

# “USO DE REGRESIÓN MULTINOMIAL LOGÍSTICA PARA DETERMINAR EL PERFIL DEL ESTUDIANTE DE LA ESPOL, SEGÚN PREFERENCIAS ALIMENTICIAS”

Betzabeth R. Yonson Sangurima <sup>1</sup>, Fernando Sandoya <sup>2</sup>

## ABSTRACT

*The suitable feeding is one of the most important factors in the intellectual development of the students since it generates or it reinforces a positive or negative attitude towards the same one. Nevertheless of the previous thing, there are few related studies to determine the profile of the student of the ESPOL according to nutritional preferences. In this investigation one sets out to discuss, that of some of the variables used by means of a made survey to the students of the ESPOL, as truely they influence in the academic yield of the student.*

## RESUMEN

*La alimentación adecuada es uno de los factores más importantes en el desarrollo intelectual de los estudiantes ya que genera o refuerza una actitud positiva o negativa hacia la misma. Sin embargo de lo anterior, hay pocos estudios relacionados para determinar el perfil del estudiante de la ESPOL según preferencias alimenticias.*

*En esta investigación se propone discutir, que de algunas de las variables utilizadas mediante una encuesta realizada a los estudiantes de la ESPOL, cuales verdaderamente influyen en el rendimiento académico del estudiante.*

## 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

<sup>1</sup>Ingeniero en Estadística Informática de la ESPOL; (e-mail: byonson@espol.edu.ec).

<sup>2</sup> Profesor de la ESPOL en el Instituto de Ciencias Matemáticas.; (e-mail: fsandoya@espol.edu.ec).

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.

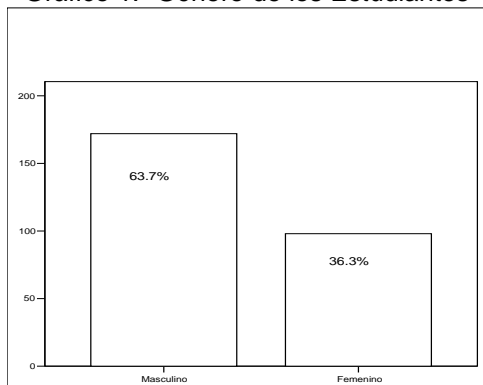
## 1. ANÁLISIS UNIVARIADO

A continuación, para las variables consideradas en el estudio se presenta un análisis descriptivo. Los detalles se presentan a continuación.

### Género del Estudiante

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.

Gráfico 1.- Género de los Estudiantes



## Rendimiento Académico

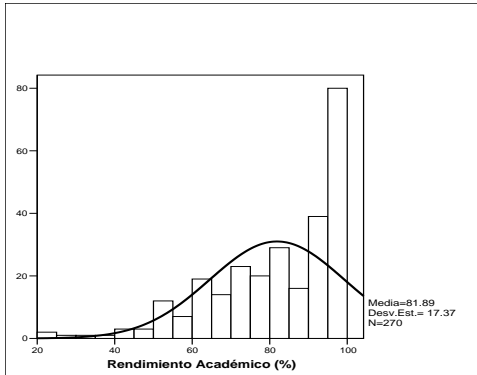
La tabla I muestra los estimadores poblacionales correspondientes al Rendimiento académico de los estudiantes entrevistados, el Rendimiento académico promedio es de 81.89%, con una dispersión de los datos alrededor de la media de 17.37% establecidos por la desviación estándar, mientras que la mediana 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 2).

Se puede 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 I.- 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

Gráfico 2.- Rendimiento Académico



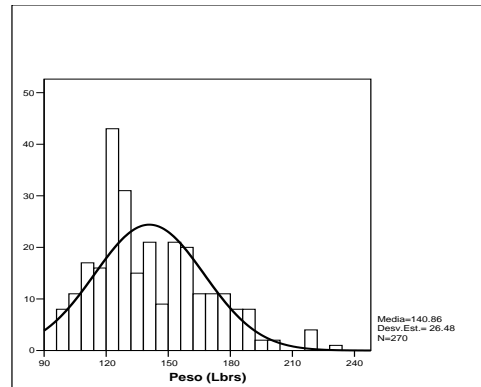
## Peso

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 3). También se puede 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 II.- 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
	50
	75
	120
	135
	160

Gráfico 3.- Peso



## Trabaja

Según los resultados obtenidos en la tabla III, 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.

Tabla III.- 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	

### Número de horas que diariamente permanece en la ESPOL

De los 270 estudiantes entrevistados el 0.7% de ellos permanecen menos de tres horas en la ESPOL y el 24.1% permanecen mas de nueve horas en la ESPOL, para más detalles véase la tabla IV.

Tabla IV.- 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
<b>Total</b>	<b>270</b>	<b>100,0</b>	

### Días a la semana que desayuna el estudiante

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. Para más detalles véase la tabla V.

Tabla V.- 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
	<b>Total</b>	<b>270</b>	<b>100,0</b>	

## 2. ANÁLISIS MULTIVARIADO

El análisis estadístico multivariado permite identificar las relaciones que existen entre dos o más variables.

### 2.1 Análisis Tablas de Contingencias

#### Rendimiento Académico vs. Edad

Tratamos de probar la siguiente hipótesis:

Ho: El rendimiento académico y la edad son independientes

H1: Rechazar Ho

Tabla VI.- 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	

Con este valor P de 0,001 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.

#### 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 VII.- 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

Con este valor P de 0,419 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.

## 2.2 Análisis de Regresión Multinomial Logística

Se identificó a la variable rendimiento académico como la variable de respuesta (dependiente), la cual será explicada en términos de las variables restantes en este estudio (independientes). Se hizo el 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 VIII, se presenta la prueba de la razón de verosimilitud del modelo final:

Tabla VIII.- Modelo ajustado

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

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.

## 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 IX, 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 IX.- Bondad de ajuste

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

## Modelo ajustado con variables que contribuyen al modelo

A continuación las tablas que muestran el modelo ajustado.

Tabla X.- Modelo ajustado

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

TABLA XI.- Bondad de ajuste

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

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.

### 3. 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.

### **3. BIBLIOGRAFÍA**

1. YONSON SANGURIMA B.,(2005) Uso de Regresión Multinomial Logística para Determinar el Perfil del Estudiante de la ESPOL, según preferencias alimenticias. Tesis de Ingeniería en Estadística e Informática, ICM-ESPOL, Guayaquil-Ecuador.

2. Luis Carlos Silva, Isabel María Barroso Utra, 2004, Regresión Logística, Madrid, La Muralla S.A., Editorial.

3. César Pérez, 2001, Técnicas Estadísticas con SPSS , Madrid, Prentice Hall.