

# FACTORES QUE INCIDEN EN LA APROBACIÓN DE UNA CARRERA UNIVERSITARIA. CASO DE ESTUDIO: CARRERAS DEL ICM

Autores: Sandra González Camba <sup>1</sup>, Fernando Guerrero Loo <sup>2</sup>

<sup>1</sup>Ingeniera En Estadística Informática, 2003, e-mail: personlight@yahoo.com

<sup>2</sup>Director de Tesis, Matemático, Escuela Superior Politécnica Nacional 1994, Postgrado en Brasil, Instituto de Matemáticas Pura y Aplicada 2000, Profesor de la ESPOL desde 2000. E-mail: alex\_fernad@yahoo.com.

## RESUMEN

En el presente trabajo se trata de determinar factores que inciden en la aprobación de materias en las carreras de Ingeniería en Estadística e Informática y Auditoría y Control de Gestión. Para ello se analizaron 15554 datos de registros de historias académicas de archivos administrativos. Se utilizó la regresión logística binaria para analizar de que manera los factores influenciaban en la aprobación o no de una materia. Los resultados del análisis fueron que el género del estudiante, el término, la materia en sí, el profesor, son variables que determinan la aprobación de la materia.

**Palabras Claves:** Aprobación de materias, Regresión Logística Binaria.

## 0. INTRODUCCIÓN

La Aprobación de materias en la ESPOL y particularmente en el ICM puede ser vista desde dos perspectivas, como profesor y como estudiante. Identificar las variables para explicar la aprobación con un modelo matemático resulta delicado, dado que entorno al tema la subjetividad juega un papel importante: nivel de atención que el estudiante tiene en las clases, que tan bueno es el profesor al dictar la materia, la dedicación que el estudiante da a la materia, el estudiante posea el material correcto, el estudiante adquiera el nivel de conocimientos que el profesor aspira, el contenido de la materia, ...etc, temas difíciles de medir. Lo que se

presenta en este trabajo es un panorama general según datos proporcionados por el CRECE de cómo el estudiante que cumple con ciertas características se desempeña en cuanto a la aprobación de materias. Se utiliza una técnica multivariada no lineal (Regresión Logística Binaria) que explicará en base a características de los estudiantes cuales son aquellos factores que influyen en él para que tenga mayor probabilidad de aprobar una materia.

Se debe aclarar que en los modelos no lineales de este tipo no se puede realizar una interpretación directa, lo que se obtiene es que un estudiante con cierta característica especial en las mismas condiciones que los demás

estudiantes tiene mayor o menor probabilidad de aprobar una materia.

La metodología que se utilizó es obtener la población de estudiantes que cursaron sus estudios durante los años 1999, 2000, 2001, junto con datos de género, factor económico P, ciudad de procedencia, tipo de colegio, término, materia, tipo de materia, profesor y horas semanales asignadas a la materia.

Para determinar las variables que expliquen la aprobación de materias se utilizó dos metodologías de investigación denominadas: Entrevista a Expertos (Profesores), Grupos de Discusión (Estudiantes), se analizó los resultados y se eligió aquellas variables medibles con datos que se podían obtener de la base de datos del Crece.

## CONTENIDO

### 1. REGRESIÓN LOGÍSTICA BINARIA.

Los modelos de regresión logística son modelos de regresión que permiten estudiar si una variable binomial depende, o no, de otra u otras variables (no necesariamente binomiales): Si una variable binomial de parámetro  $p$  es independiente de otra variable  $X$ , se cumple  $p=p|X$ , por consiguiente, un modelo de regresión es una función de  $p$  en  $X$  que a través del coeficiente de  $X$  permite investigar la relación anterior.

Para esto se puede utilizar dos tipos de modelos los Logit y Probit.

En este trabajo se utiliza el modelo Logit.

Modelo de la forma

$$Y = f(\beta_0, \beta_1 X_1, \beta_2 X_2, \dots, \beta_k X_k) + \mu$$

Donde  $f$  es la función logística. Por tanto el valor esperado de  $Y$  es:

$$E(Y) = P(Y=1) = \frac{\exp(\beta_0, \beta_1 X_1, \beta_2 X_2, \dots, \beta_k)}{1 + \exp(\beta_0, \beta_1 X_1, \beta_2 X_2, \dots, \beta_k)}$$

la estimación en modelos Logit y Probit se realiza mediante el método de Máxima Verosimilitud (MV). Además, en este tipo de modelos no resulta posible interpretar directamente las estimaciones de los parámetros, ya que son modelos no lineales.

En la práctica debe fijarse en el signo de los estimadores. Si el estimador es positivo, significará que incrementos en la variable asociada causan incrementos en  $P(Y=1)$  (aunque desconocemos la magnitud de los mismos). Por el contrario, si el estimador muestra un signo negativo, negativo, ello supondrá que incrementos en la variable asociada causarían disminuciones en  $P(Y=1)$ .

#### 1.1 ODDS RATIO

Indica cuánto más probable es el éxito que el fracaso, como parámetro característico de la distribución binomial.

$$odds = \frac{p}{1-p}$$

El odds (ventaja) representa 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. El odds asociado a un suceso es el cociente entre la probabilidad de que ocurra

frente a la probabilidad de que no ocurra, siendo  $p$  la probabilidad del suceso.

Si en la ecuación de regresión tenemos un factor dicotómico, el coeficiente  $b$  de la ecuación del factor está directamente relacionado con el odds ratio OR. Es decir es una medida que cuantifica el riesgo que representa poseer el factor correspondiente respecto a no poseerlo, suponiendo que el resto de variables del modelo permanecen constantes.

$$OR = \exp(b)$$

Cuando el coeficiente  $b$  de la variable es positivo se obtiene un odds ratio mayor que 1 y corresponde por tanto a un factor de riesgo, de poseer la característica. (En nuestro caso positivo por ser aprobación) Por el contrario, si  $b$  es negativo el odds ratio será menor que 1 (En nuestro caso no positivo).

## 2. GENERALIDADES

Inicialmente se pretendió realizar un análisis de la totalidad de los datos, pero no se lo realizó por la particularidad de que no se tiene registros en iguales intervalos de tiempo de Ing. En Estadística Informática (Historias Académicas 1999, 2000, 2001) y Auditoria y Control de Gestión (Historias Académicas 2000, 2001).

De Auditoria se tiene los registros de las materias de los primeros cuatro semestres y de Ingeniería en Estadística Informática de los ocho semestres en distintas circunstancias

(distinto término, distinto Profesor). Se presenta un análisis por carrera.

## 2 ANÁLISIS DE LOS REGISTROS DE LA CARRERA INGENIERÍA ESTADÍSTICA INFORMÁTICA:

De la Carrera Ingeniería en Estadística Informática se tiene 12172 registros de Historia Académica correspondiente a los años 1999, 2000, 2001.

### 2.1 Algoritmo de Selección:

Primero es necesario establecer el criterio de comparación de modelos y el algoritmo de selección, para escoger el mejor ajuste a los datos.

**Tabla I**  
**Algoritmo de Selección**  
**Ingeniería en Estadística Informática**

Método	Numero de Iteraciones	ULTIMA ITERACIÓN			
		Chi cuadrado	GL	Estadístico de Hosmer Lemeshow	Clasificación
Paso hacia delante(RV)	4	6,52	8	0,589	70%
Paso hacia Atrás(RV)	3	15,052	8	0,048	64%

*Realizado por: Sandra González Camba*  
*Fuente: CRECE*  
*Datos: 1999-2001*

Se eligió el modelo de pasos hacia delante (Método de la Razón de Verosimilitud), por tener mayor porcentaje de clasificación y estadístico de Hosmer-Lemeshow mas alto.

### 2.2 Test de Bondad de Ajuste

El test de Hosmer Lemeshow dio los siguientes resultados:

**Tabla II**  
**Prueba De Hosmer - Lemeshow**  
**Ingeniería en Estadística Informática**

PRUEBA DE HOSMER Y LEMESHOW			
Paso	Chi-cuadrado	gl	SIG.
4	6,52	8	0,58

*Realizado por: Sandra González Camba*  
*Fuente: CRECE*  
*Datos: 1999-2001*

No esta demás señalar que la Hipótesis Nula para el test de prueba es  $H_0 =$  El modelo Obtenido se ajusta Bien a los Datos. El criterio es No rechace  $H_0$  si el valor  $P > 0,05$ .

Como la significancia para el modelo de la técnica paso hacia delante es  $0,58 > 0,05$ ; se escogió esta técnica para modelar los datos.

### 2.3 Resultados Obtenidos

Se obtuvo los siguientes resultados:

**Tabla III**  
**Análisis de Regresión Logística**  
**Ingeniería en Estadística Informática**

	B	EXP(B)
Femenino	0,4320	1,5404
Término 1	-0,7632	0,4662
Término 2	-0,9854	0,3733
Cálculo II	-0,2562	0,7740
Estadística Matemática I	-0,1525	0,8586
Técnicas de Expresión Oral y Escrita	0,3296	1,3904
Desarrollo de Aplicaciones	0,0108	1,0109
Análisis de Variable Real	-0,2714	0,7623
Análisis Numérico	-0,2628	0,7689
Administración de Empresas	0,4524	1,5720
Cálculo I	-1,1255	0,3245
Macroeconomía	0,1480	1,1596
Contabilidad General	0,2651	1,3035
Marco Legal de la Empresa	1,0056	2,7334
Cálculo III	-0,0436	0,9574

Matemáticas Superiores	-0,5280	0,5898
Estadística Matemática II	-0,2616	0,7699
Matemáticas Discretas	-0,1411	0,8684
Álgebra Lineal	-0,9412	0,3902
Fundamentos de Computación	-0,9057	0,4042
Tratamiento Estadístico de Datos	-0,8563	0,4247
Profesor 4	0,4393	1,5516
Profesor 9	0,2086	1,2320
Profesor 10	0,2178	1,2433
Profesor 13	0,1261	1,1344
Profesor 15	0,0052	1,0053
Profesor 16	0,2755	1,3171
Profesor 25	0,2406	1,2720
Profesor 27	0,5267	1,6933
Profesor 30	0,1968	1,2175
Profesor 33	0,4102	1,5071
Profesor 45	0,3318	1,3934
Profesor 51	-0,5166	0,5965
Profesor 53	0,0240	1,0243
Profesor 55	-0,6091	0,5438
Profesor 59	-0,0275	0,9728
Profesor 63	-0,3774	0,6856
Profesor 64	-0,0717	0,9308
Profesor 67	0,2769	1,3190
Profesor 77	0,4648	1,5918
Profesor 83	0,5985	1,8193
Profesor 87	0,1220	1,1297
Profesor 89	0,2714	1,3118
Profesor 94	0,2636	1,3017
Profesor 99	-0,2547	0,7752
Profesor 100	-0,5012	0,6058
Profesor 102	0,0543	1,0558
Profesor 105	0,7592	2,1366
Profesor 111	-0,5762	0,5620
Profesor 121	-0,3828	0,6820

*Realizado por: Sandra González Camba*  
*Fuente: CRECE*  
*Datos: 1999-2001*

Los estudiantes de género femenino estadísticamente tienen mas probabilidad de aprobar una materia odds-ratio =1,5; la relación es de 3:2 por cada tres estudiantes de género femenino que aprueba hay dos

estudiantes que en las mismas condiciones reprueban.

Las Materias que ingresaron con odds-ratio mayor a uno (mas probabilidad de ser aprobadas que de ser reprobadas) ; tenemos a Marco Legal de la Empresa (odds-ratio=2,7); Administración de Empresas (odds-ratio=1,5); Técnicas de Expresión Oral y Escrita (odds-ratio=1,39); Contabilidad General (odds-ratio=1,30); Macroeconomía (odds-ratio=1,15); Desarrollo de Aplicaciones (odds-ratio=1,01); mientras mayor es el odds-ratio mayor es la probabilidad de ser aprobadas, se puede afirmar que para Marco Legal de la Empresa un estudiante tiene 2,7 veces mas probabilidad de aprobarla que de reprobarla, y así sucesivamente.

En el análisis hubo además materias que ingresaron con odds-ratio menor a uno ( menos probabilidad de ser Aprobadas ) : Cálculo I (odds-ratio=0,32), Álgebra Lineal(odds-ratio=0,39), Fundamentos de Computación(odds-ratio=0,40), Tratamiento Estadístico de Datos(odds-ratio= 0,42), Matemáticas Superiores(odds-ratio=0,58), Análisis de Variable Real (odds-ratio = 0,76), Análisis Numérico(odds-ratio = 0,76), Estadística Matemática II (odds-ratio = 0,76), Cálculo II (odds-ratio = 0,77), Estadística Matemática I (odds-ratio = 0,85), Matemáticas Discretas(odds-ratio = 0,86), Cálculo III (odds-ratio = 0,95) ; cabe anotar que mientras mas pequeño es el odds-ratio hay menos probabilidad de aprobar dicha materia, la interpretación en términos de odds-ratio se podría afirmar de la siguiente manera: para cálculo I hay una relación de paridad de 32:100 por

cada 32 estudiantes que aprueban hay 100 estudiantes que en las mismas condiciones reprueban y así sucesivamente.

La variable término I (odds-ratio = 0,46) y término II (odds-ratio=0,37), ingresaron con odds-ratio menor a uno, el segundo término tiene coeficiente mas positivo, por lo que se concluye que estadísticamente un estudiante tiene mas probabilidad de Aprobar una materia si la toma en el Primer Término.

La variable profesor hay algunos que ingresaron con odds-ratio mayores a uno y otros con odds-ratio menores a uno. Cabe notar que las materias que ingresaron con odds-ratio menor a uno, presentan un odds-ratio menor a los odds-ratio de los profesores, por lo que se podría afirmar que los estudiantes reprueban ciertas materias mas por factores de la materia misma y que por el profesor que la dictó.

Para una mejor interpretación con los resultados de las materias se realizó una Tabla de porcentajes de estudiantes reprobados por materia la que se muestra a continuación:

**Tabla IV**  
**Porcentaje de Estudiantes**  
**Reprobados Por Materia**  
**Ingeniería En Estadística Informática**

MATERIA	PORCENTAJES
Álgebra Lineal	60,47
Cálculo I	58,93
Tratamiento Estadístico de Datos	55,28
Fundamentos de Computación	51,85
Cálculo III	49,18
Cálculo II	48,82
Matemáticas Discretas	44,96

Análisis Numérico	49,43
Estadística Matemática II	49,31
Estadística Matemática I	37,41
Análisis de Variable Real	43,60
Matemáticas Superiores	42,82
Investigación de Operaciones II	37,91
Contabilidad General	36,50
Análisis Multivariado y Diseño de Experimento	44,33
Muestreo	38,99
Matemáticas Financieras	31,58
Investigación de Operaciones I	25,95
Programación Orientada a Objetos	42,72
Matemáticas Acturiales	30,84
Redes Computacionales	28,10
Microeconomía	21,21
Procesos Estocásticos	25,95
Base de Datos	23,38
Tec. Exp.Oral y Escrita	12,53
Ingeniería de Calidad	24,86
Desarrollo de Aplicaciones	16,14
Administración de Empresas	16,67
Análisis de Serie de Tiempo	28,21
Macroeconomía	12,75
Contabilidad de Costos	12,86
Utilitarios Informáticos	20,55
Análisis Financiero	11,82
Simulación Matemática	15,27
Estadística Computacional	11,92
Ciencia e Investigación	11,58
Marketing	2,59
Adm.Sist.Información	3,77
Política Empresarial	1,66
Marco Legal de la Empresa	1,96
Moneda y Banca	0,00
Investigación de Mercado	0,00

Realizado por: Sandra González Camba  
Fuente: CRECE  
Datos: 1999-2001

Las materias que mas reprueban los estudiantes de Ingeniería en Estadística Informática son Álgebra Lineal , Cálculo I, Tratamiento Estadístico de Datos, Fundamentos de Computación. Se debe notar que son

materias de los dos primeros semestres.

### 3 Análisis de los Registros de la Carrera Auditoría y Control de Gestión

Cabe anotar que de la Carrera Auditoría y Control de Gestión se tiene 3382 registros de Historia Académica correspondiente a los años 2000, 2001

#### 3.1 Algoritmo de Selección:

En la Tabla V se muestra el estadístico de Hosmer Lemeshow para los modelos de Paso Hacia delante y Paso hacia atrás, además de su porcentaje de Clasificación.

**Tabla V**  
**Algoritmo de Selección**  
**Auditoría Control de Gestión**

Método	Numero de Iteraciones	ULTIMA ITERACIÓN			
		Chi cuadrado	GL	Estadístico de Hosmer Lemeshow	Clasificación
Paso hacia delante(RV)	4	8,48	8	0,387	74,6 %
Paso hacia Atrás(RV)	3	10,55	8	0,22	73%

Realizado por: Sandra González Camba  
Fuente: CRECE  
Datos: 1999-2001

delante (Razón de Verosimilitud). Presenta un estadístico de Hosmer Lemeshow más alto y un porcentaje de clasificación mayor.

#### 3.2 Prueba de Bondad de Ajuste

Se muestra el estadístico de Hosmer-Lemeshow para la última iteración:

**TABLA VI**  
**Prueba De Hosmer – Lemeshow**  
**Auditoría Control de Gestión**

PRUEBA DE HOSMER Y LEMESHOW			
Paso	Chi cuadrado	gl	Sig.
4	8,48	8	0,387

Realizado por: Sandra González Camba  
Fuente: CRECE  
Datos: 2000-2001

El test de Hosmer-Lemeshow muestra un  $P=0,387$  que es mayor a 0,05 por lo que el ajuste a los datos es bueno.

### 3.3 Resultados Obtenidos

Del estudio de Regresión Logística se obtuvo los siguientes resultados:

**TABLA VII**  
**Modelo De Regresión Logística**  
**Auditoría Control de Gestión**

	<b>B</b>	<b>EXP(B)</b>
<b>Genero Femenino</b>	0,5283	1,6962
<b>Termino 1</b>	-0,66359	0,515
<b>Termino 2</b>	-0,87871	0,4153
<b>Matemáticas Discretas</b>	-0,90575	0,4042
<b>Matemáticas II</b>	-2,27052	0,1033
<b>Estadística I</b>	-0,94118	0,3902
<b>Matemáticas I</b>	-1,28174	0,2776
<b>Contabilidad de Costos</b>	-0,94662	0,3881
<b>Álgebra Lineal</b>	-2,5534	0,0778
<b>Fundamentos de Computación</b>	-1,57489	0,207
<b>Profesor 5</b>	-1,58445	0,2051

La variable Género femenino presenta un odds-ratio mayor a uno (1,6); un estudiante de género femenino tiene 1,6 veces más probabilidad de aprobar una materia.

Las materias Contabilidad General I, Tec. Exp. Oral, Ciencia e Investigación ingresaron con odds-ratio mayor a uno. Las materias que ingresaron con Odds-ratio menos a uno fueron Álgebra Lineal (0,07), Matemáticas II (0,10), Fundamentos de Computación (0,20), Matemáticas I (0,27), Contabilidad de Costos (0,38), Estadística I (0,39); Matemáticas Discretas (0,40). Cabe notar que las materias anteriores son en su mayoría del área de matemáticas, excepto Estadística I,

Matemáticas Discretas, Contabilidad de Costos.

La variable término I (odds-ratio = 0,51) y término II (odds-ratio=0,41), ingresaron con odds-ratio menor a uno, el segundo término tiene odds-ratio mas positivo que el primero, por lo que se concluye que estadísticamente un estudiante tiene mas probabilidad de Aprobar una materia si la toma en el Primer Término.

Se presenta la siguiente tabla de estudiantes reprobados por materia la que se muestra a continuación:

**Tabla VIII**  
**Porcentaje De Estudiantes**  
**Reprobados Por Materia**  
**Auditoria y control de Gestión**

<b>MATERIAS</b>	<b>PORCENTAJES</b>
<b>Matemáticas II</b>	60,73
<b>Álgebra Lineal</b>	59,70
<b>Fundamentos de Computación</b>	56,62
<b>Matemáticas Discretas</b>	47,90
<b>Contabilidad de Costos</b>	47,50
<b>Matemáticas I</b>	45,40
<b>Estadística I</b>	44,35
<b>Matemáticas Financieras</b>	38,96
<b>Fundamentos de contabilidad</b>	22,89
<b>Contabilidad General I</b>	16,49
<b>Organización y Métodos</b>	16,26
<b>Ética y Fundamentos de Auditoria</b>	13,97
<b>Ciencia e Investigación</b>	12,62
<b>Fundamentos de administración</b>	11,48
<b>Microeconomía</b>	9,68
<b>Auditoria Administrativa</b>	7,81
<b>Tec. Exp.Oral y Escrita</b>	6,93
<b>Utilitarios Informáticos</b>	5,77

Las materias que más reprueban los estudiantes de Auditoría y Control de Gestión son Matemáticas II, Álgebra

Lineal y Fundamentos de Computación, materias de los dos primeros semestres.

## CONCLUSIONES

El trabajo presentado tiene como objetivo investigar en base a los datos que mantiene el CRECE cuales son los factores que influyen en un estudiante para que Apruebe una materia.

Se utilizo los registros de Historias Académicas desde el año 1999 hasta el año 2001.

Las materias donde hay mayor porcentaje de estudiantes reprobados en la Carrera Ing. En Estadística Informática esta en el área de Matemáticas y Estadística. y en la Carrera Auditoria y Control de Gestión esta en el área de Matemáticas.

Las tres materias mas fuertes en Ing. En Estadística son Álgebra Lineal (60,47% de estudiantes que la toman la reprueban), Cálculo I (58,93% de estudiantes que la toman la reprueban), Tratamiento Estadístico de Datos (55,28% de estudiantes que la toman la reprueban).

Las tres materias mas fuertes en Auditoria y Control de Gestión son Matemáticas II (60,73% de estudiantes que la toman la reprueban), Álgebra Lineal (59,70% de estudiantes que la toman la reprueban), Fundamentos de Computación (56,62% de estudiantes que la toman la reprueban).

Un estudiante tiene más probabilidad de aprobar una materia si la toma en el primer término.

Los estudiantes de sexo femenino tienen mas probabilidad de aprobar una materia que sus compañeros de género masculino.

Algunos profesores ingresaron en la ecuación con coeficiente negativo por lo que los estudiantes que tomen con estos profesores tienen menos probabilidad de aprobar la materia.

Hay materias que ingresaron con coeficiente negativo por lo que un estudiante que la toma tiene menos probabilidad de aprobarlas que otras materias.

En este estudio estadístico se obvian apreciaciones subjetivas entorno al tema por lo que recomendaría se realice un estudio posterior, donde a través de una encuesta se exprese la opinión de los estudiantes y de los profesores, acerca del tema se alimente y se compare con el estudio actual.

## REFERENCIAS BIBLIOGRAFICAS

1. GONZALEZ,S. (2003). “ Análisis Estadístico de los Factores que inciden en la Aprobación de Materias en una carrera Universitaria. Caso de Estudio: Carreras del ICM”. Tesis de Grado ESPOL, Guayaquil, Ecuador.
2. Mendehall William, 1994, Estadística Matemática con Aplicaciones, Segunda Edición, Grupo Editorial Ibero América S.A. de C.V. , México.

3. Diccionario: Nuevo Espasa Ilustrado, 2003, Diccionario Enciclopédico, Editorial Espasa.
4. Francisco Abad Montes, Cádiz, 16 y 17 de Noviembre de 2000 (Area de Estadística e Investigación Operativa) IV jornadas Nacionales sobre la Enseñanza de las Relaciones Laborales.
5. Luis M. Molinero (Alce Ingeniería) , Enero 2001, La Regresión Logística I, [www. La REGRESION LOGISTICA.htm](http://www.LaREGRESIONLOGISTICA.htm)
6. Ángel Alejandro Juan Pérez (ajuamp@uoc.edu), Renatas Kizys (rkizys@uoc.edu), Luis María Manzanedo Del Hoyo (lmanzanedo@uoc.edu), Regresion Logística Binaria, [www.uoc.edu](http://www.uoc.edu)
7. Novales, A. (1993): "Econometría". McGraw-Hill.