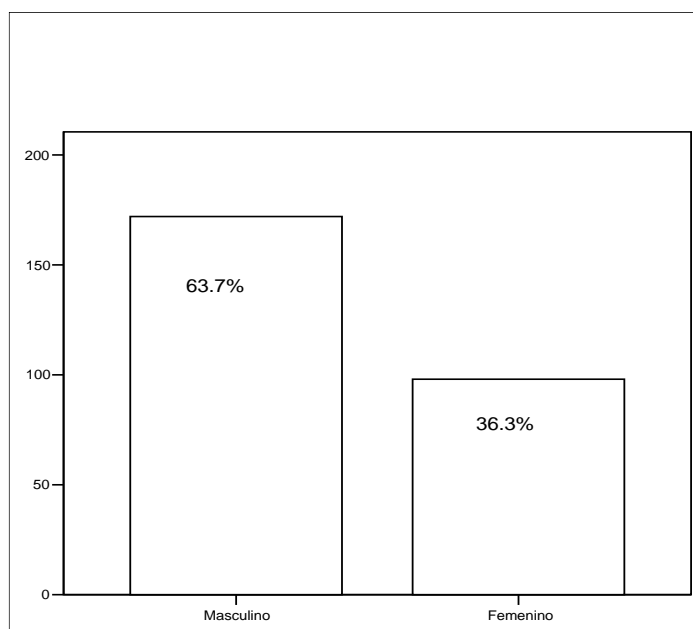
Tabla I
Distribución de Frecuencias de la variable
Género del estudiante

	Frecuencia	Porcentaje	Porcentaje acumulado
Masculino	172	63,7	63,7
Femenino	98	36,3	100
Total	270	100	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

Elaboración: Betzabeth Yonson S

Gráfico 4.1
Histograma de frecuencias de la variable
Género del estudiante



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

Elaboración: Betzabeth Yonson S

Variable X₂: Edad

La tabla II muestra los estimadores poblacionales correspondientes a la edad de los estudiantes entrevistados, la edad promedio, en años es de 21.52, con una dispersión de los datos alrededor de la media de 2.46 años establecidos por la desviación estándar, mientras que la mediana nos indica que el 50% de los estudiantes tiene una edad menor o igual a 21 años; la edad mínima observada de un estudiante es 18 años y la máxima es 35; la edad que más veces se repite entre los entrevistados es 20, es decir la Moda; la distribución de la variable es sesgada a la derecha de la curva con un sesgo de 2.071, pues el signo positivo describe la asimetría de los datos con respecto a su media (ver gráfico 4.2).

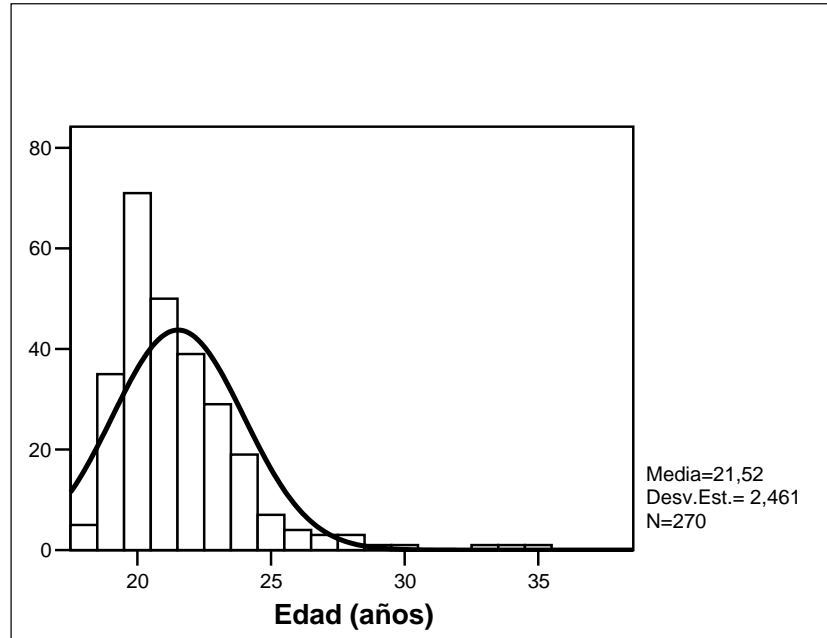
La tabla II y el gráfico 4.3 permiten apreciar que el 25% de los entrevistados tienen edades menores o iguales a 20 años y el 25% tienen edades mayores o iguales a 23 años.

Tabla II
Estadística Descriptiva de la variable
Edad

N		270
Media		21,52
Mediana		21,00
Moda		20
Desviación estándar		2,461
Varianza		6,057
Sesgo		2,071
Curtosis		7,009
Rango		17
Mínimo		18
Máximo		35
Percentiles	25	20,00
	50	21,00
	75	23,00

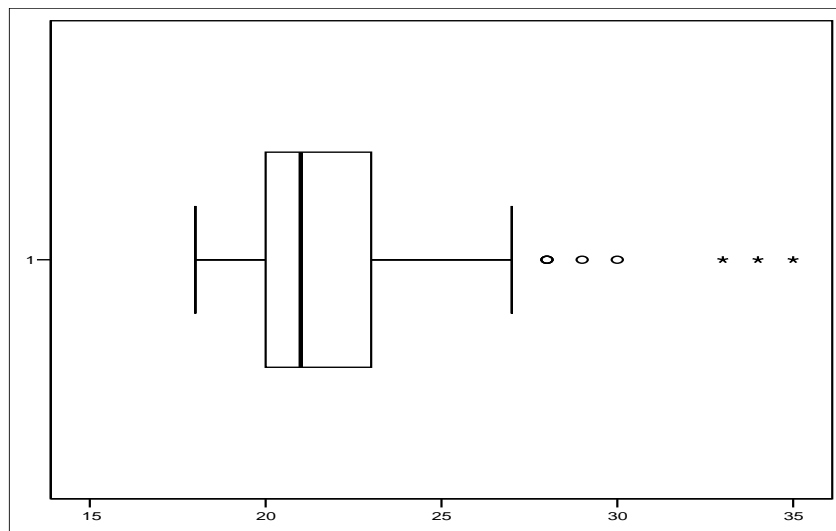
Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.2
Distribución de la variable
Edad



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.3
Diagrama de Cajas de la variable
Edad



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₃: Carrera

En el cuadro 3.3 se muestran la distribución de los estudiantes registrados seleccionados en la muestra.

Variable X₄: Factor P

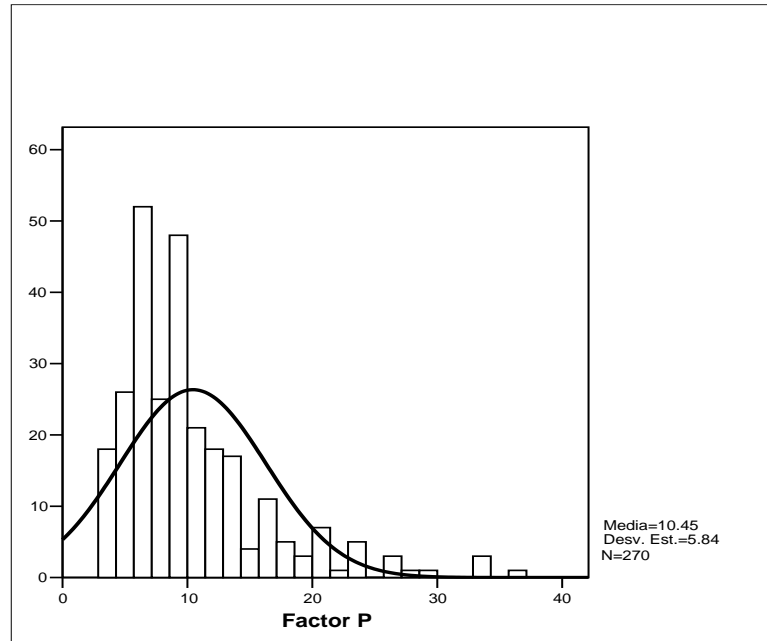
La tabla III muestra, que el Factor P promedio, es de 10.45, la distribución de la variable es asimétrica positiva, por lo tanto existe mayor concentración de los datos que se encuentran a la izquierda de la media, (ver gráfico 4.4); una dispersión de los datos alrededor de la media de 5.8 establecidos por la desviación estándar, mientras que la mediana nos indica que el 50% de los estudiantes tienen un factor P menor o igual a 9; el factor P mínimo observado de un estudiante es 3 y el máxima es 36; el factor P que más veces se repite entre los entrevistados es 7, es decir la Moda; La tabla III y el gráfico 4.5 permiten apreciar que el 25% de los entrevistados tienen factores P menores o iguales a 7 y el 25% tienen factores P mayores o iguales a 12 .

Tabla III
Estadística Descriptiva de la variable
Factor P

N	270	
Media	10.45	
Mediana	9	
Moda	7	
Desviación estándar	5.841	
Varianza	34.114	
Sesgo	1.671	
Curtosis	3.554	
Rango	33	
Mínimo	3	
Máximo	36	
Percentiles		
	25	7
	50	9
	75	12

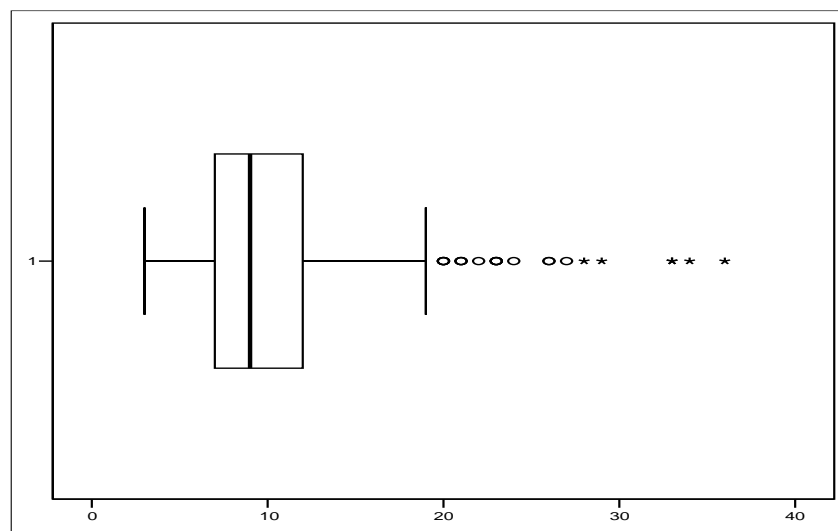
Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.4
Distribución de la variable
Factor P



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.5
Diagrama de Cajas de la variable
Factor P



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₅: Promedio

La tabla IV muestra los estimadores poblacionales correspondientes al Promedio de los estudiantes entrevistados, la media es de 7.33, con una dispersión de los datos alrededor de la media de 0.66 establecidos por la desviación estándar, mientras que la mediana nos indica que el 50% de los estudiantes tienen un Promedio menor o igual a 7.16; el Promedio mínima observado de un estudiante es 6.10 y la máxima es 9.73; el Promedio que más veces se repite entre los entrevistados es 6.80, es decir la Moda; la distribución de la variable es sesgada a la derecha de la curva con un sesgo de 1.1, pues el signo positivo describe la asimetría de los datos con respecto a su media (ver gráfico 4.6).

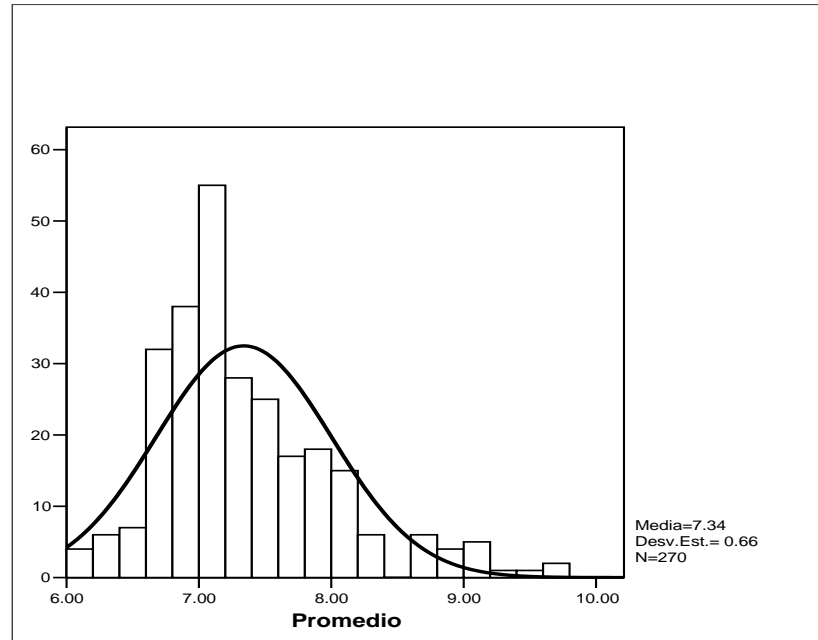
La tabla IV y el gráfico 4.7 permiten apreciar que el 25% de los entrevistados tienen Promedios menores o iguales a 6.9 y el 25% tienen Promedios mayores o iguales a 7.67.

Tabla IV
Estadística Descriptiva de la variable
Promedio

N		270
Media		7.339
Mediana		7.16
Moda		6.8
Desviación estándar		0.663
Varianza		0.44
Sesgo		1.106
Curtosis		1.429
Rango		3.63
Mínimo		6.1
Máximo		9.73
Percentiles	25	6.90
	50	7.16
	75	7.67

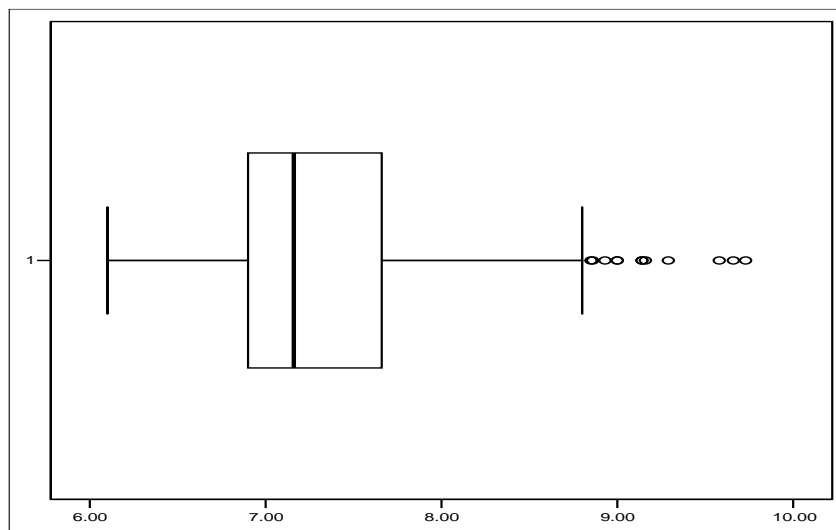
Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.6
Distribución de la variable
Promedio



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.7
Diagrama de Cajas de la variable
Promedio



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₆: Rendimiento Académico

La tabla V muestra los estimadores poblacionales correspondientes al Rendimiento académico de los estudiantes entrevistados, el Rendimiento académico promedio es de 81.89%, con una dispersión de los datos alrededor de la media de 17.37% establecidos por la desviación estándar, mientras que la mediana nos indica que el 50% de los estudiantes tienen un Rendimiento académico menor o igual a 85%; el Rendimiento académico mínimo observado de un estudiante es 22% y la máxima es 100%; la Moda es de 100%; la distribución de la variable es sesgada a la izquierda de la curva, pues el signo negativo del sesgo describe la asimetría de los datos con respecto a su media (ver gráfico 4.8).

La tabla V y el gráfico 4.9 permiten apreciar que el 25% de los entrevistados tienen Rendimiento académico menores o iguales a 70.54% y el 25% tienen Rendimiento académico igual a 100%.

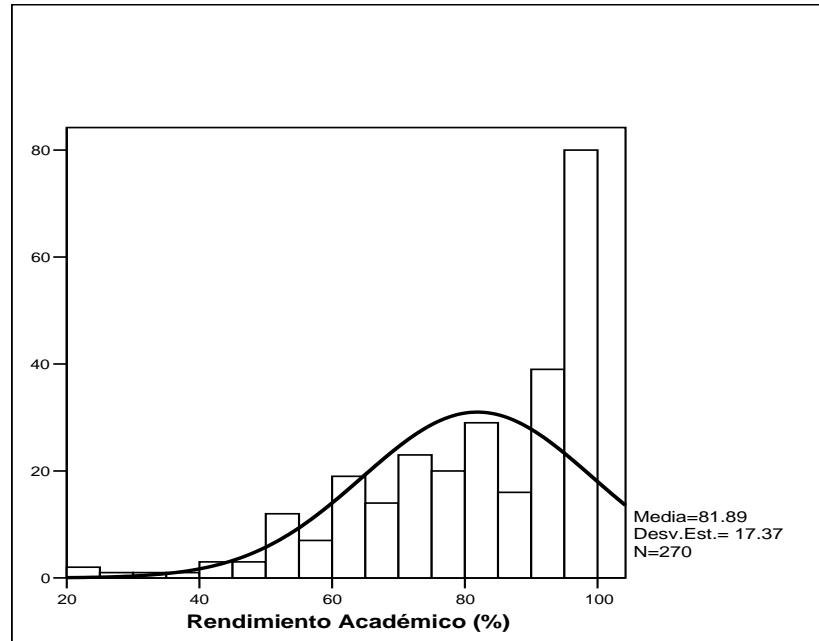
Tabla V
Estadística Descriptiva de la variable
Rendimiento Académico

N		270
Media		81.89
Mediana		84.99
Moda		100
Desviación estándar		17.371
Varianza		301.749
Sesgo		-0.903
Curtosis		0.342
Rango		78
Mínimo		22
Máximo		100
Percentiles	25	70.54
	50	84.99
	75	100

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

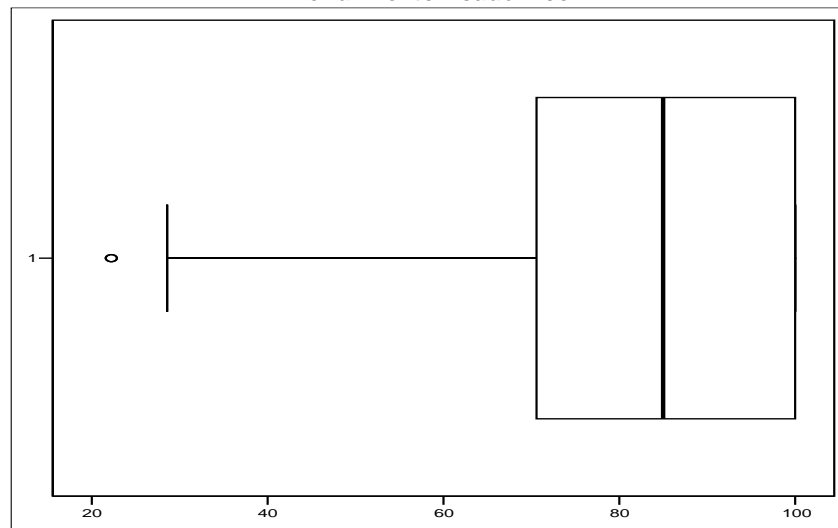
Elaboración: Betzabeth Yonson S

Gráfico 4.8
Distribución de la variable
Rendimiento Académico



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.9
Diagrama de Cajas de la variable
Rendimiento Académico



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₇: Peso

La tabla VI muestra los estimadores poblacionales correspondientes al Peso de los estudiantes entrevistados, el Peso promedio es de 140.86 lbs, con una dispersión de los datos alrededor de la media de 26.48 lbs establecidos por la desviación estándar, mientras que la mediana nos indica que el 50% de los estudiantes tiene un Peso menor o igual a 135 lbs; el peso mínimo observado de un estudiante es 97 lbs y el máximo es 230 lbs; el peso que más veces se repite entre los entrevistados es 120 lbs, es decir la Moda; la distribución de la variable es ligeramente sesgada a la derecha de la curva con un sesgo de 0.725, pues el signo positivo describe la asimetría de los datos con respecto a su media (ver gráfico 4.10).

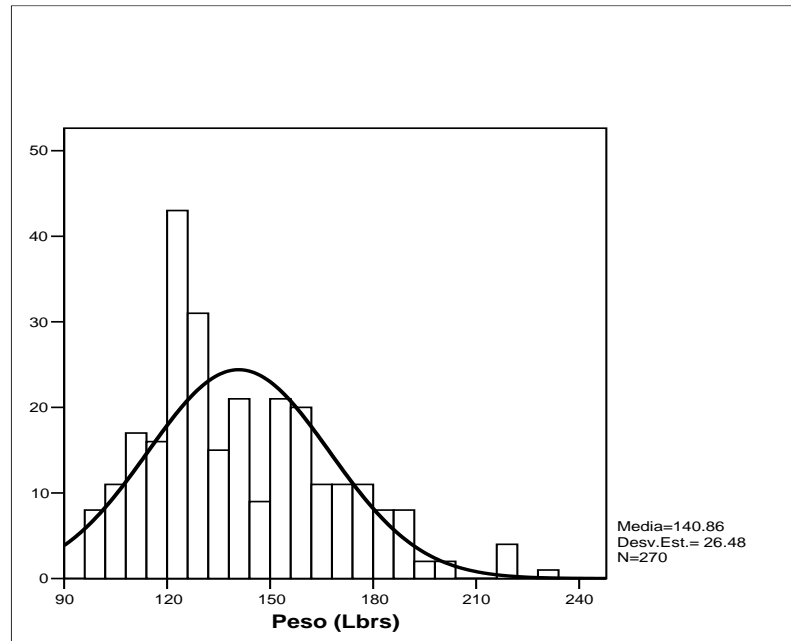
La tabla VI y el gráfico 4.11 permiten apreciar que el 25% de los entrevistados tienen pesos menores o iguales a 120 lbs y el 25% tienen pesos mayores o iguales a 160 lbs.

Tabla VI
Estadística Descriptiva de la variable
Peso

N	270
Media	140.86
Mediana	135.00
Moda	120
Desviación estándar	26.481
Varianza	701.251
Sesgo	0.725
Curtosis	0.207
Rango	133
Mínimo	97
Máximo	230
Percentiles	25 120
	50 135
	75 160

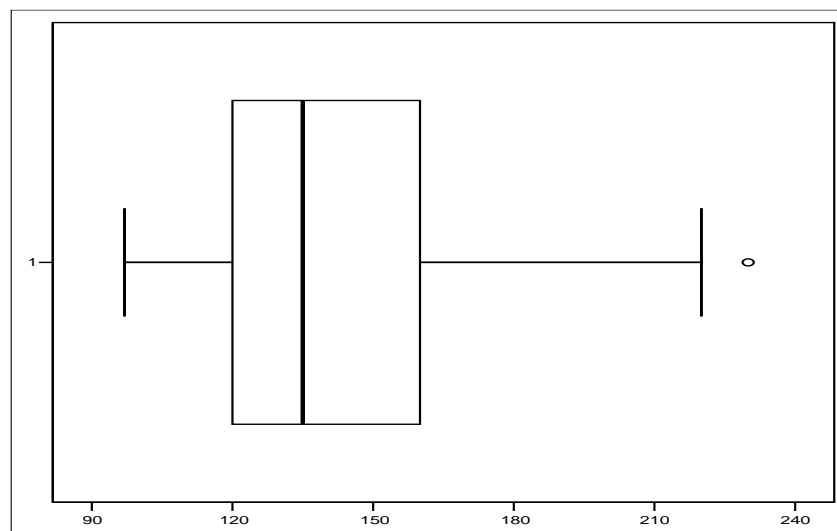
Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.10
Distribución de la variable
Peso



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.11
Diagrama de Cajas de la variable
Peso



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₈: Estatura

De acuerdo con los datos obtenidos de los 270 estudiantes entrevistados de la ESPOL, la tabla VII, muestra que el promedio de las estaturas de los estudiantes es de 167.29 cmts, con una dispersión de los datos alrededor de la media de 8.33 cmts, establecidos por la desviación estándar, mientras que la mediana nos indica que el 50% de los estudiantes tienen una estatura menor o igual a 167 cmts; la estatura mínima observada de un estudiante es de 148 cmts y la máxima es de 190; la estatura que más sobresale entre los entrevistados es de 160 cmts, es decir la Moda; el sesgo es de 0.109, pues el signo positivo describe la asimetría de los datos con respecto a su media (ver gráfico 4.12).

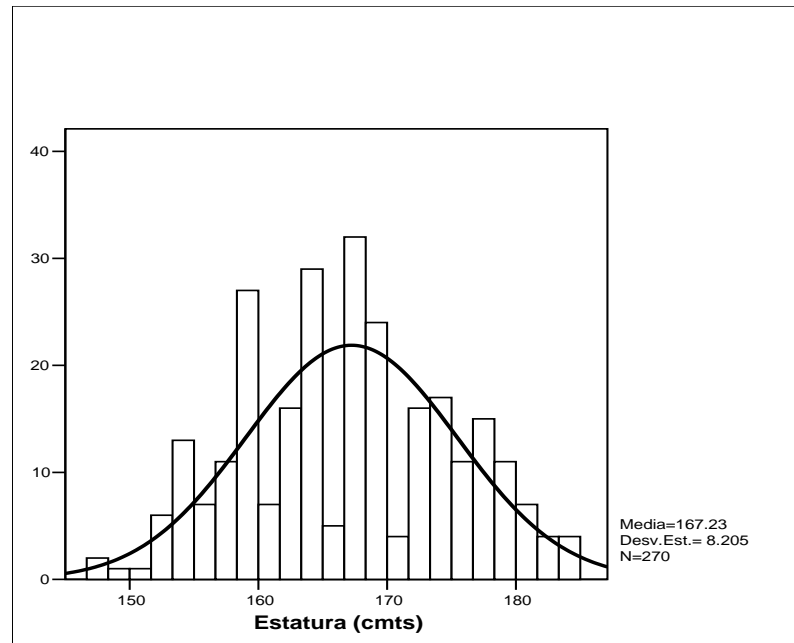
La tabla VII y el gráfico 4.13 permiten apreciar que el 25% de los entrevistados tienen estaturas menores o iguales a 160 cmts y el 25% tienen estaturas mayores o iguales a 174 cmts.

Tabla VII
Estadística Descriptiva de la variable
Estatura

N		270
Media		167,23
Mediana		167,00
Moda		160
Desviación estándar		8,205
Varianza		67,325
Sesgo		0,036
Curtosis		-0,722
Rango		37
Mínimo		148
Máximo		185
Percentiles	25	160
	50	167
	75	174

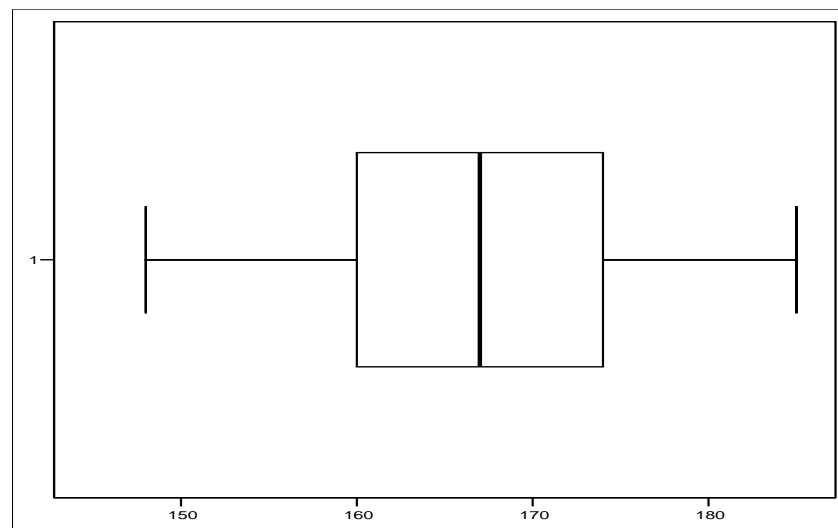
Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.12
Distribución de la variable
Estatura



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.13
Diagrama de Cajas de la variable
Estatura



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₉: Trabaja

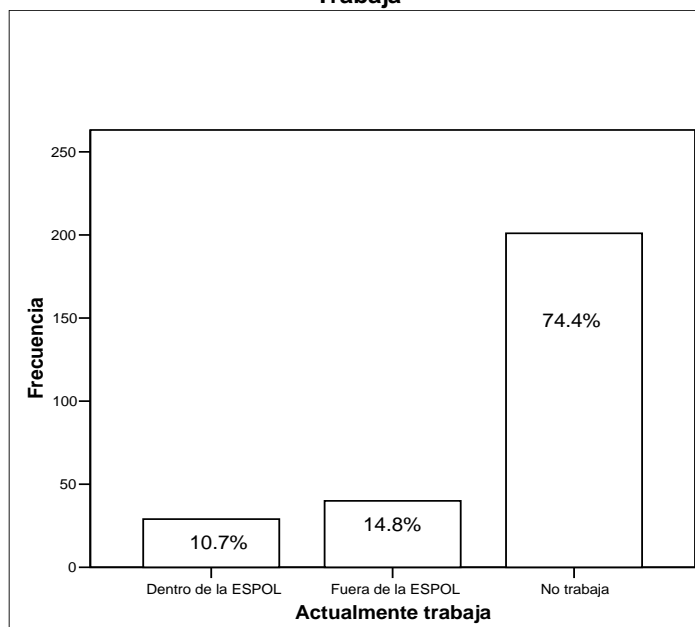
Según los resultados obtenidos en la tabla VIII, de los 270 estudiantes entrevistados se muestra que el 10.7% de ellos trabajan dentro de la ESPOL, el 14.8% trabajan fuera de la ESPOL y el 74.4% no trabajan. Como se puede observar la gran mayoría de los estudiantes, 74.4%, no trabaja. Estos resultados se pueden verificar también en el gráfico 4.15

Tabla VIII
Distribución de Frecuencias de la variable
Trabaja

	Frecuencia	Porcentaje	Porcentaje acumulado
Dentro de la Espol	29	10,7	10,7
Fuera de la Espol	40	14,8	25,6
No trabaja	201	74,4	100
Total	270	100	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.14
Histograma de frecuencias de la variable
Trabaja



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X_{10} : Número de horas que diariamente permanece en la ESPOL

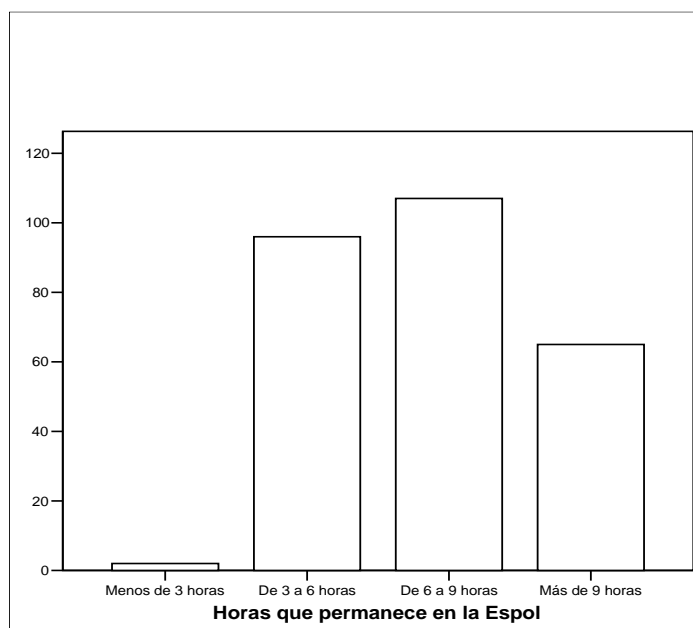
La tabla IX muestra, que de los 270 estudiantes entrevistados el 0.7% de ellos permanecen menos de tres horas en la ESPOL, el 35.6% permanecen de tres a seis horas en la ESPOL, el 39.6% permanecen de seis a nueve horas en la ESPOL y el 24.1% permanecen mas de nueve horas en la ESPOL. Estos resultados se pueden verificar también en el gráfico 4.15

Tabla IX
Distribución de Frecuencias de la variable
Horas que permanece en la Espol

	Frecuencia	Porcentaje	Porcentaje acumulado
Menos de 3 horas	2	0,7	0,7
De 3 a 6 horas	96	35,6	36,3
De 6 a 9 horas	107	39,6	75,9
Más de 9 horas	65	24,1	100,0
Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.15
Histograma de frecuencias de la variable
Horas que permanece en la Espol



Variable X_{11} : Número de días a la semana que realiza algún tipo de actividad deportiva

La tabla X muestra, que de los 270 estudiantes entrevistados el 64.4% de ellos realizan alguna actividad deportiva entre cero y un día a la semana, el 24.8% lo realizan dentro de dos a tres días a la semana, el 7.4% de cuatro a cinco días y el 3.3% de seis a siete días. Como se puede observar la gran mayoría de los estudiantes, 64.4%, practica deporte por lo menos un día a la semana. Estos resultados se pueden verificar también en el gráfico 4.16.

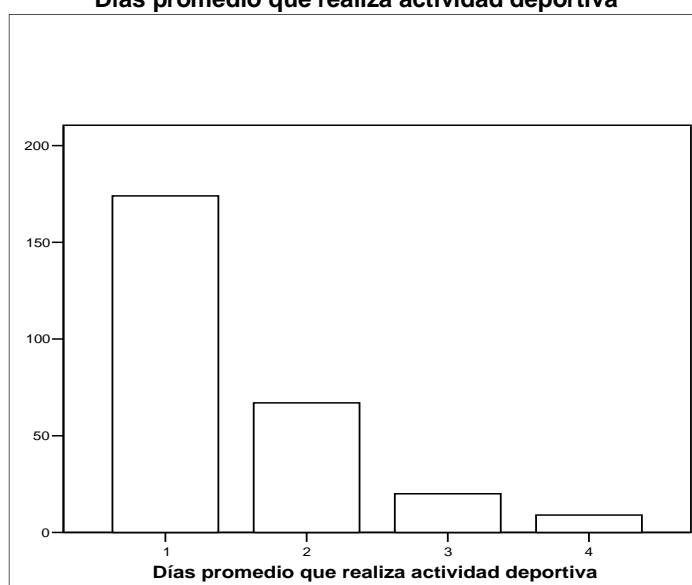
Tabla X
Distribución de Frecuencias de la variable
Días promedio que realiza actividad deportiva

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	0 a 1 día	174	64,4	64,4
2	2 a 3 días	67	24,8	89,3
3	4 a 5 días	20	7,4	96,7
4	6 a 7 días	9	3,3	100,0
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

Elaboración: Betzabeth Yonson S

Gráfico 4.16
Histograma de frecuencias de la variable
Días promedio que realiza actividad deportiva



Variable X₁₂: ¿Cómo considera la alimentación para desempeñar su rol de estudiante?

La tabla XI muestra, que de los 270 estudiantes entrevistados el 1.1% de ellos consideran nada importante la alimentación para su desempeño como estudiante, el 2.6% consideran poco importante la alimentación para su desempeño como estudiante, el 18.5% lo consideran importante, el 42.2% lo consideran muy importante y el 35.6% lo consideran extremadamente importante. Estos resultados se pueden verificar también en el gráfico 4.17.

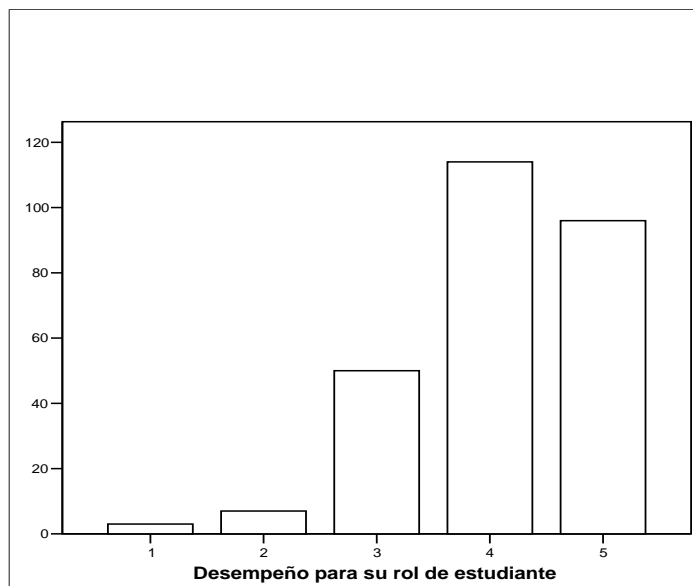
Tabla XI
Distribución de Frecuencias de la variable
Desempeño para su rol de estudiante

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	Nada Importante	3	1,1	1,1
2	Poco Importante	7	2,6	3,7
3	Importante	50	18,5	22,2
4	Muy Importante	114	42,2	64,4
5	Extremadamente Importante	96	35,6	100,0
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

Elaboración: Betzabeth Yonson S

Gráfico 4.17
Histograma de frecuencias de la variable
Desempeño para su rol de estudiante



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₁₃: Porcentaje de su alimentación es a través de comida preparada dentro de casa

La tabla XII muestra, que de los 270 estudiantes entrevistados, el promedio de su alimentación hecha en casa es de 58,93%, la distribución de la variable es ligeramente sesgada a la derecha de la media, pues el signo negativo del sesgo describe la asimetría de los datos con respecto a su media (ver gráfico 4.18); una dispersión de los datos alrededor de la media de 25.23% establecidos por la desviación estándar, mientras que la mediana nos indica que el 50% de los estudiantes, de su alimentación es hecha en casa es menor o igual al 60%; el mínimo observado de un estudiante que se alimenta dentro de casa es 5% y el máximo es 100%; la Moda es de 50%.

La tabla XII y el gráfico 4.19 permiten apreciar que el 25% de los entrevistados su alimentación hecha en casa es menor o igual a 40% y

el 25% de ellos su alimentación dentro de casa son mayores o iguales al 80%.

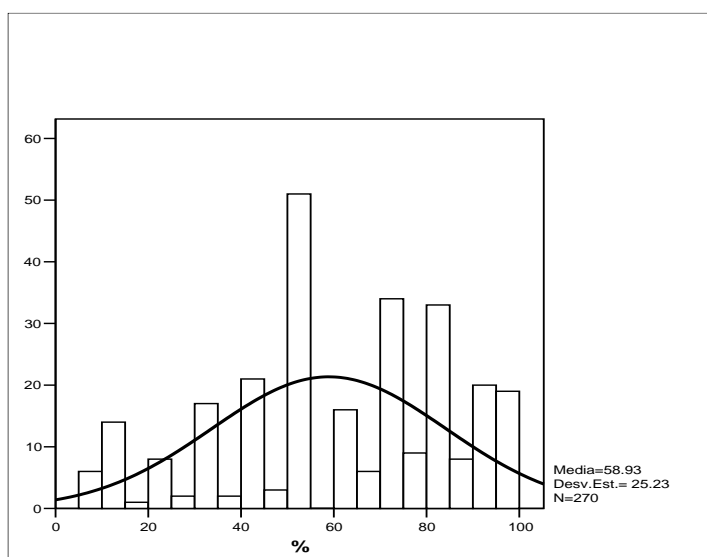
Tabla XII
Estadística Descriptiva de la variable
Porcentaje de alimentación en casa

N		270
Media		58,93
Mediana		60,00
Moda		50
Desviación estándar		25,231
Varianza		636,620
Sesgo		-0,350
Curtosis		-0,705
Rango		95
Mínimo		5
Máximo		100
Percentiles	25	40
	50	60
	75	80

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

Elaboración: Betzabeth Yonson S

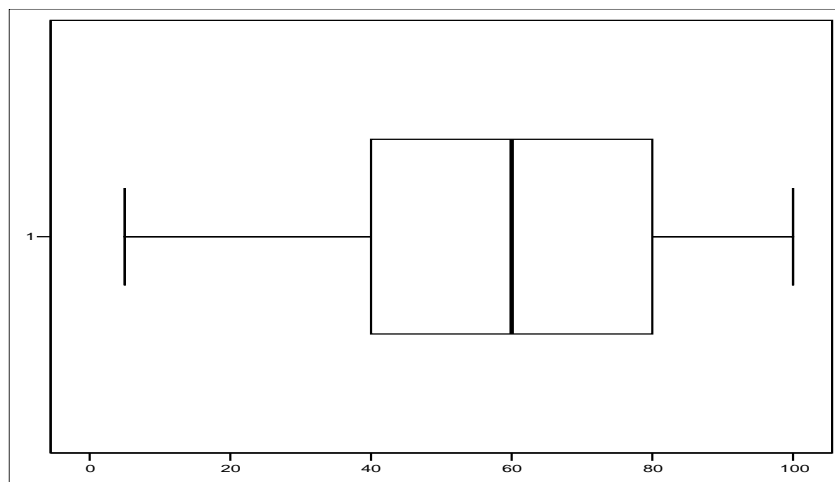
Gráfico 4.18
Distribución de la variable
Porcentaje de alimentación en casa



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

Elaboración: Betzabeth Yonson S

Gráfico 4.19
Diagrama de Cajas de la variable
Porcentaje de alimentación en casa



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₁₄: Días a la semana que desayuna el estudiante

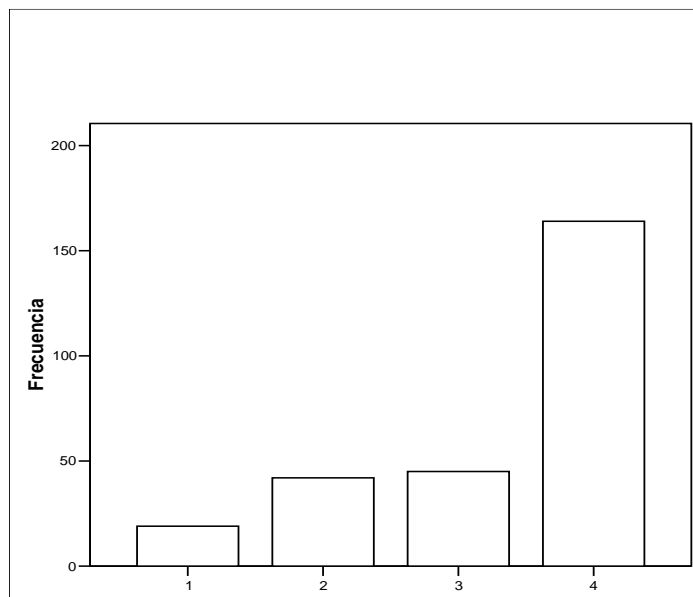
La tabla XIII muestra, que de los 270 estudiantes entrevistados el 7% de ellos desayunan entre cero y un día a la semana, el 15.6% desayunan dentro de dos a tres días, el 16.7% de cuatro a cinco días a la semana y el 60.7% de seis a siete días. Como se puede observar la gran mayoría de los estudiantes, 60.7%, desayunan. Véase también gráfico 4.20.

Tabla XIII
Distribución de Frecuencias de la variable
Días a la semana que desayuna

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	0 a 1día	19	7	7
2	2 a 3días	42	15,6	22,6
3	4 a 5días	45	16,7	39,3
4	6 a 7días	164	60,7	100
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.20
Histograma de frecuencias de la variable
Días a la semana que desayuna



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₁₅: Días a la semana que come carne roja

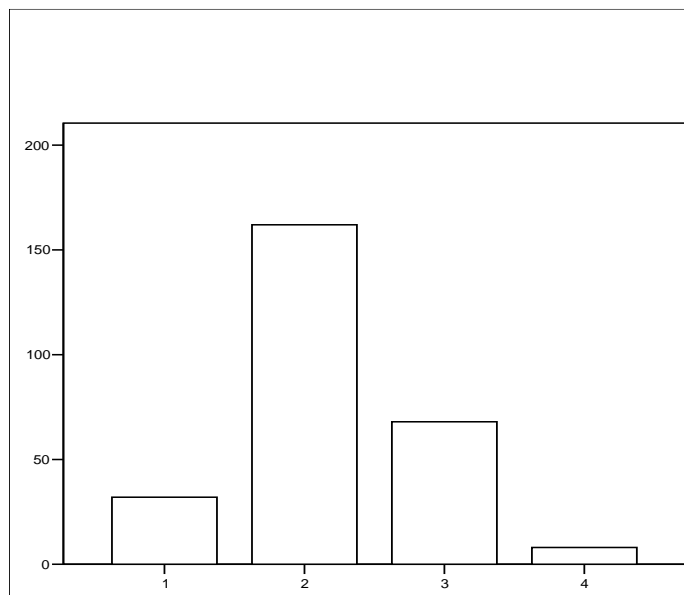
La tabla XIV muestra, que de los 270 estudiantes entrevistados el 11.9% de ellos comen carne roja entre cero y un día a la semana, el 60% comen carne roja dentro de dos a tres días a la semana, el 25.2% de cuatro a cinco días a la semana comen carne roja y el 3% de seis a siete días. Véase también gráfico 4.21.

Tabla XIV
Distribución de Frecuencias de la variable
Días a la semana que come carne roja

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	0 a 1día	32	11,9	11,9
2	2 a 3días	162	60,0	71,9
3	4 a 5días	68	25,2	97,0
4	6 a 7días	8	3,0	100,0
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.21
Histograma de frecuencias de la variable
Días a la semana que come carne roja



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₁₆: Días a la semana que come pescado

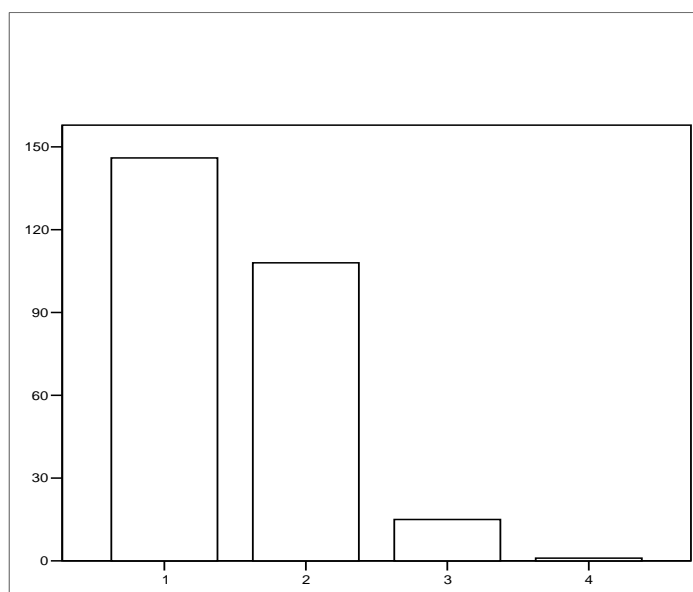
La tabla XV muestra, que de los 270 estudiantes entrevistados el 54.1% de ellos comen pescado entre cero y un día a la semana, el 40% comen pescado dentro de dos a tres días a la semana y el 4% de seis a siete días. Estos resultados se pueden verificar también en el gráfico 4.22.

Tabla XV
Distribución de Frecuencias de la variable
Días a la semana que come pescado

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	0 a 1día	146	54,1	54,1
2	2 a 3días	108	40,0	94,1
3	4 a 5días	15	5,6	99,6
4	6 a 7días	1	,4	100,0
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.22
Histograma de frecuencias de la variable

Días a la semana que come pescado

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₁₇: Días a la semana que come pollo

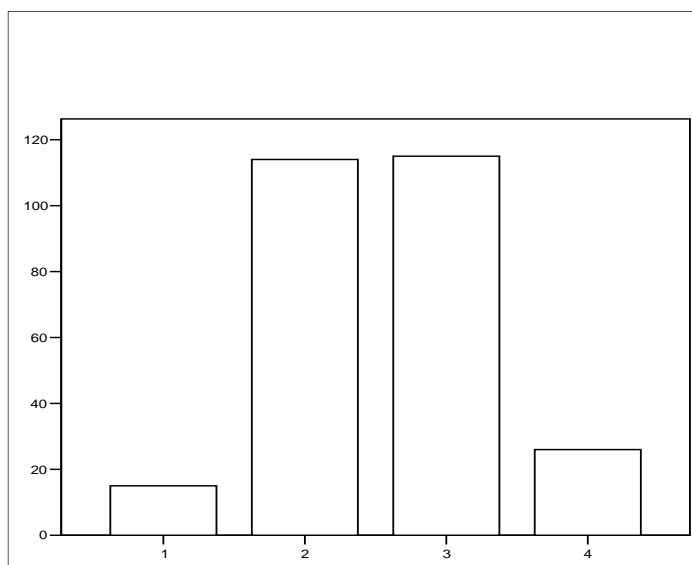
La tabla XVI muestra, que de los 270 estudiantes entrevistados el 5.6% de ellos come pollo entre cero y un día a la semana, el 42.2% come pollo dentro de dos a tres días a la semana, el 42.6% de cuatro a cinco días a la semana come pollo y el 9.6% de seis a siete días. Estos resultados se pueden verificar también en el gráfico 4.23.

Tabla XVI
Distribución de Frecuencias de la variable
Días a la semana que come pollo

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	0 a 1día	15	5,6	5,6
2	2 a 3días	114	42,2	47,8
3	4 a 5días	115	42,6	90,4
4	6 a 7días	26	9,6	100,0
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.23
Histograma de frecuencias de la variable

Días a la semana que come pollo

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₁₈: Días a la semana que toma leche

La tabla XVII muestra, que de los 270 estudiantes entrevistados el 21.1% de ellos toma leche entre cero y un día a la semana, el 24.8% toma leche dentro de dos a tres días a la semana, el 24.1% de cuatro a cinco días a la semana toma leche y el 30% de seis a siete días. Estos resultados se pueden verificar también en el gráfico 4.23.

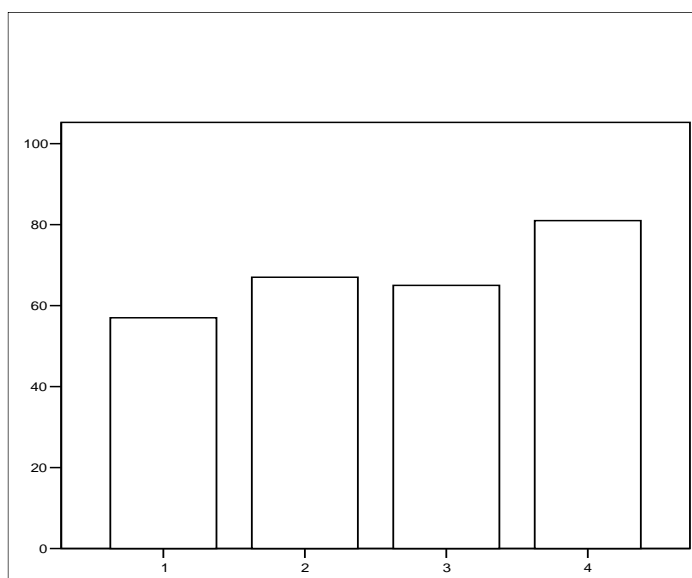
Tabla XVII
Distribución de Frecuencias de la variable
Días a la semana que toma leche

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	0 a 1día	57	21,1	21,1
2	2 a 3días	67	24,8	45,9
3	4 a 5días	65	24,1	70,0
4	6 a 7días	81	30,0	100,0
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.24
Histograma de frecuencias de la variable

Días a la semana que toma leche



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X_{19} : Días a la semana que consume vegetales

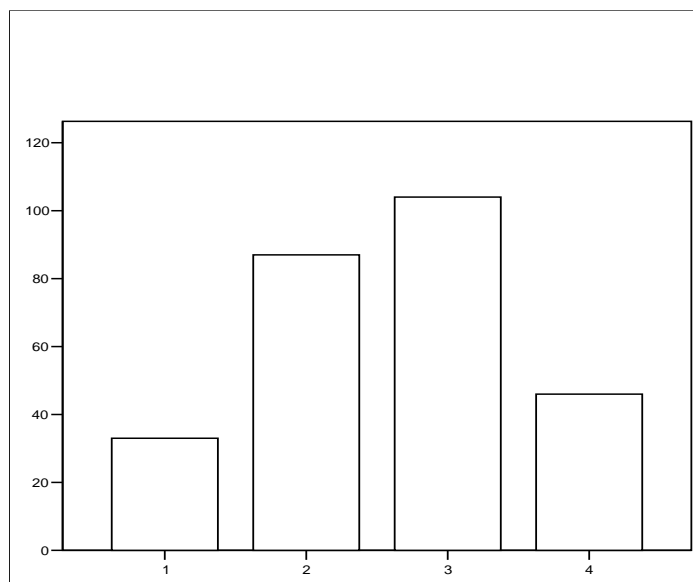
La tabla XVIII muestra, que de los 270 estudiantes entrevistados el 38.5% de cuatro a cinco días a la semana come vegetales y el 17% de seis a siete días. Estos resultados se pueden verificar en el gráfico 4.25.

Tabla XVIII
Distribución de Frecuencias de la variable
Días a la semana que consume vegetales

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	0 a 1día	33	12,2	12,2
2	2 a 3días	87	32,2	44,4
3	4 a 5días	104	38,5	83,0
4	6 a 7días	46	17,0	100,0
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.25
Histograma de frecuencias de la variable
Días a la semana que consume vegetales



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₂₀: Días a la semana que come frutas

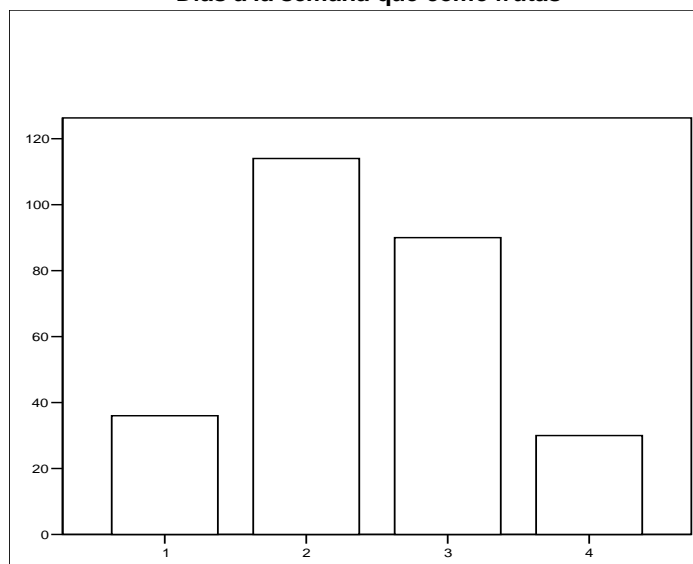
La tabla XIX muestra, que de los 270 estudiantes entrevistados el 13.3% de ellos comen frutas entre cero y un día a la semana, el 42.2% comen frutas dentro de dos a tres días a la semana, el 33.3% de cuatro a cinco días a la semana comen frutas y el 11.1% de seis a siete días. Estos resultados se pueden verificar también en el gráfico 4.26.

Tabla XIX
Distribución de Frecuencias de la variable
Días a la semana que come frutas

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	0 a 1día	36	13,3	13,3
2	2 a 3días	114	42,2	55,6
3	4 a 5días	90	33,3	88,9
4	6 a 7días	30	11,1	100,0
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.26
Histograma de frecuencias de la variable
Días a la semana que come frutas



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₂₁: Comidas que realiza en el día

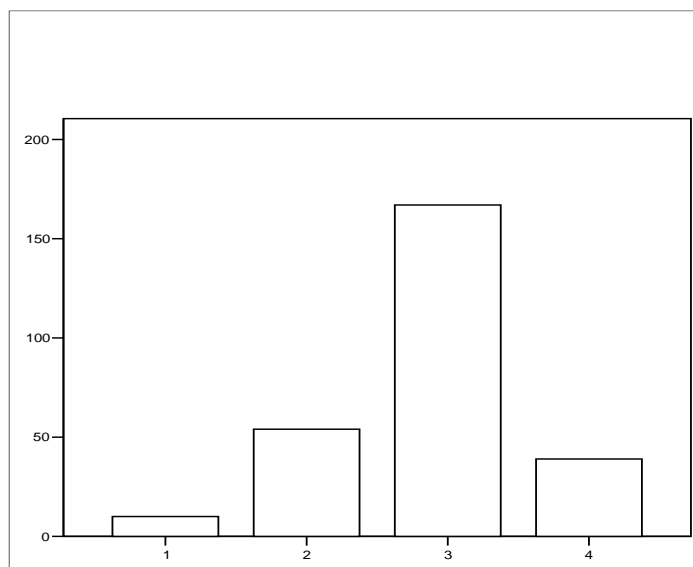
La tabla XX muestra, que de los 270 estudiantes entrevistados el 3.7% de ellos come una vez en el día, el 20% comen dos veces en el día, el 61.9% comen tres veces en el día y el 14.4% comen más de tres veces en el día. Estos resultados se pueden verificar también en el gráfico 4.27.

Tabla XX
Distribución de Frecuencias de la variable
Comidas que realiza en el día

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	Una vez	10	3,7	3,7
2	Dos veces	54	20,0	23,7
3	Tres veces	167	61,9	85,6
4	Más de tres veces	39	14,4	100
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.27
Histograma de frecuencias de la variable
Comidas que realiza en el día



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₂₂: En la hora de alimentación que tan frecuente opta por comida rápida (comida chatarra)

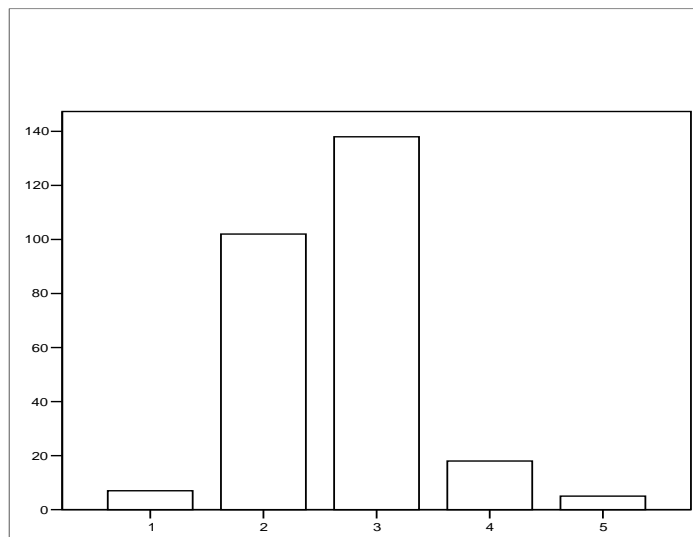
La tabla XXI muestra, que de los 270 estudiantes entrevistados el 51.1% algunas veces consumen este tipo de comidas y el 1.9% siempre la consumen. Para más detalles véase en el gráfico 4.28.

Tabla XXI
Distribución de Frecuencias de la variable
En la hora de alimentación que tan frecuente opta por comida rápida

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	Nunca	7	2,6	2,6
2	Rara vez	102	37,8	40,4
3	Algunas veces	138	51,1	91,5
4	La mayoría de las veces	18	6,7	98,1
5	Siempre	5	1,9	100
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.28
Histograma de frecuencias de la variable
En la hora de alimentación que tan frecuente opta por comida rápida



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₂₃: Horario fijo para su alimentación diaria

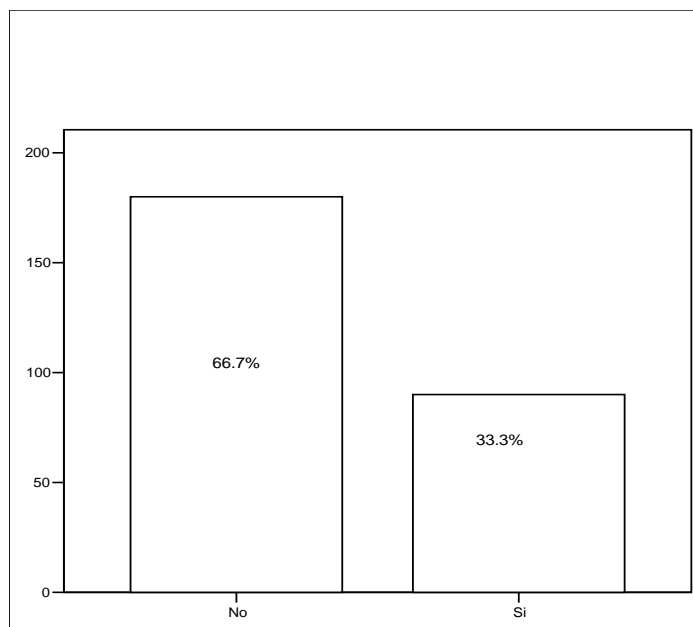
Se analiza ahora si el estudiante entrevistado tiene un horario fijo para su alimentación diaria. En la tabla XXII se puede notar de los 270 estudiantes entrevistados en el II término del 2005, el 66.7% no tienen un horario fijo para su alimentación diaria y el 33.3% si tienen un horario fijo para su alimentación diaria. Estos resultados se pueden verificar también en el gráfico 4.29

Tabla XXII
Distribución de Frecuencias de la variable
Horario fijo para su alimentación diaria

	Frecuencia	Porcentaje	Porcentaje acumulado
No	180	66,7	66,7
Si	90	33,3	100
Total	270	100	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.29
Histograma de frecuencias de la variable
Horario fijo para su alimentación diaria



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₂₄: Tipos de bebida de preferencia para satisfacer su sed

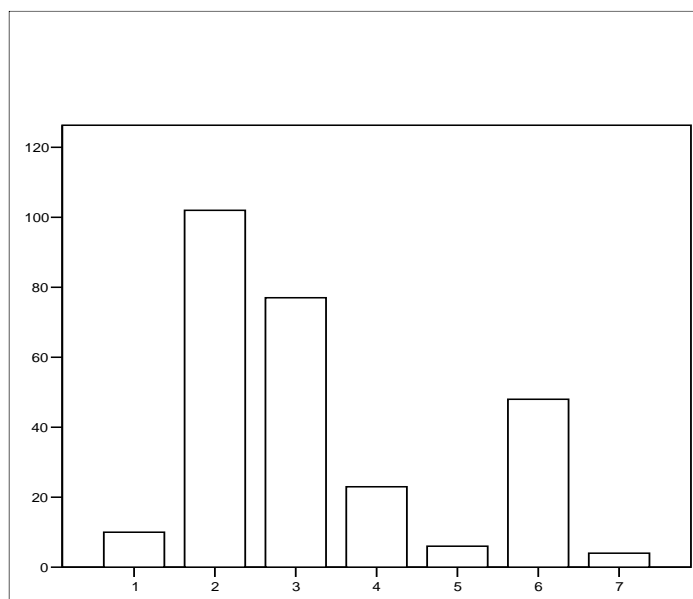
La tabla XXIII muestra, que de los 270 estudiantes entrevistados la bebida de mayor preferencia es el agua sin gas con un 37.7% seguido de los jugos preparados con un 28.5% y las gaseosas con un 17.8%, los restantes datos respecto a esta variable se presentan en el gráfico 4.30.

Tabla XXIII
Distribución de Frecuencias de la variable
Tipos de bebida de preferencia para satisfacer su sed

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	Agua con gas	10	3,7	3,7
2	Agua sin gas	102	37,8	41,5
3	Jugos preparados	77	28,5	70,0
4	Jugos envasados	23	8,5	78,5
5	Bebida con contenido alcohólico	6	2,2	80,7
6	Gaseosas	48	17,8	98,5
7	Otros	4	1,5	100,0
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.30
Histograma de frecuencias de la variable
Tipos de bebida de preferencia para satisfacer su sed



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₂₅: Frecuencia que Fuma

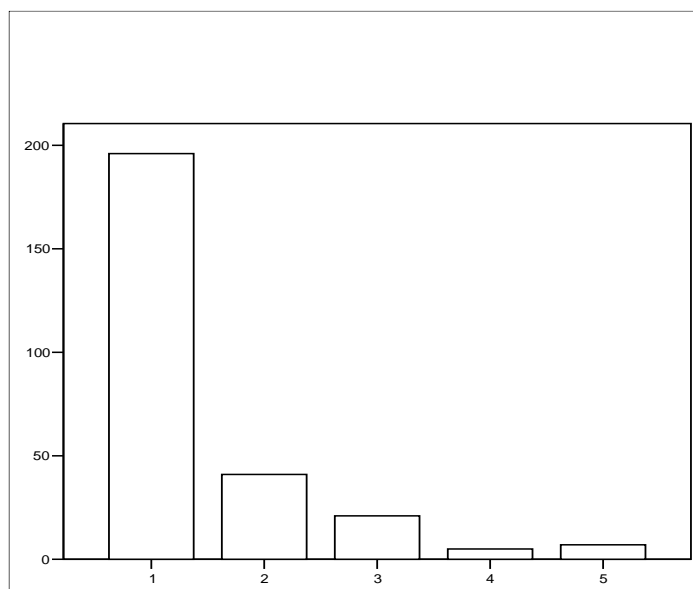
De los 270 estudiantes entrevistados, el 72.6% de ellos declaró que nunca fuman mientras que solo el 2.6% declaró que siempre fuman, para más detalles véase la tabla XXIV. Estos resultados se pueden verificar también en el gráfico 4.31.

Tabla XXIV
Distribución de Frecuencias de la variable
Frecuencia que fuma

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	Nunca	196	72,6	72,6
2	Rara vez	41	15,2	87,8
3	Algunas veces	21	7,8	95,6
4	La mayoría de las veces	5	1,9	97,4
5	Siempre	7	2,6	100,0
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.31
Histograma de frecuencias de la variable
Frecuencia que fuma



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₂₆: Frecuencia que ingiere bebida alcohólica

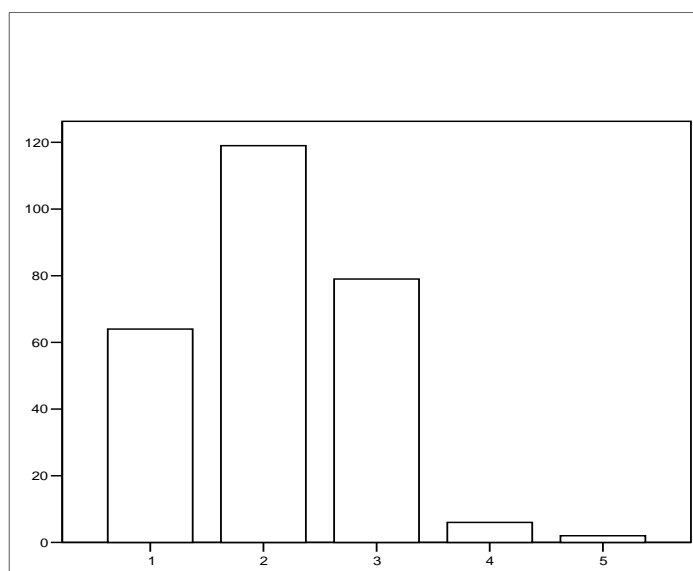
De los 270 estudiantes entrevistados, el 23.7% de ellos declaró que nunca ingieren bebida alcohólica mientras que solo el 0.7% declaró que siempre consumen este tipo de bebida, para más detalles véase la tabla XXV. Estos resultados se pueden verificar también en el gráfico 4.32.

Tabla XXV
Distribución de Frecuencias de la variable
Frecuencia que ingiere bebida alcohólica

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	Nunca	64	23,7	23,7
2	Rara vez	119	44,1	67,8
3	Algunas veces	79	29,3	97,0
4	La mayoría de las veces	6	2,2	99,3
5	Siempre	2	0,7	100,0
	Total	270	100,0	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.32
Histograma de frecuencias de la variable
Frecuencia que ingiere bebida alcohólica



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₂₇: Problema de Gastritis

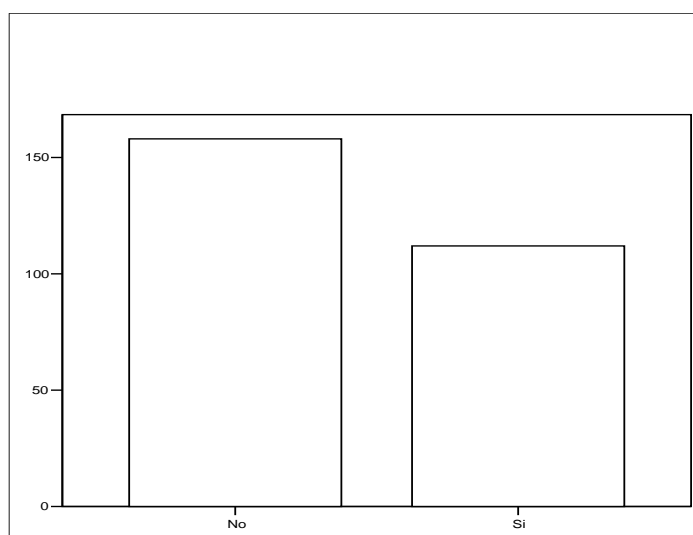
Se analiza ahora si el estudiante entrevistado a lo largo de su estancia en la ESPOL, ha pasado o pasa por el problema de gastritis, como se muestra en la tabla XXVI encontramos que el 58.5% no han padecido este problema mientras que el 41.5% lo han padecido o padecen de este problema. Estos resultados se pueden verificar también en el gráfico 4.33

Tabla XXVI
Distribución de Frecuencias de la variable
Problema de Gastritis

	Frecuencia	Porcentaje	Porcentaje acumulado
No	158	58,5	58,5
Si	112	41,5	100
Total	270	100	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.33
Histograma de frecuencias de la variable
Problema de Gastritis



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

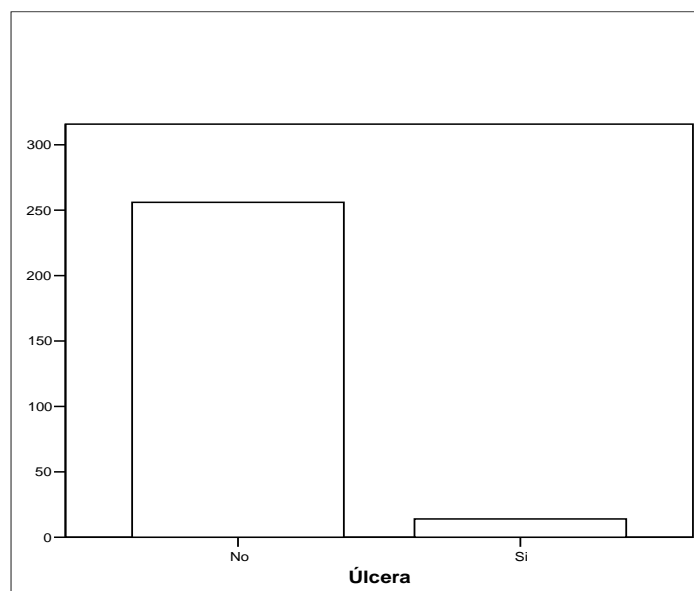
Variable X₂₈: Problema de Úlcera

Se analiza ahora si el estudiante entrevistado a lo largo de su estancia en la ESPOL, ha pasado o pasa por el problema de úlcera, como se muestra en la tabla XXVII encontramos que el 94.8% no han padecido este problema mientras que el 5.2% si han padecido o padecen de este problema. Estos resultados se pueden verificar también en el gráfico 4.34

Tabla XXVII
Distribución de Frecuencias de la variable
Problema de Úlcera

	Frecuencia	Porcentaje	Porcentaje acumulado
No	256	94,8	94,8
Si	14	5,2	100
Total	270	100	

Gráfico 4.34
Histograma de frecuencias de la variable
Problema de Úlcera



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₂₉: Problema de Garganta

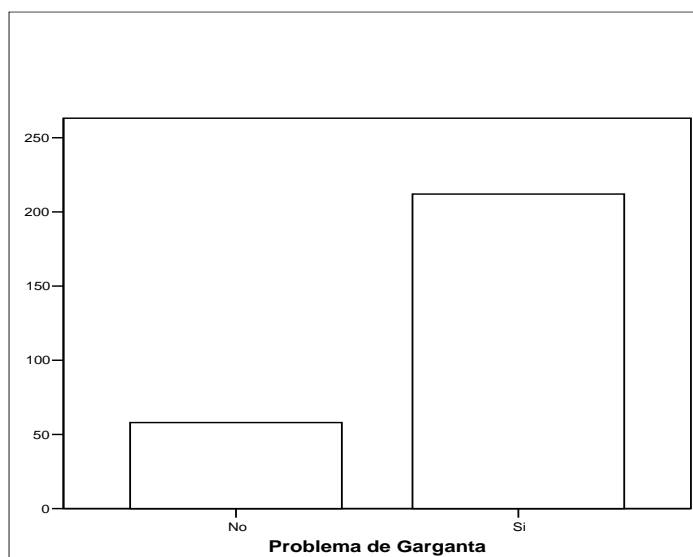
Se analiza ahora si el estudiante entrevistado a lo largo de su estancia en la ESPOL, pasa o ha pasado por los problemas de garganta, como se muestra en la tabla XXVIII encontramos que el 21.5% no han padecido este problema mientras que el 78.5% si han padecido o padecen de los problemas de garganta. Estos resultados se pueden verificar también en el gráfico 4.35

Tabla XXVIII
Distribución de Frecuencias de la variable
Problema de Garganta

	Frecuencia	Porcentaje	Porcentaje acumulado
No	58	21,5	21,5
Si	212	78,5	100
Total	270	100	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.35
Histograma de frecuencias de la variable
Problema de Garganta



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₃₀: Problema de Colesterol en Niveles Anormales

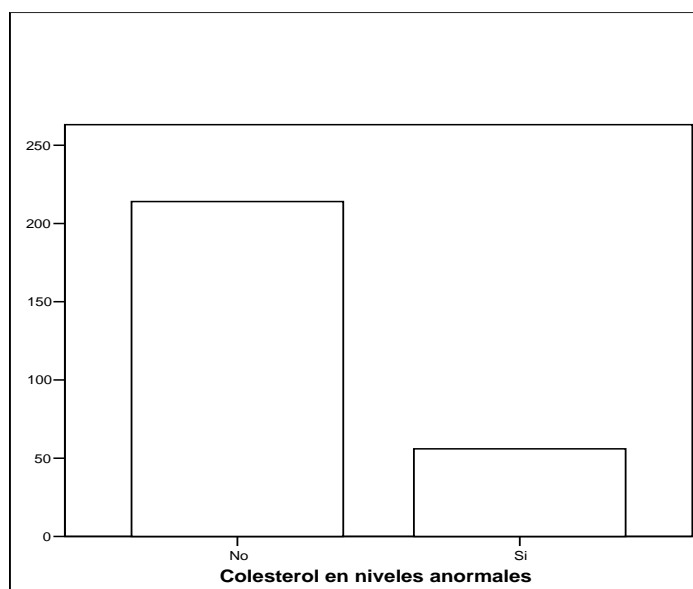
Se analiza ahora si el estudiante entrevistado a lo largo de su estancia en la ESPOL, pasa o ha pasado por problemas colesterol, como se muestra en la tabla XXIX encontramos que el 79.3% no han padecido este problema mientras que el 20.7% si han padecido o padecen de los problemas de colesterol. Estos resultados se pueden verificar también en el gráfico 4.36

Tabla XXIX
Distribución de Frecuencias de la variable
Problema de Colesterol en niveles anormales

	Frecuencia	Porcentaje	Porcentaje acumulado
No	214	79,3	79,3
Si	56	20,7	100
Total	270	100	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.36
Histograma de frecuencias de la variable
Problema de Colesterol en niveles anormales



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₃₁: Problema de Diabetes

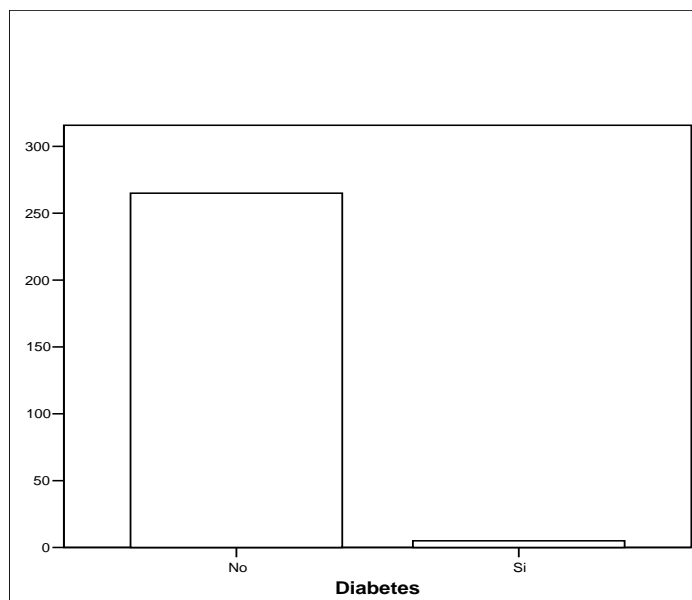
Se analiza ahora si el estudiante entrevistado a lo largo de su estancia en la ESPOL, pasa o ha pasado por los problemas de diabetes, como se muestra en la tabla XXX encontramos que el 98.1% no han padecido este problema mientras que el 1.9% si han padecido o padecen de los problemas de diabetes. Estos resultados se pueden verificar también en el gráfico 4.37

Tabla XXX
Distribución de Frecuencias de la variable
Problema de Diabetes

	Frecuencia	Porcentaje	Porcentaje acumulado
No	265	98,1	98,1
Si	5	1,9	100
Total	270	100	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.37
Histograma de frecuencias de la variable
Problema de Diabetes



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₃₂: Problema de Cálculos

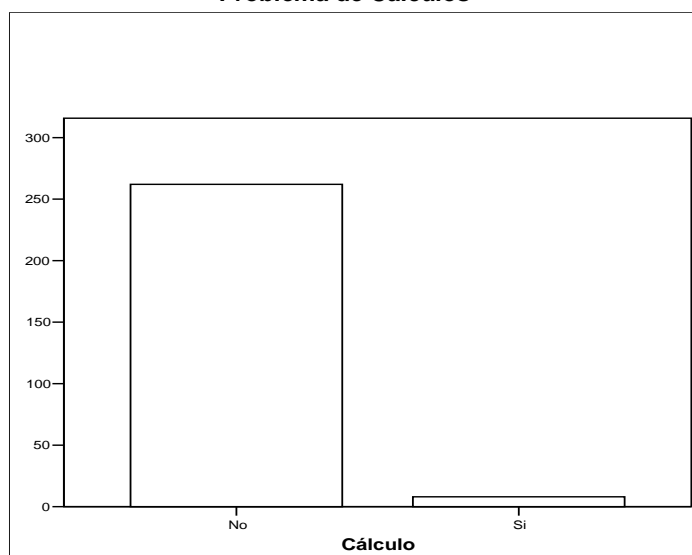
Se analiza ahora si el estudiante entrevistado a lo largo de su estancia en la ESPOL, pasa o ha pasado por los problemas de cálculo, como se muestra en la tabla XXXI encontramos que el 97% no han padecido este problema mientras que el 3% si han padecido o padecen de los problemas de cálculo. Estos resultados se pueden verificar también en el gráfico 4.38

Tabla XXXI
Distribución de Frecuencias de la variable
Problema de Cálculos

	Frecuencia	Porcentaje	Porcentaje acumulado
No	262	97,0	97,0
Si	8	3,0	100
Total	270	100	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.38
Histograma de frecuencias de la variable
Problema de Cálculos



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Variable X₃₃: Problemas Otros

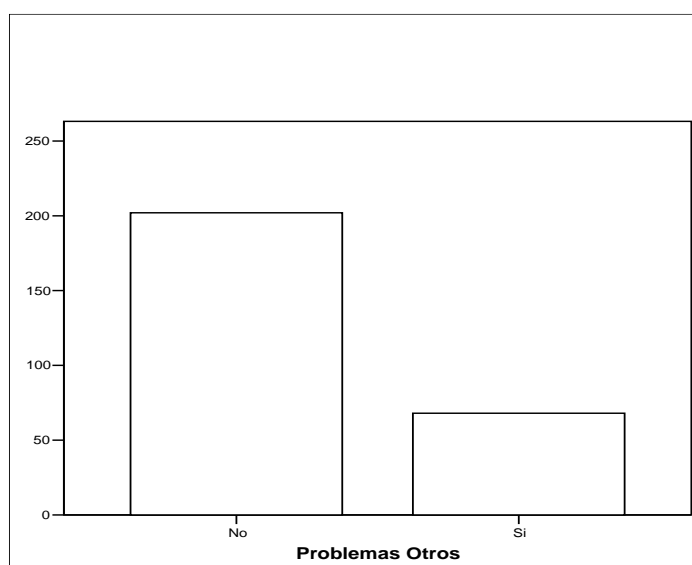
Se analiza ahora si el estudiante entrevistado a lo largo de su estancia en la ESPOL, pasa o ha pasado por los problemas otros, como se muestra en la tabla XXXII encontramos que el 74.8% no han padecido este problema mientras que el 25.2% si han padecido o padecen de problemas otros. Estos resultados se pueden verificar también en el gráfico 4.39

Tabla XXXII
Distribución de Frecuencias de la variable
Problemas Otros

	Frecuencia	Porcentaje	Porcentaje acumulado
No	202	74,8	74,8
Si	68	25,2	100
Total	270	100	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Gráfico 4.39
Histograma de frecuencias de la variable
Problemas Otros



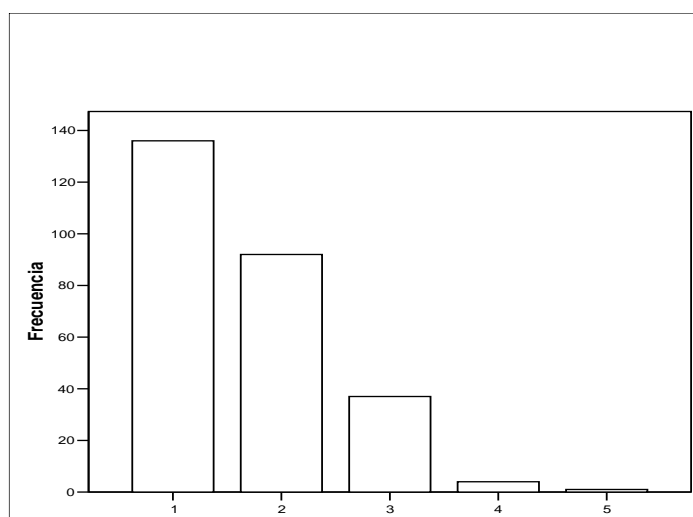
Variable X₃₄: Ausencia de clases por problemas alimenticios

De los 270 estudiantes entrevistados, el 50.4% de ellos declaró que nunca se han ausentado de clases por problemas alimenticios mientras que solo el 0.4% declaró que siempre han tenido que faltar a clases por este tipo de problemas, para más detalles véase la tabla XXXIII. Estos resultados se pueden verificar también en el gráfico 4.40.

Tabla XXXIII
Distribución de Frecuencias de la variable
Ausencia de clases por problemas alimenticios

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	Nunca	136	50,4	50,4
2	Rara vez	92	34,1	84,4
3	Algunas veces	37	13,7	98,1
4	La mayoría de las veces	4	1,5	99,6
5	Siempre	1	0,4	100,0
	Total	270	100,0	

Gráfico 4.40
Histograma de frecuencias de la variable
Ausencia de clases por problemas alimenticios



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

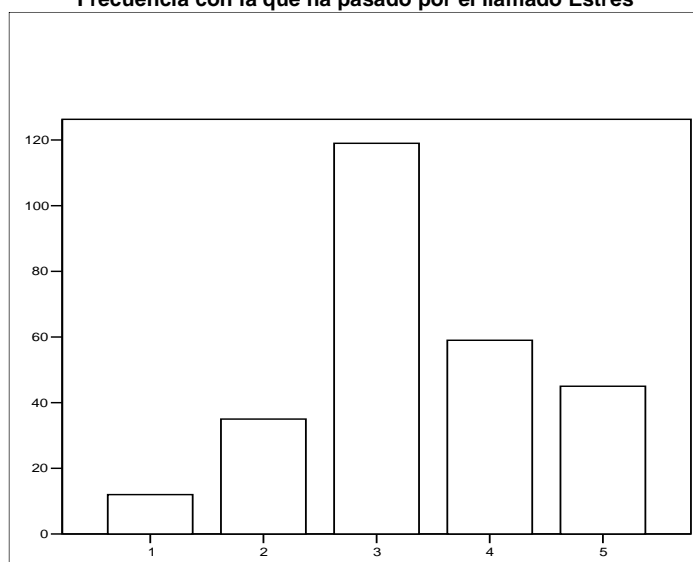
Variable X₃₅: Frecuencia con la que ha pasado por el llamado “Estrés”

De los 270 estudiantes entrevistados, el 44.1% de ellos declaró que nunca se han pasado por problemas estrés mientras que el 16.7% declaró que siempre han tenido este tipo de problemas, para más detalles véase la tabla XXXIV. Estos resultados se pueden verificar también en el gráfico 4.41.

Tabla XXXIV
Distribución de Frecuencias de la variable
Frecuencia con la que ha pasado por el llamado Estrés

Código	Rango	Frecuencia	Porcentaje	Porcentaje acumulado
1	Nunca	12	4,4	4,4
2	Rara vez	35	13,0	17,4
3	Algunas veces	119	44,1	61,5
4	La mayoría de las veces	59	21,9	83,3
5	Siempre	45	16,7	100,0
	Total	270	100,0	

Gráfico 4.41
Histograma de frecuencias de la variable
Frecuencia con la que ha pasado por el llamado Estrés



Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

CAPÍTULO V

5. ANÁLISIS MULTIVARIADO

5.1. Introducción

En este capítulo se realizará el análisis multivariado, previo análisis de cada una de las variables descritas en el capítulo anterior.

Las herramientas que se emplearán serán: Tablas de Contingencia y el Método de Regresión Logística Binaria.

La primera técnica multivariada nos indicará si existe independencia o no entre las variables de estudio, utilizando la distribución Chi cuadrado.

La segunda técnica multivariada es el método de Regresión Logística Binaria es una técnica no lineal, que nos permite explicar una variable dicotómica en términos de otras variables cualitativas o cuantitativas.

El procesamiento de las técnicas mencionadas se las realizará mediante el paquete estadístico SPSS 12.0.

5.2. Tabla de Contingencia

La tabla de contingencia es un arreglo bidimensional en la cual se van a cruzar dos variables cualitativas conjuntamente, esta tabla esta formada por i filas y j columnas , en las que se recoge la frecuencia absoluta y esperada del número de individuos para cada una de las posibles combinaciones de niveles de las dos variables. El objetivo de este análisis es determinar si las dos variables son independientes o no, para ello se construye el siguiente contraste de hipótesis:

H₀: La clasificación por renglones (variable A_i) y la clasificación por columnas (variable A_j) y son independientes

vs.

H₁: La clasificación por renglones (variable A_i) y la clasificación por columnas (variable A_j) no son independientes

Entonces se realiza las n pruebas independientes, cumpliendo:

$$\text{Frecuencia observadas de } A_i : \sum_{i=1}^l n_{ij} = n_{i.}$$

$$\text{Frecuencia observadas de } A_j : \sum_{j=1}^k n_{ij} = n_{.j}$$

Donde:

n_{ij} = número de individuos que poseen conjuntamente las modalidades entre la i -ésima fila y la j -ésima columna.

Entonces el estadístico de prueba es:

$$\chi^2 = \sum_{i=1}^k \sum_{j=1}^l \frac{(n_{ij} - E(n_{ij}))^2}{E(n_{ij})}$$

el cual se puede probar que se distribuye según una variable Ji-Cuadrado (χ^2) con $(k-1)(l-1)$ grados de libertad.

Donde $E(n_{ij}) = \frac{n_i n_j}{n}$, es la frecuencia esperada

Se rechaza la hipótesis nula H_0 a favor de H_1 con $(1-\alpha)100\%$ de confianza si:

$$\chi^2 > \chi_{\alpha}^2(k-1)(l-1).$$

5.2.1. Análisis de Contingencia Para Las Variables

Rendimiento Académico vs. Edad

El contraste de hipótesis para determinar la independencia de las variables Rendimiento académico vs. Edad es:

H_0 : El rendimiento académico y la edad son independientes

H_1 : Rechazar H_0

Tabla XXXV
Tabla de Contingencia y Ji-Cuadrado para
Rendimiento académico vs. Edad

			Rendimiento Académico			Total
			<75 %	[75 - 90)	>90%	
Edad	< 20 años	Observado	10	4	26	40
		Esperado	12.7	9.6	17.6	40
	[20-25)	Observado	65	52	91	208
		Esperado	66.3	50.1	91.7	208
	>25 años	Observado	11	9	2	22
		Esperado	7.0	5.3	9.7	22
Total	Observado	86	65	119	270	
	Esperado	86	65	119	270	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Estadístico de Prueba Ji^2	gl	valor p
18,931	4	0,001

Con este valor de p menor que un nivel de significancia 0.05, se puede concluir que existe suficiente evidencia estadística para rechazar la hipótesis nula H_0 , es decir que el rendimiento académico depende de la edad.

Rendimiento Académico vs. Género.

El contraste de hipótesis para determinar la independencia de las variables Rendimiento académico y Género es:

H_0 : El rendimiento académico y el género son independientes

H_1 : Rechazar H_0

Tabla XXXVI
Tabla de Contingencia y Ji-Cuadrado para
Rendimiento académico vs. Género

			Rendimiento Académico			Total
			<75 %	[75 - 90)	>90%	
Género Masculino	Observado		64	42	66	172
	Esperado		54.8	41.4	75.8	172
Femenino	Observado		22	23	53	98
	Esperado		31.2	23.6	43.2	98
Total	Observado		86	65	119	270
	Esperado		86.0	65.0	119	270

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S

Estadístico de Prueba Ji^2	gl	valor p
7,789	2	0,02

Con este valor de p menor que un nivel de significancia 0.05, se puede concluir que existe suficiente evidencia estadística para rechazar la hipótesis nula H_0 , es decir que el rendimiento académico depende del género del estudiante.

Días a la semana que desayuna vs. Género.

Tratamos de probar el contraste de la siguiente hipótesis:

H_0 : Días a la semana que desayuna y Género son independientes

H_1 : Rechazar H_0

Tabla XXXVII
Tabla de Contingencia y Ji-Cuadrado para
Días a la semana que desayuna vs. Género.

			Días a la semana que desayuna				Total
			0 a 1 día	2 a 3 días	4 a 5 días	6 a 7 días	
Género Masculino	Observado		9	22	31	110	172
	Esperado		12,1	26,8	28,7	104,5	172,0
Femenino	Observado		10	20	14	54	98
	Esperado		6,9	15,2	16,3	59,5	98
Total	Observado		19	42	45	164	270
	Esperado		19	42	45	164	270

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

Elaboración: Betzabeth Yonson S.

Estadístico de Prueba Ji^2	gl	valor p
5,850	3	0,119

El valor p es mayor que un nivel de significancia 0.05, por lo tanto existe evidencia estadística para no rechazar la hipótesis nula H_0 , es decir que los días a la semana que desayuna y el género del estudiante son independientes.

Peso vs. Género.

El contraste de hipótesis para determinar la independencia de las variables peso y género:

H_0 : Peso y Género son independientes

H_1 : Rechazar H_0

Tabla XXXVIII
Tabla de Contingencia y Ji-Cuadrado para
Peso vs. Género.

			Peso			Total
			<120 lbs	[120 - 180)	> 180 lbs	
Género Masculino	Observado	15	132	25	172	
	Esperado	33,1	122,9	15,9	172	
Femenino	Observado	37	61	0	98	
	Esperado	18,9	70,1	9,1	98	
Total	Observado	52	193	25	270	
	Esperado	52	193	25	270	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S.

Estadístico de Prueba Ji^2	gl	valor p
43,406	2	0,000

Con este valor de p menor que un nivel de significancia 0.05, se puede concluir que existe suficiente evidencia estadística para rechazar la hipótesis nula H_0 , es decir que el peso depende del género del estudiante.

Peso vs. Colesterol en Niveles Anormales.

El contraste de hipótesis para determinar la independencia de las variables peso y colesterol en niveles anormales:

H_0 : Peso y colesterol en niveles anormales son independientes

H_1 : Rechazar H_0

Tabla XXXIX
Tabla de Contingencia y Ji-Cuadrado para
Peso vs. Colesterol en niveles anormales.

			Peso			Total
			<120 lbs	[120 - 180)	> 180 lbs	
Colesterol	No	Observado	44	152	18	214
		Esperado	41,2	153	19,8	214
	Si	Observado	8	41	7	56
		Esperado	10,8	40	5,2	56
Total		Observado	52	193	25	270
		Esperado	52	193	25	270

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S.

Estadístico de Prueba Ji^2	gl	valor p
1,739	2	0,419

Con este valor de p, se puede concluir que no existe suficiente evidencia estadística para rechazar la hipótesis nula H_0 , es decir que el peso no depende del colesterol en niveles anormales.

Peso vs. Estatura

El contraste de hipótesis para determinar la independencia de las variables peso y estatura:

H_0 : Peso y estatura son independientes

H_1 : Rechazar H_0

Tabla XL
Tabla de Contingencia y Ji-Cuadrado para
Peso vs. Estatura

			Peso			Total
			<120 lbs	[120 - 180)	> 180 lbs	
Estatura <165 cmts	Observado	39	59	0	98	
	Esperado	18,9	70,1	9,1	98	
[165 - 175)	Observado	12	91	4	107	
	Esperado	20,6	76,5	9,9	107	
> 175 cmts	Observado	1	43	21	65	
	Esperado	12,5	46,5	6	65	
Total	Observado	52	193	25	270	
	Esperado	52	193	25	270	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

Elaboración: Betzabeth Yonson S.

Estadístico de Prueba Ji^2	gl	valor p
90,299	4	0,000

Mediante el análisis de contingencia se determinó que el valor p es menor que un nivel de significancia 0.05, por lo tanto se puede concluir que existe suficiente evidencia estadística para rechazar la hipótesis nula H_0 , es decir que el peso depende de la estatura del estudiante.

Rendimiento Académico vs. Promedio

El contraste de hipótesis para determinar la independencia de las variables Rendimiento académico y promedio es:

H_0 : El rendimiento académico y promedio son independientes

H_1 : Rechazar H_0

Tabla XLI
Tabla de Contingencia y Ji-Cuadrado para
Rendimiento académico vs. Promedio

			Rendimiento Académico			Total
			<75 %	[75 - 90)	>90%	
Promedio	<7	Observado	61	15	11	87
		Esperado	27.7	20.9	38.3	87
	[7 - 8)	Observado	25	48	70	143
		Esperado	45.5	34.4	63	143
	> 8	Observado	0	2	38	40
		Esperado	12.7	9.6	17.6	40
Total	Observado	86	65	119	270	
	Esperado	86	65	119	270	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S.

Estadístico de Prueba Ji^2	gl	valor p
118,893	4	0,000

Con este valor de p , se puede concluir que existe suficiente evidencia estadística para rechazar la hipótesis nula H_0 , es decir que el rendimiento académico y promedio no son independientes, dependen una de la otra.

Porcentaje de su alimentación es a través de comida preparada en casa vs. Género.

El contraste de hipótesis para determinar la independencia de las variables es:

H_0 : Porcentaje de su alimentación es a través de comida preparada en casa y Género son independientes

H_1 : Rechazar H_0

Tabla XLII
Tabla de Contingencia y Ji-Cuadrado para
Porcentaje de su alimentación es a través de comida preparada en casa vs. Género.

			Porcentaje de su alimentación				Total
			<25%	[25 - 50)	[50 - 75)	>75%	
Género	Masculino	Observado	16	29	78	49	172
		Esperado	18.5	28.7	68.2	56.7	172
	Femenino	Observado	13	16	29	40	98
		Esperado	10.5	16.3	38.8	32.3	98
Total	Observado		29	45	107	89	270
	Esperado		29	45	107	89	270

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

Elaboración: Betzabeth Yonson S.

Estadístico de Prueba Ji^2	gl	valor p
7.713	3	0.052

El valor p es mayor que un nivel de significancia 0.05 por lo tanto existe evidencia estadística para no rechazar la hipótesis nula H_0 , es decir que el porcentaje de alimentación del estudiante a través de comida preparada en casa y Género son independientes.

Rendimiento Académico vs. Estrés

Las nuevas categorías para la codificación de la variable: ¿Que tan frecuente, el estudiante ha pasado por el llamado Estrés? es:

Nunca = Infrecuentemente

Rara vez = Infrecuentemente

Algunas veces = Algunas veces

La mayoría de las veces = Frecuentemente

Siempre = Frecuentemente

Con el objetivo que la Prueba Ji-Cuadrado de independencia sea válida.

El contraste de hipótesis para determinar la independencia de las variables Rendimiento académico vs. Estrés es:

H_0 : El rendimiento académico y Estrés son independientes

H_1 : Rechazar H_0

Tabla XLIII
Tabla de Contingencia y Ji-Cuadrado para
Rendimiento académico vs. Estrés

			Estrés			Total
			Infrecuente	Algunas veces	Frecuente	
Rendimiento Académico	<75 %	Observado	20	40	26	86
		Esperado	15	37.9	33.1	86
	[75 - 90)	Observado	12	27	26	65
		Esperado	11.3	28.6	25.0	65
	>90%	Observado	15	52	52	119
		Esperado	20.7	52.4	45.8	119
Total	Observado	47	119	104	270	
	Esperado	47	119	104	270	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

Elaboración: Betzabeth Yonson S.

Estadístico de Prueba Ji²	gl	valor p
5.921	4	0.205

El valor p es mayor que un nivel de significancia 0.05 por lo tanto existe evidencia estadística para aceptar la hipótesis nula H₀, es decir que el Rendimiento académico y el estrés del estudiante son independientes.

Rendimiento Académico vs. Trabaja

El contraste de hipótesis para determinar la independencia de las variables Rendimiento académico y trabaja es:

H_0 : El rendimiento académico y trabaja son independientes

H_1 : Rechazar H_0

Tabla XLIV
Tabla de Contingencia y Ji-Cuadrado para
Rendimiento académico vs. Trabaja

			Trabaja			Total
			Dentro de la Espol	Fuera de la Espol	No trabaja	
Rendimiento Académico	<75 %	Observado	1	15	70	86
		Esperado	9.2	12.7	64	86
	[75 - 90)	Observado	5	8	52	65
		Esperado	7	9.6	48.4	65
	>90 %	Observado	23	17	79	119
		Esperado	12.8	17.6	88.6	119
Total	Observado	29	40	201	270	
	Esperado	29	40	201	270	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

Elaboración: Betzabeth Yonson S.

Estadístico de Prueba Ji^2	gl	valor p
18.642	4	0.001

Con este valor de p, se puede concluir que existe suficiente evidencia estadística para rechazar la hipótesis nula H_0 , es decir que el rendimiento académico y si el estudiante trabaja o no, dependen una de la otra.

Rendimiento Académico vs. Comida chatarra

Las nuevas categorías para la codificación de la variable: Que tan frecuente, el estudiante opta por comida chatarra, es:

Nunca = Infrecuentemente

Rara vez = Infrecuentemente

Algunas veces = Algunas veces

La mayoría de las veces = Frecuentemente

Siempre = Frecuentemente

Con el objetivo que la Prueba Ji-Cuadrado de independencia sea válida.

El contraste de hipótesis para determinar la independencia de las variables Rendimiento académico y Comida chatarra es:

H_0 : El rendimiento académico y Comida chatarra son independientes

H_1 : Rechazar H_0

Tabla XLV
Tabla de Contingencia y Ji-Cuadrado para
Rendimiento académico vs. Comida chatarra

			Comida chatarra			Total
			Infrecuente	Algunas veces	Frecuente	
Rendimiento Académico	<75 %	Observado	31	49	6	86
		Esperado	34.7	44.0	7.3	86
	[75 - 90)	Observado	25	34	6	65
		Esperado	26.2	33.2	5.5	65
	>90%	Observado	53	55	11	119
		Esperado	48	60.8	10.1	119
Total	Observado	109	138	23	270	
	Esperado	109	138	23	270	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005

Elaboración: Betzabeth Yonson S.

Estadístico de Prueba Ji²	gl	valor p
2.475	4	0.649

El valor p es mayor que un nivel de significancia 0.05 por lo tanto existe evidencia estadística para aceptar la hipótesis nula H_0 , es decir que el Rendimiento académico y la frecuencia por la que opta el estudiante por comida chatarra, para su alimentación son independientes.

Realiza Actividad Deportiva vs. Estrés

Las nuevas categorías para la codificación de la variable: Frecuencia, que el estudiante pasa por el llamado estrés, es:

Nunca = Infrecuentemente

Rara vez = Infrecuentemente

Algunas veces = Algunas veces

La mayoría de las veces = Frecuentemente

Siempre = Frecuentemente

La variable, Realiza Actividad Deportiva fue codificada nuevamente como lo muestra la tabla XLVI, con el objetivo que la Prueba Ji-Cuadrado de independencia sea válida.

El contraste de hipótesis para determinar la independencia de las variables Realiza Actividad Deportiva vs. Estrés es:

H_0 : Realiza Actividad Deportiva y Estrés son independientes

H_1 : Rechazar H_0

Tabla XLVI
Tabla de Contingencia y Ji-Cuadrado para
Realiza Actividad Deportiva vs. Estrés

			Estrés			Total
			Infrecuente	Algunas veces	Frecuente	
Realiza deporte	0 – 3 días	Observado	39	107	95	241
		Esperado	42.0	106.2	92.8	241
	4 – 7 días	Observado	8	12	9	29
		Esperado	5.0	12.8	11.2	29
Total	Observado	47	119	104	270	
	Esperado	47	119	104	270	

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S.

Estadístico de Prueba Ji^2	gl	valor p
2.460	2	0.292

El valor p es mayor que un nivel de significancia 0.05 por lo tanto existe evidencia estadística para aceptar la hipótesis nula H_0 , es decir que las variables, Realiza actividad deportiva y Estrés son independientes.

5.3. Análisis de Regresión Multinomial Logística

Primero se identificó a la variable rendimiento académico como la variable de respuesta (dependiente), la cual será explicada en términos de un grupo de variables independientes, como se muestra en la tabla XLVII.

TABLA XLVII
VARIABLES UTILIZADAS EN EL ANÁLISIS.

Variable Dependiente	Rendimiento académico
VARIABLES INDEPENDIENTE	Género
	Factor P
	Carrera
	Edad
	Peso
	Estatura
	Actualmente trabaja
	Horas que permanece en la ESPOL
	Días a la semana que realiza actividad deportiva
	Desempeño para su rol de estudiante
	Porcentaje de alimentación es dentro de casa
	Días a la semana que desayuna
	Días a la semana que come carne roja
	Días a la semana que come pescado
	Días a la semana que come pollo
	Días a la semana que toma leche
	Días a la semana que come vegetales
	Días a la semana que come frutas
	Cuántas comidas realiza en el día
	Frecuencia por la que opta por comida rápida
	Horario fijo para su alimentación diaria
	Tipo de bebida de preferencia para satisfacer su sed
	Frecuencia que fuma
	Frecuencia con que ingiere bebida alcohólica
	Problema de Gastritis
	Problema de Úlcera
	Problema de Garganta
	Problema de colesterol en niveles anormales
	Problema de Diabetes
	Problema de Cálculos
	Otros problemas
	Ausencia de clases por problemas alimenticios
	Frecuencia con la que ha pasado por el llamado Estrés

Se realizó un análisis de la totalidad de los datos, con el objetivo de identificar características globales de los estudiantes seleccionados en la muestra de un total de $n=270$.

A continuación en la tabla XLVIII, se presenta la prueba de la razón de verosimilitud del modelo final:

TABLA XLVIII
Modelo ajustado

Modelo	-2 Log Verosimilitud	Chi-cuadrado	gl	Sig.
Solo la intersección	576,898			
Final	276,432	300,46	182	,000

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S.

En la tabla anterior el modelo tiene una significancia menor a 0.05 lo que se puede concluir que en el modelo final los coeficientes del parámetro son cero, según la hipótesis nula.

5.3.1. Bondad de Ajuste

La tabla de bondad de ajuste, presenta dos estadísticos con distribuciones chi-cuadrado, que determinan si las variables se ajustan al modelo, como se podrá verificar en la tabla XLIX, el nivel de significancia para las dos pruebas son mayores a un nivel de significancia 0.05, por lo que se puede decir que las variables no todas

se ajustan o contribuyen al modelo, en el anexo II se muestra qué variables son las que contribuyen al modelo supuesto.

TABLA XLIX
Bondad de ajuste

	Chi-cuadrado	gl	Sig.
Pearson	377,332	356	0,209
Desviación	276,432	356	0,999

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S.

5.3.1.1 Modelo ajustado con variables que contribuyen al modelo

En el anexo III se presentó en la tabla las variables que contribuyeron y las que no al modelo, por ello se volvió a trabajar solo con las que si contribuían al modelo supuesto, en donde resultó que entre ellas habían otras que no se ajustaban al modelo, lo cual se volvieron a eliminar hasta obtener el modelo en el cual los parámetros fueran diferentes de cero. A continuación las tablas que muestran el modelo ajustado.

TABLA L
Modelo ajustado

Modelo	-2 Log Verosimilitud	Chi-cuadrado	gl	Sig.
Solo la intersección	560,603			
Final	422,13	138,472	54	0,068

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S.

TABLA LI
Bondad de ajuste

	Chi-cuadrado	gl	Sig.
Pearson	458,534	452	0,406
Desviación	408,607	452	0,929

Fuente: Perfil del estudiante politécnico según preferencias alimenticias, tesis 2005
Elaboración: Betzabeth Yonson S.

En la tabla de bondad de ajuste, la significancia son mayores que 0.05 por lo que las variables que se redujeron están adecuadamente ajustadas al modelo. En el anexo IV se muestran las variables reducidas con los respectivos coeficientes β_i , para el modelo de regresión multinomial logístico.

CAPÍTULO VI

6. CONCLUSIONES Y RECOMENDACIONES

6.1. Conclusiones

1. La media del Rendimiento académico de los estudiantes entrevistados, es 81.89%, de los cuales el 10.7% de ellos trabajan dentro de la ESPOL, el 14.8% trabajan fuera de la ESPOL y el 74.4% no trabajan. Es útil esta comparación ya que además se ha comprobado que las variables rendimiento académico y trabajo dependen la una de la otra.
2. El Peso promedio de los estudiantes entrevistados fue de 140.86 lbs, además se observó que el 79.3% de ellos no han padecido de problemas de colesterol, mientras que el 20.7%, lo padecen o han padecido de este tipo de problemas, además, según el análisis de contingencia se comprobó que no es dependiente el peso de los estudiantes, con el hecho que padezcan o no de colesterol en niveles anormales. Este porcentaje de estudiantes que afirmaron tener problemas con su nivel de colesterol resulta sorprendentemente alto, ya que implica que más de uno de cada cinco estudiantes tienen este tipo de problemas.

3. El 0.7% de los estudiantes encuestados permanecen menos de tres horas en la ESPOL, el 35.6% permanecen de tres a seis horas, el 39.6% permanecen de seis a nueve horas y el 24.1% permanecen mas de nueve horas en la ESPOL, de aquí se concluye que de manera general los estudiantes de la ESPOL pasan gran parte de su tiempo dentro de los predios del campus universitario, por lo que muchas veces su alimentación también se realiza fundamentalmente en este lugar a lo largo del día.
4. Ya que de los resultados de la encuesta se observó que el 64.4% de los estudiantes no realizan alguna actividad deportiva durante la semana y solo el 3.3% lo realizan de manera frecuente, de seis a siete días a la semana, se deduce que la practica deportiva no está muy difundida entre los estudiantes, a pesar que la ESPOL cuenta con las instalaciones deportivas adecuadas en cantidad y calidad para la realización de distintas actividades deportivas.
5. Observando los resultados de los estudiantes que han padecido por problemas de estrés, el 4.4% de ellos declaró que nunca han pasado por problemas de estrés, mientras que el 16.7% declaró que siempre lo han padecido, se puede observar que este problema está presente, frecuentemente en el convivir politécnico.
6. Si bien es cierto que la mayoría de estudiantes desayunan todos los días de la semana (esto es 60.7%), también se observó casos

en los cuales los estudiantes no desayunan ningún día de la semana (esto es 7%) lo cual es un problema ya que el desayuno debe constituir según los principales de la salud y nutrición, es la principal comida del día.

7. En cuanto a los resultados de cuantas comidas realiza en el día el estudiante, se observó que 3.7% de los estudiantes encuestados come una vez en el día, el 20% comen dos veces en el día, el 61.9% comen tres veces en el día y el 14.4% de los estudiantes comen más de tres veces en el día.
8. El 72.6% de ellos declaró que nunca fuman mientras que solo el 2.6% declaró que siempre fuman, el 23.7% de ellos declaró que nunca ingieren bebida alcohólica mientras que solo el 0.7% declaró que siempre consumen este tipo de bebida, esto indica que el hábito de fumar y beber bebidas alcohólicas no constituye un problema serio entre los estudiantes de la ESPOL.
9. De acuerdo a las enfermedades que mayormente padecen los estudiantes se encontró, la gastritis con un 41.5% quienes la padecen y las afecciones a la garganta, así lo muestran los resultados con un 78.5% de los estudiantes registrados en el primer termino del 2005.
10. El análisis de Regresión Multinomial Logística es una herramienta estadística de mejor capacidad cuando se requiere de una variable

de respuesta de más de dos categorías, en estudios como el presente es bastante útil pues, en este caso se tenía variables como rendimiento académico, promedio, etc.

11. Se concluyó mediante este análisis que las variables significativas y que aportaron en gran medida para determinar el perfil del estudiante según preferencias alimenticias, fueron: carrera, edad, si trabajan o no, como consideraban la alimentación para el desempeño de su rol de estudiante, si consumían carne roja y cuantas veces comía el estudiante en el día; descartando el resto de variables por no aportar gran influencia en el modelo según una significancia 0.05.

6.2. Recomendaciones

1. Promover y difundir hábitos alimenticios sanos dentro de la ESPOL, que hablen de los riesgos de dichas enfermedades causadas por la mala alimentación, que de alguna forma interfiere en el rendimiento académico del estudiante.
2. Promover la practica de actividades deportivas entre los estudiantes de la ESPOL, pues es sabido que el deporte es una de las mejores armas para combatir el estrés. Esto es importante ya que la mayoría

de las enfermedades gástricas están causadas en gran medida por el estrés.

3. Se recomienda que se amplíe este estudio con otras variables que puedan tener significancia tales como procedencia, si vive el estudiante en la casa de sus padres, si tiene ingresos propios, etc., que no constan en los registros del CRECE.
4. Sería interesante realizar estudios posteriores para detectar si estos problemas van en cuanto, son constantes o están disminuyendo.