

# LA DESERCIÓN ESTUDIANTIL DE LA ESPOL: EL CASO DEL INSTITUTO DE TECNOLOGÍAS

Katherine Vera<sup>1</sup>, Gaudencio Zurita<sup>2</sup>

## RESUMEN

*La investigación que se postula en el presente estudio se llevó a cabo en el segundo semestre del año 2005, cuya población objetivo son los estudiantes desertores de las carreras que se dictan en el Instituto de Tecnologías de la ESPOL en el periodo considerado ( 1994 – 2003 ) El objetivo de la investigación es conocer las razones que llevaron al estudiante a retirarse del Instituto de Tecnologías, para esto se aplicó un cuestionario, un error de diseño de 5.3% y un nivel de confianza del 95%, obteniéndose un tamaño de muestra de 171 estudiantes desertores. Se conoce además, que 13.30% del total de estudiantes que ingresaron a las carreras que se dictan en el Instituto de Tecnologías en el período (762 estudiantes), desertaron de las mismas. Se determinará las características relevantes al problema de deserción de los entes de la muestra en base a técnicas estadística; se presentan los Análisis Univariados y Bivariados, Análisis de Componentes Principales y Correlación Canónica.*

## ABSTRACT

*The investigation that is solicited in the present research was carried out during the second term of 2005, which Objective Population is the students who deserted from the Technology Schools among 1994 – 2003. Its objective is to know the reasons that made him or her to forsake. In order to this, was applied a questionnaire, using a error of 5.3% and a confidence level of 95%, obtaining a sample size of 171 defector students. We know also, that the 13.30% from the amount of students that joined to the Technology Schools (5240 students) are deserters by now. It will distinguish the relevants characteristics to the problem of the elements of the element samples with the support of statistical techniques, in the following lines we present the Univariate and Bivariate Analysis, the Principal Components Analysis and Canonical Correlations.*

## 1. INTRODUCCIÓN

La Deserción Estudiantil es uno de los principales problemas de la educación tanto superior como básica, para el presente estudio se considera “desertor” a un estudiante que no se ha registrado en alguna materia durante tres semestres consecutivos. La investigación que se postula, tiene como objetivo general conocer los principales factores que llevaron a los estudiantes a desertar de las carreras que se dictan en el Instituto de Tecnologías de la ESPOL y sus características antes, durante y después de su deserción.

El marco muestral de esta investigación corresponde a un listado proporcionado por el CRECE [15], donde constan todos los estudiantes que habiendo tomado materias en la ESPOL, y han dejado de estudiar. En el Instituto de Tecnologías en un período considerado desde 1993 a 2003, han desertado 171 estudiantes, excluyendo a aquellos que han sido sancionados y retirados de la ESPOL y aquellos que no han tomado materia alguna dentro de la Politécnica.

## 2. ANÁLISIS UNIVARIADO

En esta sección se presenta el análisis estadístico individualizado de las variables observadas en el cuestionario mencionado anteriormente, además de los estimadores de los parámetros de cada variable. En el caso de las variables cuantitativas, se estiman parámetros tales como media, curtosis, sesgo, varianza, desviación estándar, error estándar, cuartiles y moda.

### Características del Entrevistado

En esta sección se analizan las características personales del desertor entrevistado.

*Género.*- La mayoría de los desertores entrevistados, son de género masculino, es decir el 76%; frente a un 24% del género femenino. Del total de estudiantes desertores (762) el 84.52% son de género masculino y el 15.48% son mujeres.

*Edad del Entrevistado al momento de desertar.*- De los entrevistados al momento de desertar el 19% tenía edades entre 17 y 18 años al momento de desertar de la ESPOL; el 52% entre 19 y 20 años; el 20% entre 21 y 22 años; el 5% entre 23 y 24 años y el 5% tenía edades mayores o iguales a 25 años. La edad promedio estimada es  $19.83 \pm 0.131$  años, las mismas que varían entre 26 y 17 donde se observa un rango estimado de 9. El estimador de la

<sup>1</sup> Ingeniera en Estadística Informática.

<sup>2</sup> M.Sc. Matemáticas, M.Sc. Estadística, Profesor ICM, ESPOL, Director del Centro de Estudios e Investigaciones Estadísticas ICM – ESPOL ;email: gzurita@espol.edu.ec

desviación estándar es 1.708 y del error estándar 0.131, véase el Cuadro I.

<b>Cuadro I</b> <b>Estimadores Poblacionales de la</b> <b>Edad de los Entrevistados al</b> <b>desertar</b>	
n	171
Media	19.83
Mediana	20
Moda	19
Varianza	2.918
Desviación Estándar	1.708
Error Estándar	0.131
Sesgo	1.020
Curtosis	1.806

Mediante la Prueba de Kolmogorov-Smirnov, se verifica si la función de densidad de la variable edad puede ser modelada como una normal con media  $\mu = 19.83$  y varianza  $\sigma^2 = 2.9$ . (Véase Cuadro II)

<b>Cuadro II</b> <b>Prueba de Bondad de Ajuste, Kolmogorov-Smirnov:</b> <b>Edad del Entrevistado al desertar</b>
<b>H<sub>0</sub>:</b> La edad de los desertores tiene una distribución N(19.83 , 2.9) <b>vs.</b> <b>H<sub>1</sub>:</b> No es verdad H <sub>0</sub> Valor p = 0.000

Según los resultados obtenidos mediante la prueba de bondad de ajuste de Kolmogorov-Smirnov, el valor p es 0.000, lo cual significa que podemos decir que existe evidencia estadística para rechazar la hipótesis nula.

*Después de su retiro de la ESPOL.-* La proporción estimada de los entrevistados que afirman que después de su retiro de la ESPOL estudiaron una carrera diferente a la que tenían, es de 0.36, mientras que para los entrevistados que expresaron no haber continuado estudiando, la proporción es de 0.38; es decir que 380 de cada mil desertores de las carreras del Instituto de Tecnologías, no continuaron estudiando. Los entrevistados que estudian una carrera afín a la que estudiaba en la ESPOL alcanzan una proporción del 0.18 frente a los estudiantes que siguieron la misma carrera que alcanza el 0.08.

*Nivel Educativo Formal.-* Del los entrevistados la mayoría (48%) posee el mismo nivel educativo que tenía cuando se retiró de la ESPOL; el 15% de los desertores entrevistados afirman haber terminado los estudios de Pregrado, mientras que el 2% se encuentra cursando estudios de postgrado.

## Trayectoria Académica

En esta sección se consultó a los Entrevistados sobre su Trayectoria Académica antes de ingresar a la ESPOL.

*Tipo de Colegio.-* La variable “Tipo de Colegio”, indica que del total de entrevistados, la mayoría (53%) estudió en un Colegio Particular, donde el 27% en uno Religioso y el 26% restante en un Colegio Laico, mientras que el 32% estudió en un Colegio Fiscal, el 14% proviene de un Colegio Técnico Experimental y el 1% en otro tipo de colegio.

*Especialización en el Colegio.-* La mayoría de los entrevistados obtuvo su título de grado como bachiller en Informática, mientras que el 24% lo obtuvo en ciencias Físicas Matemáticas, además el 8% se graduó en ciencias Químico Biológicas, y el 18% estudió una especialidad Técnica.

## Trayectoria académica en la ESPOL

Además de conocer que la Trayectoria del Desertor antes de la ESPOL es importante conocer como se desempeñó dentro de la misma y lo que se analiza en esta sección.

*Motivo de ingreso a la ESPOL.-* Se observa que la mayoría de los entrevistados ingresaron a la ESPOL por decisión propia es decir el 70%, a continuación, el 17% por influencia familiar, además el 11% decidió ser estudiante de la ESPOL porque la carrera que tomó era de acuerdo a su especialización y el 1% por situación económica.

*Opción de Ingreso a la ESPOL.-* Para la variable “Opción de Ingreso”, se obtienen los siguientes resultados: el 83% de los entrevistados, la mayoría, ingresó a la ESPOL mediante el curso prepolitécnico y el 17% a través del Examen de ingreso.

*Factor P.-* De los entrevistados, el 55%, es decir la mayoría, tienen un Factor P entre 6 y 10, el 23% entre 11 y 15, además el 11% tienen un factor P menor o igual a 5 ; el 6% entre 16 y 20 y el 5% tiene un Factor P mayor que veinte. Estas cifras se observan en la Tabla I.

<b>Tabla I</b> <b>Distribución de Frecuencias</b> <b>Factor P</b>	
<b>Factor P</b>	<b>Frecuencia Relativa</b>
[0-5]	0.11
(5-10]	0.55
(10-15]	0.23
(15-20]	0.06
Más de 20	0.05
<b>Total</b>	<b>1.00</b>

En el Cuadro III se presentan las estimaciones de los parámetros correspondientes al Factor P de los entrevistados: el Factor P promedio estimado es  $9.59 \pm 0.37$ , mismos que varían entre 0 y 25 donde se observa un rango estimado de 25. El estimador de la desviación estándar es 4.86.

<b>Cuadro III</b> <b>Estimadores Poblacionales de Factor P de los entrevistados</b>	
n	171
Media	9.59
Mediana	7
Varianza	23.58
Desviación Estándar	4.86
Error Estándar	0.37
Sesgo	1.21
Curtosis	1.35

*Promedio en la ESPOL.*- Nótese en la Tabla II, que el 15% de los entrevistados, tienen un promedio de 0, es decir no aprobaron ninguna materia, el 45%, tienen un promedio mayor o igual que 6 y menor o igual que 7, el 36%, mayor que 7 y menor o igual a 8, y el 4% tiene un promedio mayor que 8.

<b>Tabla II</b> <b>Distribución de Frecuencias</b> <b>“Promedio en la ESPOL”</b>	
Promedio	Frecuencia Relativa
Ninguna materia aprobada	0.10
[6.0-6.5)	0.20
[6.5-7.0)	0.36
[7.0-7.5)	0.23
[7.5-8.0)	0.06
[8.0-8.5)	0.02
[8.5-9.0)	0.02
<b>Total</b>	<b>1.00</b>

En el Cuadro IV se presentan las estimaciones de los parámetros correspondientes al Promedio de notas en la ESPOL de los entrevistados: el promedio estimado es  $6.20 \pm 0.16$ , mismos que varían entre 0 y 9.11 donde se observa un rango estimado de 9.10. El estimador de la desviación estándar es 2.5640.

El estimador del sesgo es de -1.841, es decir que la mayor concentración de datos está a la derecha de la media.

<b>Cuadro IV</b> <b>Estimadores Poblacionales de</b> <b>“Promedio en la ESPOL”</b>	
n	171
Media	6.20
Varianza	4.59
Desviación Estándar	2.14
Error Estándar	0.16
Sesgo	-2.34
Curtosis	4.32

Mediante la Prueba de Kolmogorov-Smirnov, se demostrará que la función de densidad de esta variable no puede ser modelada como una normal.

<b>Cuadro V</b> <b>Prueba de Bondad de Ajuste, Kolmogorov-Smirnov: Promedio en la ESPOL</b>
<b>H<sub>0</sub>:</b> El promedio de notas de los entrevistados tiene una distribución $N(6.20, 4.59)$
<b>vs.</b>
<b>H<sub>1</sub>:</b> No es verdad H <sub>0</sub>
Valor p = 0.000

Según los resultados obtenidos mediante la prueba de bondad de ajuste, podemos decir que existe evidencia estadística para rechazar la hipótesis nula, es decir que la distribución de las edades no sigue una distribución normal con media 6.20 y desviación estándar 4.59.

*Número de Materias Tomadas en la ESPOL.*- De los entrevistados, el 37%, han tomado menos de 10 materias; el 34%, 10 materias o más y menos de 15, el 23% tomaron un número de materias mayor o igual a 15 y menor que 20, y el 6% cursaron más de 20 materias. Cifras presentadas Tabla III.

<b>Tabla III</b> <b>Distribución de Frecuencias Número de materias tomadas</b>	
Número de materias tomadas	Frecuencia Relativa
Menos de 10	0.37
[10-15)	0.34
[15-20)	0.23
20 ó más	0.06
<b>Total</b>	<b>1.00</b>

En el Cuadro VI se presentan las estimaciones de los parámetros correspondientes al número de materias tomadas por los entrevistados: el promedio de materias tomadas estimado es  $11.42 \pm 0.37$ , mismas que varían entre 3 y 22 materias. El estimador de la desviación estándar es 20.91.

Cuadro VI Estimadores Poblacionales de Número de materias tomadas	
n	171
Media	11.42
Varianza	24.46
Desviación Estándar	4.9
Error Estándar	0.37
Sesgo	0.257
Curtosis	-0.882

### Referentes Familiares

En esta sección se analizan los datos recopilados referentes al entorno familiar en el que se desenvolvía el desertor mientras era estudiante de la ESPOL.

*Con quien vivía mientras estudiaba en la ESPOL.*- Del total de entrevistados el 41% expresaron que vivían con ambos padres mientras estudiaban en la ESPOL, el 18% solo con uno de los padres, el 7% vivía con amigos, el 6% vivía solo y el 28% vivía con otros.

*Número de hermanos dependientes económicamente de la familia.*- Para la variable “Número de hermanos dependientes económicamente de la familia” se obtuvo que, el 1% de los entrevistados eran los únicos hijos dependientes económicamente de la familia, el 22% tenían un hermano, el 38% pertenecían a familias con 2 hermanos dependientes económicamente, el 26% con 3 hermanos, el 4% con 4 hermanos y el 2% con 5.

*Ingreso promedio mensual familiar.*- De la Tabla IV, se puede decir que el 29% de las familias de los entrevistados poseía ingresos mensuales promedio de menos de 200 dólares mensuales mientras el entrevistado era estudiante de la ESPOL, el 38% tenía ingresos promedio mensuales de entre 200 y 500 dólares, también, el 20% poseía ingresos mensuales promedio de entre 500 y 1000 dólares, el 9% entre 1000 y 1500 dólares y el 5% más de 1500 dólares mensuales.

Tabla IV Distribución de Frecuencias de “Ingreso promedio mensual de la familia”	
Ingreso mensual	Frecuencia Relativa
\$ 0-200	0.29
\$ 200-500	0.38
\$ 500-1000	0.20
\$ 1000-1500	0.09
Más de \$ 1500	0.05
<b>Total</b>	<b>1.00</b>

Medio de transporte que utilizaba el entrevistado para llegar a la ESPOL.- La mayoría de los entrevistados se trasladaban en transporte público a la ESPOL, desde sus hogares, mientras que el 16% lo hacía en su propio

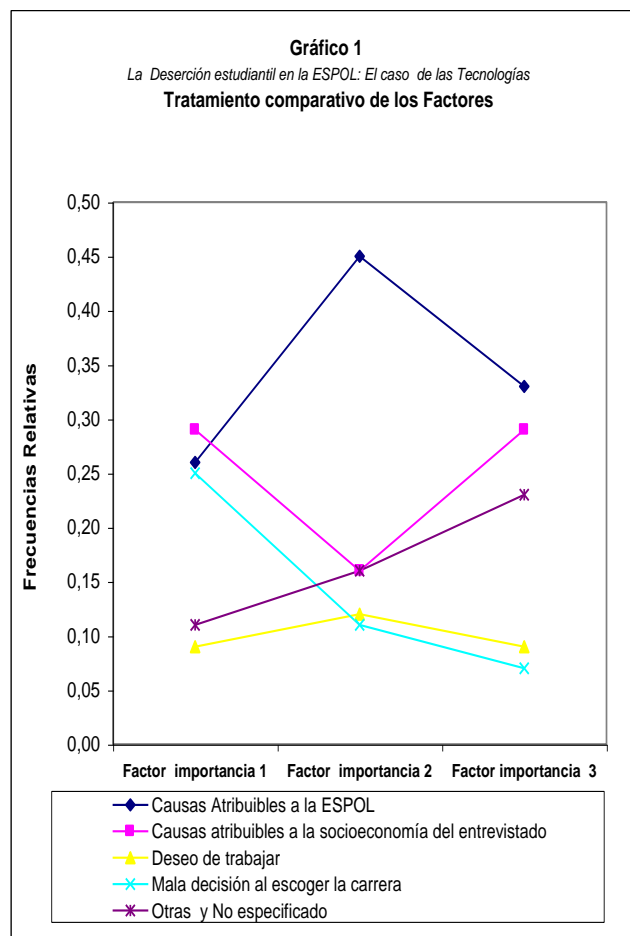
vehículo y el 6% contrataba los servicios de un expreso para trasladarse a la Politécnica.

### Factores de Deserción

En el Gráfico 1 se observa el Histograma de Frecuencia Simultáneo de los grupos de factores que llevan a un estudiante a desertar de la ESPOL y la proporción en que han sido asignados al orden de importancia según los desertores entrevistados.

Se puede apreciar que el primer Factor de deserción, la más alta proporción corresponde a “Causas atribuibles a la Socioeconomía del Desertor” seguido por las “Causas atribuibles a la ESPOL”. Para el segundo Factor de Deserción, la mayor proporción pertenece a las “Causas atribuibles a la ESPOL” y luego a las “Causas atribuibles a la Socioeconomía del Desertor” y “Otro” factor.

En el caso del tercer Factor de deserción, aproximadamente el 45% de los entrevistados atribuye su decisión de desertar a la ESPOL, al igual que el caso anterior.



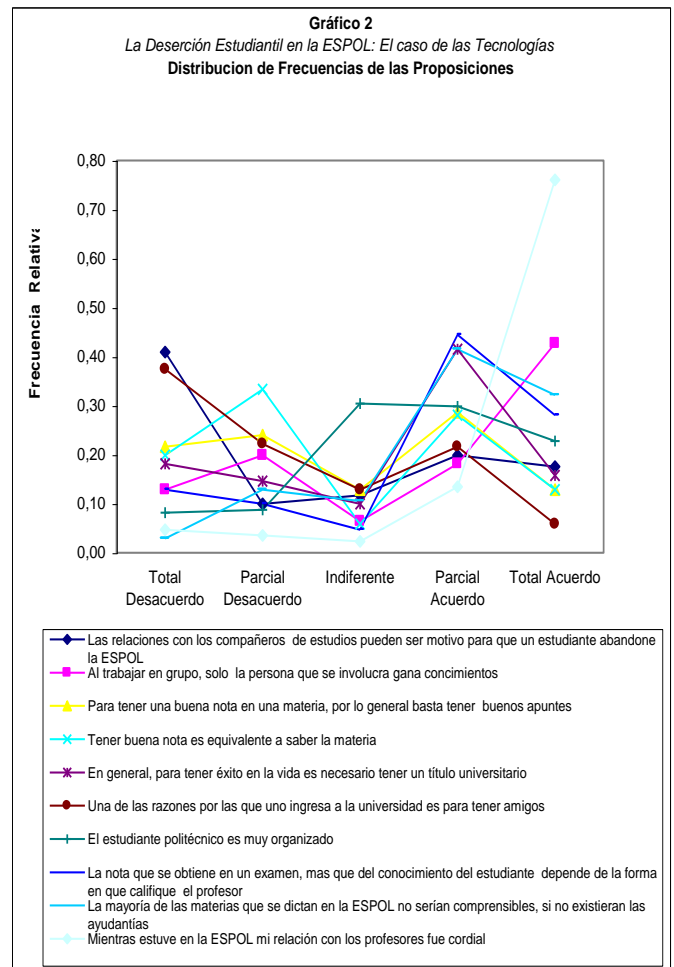
### Pensamiento del Desertor

La Sección “Pensamiento del Desertor” encierra diez proposiciones donde se busca conocer la opinión del Desertor sobre algunos aspectos. Las proposiciones mencionadas son:

- “Las relaciones con los compañeros de estudio pueden ser motivo para que un estudiante abandone la ESPOL”
- “Al trabajar en grupo, solo la persona que se involucra gana conocimientos”
- “Para tener buena nota en una materia, por lo general basta tener buenos apuntes”
- “Tener buena nota es equivalente a saber la materia”
- “En general, para tener éxito en la vida, es necesario tener un título universitario”
- “Una de las razones por las que uno ingresa a la universidad es para tener amigos”
- “El estudiante politécnico es muy organizado”
- “La nota que se obtiene de un examen, mas que del conocimiento del estudiante depende de la forma en que califique el profesor”
- “La mayoría de las materias que se dictan en la ESPOL no serían comprensibles, si no existieran las ayudantías”
- “Mientras estuve en la ESPOL mi relación con los profesores fue cordial”.

En el Gráfico 2 se presenta el Histograma de Frecuencia Simultáneo de las Proposiciones descritas anteriormente, donde se observa que aproximadamente e 45% de los desertores están en desacuerdo con la proposición “La relación con los compañeros no es motivo para que un estudiante abandone la ESPOL”, mientras que la misma proposición en la opción total acuerdo esta por debajo del 20%, también se observa que en la opción parcial acuerdo más del 30% no considera que tener buena nota es equivalente a saber la materia.

En la opción de indiferencia observamos que en la proposición “Mientras estuve en la ESPOL mi relación con los profesores fue cordial” tiene menos del 1%, mientras que en la opción total acuerdo sobrepasa el 70% superando a las otras afirmaciones, comportamiento similar tienen las proposiciones “Al trabajar en grupo, solo la persona que se involucra gana conocimientos” y “La nota que se obtiene en un examen mas que del conocimiento del estudiante depende de la forma en que califique el profesor” ambas con menos del 15% en la opción de indiferencia y en la opción total acuerdo con más del 40%.



### 3. ANÁLISIS MULTIVARIADO

En este capítulo se realiza el Análisis Multivariado para los datos correspondientes a las variables consideradas de manera simultánea.

**Análisis Bivariado.-** Una tabla bivariada es un arreglo ordenado de r filas y c columnas, donde las filas corresponden a los valores que toma la variable aleatoria discreta X y las columnas a los valores que toma la variable aleatoria discreta Y. El objeto de esta técnica es conocer la “Distribución Conjunta” entre cada par de valores posibles que pueden tomar las variables aleatorias X y Y. Es decir:

$$f(x_i, y_j) = P(X = x_i, Y = y_j)$$

donde  $f(x_i, y_j)$  es la probabilidad de que la variable X tome el valor  $x_i$  al mismo tiempo que Y toma el valor  $y_j$ .

Además de las Tablas Bivariadas, se presentan las tablas de Distribución Condicional:  $P(Y=y_j|X=x_i)$  y  $P(X=x_i|Y=y_j)$ , en donde para el primer caso, los valores de la intersección de la i-ésima fila con la j-ésima columna corresponderán al resultado de  $f_{i,j}|f_{i.}$  que es la probabilidad condicional de que Y tome el valor de  $y_j$  dado que X toma el valor de  $x_i$ . Para el caso en que la Tabla de Distribución Condicional corresponda a  $P(X=x_i|Y=y_j)$  los valores de la intersección de la i-ésima fila con la j-ésima columna corresponderán al resultado de  $f_{i,j}|f_{.j}$  que es la probabilidad condicional de que la variable X tome el valor de  $x_i$ , dado que Y toma el valor de  $y_j$ .

A continuación se presentan algunas tablas bivariadas cuya información resulta relevante para el presente trabajo.

*Después de la ESPOL vs. Nivel de Educativo Actual.* - En la Tabla V se observa que 64 de cada mil entrevistados continuó estudiando después de retirarse de la ESPOL, pero mantiene el mismo nivel educativo, es decir, el 6.4% se retiró nuevamente de la universidad donde tomó la misma carrera que cursaba en la ESPOL. El 0.6% estudió una carrera afín a la que seguía en la ESPOL, pero mantiene el mismo nivel educativo que tenía al momento de desertar, el 2.9% estudió una carrera diferente y mantiene el mismo nivel educativo formal, finalmente el 38% de los entrevistados no continuaron estudiando y por ende conservan en nivel educativo formal que obtuvieron como estudiantes de la ESPOL.

Después de la ESPOL	Nivel Educativo Actual					Marginal de "Después de la ESPOL"
	Igual al que tenía cuando se retiró de la ESPOL	Sigue estudiando pero no se ha graduado	Termino sus estudios de pregrado	estudios de pregrado y está estudiando	Termino sus estudios de postgrado	
Estudió la misma carrera que en la ESPOL	0.064	0.012	0.006	0.000	0.000	<b>0.082</b>
Estudió otra carrera afín	0.006	0.111	0.047	0.012	0.000	<b>0.175</b>
Estudió una carrera diferente	0.029	0.234	0.082	0.012	0.006	<b>0.363</b>
No continuó estudiando	0.380	0.000	0.000	0.000	0.000	<b>0.380</b>
<b>Marginal de "Nivel Educativo Actual"</b>	<b>0.480</b>	<b>0.357</b>	<b>0.135</b>	<b>0.023</b>	<b>0.006</b>	<b>1,000</b>

*Factor motivo de deserción en orden de importancia y Edad de deserción.* - El 21.10% de los entrevistados son de género masculino y consideran como primer factor motivo de su deserción a las Causas atribuibles a su socioeconomía, mientras que el 21%, se retiraron por causas, a su parecer, atribuibles a la ESPOL. En el caso del género femenino el 7.6% se retiraron por causas atribuibles a la ESPOL, siendo este el mayor porcentaje para este género.

Factor de Deserción importancia 1	Edad de Deserción		Marginal de "Primer Factor"
	Masculino	Femenino	
Causas atribuibles a la socioeconomía del desertor	0.211	0.064	<b>0.275</b>
Causas atribuibles a la ESPOL	0.210	0.076	<b>0.287</b>
Deseo de Trabajar	0.205	0.041	<b>0.246</b>
Mala decisión al escoger la carrera	0.070	0.023	<b>0.094</b>
Otro	0.064	0.035	<b>0.099</b>
<b>Marginal de "Motivo de Ingreso"</b>	<b>0.760</b>	<b>0.240</b>	<b>1.000</b>

### Análisis de Componentes Principales

El análisis de Componentes Principales es, en general, un instrumento de apoyo para otras técnicas multivariadas, que permiten la reducción de la cantidad de variables o para agrupar las observaciones con las que se trabaja. Su objetivo principal es explicar la mayor proporción de la varianza de un conjunto de n variables observables por medio de un conjunto de p variables no observables donde  $p \leq n$ .

Los Componentes Principales se presentan algebraicamente como combinaciones lineales de las p variables aleatorias observadas y geoméricamente estas combinaciones lineales representan la creación de un nuevo sistema de coordenadas obtenidas al rotar el sistema original. Permite describir la estructura de interrelación de variables originales consideradas simultáneamente, determinando así q combinaciones

lineales de  $p$  variables observables que contengan la mayor parte de la variación total, y así resumir y reducir los datos disponibles.

Sea  $\mathbf{X}^T = [X_1, X_2, \dots, X_p]$  un vector aleatorio  $p$  - variado, donde cada una de las variables que la componen son variables aleatorias observables. El vector  $p$ -variado  $\mathbf{X}$  a  $\Sigma$  tiene como matriz de varianzas y covarianzas, y sean  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$  los valores propios correspondientes a  $\Sigma$ .

Se representan las componentes principales por medio de las siguientes combinaciones lineales:

$$\begin{aligned} Y_1 &= \mathbf{a}_1^T \mathbf{X} = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \\ Y_2 &= \mathbf{a}_2^T \mathbf{X} = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p \\ &\vdots \\ Y_p &= \mathbf{a}_p^T \mathbf{X} = a_{p1}X_1 + a_{p2}X_2 + \dots + a_{pp}X_p \end{aligned}$$

Donde  $Y_1, Y_2, \dots, Y_p$  son las Componentes Principales; se supone que no están correlacionadas entre sí, son ortonormales y se cumple además que:

$$\text{Var}(Y_1) \geq \text{Var}(Y_2) \geq \dots \geq \text{Var}(Y_p) \geq 0.$$

Se puede demostrar que:

$$\text{Var}(Y_j) = \mathbf{a}_j^T \Sigma \mathbf{a}_j \quad \text{para } j=1, 2, \dots, p;$$

$$\text{Cov}(Y_i, Y_j) = \mathbf{a}_i^T \Sigma \mathbf{a}_j = 0 \quad \text{para } i \neq j$$

Se debe cumplir además:  $\|\mathbf{a}_i\| = 1$  para  $i=1,2,\dots, p$  y  $\langle \mathbf{a}_i, \mathbf{a}_j \rangle = 0$  para  $i \neq j$ . Donde  $\|\mathbf{a}_i\|$  es la norma del vector  $\mathbf{a}_i$  y  $\langle \mathbf{a}_i, \mathbf{a}_j \rangle$  es el Producto Interno de dos vectores en  $\mathfrak{R}^p$ .

En general, la  $i$ -ésima Componente Principal es la combinación lineal que maximiza la varianza de  $Y_i = \mathbf{a}_i^T \mathbf{X}$ , sujeta a que la norma del vector  $\mathbf{a}_i$  sea unitaria y la  $\text{Cov}(Y_i, Y_k) = 0$  para  $k \neq i$ .

Como resultado obtenemos que  $\Sigma$  es la matriz de covarianzas asociadas con el vector aleatorio  $\mathbf{X}^T = [X_1, X_2, \dots, X_p]$ ,  $\Sigma$  tiene los pares de valores propios y sus correspondientes vectores propios ( $(\lambda_1, \mathbf{e}_1), (\lambda_2, \mathbf{e}_2), \dots, (\lambda_p, \mathbf{e}_p)$ ) donde  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ .

Entonces, se puede probar que la  $i$ -ésima componente principal viene dada por:

$$Y_i = \mathbf{e}_i^T \mathbf{X} = e_{i1}X_1 + e_{i2}X_2 + \dots + e_{ip}X_p,$$

para  $i=1, 2, \dots, p$

Además, se puede probar de igual manera que la  $\text{Var}(Y_i) = \mathbf{e}_i^T \Sigma \mathbf{e}_i = \lambda_i$ , para  $i=1, 2, \dots, p$  y la  $\text{Cov}(Y_i, Y_k) = \mathbf{e}_i^T \Sigma \mathbf{e}_k$ , para  $i \neq k$ .

Bajo estas condiciones, el porcentaje de la varianza total contenida por la  $i$ -ésima componente principal, o su explicación viene dado por  $\frac{\lambda_i}{\sum_{i=1}^p \lambda_i} \times 100\%$ .

Para verificar si la técnica de componentes principales puede ser aplicada a un grupo de datos, se debe aplicar la prueba de Significancia Estadística de Bartlett en la cual se plantea el contraste de hipótesis que aparecen en el Cuadro VII.

**Cuadro VII**  
**Prueba de Bartlett**

$$H_0 : \Sigma = \begin{bmatrix} \sigma_{11} & 0 & \dots & 0 \\ 0 & \sigma_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{pp} \end{bmatrix}$$

vs.

$H_1 : \text{No es verdad } H_0$

Estadístico de Prueba= 818.353  
Valor p=0.000

Al aplicar la prueba a los datos de este estudio el valor p obtenido es de 0.000, de manera que existe evidencia estadística para rechazar la hipótesis nula, es decir algunos valores de las covarianzas,  $\sigma_{ij}$  son diferentes de 0 para  $i \neq j$ , y podemos afirmar que no existe independencia entre las variables de la matriz de datos de los entrevistados y se concluye que se puede proceder con el análisis de componentes principales.

Debido a que algunas de las características, no se encuentran en la misma escala, se aplicará la técnica de Componentes Principales utilizando los datos estandarizados. La estandarización significa que a cada dato observado se le resta la media estimada y se lo divide para la desviación estándar estimada de las variables; obteniéndose  $Z_1, Z_2, \dots, Z_p$ , correspondientes a las variables  $X_1, X_2, \dots, X_p$  estandarizadas.

$$Z_i = \frac{X_i - \bar{X}}{S_i}$$

Al estandarizar la matriz de datos originales y aplicar el método de Componentes Principales, con seis Componentes Principales se explica el 51% de la varianza total. En este caso se puede concluir que la Técnica de Componentes Principales no proporciona una reducción significativa de las variables.

## Análisis de Correlación Canónica

El Análisis de Correlación Canónica es una técnica estadística multivariada que permite identificar y cuantificar el grado de asociación lineal entre dos conjuntos de variables aleatorias observadas.

El primer grupo de variables se representa por el vector p-variado  $\mathbf{X}^{(1)}$  y el segundo de q variables se representa por el vector  $\mathbf{X}^{(2)}$ , donde  $p \leq q$ . Es decir:

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_p \\ \cdots \\ X_1 \\ \vdots \\ X_q \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{(1)} \\ \mathbf{X}^{(2)} \end{bmatrix}$$

Para esto determinamos p pares de variables “canónicas” de tal forma que expliquen la relación entre un primer conjunto p-variado,  $\mathbf{X}^{(1)}$  y un segundo q-variado,  $\mathbf{X}^{(2)}$ ;

De tal modo que se determina la correlación canónica mayor entre una combinación lineal de variables de un conjunto  $\mathbf{X}^{(1)}$  y una combinación lineal de variables del otro  $\mathbf{X}^{(2)}$ ; para determinar la más alta entre todos los pares de variables U, V, donde, los pares de combinaciones lineales son las variables canónicas y sus correlaciones las “Correlaciones Canónicas”, siendo:

$$\begin{aligned} U &= \mathbf{a}^T \mathbf{X}^{(1)} \\ V &= \mathbf{b}^T \mathbf{X}^{(2)} \end{aligned}$$

Que cumplen:

$$\begin{aligned} \text{Var}(U) &= \mathbf{a}^T \sum_{ii} \mathbf{a} \quad \text{Var}(V) = \mathbf{b}^T \sum_{jj} \mathbf{b} \quad \text{y} \\ \text{Cov}(U, V) &= \mathbf{a}^T \sum_{ij} \mathbf{b} \end{aligned}$$

donde,

$$\begin{aligned} \mathbf{a}_1^T &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \end{bmatrix} \\ \mathbf{b}_1^T &= \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1q} \end{bmatrix} \end{aligned}$$

Se deben hallar los  $\mathbf{a}$  y  $\mathbf{b}$  tal que:

$$\text{Corr}(U, V) = \frac{\mathbf{a}^T \sum_{ij} \mathbf{b}}{\sqrt{\mathbf{a}^t \sum_{ii} \mathbf{a}} \sqrt{\mathbf{b}^t \sum_{jj} \mathbf{b}}}$$

Se define lo siguiente:

El primer par de variables canónicas, que es el par de combinaciones lineales  $(U_1, V_1)$  tiene varianza unitaria y maximiza la correlación entre ambas.

El segundo par de variables canónicas, que es el par de combinaciones lineales  $(U_2, V_2)$  tiene varianza unitaria y maximiza la correlación entre ambas y además no está

correlacionada con el primer par de variables canónicas; para el caso del i-ésimo par  $(U_i, V_i)$

$$\begin{aligned} U_i &= a_{i1} X_1^{(1)} + a_{i2} X_2^{(1)} + \cdots + a_{ip} X_p^{(1)} \\ V_i &= b_{i1} X_1^{(2)} + b_{i2} X_2^{(2)} + \cdots + b_{iq} X_q^{(2)} \end{aligned}$$

Este análisis parte de la necesidad de encontrar asociación entre dos conjuntos de variables, para elegir cuales serán los conjuntos de estudio se considerará las combinaciones entre las secciones que conforman el cuestionario., de la siguiente manera: para el primer conjunto de variables aleatorias que conforman el vector  $\mathbf{X}^{(1)}$ , son las variables referentes a la sección “Pensamiento del desertor”, con un total de 10 variables y las concernientes a la “Trayectoria Académica del Entrevistado”, con un total de 13 variables, el vector  $\mathbf{X}^{(2)}$

### Grupo 1: “Pensamiento del Desertor”, $\mathbf{X}^{(1)}$

- Las relaciones con los compañeros de estudios son motivo de deserción.
- Al trabajar en grupo solo el que se involucra adquiere conocimientos
- Tener una buena nota, tener buenos apuntes.
- Tener buena nota es equivalente a saber la materia.
- Parar tener éxito en la vida es necesario tener un Título universitario.
- Una de las razones para ingresar a la universidad es para tener amigos.
- El estudiante politécnico es muy organizado.
- La nota que se obtiene en un examen, depende de la forma en que califique el profesor.
- Las materias no serían comprensibles sin las ayudantías.
- La relación con los profesores fue cordial.

### Grupo 2: “Trayectoria Académica”, $\mathbf{X}^{(2)}$

- Edad al terminar la primaria
- Número de Hermanos dependientes económicamente de la familia
- Ingreso mensual familiar promedio, cuando estudiaba en la ESPOL
- Número de veces que tomó el curso prepolitécnico
- Número de veces que tomó el examen de ingreso
- Factor Socioeconómico (Factor P)
- Número de materias tomadas en la ESPOL
- Número de materias aprobadas en la ESPOL
- Número de materias reprobadas en la ESPOL
- Eficiencia (Proporción de materias tomadas que fueron aprobadas por el estudiante en la ESPOL)
- Promedio General
- Edad al desertar

Se procede a determinar los coeficientes de las 10 variables canónicas entre los vectores denominados “Pensamiento del Desertor” y “Trayectoria Académica”. Presentados en la Tabla VII.



<b>Tabla VII</b> <b>Correlación Canónica</b>	
<b>Par de Variables</b>	<b>Correlación Canónica</b>
1	0.501
2	0.423
3	0.371
4	0.342
5	0.276
6	0.212
7	0.182
8	0.127
9	0.080
10	0.022

Se consideran las Correlaciones Canónicas entre 0.5 y 1 ó -0.5 y -1. Como se puede observar en la Tabla VII el valor más alto de correlación es 0.501 que según el criterio anterior se puede considerarse correlación entre los grupos de variables.

Escogemos el primer par de variables canónicas (U1, V1), donde hay una correlación canónica de 0.501, lo que indica que existe una leve relación lineal entre este par de variables

Se tiene que las características que aportan mayores pesos para la variable canónica U1 son:

- Ingreso a la universidad para tener amigos
- Relación con los profesores fue cordial

La variable canónica V1 se presenta en la Tabla IX, las características que aportan mayor ponderación son:

- Número de hermanos
- Número de exámenes de ingresos tomados
- Ingreso promedio mensual
- 

<b>Tabla IX</b> <b>Coefficientes de las Variables Canónicas</b> <b>Trayectoria académica estudiantil del desertor</b>	
<b>Vector</b>	<b>Coefficientes</b>
<b>Trayectoria académica estudiantil</b>	<b>V<sub>1</sub></b>
Edad al terminar la primaria	0.265
Número de hermanos	-0.521
Ingreso promedio mensual	0.330
Número de prepolitécnico tomados	0.021
Número de examen de ingreso tomados	0.410
Factor P	-0.032
Número de materias Tomadas	0.267
Número de materias Aprobadas	0.154
Número de materias Reprobadas	0.218
Promedio general	0.058
Eficiencia estudiantil	-0.071
Edad al desertar	0.011

<b>Tabla VIII</b> <i>Deserción Estudiantil en la ESPOL: El caso de las Tecnologías</i> <b>Coefficientes de las Variables Canónicas</b> <b>Conjunto de proposiciones planteadas al desertor</b>	
<b>Vector</b>	<b>Coefficientes</b>
<b>Conjunto de proposiciones planteadas al entrevistado</b>	<b>U<sub>1</sub></b>
Relación con los compañeros de estudio	0.227
Trabajar en grupo	0.397
Tener buenos apuntes	-0.048
Saber la materia	0.069
Tener éxito en la vida	-0.019
Ingresó a la universidad para tener amigos	0.744
El estudiante politécnico es muy organizado	-0.215
Nota que se obtiene en el examen	0.104
Existencia de las ayudantías	0.299
Relación con los profesores fue cordial	0.408

#### 4. CONCLUSIONES

El estudiante desertor de las carreras que se dictan en el Instituto de Tecnologías por lo general es de género masculino, con edad de deserción entre 19 y 21 años, que culminó la primaria en una escuela Particular a la edad de 11 años. el colegio del que proviene es particular y la especialización Informática.

Un considerable porcentaje de entrevistados, el 70% expresaron que su ingreso a la Politécnica fue por decisión propia, mediante la aprobación del Curso Prepolitécnico el cual tomó sólo una vez y no rindió el Examen de Ingreso. El Factor P asignado al desertor es menor o igual a 10.

El estudiante toma menos de 10 materias en la ESPOL, permanece en la Politécnica menos de 3 semestres y nunca obtuvo alguna Representación Estudiantil.

Al preguntar al entrevistado sobre los factores que lo llevaron a desertar de la ESPOL, el Factor de importancia 1, corresponde a la “Causas Atribuibles a la socioeconomía del desertor” según lo indicado por el desertor, seguido por la “Causas Atribuibles a la

ESPOL". Para el caso del Factor de importancia 2, el mayor porcentaje de respuesta se observó en la opción "Causas Atribuibles a la ESPOL".

El desertor opina que las relaciones con los compañeros de estudio no pueden ser motivo para que un estudiante abandone la ESPOL y piensa que al trabajar en grupo solo el que realmente se involucra gana conocimientos.

Opina también que sin las ayudantías la mayoría de las materias que se dictan en la ESPOL no serían comprensibles. Por lo general la relación entre el desertor y los profesores fue cordial.

## 5. BIBLIOGRAFIA

- [1]. **Vera, K.**, "Deserción Estudiantil en la ESPOL: Caso de las carreras del Instituto de Tecnologías, Tesis de Ingeniería en Estadística e Informática", ICM, ESPOL, 2005
- [2]. **ESPOL**, (2005) "Guía de Admisión 2005", Grandes Hitos, Guayaquil, Ecuador.
- [3]. **Meléndez, M.**, (1994) "Reprobación y Deserción Estudiantil en el ITParral: un Estudio de Caso", México (Pág.1), <http://www.uasnet.mx/centro/deptos.html>, última visita: 28 de Enero de 2005.
- [4]. **Tinto, V.**, "Una reconsideración de las teorías de deserción estudiantil", Handbook of theory and research, (Pág. 359-384), Agathon Press, New York, USA, 1986.
- [5]. **Freund, J., Miller, I., Miller, M.** (2000), "Estadística Matemática con Aplicaciones", Editorial Pearson Educación, México D.F., México.
- [6]. **Zurita, G., Maura, C.** (2003), "Estudio Estadístico acerca de la incidencia de Internet en la Educación Secundaria: El caso Urbano-Fiscal de la Provincia del Guayas", Revista Tecnológica Vol. 16, N°1, Ecuador, Guayaquil.
- [7] **Martínez, W.; Martínez, A.** (2002) "Computational Statistics Handbook with Matlab", Boca Raton, Chapman & Hall/CRC
- [8] **Pérez, C.** (2004), "Técnicas estadísticas con SPSS", Editorial Pearson Prentice Hall, Madrid, España.
- [9] **Rencher, A.**(1998), "Multivariate Statistical Analysis and Applications" New York: Wiley series in Probability and Statistics.
- [10] **CRECE – ESPOL (Centro de Registros Calificaciones y Estadística)**, Archivos históricos de los estudiantes de las Facultades de Ingeniería de los años 1994 al 2003, Guayaquil, Ecuador.
- [11] **Teobaldo, M.** (1996), "Evaluación de la Calidad Educativa en el primer año de universitario" Universidad de Buenos Aires, UBA, Buenos Aires, Argentina.
- [12] **Ricaurte, S.** (2002), "Un Análisis Estadístico de las carreras Autofinanciadas de la ESPOL", Tesis de grado ESPOL, Guayaquil, Ecuador.
- [13] **Fernández, A.** (1995), "La universidad pública en tiempos de incertidumbre", Universidad de Buenos Aires, UBA, Buenos Aires, Argentina.
- [14] **Bartlett, M.** (1957), "A note on tests of significance in Multivariate Analysis", Proceedings of the Cambridge Philosophical Society, 34, 33-40.
- [15] **Mendenhall, W., Wackerly, D. Y Scheaffer, R.** ; (1994). "Estadística Matemática con aplicaciones", 2a Edición; Editorial Iberoamericana, México, México.
- [16] **Wilkinson, L.** (1998). "SYSTAT 7.0 for Windows", SYSTAT PRODUCTS SPSS INC, Chicago, USA.
- [17] **Johnson, R. Y Wichern, W.** (1998). "Applied Multivariate Statistical Analysis". Prentice Hall, Uper Saddle river, New Jersey, USA.
- [18] Software Estadístico SPSS (1999), (Social Purpose Statistical System), version 10 para Windows.