

VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES.

Variables de incidencia sobre las pruebas de admisión a las universidades venezolanas.

Doray Contreras¹

Universidad de los Andes, Venezuela

Jimmy Corzo²

Universidad Nacional de Colombia

ABSTRACT

We use three regression models with categorical variables the first of them establishes that excepting four provenance states, applicants from public schools of the others states tend to have better scores than those in private schools. The second reveals that when the mother has a university or secondary education, candidates of the age group under 19 years obtain better scores, but paradoxically a level of higher education of the mother tends to be adverse to applicants aged between 19 and 22 years. The third model shows that the age and gender strongly influence the values of the scores, identifying that men have higher scores than women.

Key Words: regression model, ranking education, admissions test, aspirant profile.

¹Profa. Agregado Facultad de Humanidades. Departamento de Ciencias. Universidad de los Andes -Táchira. Venezuela desde 2000. Magister en Matemática mención Educación. E-mail cdoray@ula.ve

²Prof. Asociado Facultad de Ciencias. Departamento de Estadística. Universidad Nacional de Colombia. Sede Bogotá desde 1979.
Doctor en Ciencias Naturales. Universidad de Dortmund (Alemania)
E-mail jacorzos@unal.edu.co

RESUMEN

Se utilizan tres modelos de regresión con variables categóricas, el primero de los cuales permitió establecer que con excepción de cuatro estados de procedencia, los aspirantes de planteles públicos tienden a obtener mejores puntajes que los de privados. El segundo develó que cuando la madre tiene preparación académica universitaria a secundaria, los aspirantes menores a 19 años obtienen mejores puntajes, aunque un nivel de instrucción alto es contraproducente para los aspirantes entre 19 y 22 años. El tercero estableció que la edad y el género influyen decisivamente en los valores de los puntajes, develando que los hombres obtienen mejores puntuaciones que las mujeres.

Palabras Clave: Modelo de Regresión, Educación superior, pruebas de admisión, perfil aspirante.

1. INTRODUCCIÓN

El proceso de admisión al sistema de educación superior en Venezuela, constituye una oportunidad de postulación para los estudiantes del último año de educación media, y los ya bachilleres o población flotante que no han podido ingresar en su respectivo año de grado pero aún aspiran a obtener un cupo en la universidad.

El proceso de admisiones se realiza anualmente de la siguiente manera: los aspirantes registran su género, edad, código de entidad federal de residencia o procedencia geográfica, la proveniencia de plantel público o privado, ocupación del jefe de familia, nivel de instrucción de la madre, nivel de instrucción del padre, fuente de ingreso familiar, condición de alojamiento, número de personas que conforman el grupo familiar y tiempo de traslado de residencia a la institución y de tres opciones de carreras que quisiera cursar; hasta 2007 los aspirantes presentaban una prueba de razonamiento verbal y matemático y según el puntaje obtenido el aspirante obtenía un cupo

para la carrera que había escogido en alguna de las opciones (primera, segunda y tercera); actualmente después del registro la Oficina de Planificación del Sector Universitario (OPSU) distribuye a los aspirantes a las carreras solicitadas por ellos basándose en los promedios de asignaturas afines dichas carreras y en la capacidad u oferta de plazas reportadas por las instituciones de educación superior.

El objetivo de éste estudio es identificar la influencia de ciertas variables sociales, demográficas y económicas, o interacciones entre ellas, sobre puntajes obtenidos por los aspirantes en las pruebas para conseguir un cupo en las universidades venezolanas, en el período comprendido entre los años 2003 y 2007 inclusive. Para lo anterior se utilizó el modelo de regresión lineal general con regresores categóricos con el cual se pudo identificar los factores que mayor incidencia tienen sobre los puntajes en las pruebas de admisión.

2. DESCRIPCIÓN DE LA BASE DE DATOS

La base de datos, fue proporcionada por el Oficina de Planificación del Sector Universitario, adscrita al Consejo Nacional de Universidades, OPSU-CNU, Caracas, Venezuela. Las características específicas del grupo objetivo de estudio están referidas a los aspectos demográficos y socioeconómicos de los aspirantes a un cupo en las universidades Venezolanas, y registrados en la OPSU entre el año 2003 y el año 2007. Los datos iniciales contenían los registros de 2'144.136 aspirantes, y a través de un proceso de depuración, se estableció una base de datos definitiva con 1'931.400 aspirantes. Este proceso de depuración consistió en colocar algunos filtros con el fin de desechar algunos registros o valores que no se ajustan a la realidad, tales como

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

estudiantes que tenían un registro de edad no apta para estar en este nivel educativo como es el caso de individuos que aparecían por ejemplo con 1, 2 o 120 años de edad, o individuos con registros de año de grado por ejemplo de 1900 o 2017, también se eliminaron aquellos registros con datos faltantes.

La base de datos contiene 9 variables categóricas que se describen a continuación:

Género.

1. masculino
2. femenino

Edad.

1. <= 16
2. 17-18
3. 19-22
4. > 22

Dependencia del Plantel.

1. Público
2. Privado

Código de entidad federal de residencia del aspirante.

- | | | | |
|-------------------------|--------------------|--------------------|-----------------------|
| 1. Distrito Capital(11) | 2. Miranda(12) | 3. Vargas(13) | 4. Aragua(21) |
| 5. Carabobo(22) | 6. Cojedes(23) | 7. Apure(31) | 8. Guárico(32) |
| 9. Falcón(41) | 10. Lara(42) | 11. Portuguesa(43) | 12. Yaracuy(44) |
| 13. Zulia(51) | 14. Barinas(61) | 15. Mérida(62) | 16. Táchira(63) |
| 17. Trujillo(64) | 18. Anzoátegui(71) | 19. Monagas(72) | 20. Nueva Esparta(73) |
| 21. Sucre(74) | 22. Bolívar(81) | 23. Amazonas(82) | 24. Delta Amacuro(83) |

Ocupación del jefe de la familia.

1. Profesión universitaria, propietarios o gerente de empresas, ejecutivos de alto nivel.
2. Técnicos superiores universitarios, productores, comerciantes y gerentes medianos, artistas.
3. Empleados de organismos públicos y privados, técnicos medios, pequeños empresarios o productores.
4. Obreros y trabajadores del sector informal con primaria completa, chóferes, plomeros, electricistas y similares.
5. Obreros del sector informal no especializados, con primaria incompleta, jornaleros, trabajadores con ingreso inestable.

Nivel de instrucción de la madre/padre.

1. Profesión universitaria
2. Secundaria completa
3. Secundaria incompleta
4. Primaria completa
5. Analfabeta

Fuente de ingreso de la familia.

1. Fortuna heredada o adquirida.
2. Honorarios profesionales, ganancias por inversión.
3. Salario o remuneración mensual o quincenal.
4. Trabajos ocasionales, inestables, y a destajo.
5. Ayudas, donaciones o subsidios.

Condiciones de alojamiento.

1. Vivienda de lujo y máxima comodidad.
2. Casa o apartamento ubicada en zonas.
3. Residenciales elegantes.
4. Casa o apartamento en buenas condiciones.
5. Vivienda de construcción sólida con incomodidad.
6. Rancho o vivienda en pésimas condiciones.
- 7.

Año en que presentó la prueba:

1. 2003
2. 2004
3. 2005
4. 2006
5. 2007

3. MATERIALES Y MÉTODOS

En esta sección se introducen las dos metodologías utilizadas en desarrollo del trabajo. La de análisis que describe la forma como se organizó la base de datos; las consideraciones tomadas en cuenta para el análisis; y la metodología estadística que incluye la presentación del modelo estadístico utilizado y los supuestos requeridos para su adecuada aplicación.

3.1 Metodología de análisis.

La población objeto de este estudio puede definirse como el conjunto de todos los aspirantes que presentaron pruebas para ser admitidos a las universidades venezolanas en el período comprendido entre los años 2003 y 2007 inclusive. Sin embargo, hay una población o universo más grande conformado por todos aquellos que podrían haber aspirado a un cupo en la universidad pero que nunca se presentaron. Con esta óptica, el conjunto de datos disponible es una muestra aleatoria pero no probabilística de ese universo desconocido de posibles aspirantes.

Este entorno de universo y muestra justificaría el uso de pruebas de significancia estadística para establecer los grados de influencia de algunas de las variables disponibles en la base de datos, sobre los puntajes de las pruebas. Sin embargo, el gran volumen de información disponible conduce a un problema estructural de la estadística descubierto en el contexto de la minería de datos, pues con grandes volúmenes de información las pruebas de significancia estadística tienden a rechazar siempre la hipótesis nula. Por lo anterior, los supuestos de los modelos y la justificación del uso de un solo modelo para todos los períodos fueron hechos en muestras de la población objeto.

Posteriormente se estimaron los parámetros de los modelos sobre los puntajes globales de todos los aspirantes en las pruebas durante todos los períodos. Los resultados en todos los casos van acompañados de las desviaciones estándar de los parámetros, los valores test y los valores p usados como indicadores de la calidad del modelo.

3.2 Metodología estadística.

Como la variable de estudio es de carácter continuo por tratarse de puntajes entre cero y cien, y las variables explicativas son categóricas, se utilizará para el análisis un modelo de regresión lineal simple con regresores categóricos.

Antes de introducir dicho modelo, se define el modelo de regresión lineal general clásico (MRL), el cual tiene lugar cuando se supone que los valores de una variable dependiente Y_i , dependen de los valores de otras variables independientes $X_{i1}, X_{i2}, \dots, X_{ik}$, $i = 1, \dots, n$, por medio de una ecuación lineal de la siguiente manera

$$Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik} + \varepsilon_i, \quad i = 1, \dots, n.$$

Los $\beta_1, \beta_2, \dots, \beta_k$ son parámetros desconocidos que serán estimados y ε_i variables aleatorias independientes o errores aleatorios con distribución normal estándar de media cero y varianza σ^2 . Un parámetro β_i se interpreta como el aporte de la variable X_i a la respuesta Y_i del individuo i .

El modelo lineal también se puede formular en forma más compacta con notación matricial definiendo los siguientes vectores y matrices:

$$Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & X_{11} & \dots & X_{1k} \\ \vdots & \vdots & & \vdots \\ 1 & X_{n1} & \dots & X_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Con esta notación el modelo se puede escribir en forma condensada matricialmente como sigue:

$$Y = X \beta + \varepsilon$$

La validez del modelo completo se evalúa probando la hipótesis $H_0: \beta_0 = \beta_1 = \dots = \beta_k = 0$, frente a la alternativa $H_1: \beta_i \neq 0$ para al menos un i y los resultados se organizan como en la Tabla 1, llamada tabla de análisis de varianza la cual tiene la siguiente estructura:

Tabla 1. Análisis de Varianza para el modelo lineal general

Fuente de variación	gl	Sumas de cuadrados	Cuadrados medios	F	Significancia
Modelo	k	$SCM = \beta^t X^t Y$	$CMM = SCM/k$	$\frac{CMM}{CME}$	Valor p
Error	$n - k$	$SCE = Y^t Y - \beta^t X^t Y$	$CME = SCE/n - k$		
Total	n	$SCT = Y^t Y$			

En la Tabla 1 el valor F se utiliza para verificar si el aporte de la suma de cuadrados del modelo a la suma de cuadrados total es significativo o no. Así, un valor grande de F con significancia menor o igual al 5% (o al 1%) será indicador de que el modelo explica adecuadamente la variabilidad de la respuesta Y_i . Cuando la hipótesis H_0 se rechaza con un valor grande de F, se concluye que el modelo da cuenta de la variabilidad de la repuesta Y .

Nótese que en el MRL todas las variables involucradas son de carácter continuo, razón por la que una variable categórica como la entidad territorial, o el tipo de plantel de donde procede un aspirante, no pueden utilizarse para explicar el comportamiento de los puntajes. Entonces se hace necesario introducir algunas modificaciones en el MRL que permita variables categóricas. El caso más sencillo de regresión con variables categóricas ocurre por ejemplo si queremos explicar el comportamiento de los puntajes a partir de una variable como el tipo de plantel de donde egresó el aspirante que es una variable categórica. En este caso el modelo se construye definiendo la siguiente variable indicadora del tipo de plantel:

$$X_1 = \begin{cases} 1 & \text{si el aspirante viene de un plantel público} \\ 0 & \text{si el aspirante viene de un colegio privado} \end{cases}$$

Entonces el modelo se puede escribir como en la siguiente ecuación:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \varepsilon_i, \quad i = 1, \dots, n$$

En el anterior modelo los parámetros se pueden interpretar como sigue: si el i -ésimo aspirante es egresado de un plantel público, entonces $X_{i1} = 1$, y por lo tanto el modelo se transforma en:

$$Y_i = (\beta_0 + \beta_1) + \varepsilon_i, \quad i = 1, \dots, n,$$

Indicando que $\beta_0 + \beta_1$ es el puntaje de un aspirante de plantel público. Por otra parte, cuando el i -ésimo aspirante viene de un plantel privado, entonces $X_{i1} = 0$, y el modelo toma la forma:

$$Y_i = \beta_0 + \varepsilon_i, \quad i = 1, \dots, n,$$

de manera que β_0 es el puntaje de un aspirante de colegio privado.

3.3 Modelo para el análisis de la procedencia y el tipo de plantel.

El primer modelo se construirá para determinar la influencia sobre los puntajes según el estado de procedencia del aspirante y según el tipo de plantel público y privado. Para definir las p variables indicadoras de los estados es necesario codificarlos como se muestra en la

Tabla 2, contiene un listado de los 24 estados venezolanos con los

<i>j</i>	<i>nombre del estado</i>	<i># de aspirantes</i>	<i>j</i>	<i>nombre del estado</i>	<i># de aspirantes</i>
1	Distrito Capital (11)	207872	13	Zulia(51)	194740
2	Miranda(12)	145432	14	Barinas(61)	44936
3	Vargas(13)	20548	15	Mérida(62)	80256
4	Aragua(21)	125810	16	Táchira(63)	89064
5	Carabobo(22)	170072	17	Trujillo(64)	43592
6	Cojedes(23)	21221	18	Anzoátegui(71)	78868
7	Apure(31)	24923	19	Monagas(72)	51540
8	Guárico(32)	47874	20	Nueva Esparta(73)	24705
9	Falcón(41)	58691	21	Sucre(74)	47464
10	Lara(42)	118797	22	Bolívar(81)	89766
11	Portuguesa(43)	47828	23	Amazonas(82)	5578
12	Yaracuy(44)	38755	24	Delta Amacuro(83)	7022

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

códigos de las entidades federales, los códigos asignados denotados por *j* al lado izquierdo y el número de aspirantes por estado.

Tabla 2. Códigos y número de aspirantes por estado

<i>j</i>	<i>nombre del estado</i>	<i># de aspirantes</i>	<i>j</i>	<i>nombre del estado</i>	<i># de aspirantes</i>
1	Distrito Capital (11)	207872	13	Zulia(51)	194740
2	Miranda(12)	145432	14	Barinas(61)	44936
3	Vargas(13)	20548	15	Mérida(62)	80256
4	Aragua(21)	125810	16	Táchira(63)	89064
5	Carabobo(22)	170072	17	Trujillo(64)	43592
6	Cojedes(23)	21221	18	Anzoátegui(71)	78868
7	Apure(31)	24923	19	Monagas(72)	51540
8	Guárico(32)	47874	20	Nueva Esparta(73)	24705
9	Falcón(41)	58691	21	Sucre(74)	47464
10	Lara(42)	118797	22	Bolívar(81)	89766
11	Portuguesa(43)	47828	23	Amazonas(82)	5578
12	Yaracuy(44)	38755	24	Delta Amacuro(83)	7022

Entonces para construir el modelo es necesario definir las siguientes 23 variables indicadoras del estado de procedencia:

$$X_j = \begin{cases} 1 & \text{si el aspirante viene del estado } j \\ 0 & \text{si el aspirante no viene del estado } j \end{cases}, \quad j = 1, \dots, 23$$

Nótese que la definición de estas variables excluye la posibilidad de que todas o cualquier subconjunto de ellas tomen simultáneamente el valor uno. Así por ejemplo $X_1 = 1$ indica que el aspirante viene del Distrito Capital, mientras que $X_{18} = 1$ indica que el aspirante viene de Anzoátegui, y así sucesivamente. Solo se definieron 23 variables indicadoras debido a que cuando todas ellas tomen el valor cero, queda automáticamente definido el código para el estado de Delta Amacuro.

También es necesario definir una variable indicadora del tipo de plantel así;

$$Z = \begin{cases} 1 & \text{si el el plantel donde se graduó el aspirante es público} \\ 0 & \text{si el el plantel donde se graduó el aspirante es privado} \end{cases}$$

Con estas variables indicadoras el modelo para los puntajes Y en función del estado de procedencia para el i -ésimo aspirante se puede escribir así:

$$Y_i = \alpha_0 + \alpha_1 X_{i1} + \alpha_2 X_{i2} + \dots + \alpha_{23} X_{i23} + \alpha_{24} Z_i + \sum_{j=1}^{23} \gamma_j X_{ij} Z_i + \varepsilon_i, \quad i = 1, \dots, n$$

donde los productos $X_{ij} Z_i = 1$ cuando el aspirante i se graduó en un plantel público y viene del estado j , y representan las interacciones

Estado de Procedencia por tipo de Plantel. Por ejemplo, $X_{i,16}Z_i = 1$, significa que el aspirante i viene de un plantel público del estado del Táchira.

Para interpretar los parámetros nótese por ejemplo que cuando $X_{ij} = 0$ para todo j y $Z_i = 0$, el parámetro β_0 representa el aporte de un aspirante del estado de Delta Amacuro egresado de colegio privado al puntaje global. Así mismo, si $X_{ij} = 0$ para todo j y $Z_i = 1$, la suma $\beta_0 + \beta_{24}$ representa el aporte de un aspirante del estado de Delta Amacuro egresado de plantel público al puntaje global. En general, cuando todas las variables toman el valor cero excepto la j -ésima, y $Z_i = 1$, el valor de $\beta_0 + \beta_j + \beta_{24}$ corresponde al aporte de aspirantes de planteles públicos del estado j al puntaje global. Además, si por ejemplo $X_{i1} = X_{i2} = 1$, y $X_{ij} = 0$, para las restantes $j=3, \dots, 23$, y $Z_i = 1$, entonces $\beta_0 + \beta_1 + \beta_2 + \beta_{24}$ representa el aporte de los aspirantes de planteles públicos del Distrito Capital y del estado de Miranda al puntaje global.

3.4 Modelo para el análisis del nivel de instrucción de la madre y la edad.

El segundo modelo que se propone es para estudiar la incidencia del nivel de instrucción de la madre y la edad en los puntajes obtenidos en las pruebas. Para esto, siguiendo la metodología establecida anteriormente, se definen las siguientes variables auxiliares para los niveles de instrucción de la madre:

$$W_j = \begin{cases} 1 & \text{si la madre del aspirante tiene el nivel de instrucción } j \\ 0 & \text{si no} \end{cases}, \quad j = 1, \dots, 4$$

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

Y las variables para los grupos etarios

$$E_k = \begin{cases} 1 & \text{si el aspirante esta en el grupo de edad } j \\ 0 & \text{si el aspirante no esta en el grupo de edad } j \end{cases}, \quad k = 1,2,3$$

De manera que la interacción edad nivel educativo de la madre se puede representar por

$$E_k W_j = \begin{cases} 1 & \text{si el aspirante es del grupo de edad } k \\ & \text{y su madre tiene el nivel de instrucción } j \\ 0 & \text{en cualquier otro caso} \end{cases}, \quad \begin{matrix} j = 1, \dots, 4 \\ k = 1,2,3 \end{matrix}$$

Entonces el modelo para el análisis de la influencia del nivel de instrucción de la madre y la edad en los puntajes de los aspirantes es el siguiente:

$$Y_i = \beta_0 + \beta_1 W_{i1} + \beta_2 W_{i2} + \beta_3 W_{i3} + \beta_4 W_{i4} + \beta_4 S_i + \sum_{j=1}^4 \sum_{k=1}^3 \zeta_j E_{ik} W_{ij} + \varepsilon_i, \quad i = 1, \dots, n$$

3.5 Modelo para el análisis de la edad y el género del aspirante.

El último modelo se construirá para identificar la influencia de la edad y el género de los aspirantes sobre los puntajes globales obtenidos en las pruebas. Para esto se definen las siguientes variables auxiliares del género:

$$S = \begin{cases} 1 & \text{si el aspirante es hombre} \\ 0 & \text{si el aspirante es mujer} \end{cases}$$

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

Entonces las interacciones se obtienen por medio de los siguientes productos

$$E_j S = \begin{cases} 1 & \text{si el aspirante es un hombre del grupo de edad } j \\ 0 & \text{en cualquier otro caso} \end{cases}, \quad j = 1, \dots, 4$$

De manera que el modelo se puede escribir así:

$$Y_i = \delta_0 + \delta_1 E_{i1} + \delta_2 E_{i2} + \delta_3 E_{i3} + \delta_4 S_i + \sum_{j=1}^3 \tau_j E_{ij} S_i + \varepsilon_i, \quad i = 1, \dots, n$$

3.6 Contraste de Levene para la igualdad de varianzas.

Uno de los supuestos importantes para la validez del modelo es el de la homogeneidad de varianzas de los errores en los grupos analizados. Para validarlo se utilizó el test de Levene, con el cual se prueba de la hipótesis de igualdad de varianzas de los errores en k grupos. El test consiste en un Análisis de Varianza (ANOVA) de los valores absolutos de las desviaciones de los valores muestrales respecto a un estadístico de centralidad que generalmente es el promedio. Para su cálculo se procede de la siguiente manera:

1. Calcular las diferencias en valor absoluto entre las observaciones y el promedio por grupo:

$$D_{ij} = |X_{ij} - \bar{X}_j|,$$

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

donde X_{ij} es el puntaje global del aspirante i proveniente del estado j y \bar{X}_j es el puntaje promedio del estado j .

2. Calcular la media de las diferencias en el grupo j : $\bar{D}_j = \frac{\sum_i D_{ij}}{n_j}$,
 donde n_j es el número de aspirantes en el grupo j .

3. Calcular la media de total de las diferencias: $\bar{D}_t = \frac{\sum_{i=1}^n \sum_{j=1}^k D_{ij}}{N}$,
 donde N es el número de aspirantes que se presentaron a la prueba.

4. Calcular la suma de cuadrados de las diferencias intra grupo:

$$CM_{intra-grupos} = \sum_{i=1}^n \sum_{j=1}^k (D_{ij} - \bar{D}_j)^2$$

5. Calcular la suma de cuadrados inter grupo:

$$CM_{inter-grupos} = \sum_{j=1}^k n_j (\bar{D}_j - \bar{D}_t)^2$$

6. Calcular los grados de libertad:

$gl_{inter-grupos} = k - 1$, siendo k los 24 estados de Venezuela.

$gl_{intra-grupos} = \sum_{j=1}^k (n_j - 1)$, siendo n_j el tamaño muestral del grupo j .

7. Calcular la media cuadrática inter-grupos: $CM_{inter-grupos} = \frac{SC_{inter}}{gl_{inter}}$

8. Calcular la media cuadrática intra-grupos: $CM_{intra-grupos} = \frac{SC_{intra}}{gl_{intra}}$

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

9. Calcular la $F = \frac{CM_{inter-grupos}}{CM_{intra-grupos}}$

10. La información obtenida para hacer la prueba de la hipótesis de igualdad de varianzas se organiza en la siguiente tabla:

Tabla 3. Contraste de Levene para igualdad de Varianzas

F	gl_{inter}	gl_{intra}	Sig.
$\frac{CM_{inter-grupos}}{CM_{intra-grupos}}$	$k - 1$	$\sum_{j=1}^k (n_j - 1)$	Valor p

donde k es el número de grupos. La hipótesis de igualdad de varianzas entre los estados grupos se rechaza cuando el valor p es menor o igual que el nivel de significancia del 5%.

3.7 Validación de los supuestos en los modelos utilizados.

Las pruebas de Levene realizadas en muestras de la población objeto sobre los puntajes globales para validar el supuesto de homogeneidad de varianzas en los tres modelos, concluyeron que en las muestras el supuesto de homogeneidad de varianzas no se podía garantizar, aunque al calcular las varianzas en la población completa no se observaron mayores diferencias. Posteriormente, para estabilizar la varianza se calculó el logaritmo de los puntajes globales y se volvieron a realizar las pruebas concluyéndose que para esta nueva variable si se puede garantizar la homogeneidad de las varianzas en las muestras. Ante esta disyuntiva se estimaron los modelos con los puntajes originales y con los logaritmos de los puntajes, observándose que los resultados no difieren de los modelos con logaritmos a los modelos con los puntajes originales. Por esta razón y por simplicidad en las interpretaciones se decidió utilizar para el análisis los resultados con los puntajes originales.

En cuanto al supuesto de normalidad se validó por inspección de los gráficos Q y Q calculados sobre todos los valores de los puntajes en la población, teniendo en cuenta que el volumen de información 1'931.400 aspirantes en los cinco años analizados facilita el cumplimiento de este requisito.

3.8 Elección de un solo modelo para todo el período analizado.

Por disponer de una serie de datos longitudinales, lo más natural para su tratamiento es tener en cuenta la dinámica de los resultados de las pruebas en el tiempo. Una forma de identificar dicha dinámica es establecer si las diferencias entre puntajes promedio por año difieren

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

significativamente. Para esto es necesario verificar primero la homogeneidad de las varianzas a través de los años. Lo anterior se verificó con la prueba e Levene anteriormente descrita a partir de una muestra del 1% de los aspirantes de los cinco años en análisis.

En la Tabla 4 se muestran las proporciones de aspirantes por año en la población y en la muestra que será utilizada para la verificar la homogeneidad de varianzas y de promedios a través de los años. Se observa que dichas proporciones se mantienen para todos los años, razón por la que se consideró que la muestra representa adecuadamente los años del período analizado.

Tabla 4. Distribución de aspirantes en la muestra y en la población

Año de presentación de la prueba de selección	Porcentaje de aspirantes en la muestra	Porcentaje de aspirantes en la población
2003	19,8	20,0
2004	21,9	21,8
2005	17,5	17,5
2006	21,2	21,5
2007	19,7	19,1
Total	100,0	100,0

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

Los resultados de la prueba de Levene indican que el valor p obtenido es mucho mayor que el nivel de significancia prestablecido en 5% (Tabla 5).

Tabla 5. Prueba de Levene sobre la igualdad de las varianzas error

F	gl1	gl2	Sig.
1,302	4	19086	,267

Se concluye que las varianzas no difieren de año a año, y por lo tanto se puede proceder a realizar una prueba de comparación de medias.

En la Tabla 6 se presentan las medias y desviaciones estándar por año. Se observa que las medias y las desviaciones estándar son muy parecidas con excepción del año 2007, en el cual la media es ligeramente menor que las de los otros años.

Tabla 6. Promedios y desviaciones estándar por año

Año en que presentó la prueba	Media	Desviación típica	N
2003	50,3324	8,12867	3781
2004	50,2603	8,32394	4180
2005	50,3839	8,43582	3333
2006	50,0059	8,35195	4045
2007	49,7995	8,13165	3752
Total	50,1517	8,27582	19091

En la Tabla 7 se muestran los resultados del análisis de varianza realizado para comparar los promedios por años. Se puede observar

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**
 www.revistaorbis.org.ve / núm 23 (año 9) 159-205

que a un nivel de significancia ligeramente menor del 1% no hay evidencia de que los promedios entre años sean diferentes.

Tabla 7. Análisis de varianza para comparación de puntajes promedio por año

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Inter-años	903,888	4	225,972	3,301	,010
Intra-años	1306553,678	19086	68,456		
Total	1307457,566	19090			

El anterior análisis justifica el uso de un solo modelo para todos los años.

4. RESULTADOS Y ANÁLISIS.

En esta sección se presentan los resultados de los tres análisis seleccionados. La primera de interés general para el sistema de admisiones a las universidades venezolanas: la interacción estado de procedencia con tipo de plantel; la segunda, de interés cultural académico y corresponde a la interacción del nivel de instrucción de la madre con la edad de los aspirantes; la tercera de interés demográfico correspondiente a la interacción del grupo etario con género.

4.1 Análisis de la influencia del Estado de Procedencia y Tipo de Plantel sobre los puntajes.

La Tabla 8 muestra las estadísticas descriptivas para la combinación de los factores estado de procedencia de los aspirantes y Tipo de Plantel donde se graduaron para todo el 1'931.400 de aspirantes. Se observa que con excepción de Cojedes, estado en el que coinciden los promedios, en los restantes 23 estados los aspirantes que provienen de los planteles privados obtuvieron puntajes superiores a los de aspirantes que provienen de los planteles públicos; lo anterior también ocurre con el puntaje promedio total de todos los estados, que es de 49,35 en los planteles públicos, y de 51,52 en los privados. En el Gráfico 1 se visualiza esta situación.

Tabla 8. Estadísticas descriptivas del puntaje global por tipo de plantel

Estado	Desv.		Desv.	
	Media Oficiales	Estándar Oficiales	Media Privados	Estándar Pr
Distrito Capital	48.80	7.70	51.30	8.89
Miranda	49.33	7.94	51.61	8.48
Vargas	49.31	8.29	51.25	8.56

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING
ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

Aragua	49.19	7.61	51.73	8.96
Carabobo	49.19	7.65	51.35	8.83
Cojedes	49.97	8.11	50.01	8.23
Apure	49.69	7.82	52.40	8.99
Guárico	49.55	7.77	50.39	8.13
Falcón	50.73	8.98	51.33	8.45
Lara	49.50	7.93	52.16	9.05
Portuguesa	49.73	8.14	51.07	8.63
Yaracuy	49.52	7.85	52.27	8.75
Zulia	48.88	7.52	51.94	8.58
Barinas	49.74	7.70	50.88	8.81
Mérida	49.77	8.01	51.81	8.72
Táchira	49.79	8.17	50.79	8.42
Trujillo	49.13	7.59	52.15	8.94
Anzoátegui	49.14	7.61	51.12	8.58
Monagas	49.65	7.98	52.08	9.28
Nueva Esparta	48.19	7.11	53.26	9.75
Sucre	49.85	7.98	50.48	8.67

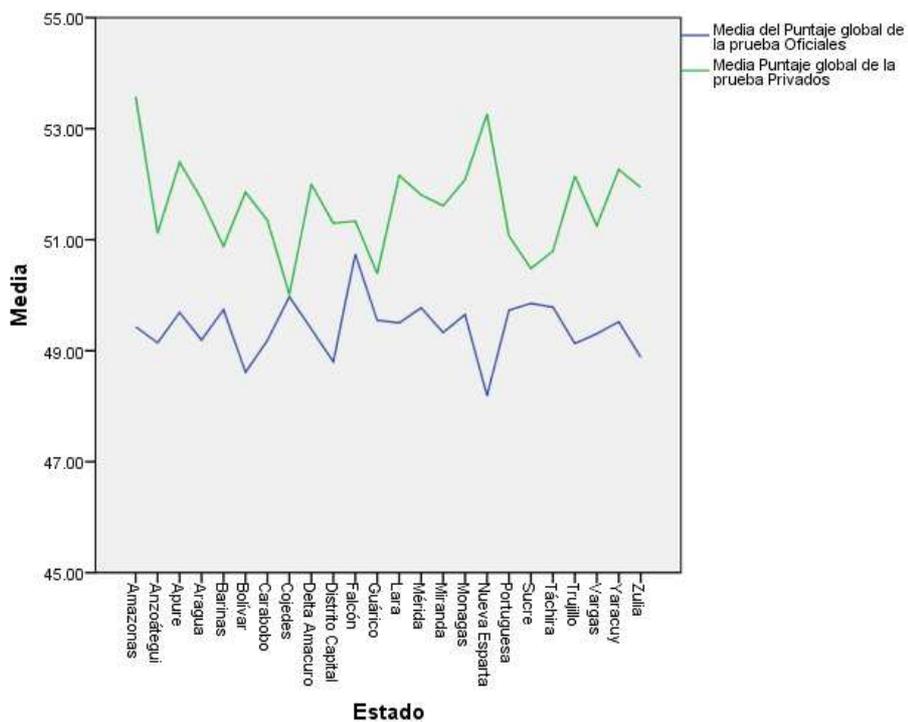
Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

Bolívar	48.61	7.21	51.86	9.27
Amazonas	49.43	7.55	53.58	10.06
Delta Amacuro	49.39	7.50	52.00	9.34
Total	49.35	7.83	51.52	8.79

Gráfico 1. Comparación de puntajes promedio en los estados por tipos de planteles



En cuanto a las desviaciones estándar, se puede decir que, con excepción del estado de Falcón, los planteles privados tienen mayores desviaciones estándar que los públicos, es decir, los puntajes de aspirantes de planteles públicos son más homogéneos, siendo en todo caso, menores que los puntajes de aspirantes de planteles privados. Esta situación se ilustra en el Gráfico 2. Por otra parte, el hecho de que los puntajes de aspirantes de planteles privados tengan desviaciones estándar mayores, indica que puede haber planteles privados con puntajes muy superiores al promedio, pero también que algunos planteles privados, especialmente de aquellos estados con mayores dispersiones, pueden tener puntajes promedio más bajos que los de algunos públicos.

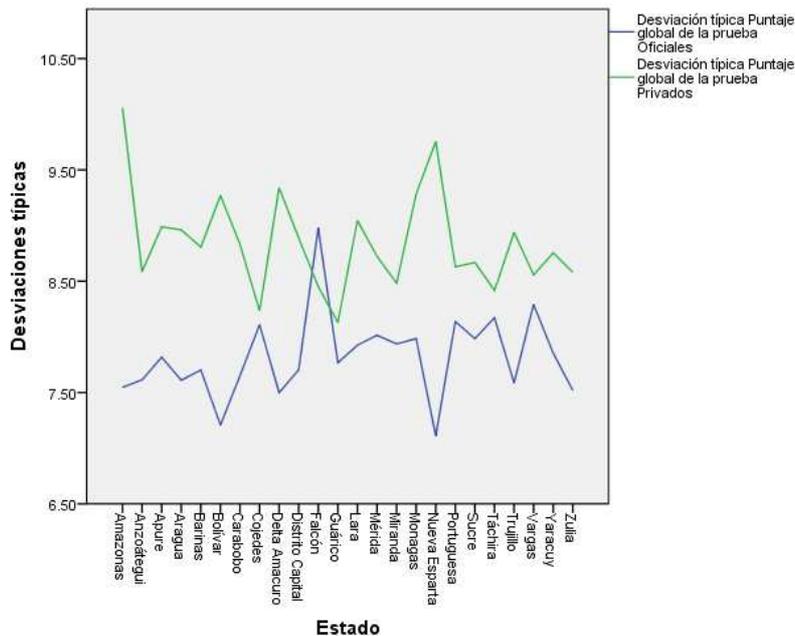
Ejemplos de la situación anterior pueden ocurrir en estados como Amazonas, Bolívar, Delta Amacuro y Nueva Esparta, que son los estados con las mayores dispersiones en los privados y las menores en los públicos. Como se anunció arriba, se distingue de todo el conjunto el estado de Falcón, en el cual la desviación estándar de los puntajes de planteles oficiales es mucho mayor que la de todos los demás planteles oficiales, y también mayor que la de los colegios privados.

Gráfico 2. Comparación de las desviaciones típicas en los estados por tipos de planteles

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205



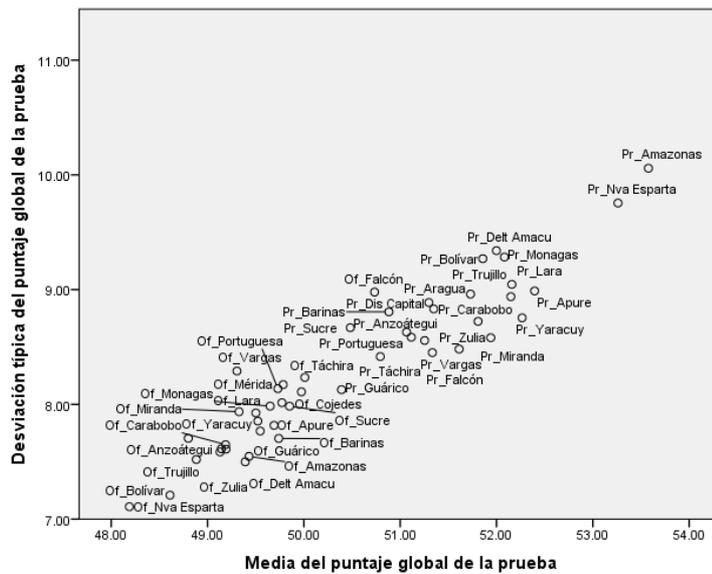
En el Gráfico 3 se muestran los promedios y las desviaciones estándar para grupo en la interacción estado por tipo de plantel calculados para el 1'931.400 aspirantes. En general se nota una tendencia a los puntajes bajos con bajas desviaciones estándar en los planteles oficiales, y puntajes altos con desviaciones estándar grandes en los colegios privados de los demás estados. En especial se nota que los colegios privados de Amazonas y Nueva Esparta tienen los mayores puntajes promedio, pero a su vez también son los que mayor dispersión tienen. Esta información confirma que existe una interacción clara entre el estado de procedencia y el tipo de plantel.

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

Gráfico 3. Puntajes promedio por grupos en la interacción Estado por tipo de plantel



4.1.1 Modelo estimado para Estado de Procedencia y Tipo de Plantel.

Como se explicó en el párrafo 3.3, el modelo de regresión para los puntajes en función de la procedencia y el tipo de plantel se escribe de la siguiente manera:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_{23} X_{i23} + \beta_{24} Z_i + \sum_{j=1}^{23} \gamma_j X_{ij} Z_i + \varepsilon_i, \quad i = 1, \dots, n.$$

El modelo tiene 48 parámetros (uno de intercepto, 23 para los estados, 23 para las interacciones y uno para tipo de plantel) cuyos valores calculados a partir del 1`931.400 aspirantes con sus desviaciones estándar el valor t y la significancia estadística de dicho valor se presentan en la Tabla 9 para los estados y el tipo de plantel cuando se encuentran solos en el modelo es decir β_1 hasta β_{23} y en la Tabla 10 para las interacciones entre estos dos factores, o sea γ_1 hasta γ_{23} .

Se observa en la Tabla 9 que el Tipo de Plantel oficial aporta negativamente a los valores de los puntajes y que con excepción de los estados Nueva Esparta y Amazonas (cód. 73 y 82 en la lista) la procedencia de cualquier estado disminuye los puntajes.

Tabla 9. Parámetros del modelo que tienen valores significativos en el modelo para el tipo de plantel y los estados de procedencia

Parámetro β_j correspondiente a:	Valor de β_j	Error típ.	t	Sig.
Intersección	52,000	,194	268,595	0,000
[TipoPlantel=1]	-2,608	,224	-11,635	0,000
[EntidadFederal=23]	-1,990	,230	-8,665	0,000
[EntidadFederal=32]	-1,609	,206	-7,826	0,000

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

[EntidadFederal=74]	-1,518	,214	-7,086	0,000
[EntidadFederal=63]	-1,206	,199	-6,055	0,000
[EntidadFederal=73]	1,261	,212	5,938	0,000
[EntidadFederal=61]	-1,117	,207	-5,400	0,000
[EntidadFederal=43]	-,933	,208	-4,484	0,000
[EntidadFederal=71]	-,884	,198	-4,456	0,000
[EntidadFederal=82]	1,577	,363	4,346	0,000
[EntidadFederal=11]	-,702	,195	-3,593	0,000
[EntidadFederal=13]	-,746	,211	-3,534	0,000
[EntidadFederal=22]	-,651	,196	-3,322	0,001
[EntidadFederal=41]	-,666	,206	-3,237	0,001
[EntidadFederal=12]	-,388	,197	-1,974	0,048

En la

Tabla **10** se observa que, excluyendo los estados Nueva Esparta, Amazonas, Bolívar y Zulia, todas las demás interacciones entre tipo de plantel y estado de procedencia aportan positivamente al modelo. Resulta interesante notar que los estados de Nueva Esparta y Amazonas contribuyen por sí solos a subir los puntajes estimados, pero pareciera que la interacción con los planteles oficiales hace que bajen sus puntajes. La razón de esto puede estar en que estos dos estados

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

son los que tiene los mayores promedios en los puntajes de colegios privados.

Tabla 10. Parámetros de las interacciones Tipo de Plantel y Estado de Procedencia con valores significativos para el modelo

Parámetro γ_j correspondiente a la interacción ZX_j	Valor de γ_j	Error típ.	t	Sig.
[TipoPlantel=1] * [EntidadFederal=73]	-2,464	,249	- 9,889	,000
[TipoPlantel=1] * [EntidadFederal=23]	2,573	,264	9,760	,000
[TipoPlantel=1] * [EntidadFederal=41]	2,009	,238	8,446	,000
[TipoPlantel=1] * [EntidadFederal=74]	1,978	,246	8,052	,000
[TipoPlantel=1] * [EntidadFederal=32]	1,764	,239	7,389	,000
[TipoPlantel=1] * [EntidadFederal=63]	1,600	,232	6,911	,000
[TipoPlantel=1] * [EntidadFederal=61]	1,464	,240	6,102	,000
[TipoPlantel=1] * [EntidadFederal=43]	1,272	,241	5,284	,000
[TipoPlantel=1] * [EntidadFederal=82]	-1,538	,398	- 3,866	,000
[TipoPlantel=1] * [EntidadFederal=81]	-,639	,231	- 2,769	,006
[TipoPlantel=1] * [EntidadFederal=71]	,637	,232	2,749	,006
[TipoPlantel=1] * [EntidadFederal=13]	,660	,252	2,622	,009
[TipoPlantel=1] *	,572	,233	2,450	,014

[EntidadFederal=62]				
[TipoPlantel=1] *	-,449	,227	-	,048
[EntidadFederal=51]			1,975	
[TipoPlantel=1] *	,447	,228	1,963	,050
[EntidadFederal=22]				

Entonces la ecuación para el modelo estimado de la incidencia del estado y tipo de plantel sobre los puntajes se puede escribir de la siguiente manera:

$$\begin{aligned} \hat{Y}_i = & 52 - 0,702 X_{i1} - 0,388 X_{i2} - 0,746 X_{i3} - 0,651 X_{i5} - 1,99 X_{i6} \\ & - 1,609 X_{i8} - 0,666 X_{i9} - 0,933 X_{i11} - 1,117 X_{i14} \\ & - 1,206 X_{i16} - 0,884 X_{i18} + 1,261 X_{i20} - 1,518 X_{i21} \\ & + 1,557 X_{i23} - \mathbf{2,608 Z_i} + 0,660 Z_i \times X_{i3} + 0,447 Z_i \times X_{i5} \\ & + 2,573 Z_i \times X_{i6} + 1,764 Z_i \times X_{i8} + 2,009 Z_i \times X_{i9} + 1,272 Z_i \\ & \times X_{i11} - 0,449 Z_i \times X_{i13} + 1,464 Z_i \times X_{i14} + 0,572 Z_i \times X_{i15} \\ & + 1,6 Z_i \times X_{i16} + 0,637 Z_i \times X_{i18} - 2,462 Z_i \times X_{i20} \\ & + 1,978 Z_i \times X_{i21} - 0,639 Z_i \times X_{i22} - 1,538 Z_i \times X_{i23} \end{aligned}$$

Para ilustrar el funcionamiento del modelo obsérvese que por ejemplo el puntaje estimado de un aspirante de colegio privado del estado Guárico (cód. 32) es $\hat{Y}_i = 52 - 1,609$ debido a que $X_{i8} = 1$ y $X_{ij} = 0$ para $j = 1, \dots, 7, 9, \dots, 23$ y $Z_i = 0$.

También por ejemplo, el puntaje estimado de un aspirante de un plantel público del estado Delta Amacuro (cód. 83) es $\hat{Y}_i = 52 - 2,608$ puesto que $X_{ij} = 0$ para todo $j = 1, \dots, 23$ y $Z_i = 1$. Para ilustrar cómo funciona la interacción en el modelo obsérvese por ejemplo que el puntaje estimado de un aspirante de plantel público para el estado Cojedes (cód. 23) los valores de las variables son: $X_{i6} = 1$, $X_{ij} = 0$ para $j = 1, \dots, 5, 7, \dots, 23$ y $Z_i = 1$, con lo cual $\hat{Y}_i = 52 - 2,608 - 1,99 +$

2.573 = 49.975. Las cifras corresponden a los efectos negativos del plantel público -2.608 y del estado por sí solo -1.99, y al efecto positivo de la interacción 2.573. Nótese que la diferencia $2.573 - 1.99 = 0.583$ es el efecto del tipo de plantel oficial después de quitar la interacción con el estado de procedencia, pero ésta lo único que hace es menguar el efecto negativo de los planteles oficiales.

Volviendo a la ecuación del modelo completo se observa que con excepción de los estados Zulia, Nueva Esparta, Bolívar y Amazonas todas las demás interacciones son positivas. Sin embargo esto no significa que sus contribuciones totales al modelo del estado con su interacción sean positivas debido a que todas las diferencias $\Delta_j = \gamma_j - \beta_j$ que son las que dan cuenta del efecto del plantel después de eliminar la interacción con el estado de procedencia son positivas y menores que 2.608 y por tanto $\Delta_j - 2.608 \leq 0$ con lo cual disminuye siempre el valor del puntaje estimado en los planteles públicos.

Claramente, se observa en el modelo anterior que el tipo de plantel oficial aporta negativamente a los valores de los puntajes de los estados, a excepción de los estados Nueva Esparta y Amazonas que aportan positivamente. Lo cual indica en términos generales que la procedencia de cualquier estado disminuye los puntajes.

Por otra parte se observa que en las interacciones del tipo plantel con la procedencia, en su mayoría aportan de manera positiva a los puntajes, con excepción de los estados Nueva Esparta, Amazonas, Bolívar y Zulia, llama la atención los estados Nueva Esparta y Amazonas, los cuales por si solos aportan de manera positiva, pero cuando interactúan con el tipo de plantel oficial, lo hacen de manera negativa, este fenómeno se debe a que estos estados son los que tienen mayores promedios en los puntajes de planteles privados. Tal como se observa en la Tabla 8.

4.2 Análisis de la influencia nivel de instrucción de la madre y los grupos etarios sobre los puntajes.

La Tabla 11 muestra la media y la desviación estándar del puntaje global para la combinación de los factores nivel de instrucción de la madre y la edad del aspirante. Se observa que los aspirantes que tienen madre con un grado de instrucción de profesión universitaria tienen puntajes promedio más altos y que los aspirantes cuya madre es analfabeta tienen puntajes promedio más bajos en todos los grupos de edad. También puede observarse que el grupo etario correspondiente a los aspirantes con edades entre 19 y 22 años, tienen los puntajes más bajos en comparación a los demás grupos; esta situación se visualiza en Gráfico 4.

Tabla 11. Medias y desviaciones estándar del puntaje global del nivel de instrucción de la madre por grupos etarios

Nivel de Instrucción de la Madre	17 - 18		19 - 22		> 22			
	Media	Desv. Estándar	Media	Desv. Estándar	Media	Desv. Estándar		
1 Profesión universitaria	53,95	9,01	54,34	9,07	51,52	8,92	53,44	9,03
2 Secundaria completa	52,04	8,43	52,07	8,43	50,29	8,31	51,14	8,32

3	49,84	7,93	49,91	7,80	48,85	7,84	49,44	7,80
Secundaria incompleta								
4 Primaria completa	49,16	7,82	49,23	7,65	48,21	7,70	48,97	7,76
5 Analfabeta	48,81	8,05	48,81	7,71	47,74	7,63	48,32	7,80
Total	50,94	8,44	50,92	8,35	49,05	8,04	50,21	8,25

En cuanto a las desviaciones estándar, se observa en la Tabla 11 que dentro de cada nivel de instrucción las desviaciones estándar se mantienen más o menos estables para todos los grupos de edad. Por otra parte, dentro de cada grupo de edad, se notan cambios considerables en las desviaciones estándar para los niveles de instrucción; por ejemplo en el nivel de instrucción universitario de la madre se encuentran las mayores desviaciones estándar en todos los grupos de edad y en el nivel de instrucción primaria completa se encuentran las menores desviaciones con excepción del grupo de edad entre 19 y 22 años.

Por otra parte, en el Gráfico 5 de las desviaciones estándar y los promedios se nota una tendencia a puntajes bajos con bajas desviaciones estándar en aspirantes cuyas madres tienen baja preparación académica y a puntajes altos con desviaciones estándar altas en los aspirantes cuyas madres tienen un nivel de instrucción universitaria. Lo anterior devela la existencia de una interacción entre el nivel de instrucción de la madre y los grupos de edad.

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

Gráfico 4. Comparación de puntajes promedios en el nivel de instrucción de la madre por edad

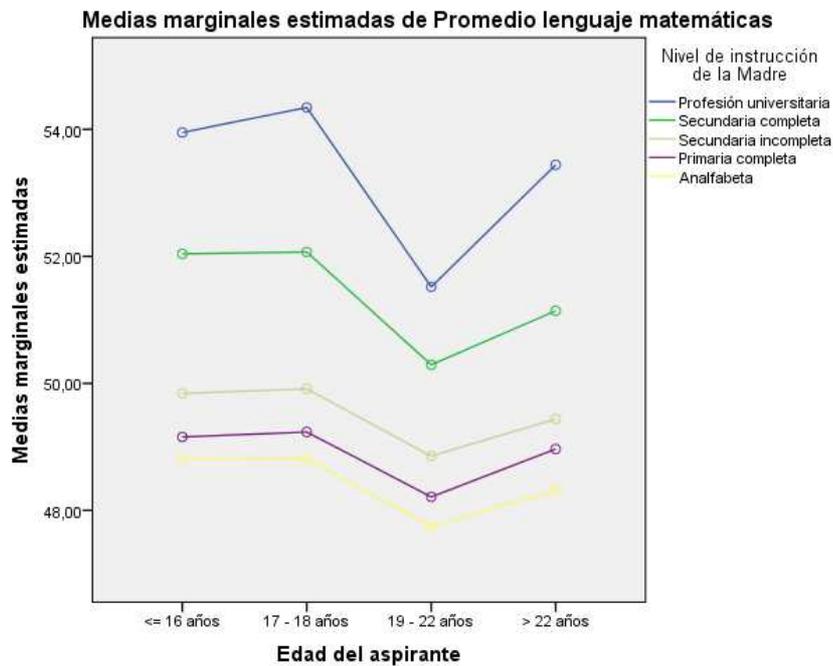
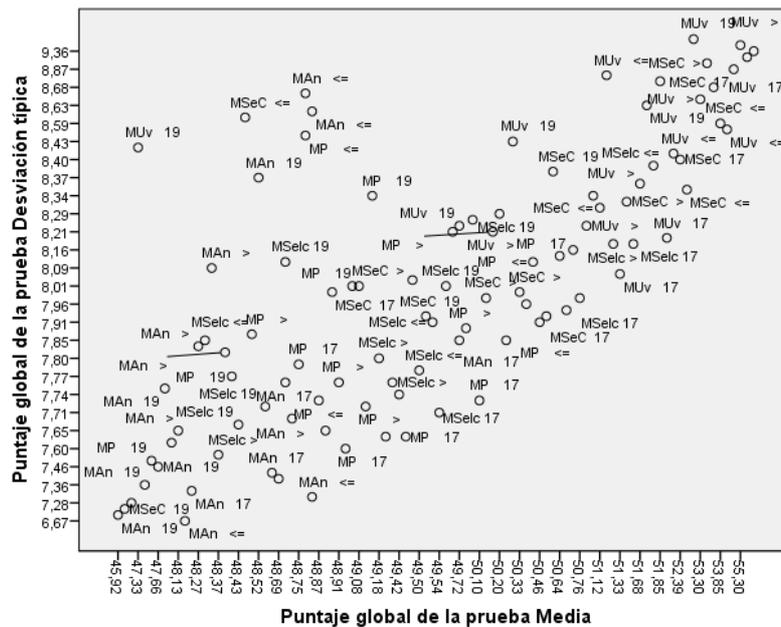


Gráfico 5. Dispersión de los puntajes promedio de los niveles de instrucción de la madre por grupos etarios.

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205



4.2.1 Modelo estimado para el análisis de la incidencia del nivel de instrucción de la madre y los grupos etarios sobre los puntajes de los aspirantes.

En el parágrafo 3.4 se introdujo el siguiente modelo para representar la incidencia del nivel de instrucción de la madre y los grupos etarios sobre los puntajes:

$$Y_i = \beta_0 + \beta_1 W_{i1} + \beta_2 W_{i2} + \beta_3 W_{i3} + \beta_4 W_{i4} + \beta_5 E_{i1} + \beta_6 E_{i2} + \beta_7 E_{i3} + \sum_{j=1}^4 \sum_{k=1}^3 \zeta_{kj} E_{ik} W_{ij} + \epsilon_i, \quad i = 1, \dots, n$$

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

El modelo tiene 19 parámetros cuyos valores con sus desviaciones estándar el valor *t* y la significancia estadística de dicho valor se presentan en la Tabla 12, calculados partir del 1`931.400 aspirantes:

Tabla 12. Parámetros del modelo para el nivel de instrucción de la madre y la edad del aspirante

Parámetro	B	Error típ.	t	Sig.
Intersección	48,320	,032	1516,351	0,000
[NIM=1]	5,120	,036	140,462	0,000
[NIM=2]	2,822	,035	81,336	0,000
[NIM=3]	1,119	,034	32,600	,000
[NIM=4]	,646	,034	18,982	,000
[EDAD=1]	,490	,233	2,099	,036
[EDAD=2]	,489	,103	4,764	,000
[EDAD=3]	-,576	,103	-5,614	,000

Se observa en la Tabla 12 que los valores de los parámetros correspondientes a los niveles de instrucción de la madre son todos positivos, y que con excepción de los aspirantes con edades comprendidas entre 19 y 22 años, los parámetros de los grupos etarios también son positivos; lo anterior indica la importancia del nivel de instrucción de la madre y resalta especialmente que los aspirantes

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

cuyas madres tienen profesión universitaria dan cuenta de los mayores puntajes en las pruebas.

En la Tabla 13 se muestran los valores de los parámetros para la interacción entre el nivel de instrucción de la madre y el grupo etario. Se observa que, las interacciones entre el nivel de instrucción de la madre con profesión universitaria y Secundaria completa con los grupos etarios entre 17 y 18 años aportan positivamente al modelo. Paradójicamente cuando estos dos niveles de instrucción de la madre interactúan con la edad de aspirantes entre 19 y 22 años su aporte es negativo, al igual que la interacción de esta edad con el nivel académico primaria completa.

Tabla 13. Parámetros de las interacciones Nivel de Instrucción de la Madre y la Edad del aspirante

Parámetro	B	Error típ.	t	Sig.
[NIM=1] * [EDAD=2]	,415	,114	3,635	,000
[NIM=1] * [EDAD=3]	- 1,345	,132	- 10,188	,000
[NIM=2] * [EDAD=2]	,437	,111	3,951	,000
[NIM=2] * [EDAD=3]	-,274	,118	-2,331	,020
[NIM=4] * [EDAD=2]	-,222	,109	-2,032	,042

El modelo estimado para la explicación de los puntajes a partir del nivel de instrucción de la madre y los grupos etarios es el siguiente:

$$\begin{aligned}\hat{Y}_i = & 48,32 + 5,12 W_{i1} + 2,822 W_{i2} + 1,119 W_{i3} + 0,646 W_{i4} + 0,49E_{i1} \\ & + 0,489E_{i2} - 0,576E_{i3} + 0,415W_{i1} \times E_{i2} - 1,345W_{i1} \times E_{i3} \\ & + 0,437W_{i2} \times E_{i2} - 0,274W_{i2} \times E_{i3} - 0,222W_{i4} \times E_{i2}\end{aligned}$$

Para ilustrar el funcionamiento del modelo obsérvese que por ejemplo el puntaje estimado de un aspirante del grupo de edad de 17 a 18 años y cuya madre tiene nivel de instrucción profesión universitaria es $\hat{Y}_i = 48,32 + 5,12 + 0,48 = 53,92$ debido a que $W_{i1} = 1$ y $W_{ij} = 0$ para $j = 2,3,4$ y $E_{i2} = 1$ y $E_{ij} = 0$ para $j = 1,3$.

En el caso de un aspirante mayor a 22 años y con madre analfabeta, obtendrá un puntaje promedio estimado de 48,32 puntos, el modelo estimado es $\hat{Y}_i = 48,32$ debido a que $W_{ij} = 0$ para $j = 1, \dots, 4$ y $E_{ij} = 0$, $j = 1, \dots, 3$.

4.3 Análisis de la influencia de la Edad y el Género sobre los puntajes.

En la Tabla 14 que contiene los puntajes promedio en los grupos de edad por género se observa claramente que los puntajes promedio de los hombres son mayores que los de las mujeres en todos los grupos de edad. Esta situación se ilustra en el Gráfico 6, en el cual además se nota un marcado descenso en los puntajes en el grupo de edad de 19 a 22 años para hombres y mujeres.

Cita / Citation:

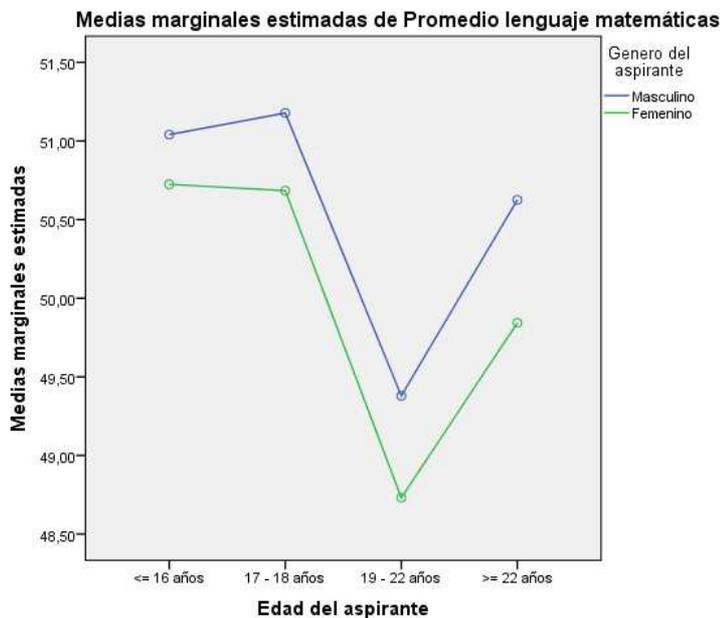
Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

Tabla 14. Puntajes promedio en los grupos de edad por género

		Género del aspirante			
		Masculino		Femenino	
		Puntaje global de la prueba		Puntaje global de la prueba	
		Media	Desviación típica	Media	Desviación típica
Edad del aspirante	<= 16 años	51,04	8,69	50,72	8,28
	17 - 18 años	51,18	8,54	50,68	8,19
	19 - 22 años	49,38	8,31	48,73	7,80
	> 22 años	50,63	8,51	49,84	8,03
	Total	50,60	8,51	49,88	8,05

Gráfico 6. Puntajes promedio en los grupos de edad por género



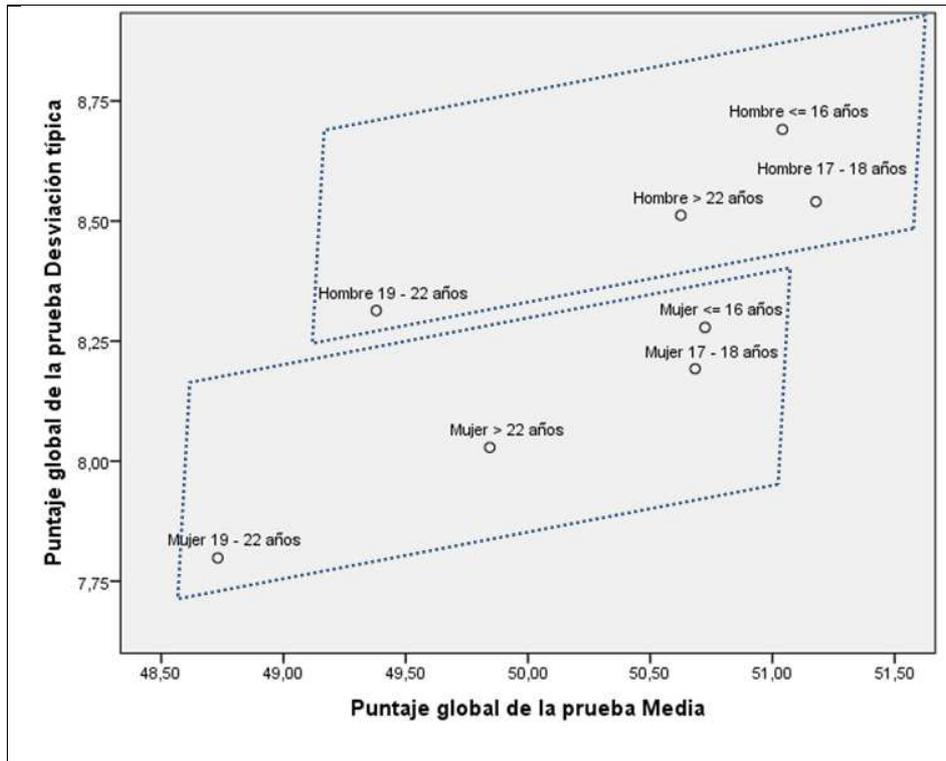
En el Gráfico 7 se ve la tendencia a aumentar la desviación estándar con el aumento los promedios tanto para mujeres como para hombres. Se notan también, que los hombres tienden a tener promedios más altos que las mujeres como se había anotado, pero las dispersiones entre los hombres son mayores en todos los grupos de edad; también se observa que tanto hombres como mujeres jóvenes (los dos grupos menores de 16 y entre 16 y 18) tienden a obtener los puntajes más altos con altas desviaciones estándar; después con puntajes un poco más bajos y menores desviaciones estándar se posiciona el grupo mayores de 22 años, y con puntajes mucho menores y así mismo menores desviaciones estándar se posicionan los grupos de hombres y mujeres entre 19 y 22 años.

Gráfico 7. Dispersión de la interacción de grupos etarios por edades

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205



4.3.1 Modelo estimado para edades y género.

En la sección 3.5 se introdujo el siguiente modelo para analizar la influencia de las edades y el género sobre los puntajes:

$$Y_i = \delta_0 + \delta_1 E_{i1} + \delta_2 E_{i2} + \delta_3 E_{i3} + \delta_4 S_i + \sum_{j=1}^3 \tau_j E_{ij} S_i + \varepsilon_i, \quad i = 1, \dots, n$$

Cita / Citation:

Doray del Carmen Contreras y Jimmy Corzo (2012). **VARIABLES INFLUENCING ADMISSION TESTS TO VENEZUELAN UNIVERSITIES**

www.revistaorbis.org.ve / núm 23 (año 9) 159-205

El modelo tiene ocho parámetros cuyas estimaciones calculadas a partir del 1`931.400 aspirantes se muestran en la Tabla 15 con sus desviaciones estándar, y los indicadores valor *t* y valor *p*.

Tabla 15. Parámetros del modelo que tienen valores significativos en el modelo para el la edad con el género

Parámetro	B	Error típ.	t	Sig.
Intersección	49,844	,009	5706,311	0,000
[SEXO=1]	,781	,014	57,541	0,000
[EDAD=1]	,880	,057	15,506	0,000
[EDAD=2]	,840	,026	31,768	0,000
[EDAD=3]	-1,113	,034	-32,790	0,000
[SEXO=1] * [EDAD=1]	-,465	,091	-5,136	0,000
[SEXO=1] * [EDAD=2]	-,287	,041	-6,999	0,000
[SEXO=1] * [EDAD=3]	-,133	,051	-2,623	0,009

El modelo estimado es el siguiente:

$$\hat{Y}_i = 49,844 + 0,781S_1 + 0,88E_{i1} + 0,840E_{i2} - 1.113E_{i3} - 0,465S_1 \times E_{i1} \\ - 0,287S_1 \times E_{i2} - 0,133S_1 \times E_{i3}$$

A manera de ilustración se muestra la operatividad del modelo con un ejemplo supongamos que un aspirante i es hombre y se encuentra en el grupo etario de 18 a 19 años de edad; el puntaje estimado para este estudiante es $\hat{Y}_i = 49,844 + 0,781 + 0,840 - 0,287 = 51,752$ debido a que $S_i = 1$ y también, $E_{i2} = 1$ y $E_{ij} = 0$ para $j = 1,3$.

Se observa en el modelo que los aspirantes del sexo masculino por si solo contribuye a aumentar el puntaje estimado en 0.781, mientras que cuando interactúa con los tres grupos de edad que quedaron en el modelo disminuyen el valor estimado del puntaje.

Asimismo en el modelo anterior dos de los grupos etarios aportan positivamente a los valores de los puntajes, mientras que el grupo de aspirantes con edades comprendidas entre 19 y 22 años tiende a disminuir el puntaje. Era de esperarse como se puede observar en el Gráfico 6; **Error! No se encuentra el origen de la referencia.** en el que se ve que los valores de los puntajes descienden en este tercer grupo etario. Sin embargo en términos generales la edad del aspirante contribuye a aumentar los puntajes en el modelo.

5. CONCLUSIONES Y DISCUSIÓN.

5.1 CONCLUSIONES.

Para el estado de procedencia el modelo permite concluir que con excepción de los estados Amazonas y Nueva Esparta, todos los demás estados, por sí solos, muestran tendencia a incidir negativamente sobre los puntajes; Así mismo, el tipo de plantel oficial por sí solo entra al modelo con aportes negativos sobre los puntajes. Sin embargo, cuando los aspirantes provienen planteles oficiales de la mayoría de los estados (exceptuando Nueva Esparta, Amazonas, Bolívar y Zulia) la incidencia es para mejorar los puntajes.

Los niveles de instrucción de la madre universitario y secundario son los que más contribuyen para la obtención de puntajes altos, y especialmente cuando interactúan con el grupo de edad de 17 18 años. En cambio, cuando el nivel de formación universitario interactúa con el grupo de edad de 19 a 22 disminuyen los puntajes de manera dramática indicando que entre los más adultos el nivel de instrucción de la madre tiende a ser contraproducente.

Entre más adultos los aspirantes, mayor es el puntaje exceptuando el grupo de edad de 19 a 22 años en el cual los puntajes tienden a disminuir.

5.2 Discusión.

¿Puede pensarse que el hecho de que la interacción plantel oficial estado aporte positivamente al incremento de los puntajes es un indicador de calidad de la educación impartida en los planteles públicos?

¿es realmente contraproducente el nivel más alto de formación de la madre en aspirantes de mayores edades? o es un problema de la base

de datos debido a que casi el 80% de los aspirantes son de este grupo de edad?

¿Se puede concluir realmente que en la sociedad venezolana hay un problema de dominio de género derivado de la conclusión de que los puntajes de las mujeres son en general, menores que los de los hombres?

6. REFERENCIAS BIBLIOGRÁFICAS.

Asamblea Nacional Constituyente. (2000). **Constitución de la República Bolivariana de Venezuela 1999**. Lima. Sairam editores S.R.L.,

Fuenmayor Toro, Luis y Vidal, Yasmila Y. (2001). **La Admisión Estudiantil a las Universidades Públicas Venezolanas: Causas de las iniquidades**". **Revista de Pedagogía**, Vol. XXII, N° 64, 219-241, Caracas.

Hosmer & Lemeshow (2000), Applied Logistic Regression Wiley.

Interscience. Hair, Et Al (1999), Análisis Multivariante. Prentice Hall. Madrid.

OPSU (2007). **Características del Sistema Nacional de Ingreso y Prosección en la Educación Superior Venezolana**.

Steel/Torrie (1997), Bioestadística Principios y Procedimientos 2da. Edición. Mc Graw Hill México.

Zerpa I. (2004). **Los Derechos del Estudiante Universitario, Ideas para Mejorar Nuestra Educación Superior**. Consejo Nacional de Universidades-Oficina de Planificación del Sector Universitario.