

Análisis de la relación entre resultados en la Herramienta Diagnóstica al Ingreso (HDI) y rendimiento académico en primer año de las carreras de la Facultad de Ingeniería.

1. Metodología

Se utilizaron los Modelos Lineales Generalizados (GLM) para estudiar la relación entre el rendimiento en la HDI (puntaje en Matemática, Física y Química: “puntaje MQF”) y el rendimiento en los cursos de primer año.

Los Modelos Lineales Generalizados fueron propuestos por Nelder y Wedderburn en 1972 y son una extensión a los Modelos Lineales, permitiendo que la distribución de la variable dependiente pueda pertenecer a cualquier distribución de la familia exponencial, además los supuestos de estos modelos son menos restrictivos, permitiendo errores no homocedásticos (McCullagh y Nelder, 1989).

Con respecto al rendimiento en la HDI se consideró solamente el puntaje MQF (puntaje Matemática + puntaje Física + puntaje Química), no incluyendo el puntaje obtenido en las múltiple opción de Comprensión Lectora. Esta decisión se basó en los resultados obtenidos en los años 2005 y 2006 (informes HDI 2005 y 2006) en donde se observa que la mayoría de los estudiantes obtienen la suficiencia en las preguntas de repuesta múltiple opción de esta componente. Los bajos índices de dificultad que presentan usualmente los ítems de las preguntas múltiple opción de Comprensión Lectora indican que no son los más adecuados para modelar predicciones, aunque son útiles para identificar rápida y claramente a aquellos estudiantes que tienen grandes dificultades en esta área.

Se utilizaron dos modelos para evaluar la relación mencionada (MQF vs. rendimiento en los cursos de primer año).

En el Modelo 1, la variable independiente (predictora) fue el *puntaje MQF* mientras que en el Modelo 2 esta variable fue desglosada en todos sus componentes (Matemática+Física+Química) de modo de ver el efecto de cada componente por separado (Tabla 1).

La variable dependiente para ambos modelos fue la variable “*Llegó50*” la cual toma valor 1 cuando el estudiante alcanza a aprobar el 50% de los cursos y 0 cuando no alcanza a aprobar la mitad de las asignaturas. De esta forma la variable dependiente tiene distribución binomial por lo que aplicamos GLM mediante una regresión logística.

Se utilizó un tercer modelo en donde los Modelos Lineales (ML) fueron utilizados para relacionar el rendimiento en la HDI y el promedio de la nota obtenida en los exámenes.

Tabla 1. Variables independientes y variable dependiente de los modelos utilizados para estudiar la relación entre el rendimiento en la HDI y en los cursos de primer año.

	Variable/s independiente/s	Variable dependiente	Modelo utilizado
MODELO 1	Puntaje MQF	Llegó 50	GLM
MODELO 2	Puntaje Matemática + Puntaje Física + Puntaje Química	Llegó 50	GLM
MODELO 3	Puntaje MQF	Promedio nota exámenes	ML

Con respecto a la variable “*Llegó50*” cabe aclarar que el 50% de las asignaturas se escoge sobre la de base del total de asignaturas a las cuales el estudiante se inscribió, y no al total de asignaturas de primer año (9 para el modelo). Lo importante a destacar es que la elección de la variable *Llegó50* se basa, entre otras cosas, en una decisión del estudiante, previa a comenzar a cursar y que implica diversas razones.

Se trabajó con las generaciones 2005 y 2006 por separado dadas las diferencias que existen entre las dos pruebas en la componente MQF (Tabla 2).

Tabla 2. Componentes de Matemática, Física y Química, puntaje total y nivel de suficiencia de las HDI 2005 y 2006. MO: preguntas múltiple opción, A: preguntas abiertas

	Matemática	Física	Química	Puntaje total	Suficiencia
HDI 2005	12 MO	11 MO	6 MO	29 puntos	17 puntos
HDI 2006	13 MO y 2 A	12 MO	10 MO	37 puntos	24 puntos

Las asignaturas de primer año que se tuvieron en cuenta en este estudio fueron:

- Cálculo 1 semestral y anual
- Geometría y Álgebra Lineal (GAL) 1 semestral y anual
- Física 1
- Cálculo 2
- GAL 2
- Física 2 (opcional para los estudiantes de Computación)
- Programación 1 (sólo para los de Computación, opcional para las otras carreras)
- Matemática Discreta (sólo Computación, opcional para las otras carreras)
- Economía

No se consideraron en el presente modelo, debido a que no estaban disponibles los datos en el momento de la elaboración del informe, las siguientes asignaturas:

- Taller de Diseño, Comunicación y Representaciones Gráficas (Eléctrica, Industrial y Naval)
- Taller de Expresión (Civil)
- Estática Aplicada (Naval)
- Mecánica de los Materiales (Naval)

Esto significa que en nuestro modelo el máximo de asignaturas que puede cursar un estudiante en primer año es 9, contando las opcionales.

2. Resultados

2.1. HDI 2005

2.1.1 Población de estudio

El número de estudiantes tomados en cuenta en el modelo fue de 728. No fue considerada para el modelo la población que se considera desertora (de acuerdo al concepto de deserción utilizado en Míguez y cols., 2005¹) y que comprende a 196 estudiantes (17.3%) que ingresaron en el 2005 y no presentaron ninguna actividad en el 2005 ni en el 2006.

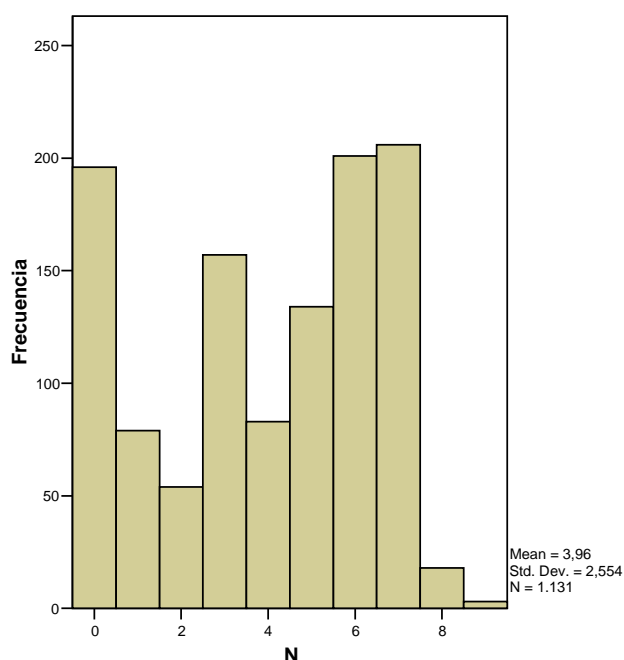


Figura 1. Distribución del número de asignaturas del primer año cursadas (N) por los estudiantes de la generación 2005.

¹ Míguez, M.; Loureiro, S. y Otegui, X. (2005) *Aprendizaje, Enseñanza y Desempeño Curricular en la Facultad de Ingeniería: análisis cuantitativos y cualitativos*, Serie Análisis de Datos, Unidad de Enseñanza, Facultad de Ingeniería, Montevideo.

En la Tabla 4 se muestra que las variables Matemática (***) , Física (**) y Química (*) son significativas (p-valor asociado a el estadístico z, tabla 2).

Las tres variables son útiles para explicar la relación encontrada, siendo Matemática la que más explica la relación seguida de Física y Química.

2.1.4. Modelo 3 – Regresión Lineal: MQF vs. Nota Promedio Exámenes.

A pesar de que pocos estudiantes han rendido una considerable cantidad de exámenes de las materias de primer año (Figura 3), se exploró la relación entre el promedio de nota en los exámenes y el puntaje MQF.

Se tomó en cuenta la nota del último examen rendido (el examen pudo haberse rendido una o más veces). Se escogió este criterio por ser el más simple, aunque el número de veces que el estudiante rinde un examen es un factor que podría ser considerando a la hora de evaluar el rendimiento. El modelo utilizado es una regresión lineal simple. Antes de aplicar el modelo se verificaron las hipótesis que asume el modelo lineal (supuesto de normalidad y de homogeneidad de varianza).

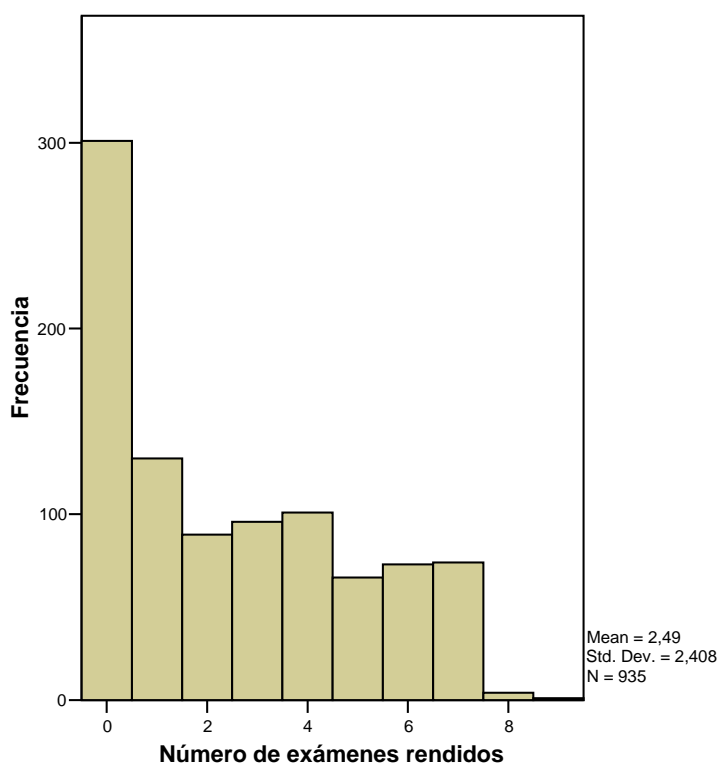


Figura 3. Distribución del número de exámenes rendidos por los estudiantes (no desertores) .

Las Tablas 5, 6 y 7 muestran la salida del modelo lineal que relaciona el puntaje MQF con el promedio de las notas de los exámenes.

Tabla 5.

R	R cuadrado	R cuadrado ajustado	Error estándar
,511	,262	,260	2,06135

Tabla 6.

	Suma de cuadrados	Gl	Mean Square	F	p-valor
Regresión	931,693	1	931,693	219,265	,000
Residual	2630,237	619	4,249		
Total	3561,930	620			

Tabla 7.

	Coeficientes		t	p-valor
	B	Std. Error		
Término indep.	,264	,236	1,117	,265
Puntaje MQF	,271	,018	14,808	,000 (*)

Análisis de los coeficientes

La expresión del modelo es:

$$\text{Promedio Nota Exámenes} = 0,264 + 0,271(\text{MQF})$$

El R cuadrado (tabla 5) indica que el modelo explica el 26% de la varianza de la variable *Promedio Nota Exámenes*, lo que es significativo (F y p-valor , tabla 6).

La tabla 7 muestra que el estimador del coeficiente de la variable MQF es significativo (*), informa cuánto aumenta la variable dependiente por cada unidad de la variable independiente.

Los p asociados a los t indican si se debe rechazar la hipótesis nula: “no hay relación lineal entre la variable independiente y la dependiente”. En este caso, para el coeficiente de MQF se rechaza la hipótesis nula.

El R (0,511) (Tabla 5) corresponde a la correlación de Pearson entre las dos variables, que resultó ser significativa ($p < 0,001$).

De esta forma se puede afirmar que la regresión lineal es adecuada para estudiar la relación entre el comportamiento del rendimiento en MQF y el promedio de las notas de los exámenes. La figura 4 muestra gráficamente dicha relación. Se señalan dos grupos (A y B) de estudiantes que se alejan del patrón general de la regresión.

Grupo A (6 casos):

Notas muy bajas en la HDI (de 0 a 4 puntos) y promedios aceptables en las calificaciones de los exámenes (de 3 a 5,5)

Grupo B (6 casos):

Notas medias-bajas en la HDI (8 a 17) y muy buen rendimiento en exámenes (10 y 11).

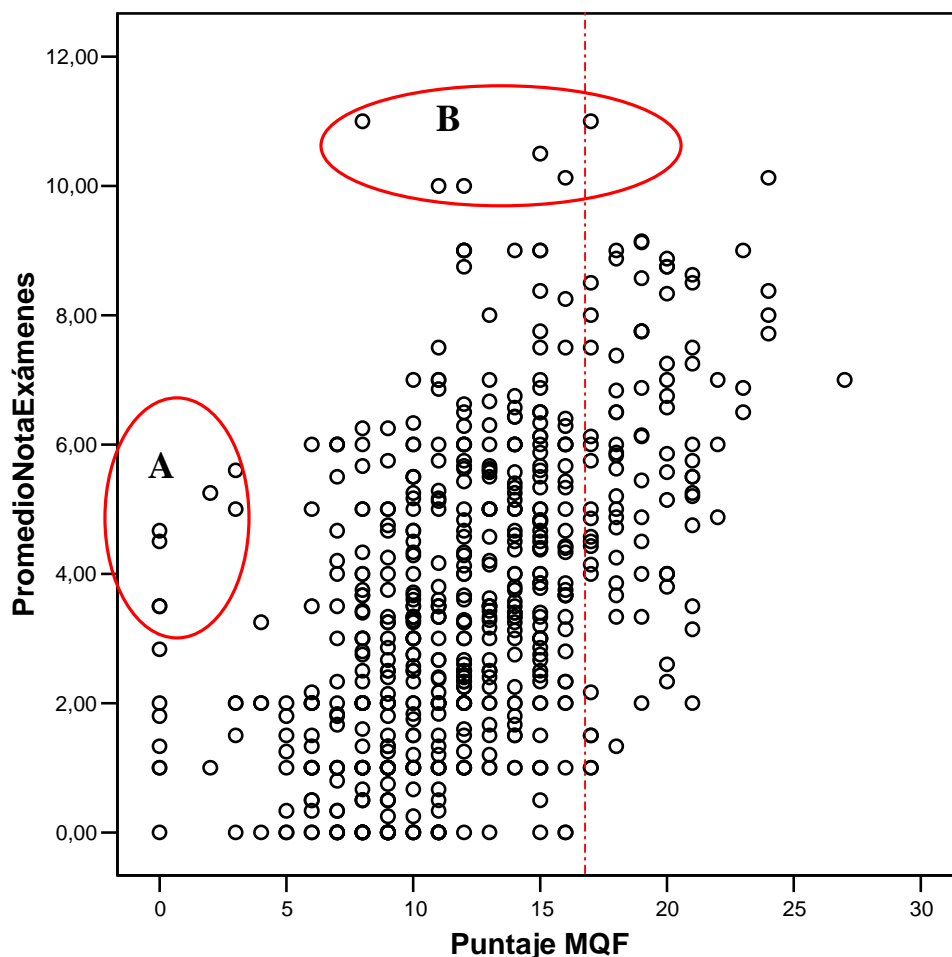


Figura 4. Relación entre el puntaje MQF y Promedio Nota Exámenes. Se señalan los dos grupos (A y B) de casos atípicos mencionados en el texto y la línea punteada indica el nivel de suficiencia HDI 2005.

Descripción de los casos atípicos

Variables de base

En cuanto a las variables de base se observaron ciertas diferencias comparando a los grupos A y B con el universo.

Sexo

Grupo A: varones: 83%, mujeres: 17%

Grupo B: varones: 83%, mujeres: 17%

Universo: varones: 70%, mujeres: 30%

Procedencia Geográfica

Grupo A: Montevideo: 50%, Interior: 50%

Grupo B: Montevideo: 66%, Interior: 33%

Universo: Montevideo: 55%, Interior 45%

Instituto de origen

Grupo A: Público: 66%, Privado: 33%

Grupo B: Público: 66%, Privado: 33%

Universo: Público: 58%, Privado: 35%, UTU: 7%

Tener en cuenta que los grupos A y B están integrados por muy pocos estudiantes (6 estudiantes en cada grupo).

Se están realizando entrevistas en profundidad a los grupos de estudiantes con comportamiento atípico. Estas entrevistas permiten tener una aproximación cualitativa a las posibles variables que podrían explicar dichos comportamientos.

Al momento se han entrevistado a 6 de los 12 estudiantes (3 de cada grupo), habiendo identificado algunas variables explicativas relevantes:

- Todos ellos son estudiantes que dedican un tiempo muy considerable a la Facultad, uno de ellos señala de modo muy categórico

“Yo cuando no estoy en Facultad estoy estudiando, estudio todo el día”.

- No trabajan.
- Poseen un contexto socio-económico que les permite dedicarse a su carrera a tiempo completo.

En particular pueden señalarse algunas características adicionales:

Uno de los estudiantes ha cursado en la modalidad de trayectos diferenciados, donde ha adquirido estrategias de estudio en grupo, estudia con un grupo de estudiantes a los que denomina *“los anuales 2005”*, estos estudiantes continúan estudiando juntos, aún aquellos que han optado por distintas carreras de ingeniería.

En un caso el bajo rendimiento en la HDI se explica básicamente por la variable Instituto de origen, el estudiante (Grupo A) cursó un bachillerato técnico en UTU y señala no haber visto allí los conocimientos evaluados en la HDI en su componentes MQF.

En el caso de dos estudiantes (Grupo A y B), la mejora en el rendimiento puede ser explicada por la preparación de cursos paralelos y de exámenes con profesor particular.

Dos estudiantes (Grupo A y B) que habían salvado las materias de sexto año liceal en diciembre dicen *“no recordar nada”* en marzo cuando rindieron la HDI (tampoco repasaron para la instancia de la prueba diagnóstica).

Cuatro de seis estudiantes manifiestan haber recurrido al menos una vez, por lo que a partir de las entrevistas se hace evidente la necesidad de considerar la

variable “recurso” en los estudios estadísticos. En este mismo sentido algunos estudiantes con bajo rendimiento en la HDI llegan a salvar exámenes luego de rendirlos varias veces, por lo que como ya se manifestó este es un factor más a ser considerado en los modelos explicativos en estudios posteriores.

2.2. HDI 2006

2.2.1 Población de estudio

Para la generación 2006 se tomó en cuenta a todos los estudiantes que rindieron la HDI (898) dado que al haber transcurrido sólo un año desde que ingresaron a la Facultad no se les considera desertores a pesar de que no hayan presentado actividad durante el año lectivo.

La figura 5 muestra la distribución del número de asignaturas cursadas para la generación 2006. Al igual que en 2005, vemos que la mayoría de los estudiantes cursa alguna asignatura y que un número alto cursa 6 y 7 (Figura 5). Llama la atención en relación con 2005 el bajo número de estudiantes que no cursó ninguna asignatura (28 estudiantes: 3,1%).

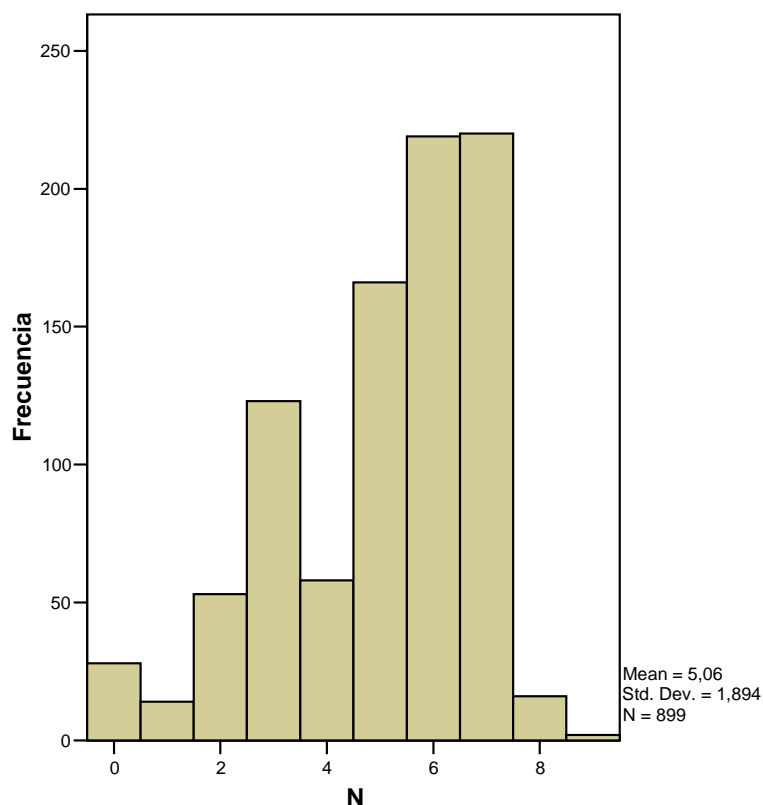


Figura 5. Distribución del número de asignaturas cursadas (N) en primer año de la carrera para la generación 2006.

2.2.2. Modelo 1 - Regresión logística: MQF vs. Llegó50.

Tabla 8. Salida del modelo 1.

	Valor Estimado de los Coeficientes	Valor z	p-valor
Término indep.	-3,00958	-11,14	<2e-16***
MQF	0,17124	11,10	<2e-16***

Códigos de significación: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 ''

Devianza nula: 1237,8 en 898 grados de libertad

Devianza residual: 1081,4 en 897 grados de libertad

Número de iteraciones Fisher Scoring: 4

La expresión del modelo es:

$$\ln(p/1-p) = -3,00958 + 0,17124(\text{MQF})$$

Observamos que para 2006 la variable MQF también es significativa para el modelo (***) (p-valor, Tabla 8).

La figura 6 muestra la representación gráfica del modelo.

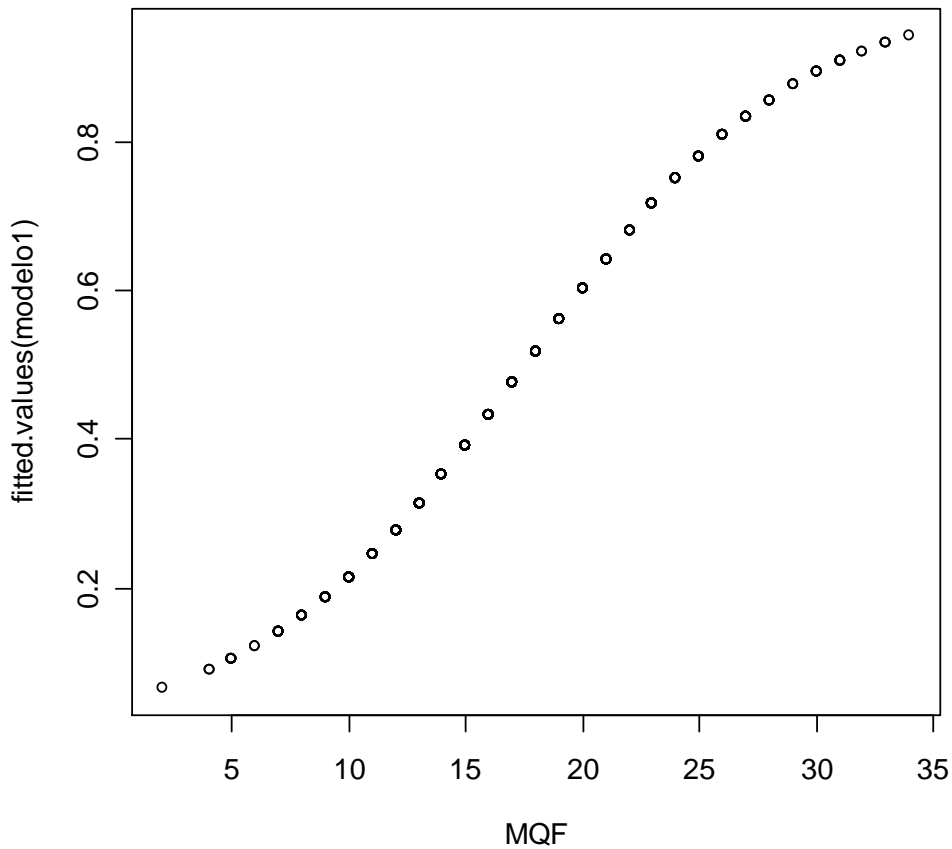


Figura 6. Probabilidad de aprobar el 50% de los cursos de primer año en función del puntaje MQF.

2.2.4. Modelo 3 – Regresión lineal: MQF vs. Nota Promedio Exámenes.

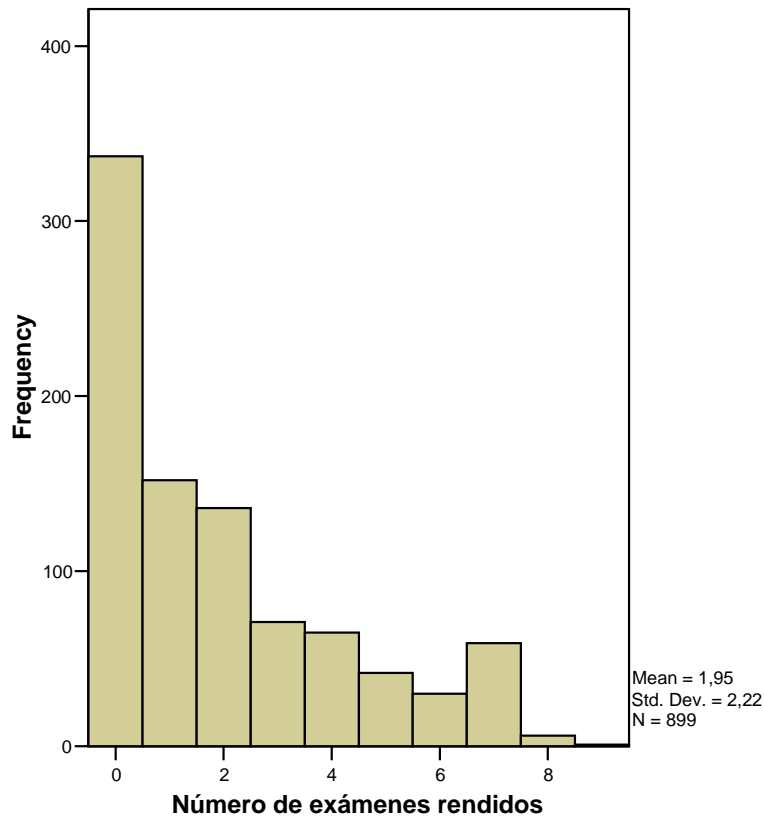


Figura 7. Distribución del número de exámenes rendidos para estudiantes de generación 2006.

Al igual que en 2005 se utilizó una regresión lineal simple para estudiar esta relación. Las tablas 10, 11 y 12 muestran las salidas del modelo.

Tabla 10.

R	R cuadrado	R cuadrado ajustado	Error estándar
0,535	0,286	,285	2,42019

Tabla 11.

	Suma de cuadrados	Grados de libertad	Mean Square	F	p-valor
Regresión	2102,693	1	2102,693	358,984	,000
Residuos	5254,037	897	5,857		
Total	7356,730	898			

Tabla 12.

	Coeficientes		t	p-valor
	B	Std. Error		
Término indep.	-2,033	,252	-8,068	,000
Puntaje MQF	,276	,015	18,947	,000

Análisis de coeficientes

La expresión del modelo es:

$$\text{Promedio Nota Exámenes} = -2,033 + 0,276(\text{MQF}).$$

El R cuadrado (Tabla 10) muestra que la varianza explicada por el modelo es del 29%, lo que es significativo. El R (0,535) que corresponde a la correlación de Pearson resultó significativo ($p < 0,001$).

La tabla 11 muestra que el estimador del coeficiente de la variable MQF es significativo(*), informa cuánto aumenta la variable dependiente por cada unidad de la variable independiente.

Como se mencionó para 2005, el p-valor asociado a los t indica si se debe rechazar la hipótesis nula “no hay relación lineal entre la variable independiente y la dependiente” (Tabla 12). En este caso, para el coeficiente de MQF debe rechazarse la hipótesis nula.

La figura 8 representa la relación entre las 2 variables estudiadas.

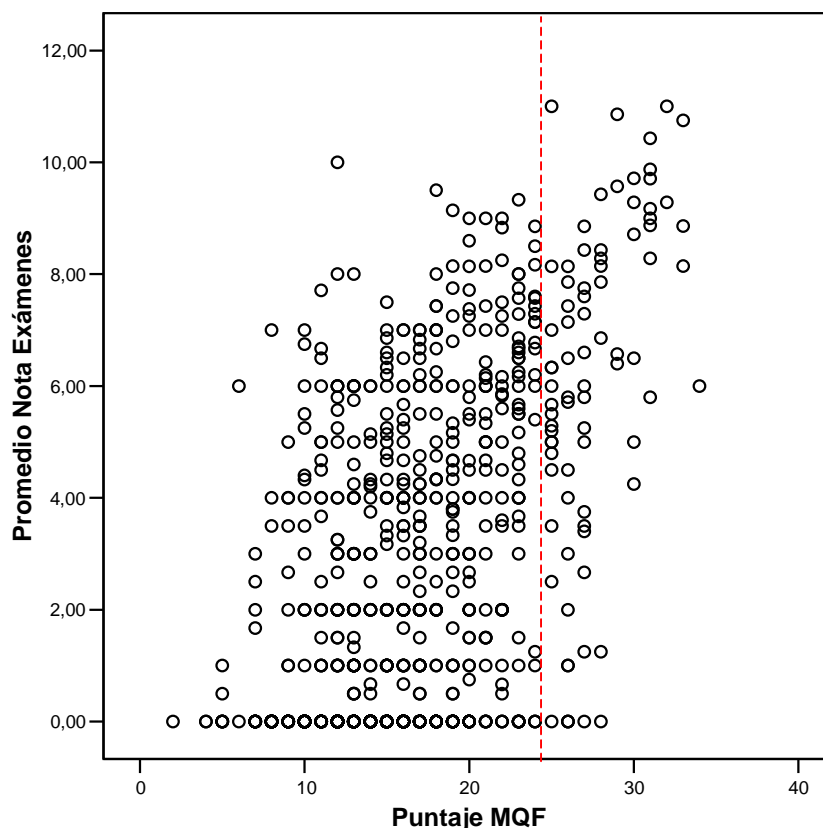


Figura 8. Relación entre el puntaje MQF y el Promedio Notas Exámenes de los estudiantes generación 2006. La línea punteada indica el nivel de suficiencia de la HDI 2006.

A pesar de haber casos aislados que se alejan del patrón de la regresión no se observan grupos con casos atípicos como ocurre en 2005.

2.2.5. Relación Idea Principal (IP) con comportamiento en los cursos de primer año (Llegó50)

Además de evaluar el desempeño en comprensión lectora mediante cinco preguntas múltiple opción, en la HDI se evaluó la capacidad de identificar la idea principal de un texto breve. La corrección de las respuestas abiertas de los estudiantes evaluó si se captaba (“Presencia”) o no (“Ausencia”) la idea principal del texto.

El test de Chi-cuadrado mostró una relación significativa entre IP y comportamiento en los cursos ($\chi^2 = 8,153$, $p < 0,01$), la que puede visualizarse en la siguiente tabla.

Tabla 13. Tabla de contingencia que relaciona las variables Idea Principal y Llegó50.

		Idea Principal		Total
		Ausencia	Presencia	
AprobóMitadCursos	No	65	16	81
	Si	40	28	68
Total		105	44	149

3. Conclusiones

Los resultados obtenidos para las generaciones 2005 y 2006 nos permiten afirmar que existe una relación entre el rendimiento en la HDI (puntaje MQF) y el comportamiento en los cursos de primer año; en el caso de la generación 2006 también con la componente IP.

En ambas generaciones los GLM mostraron que la variable MQF es adecuada para explicar el comportamiento en los cursos de primer año, observándose mayores probabilidades de llegar a aprobar la mitad de los cursos de primer año a mayores puntajes de MQF.

Se observó una relación decreciente de los componentes de MQF para explicar el comportamiento en los cursos de primer año en el siguiente sentido: Matemática, Física y Química. Matemática es la componente que más explicó el comportamiento en los cursos de primer año mientras que Química fue la que menos.

Las predicciones del modelo nos permiten evaluar cuál es la probabilidad de aprobar el 50% de los cursos de acuerdo al puntaje obtenido en MQF, del mismo modo que, a partir de probabilidades esperadas de “éxito” se puede

estimar cuál es el puntaje mínimo que el estudiante debe obtener en MQF. A partir de lo anterior se podrían realizar recomendaciones a los estudiantes de acuerdo a su rendimiento en la HDI (ver más abajo).

Los modelos lineales, por su parte, también mostraron relaciones significativas entre el rendimiento en la HDI y el promedio de la nota en los exámenes. Esto constituye una primera aproximación para evaluar la relación puntaje MQF – promedio nota exámenes dado que aún hay varios estudiantes de las generaciones en estudio que no rindieron muchos de los exámenes de primer año.

Los resultados obtenidos permiten establecer la utilidad de la HDI para predecir el rendimiento estudiantil en el primer año de las carreras. También, como se discute a continuación, resulta relevante para analizar y confirmar la composición relativa de la HDI, complementando la validación de la herramienta diseñada. (Ver informes HDI)

Finalmente permite sugerir algunas formas de determinar de manera prudente recomendaciones más específicas acerca del cursado en primer año; no establece la obligatoriedad de las recomendaciones, ni fundamenta de ningún modo una prueba de ingreso.

3.1. Composición de la prueba

Como mostraron los modelos, las componentes de Matemática y Física son las que explican mejor el comportamiento de los estudiantes en los cursos del primer año; la componente Química resultó significativa con un nivel de significación menor (p -valor = 0.04714 para 2005 y p -valor = 0,0112 para 2006). Esto se encuentra dentro de lo esperado dado que en primer año de Facultad de Ingeniería no hay asignaturas relacionadas directamente con Química. Esta componente podría ser útil para predecir el comportamiento de los estudiantes en otras etapas de la carrera o en asignaturas de química cursadas en la Facultad de Química por los estudiantes de las carreras Ingeniería Química y en Alimentos.

Se puede observar que la componente Idea Principal fue significativa para 2006. En el caso de esta variable se recomienda continuar aplicándola dado que brinda información profunda sobre la comprensión lectora, que es una habilidad cognitiva imprescindible para el buen desempeño académico universitario en las diferentes áreas de conocimiento.

En referencia al cuestionario de Motivación y de Estrategias de Aprendizaje, se están realizando actualmente análisis multivariados. Se han encontrado relaciones interesantes de respuestas a determinadas preguntas y el comportamiento de los estudiantes en las carreras (ver informes HDI 2005, 2006 y 2007).

De esta manera la HDI 2008 se sugiere debe estar integrada por:

13 preguntas de Matemática múltiple opción

2 preguntas de Matemática abiertas

12 preguntas de Física

10 preguntas de Química

5 preguntas de comprensión

Idea Principal.

Cuestionario de Motivación y de Estrategias de Aprendizaje.

3.2. Recomendaciones

En función de los estudios realizados, de acuerdo a los resultados obtenidos en cuanto al poder predictivo de la HDI y considerando además los datos adicionales aportados por el análisis cualitativo que confirman lo establecido por el análisis cuantitativo, se pueden establecer algunas sugerencias específicas para el cursado al ingreso a Facultad de Ingeniería.

Como primera aproximación y de acuerdo a las predicciones del modelo obtenido se determinó el puntaje mínimo necesario de MQF para tener una probabilidad mayor o igual al 30% y al 50% de aprobar la mitad de los cursos del primer año, obteniéndose los siguientes resultados:

1. $p=30\%$

El puntaje que supera esta probabilidad es de 13 puntos en 37 ($p=0.3261016$). En 2006 el número de estudiantes que no superó dicho puntaje fue 236.

2. $p=50\%$

El puntaje que supera esta probabilidad es de 18 puntos en 37 ($p=0.5181722$). El número de estudiantes que queda por debajo de este umbral en 2006 fue 491.

De acuerdo a los resultados obtenidos hasta el momento y a los estudios de relación entre los resultados de la HDI y el rendimiento académico de los estudiantes en las asignaturas de primer año (cursos y exámenes), se sugiere a los Señores Consejeros reflexionar acerca de las recomendaciones específicas a dar a los ingresantes en función de sus resultados en la HDI.

Algunas posibles alternativas serían:

A- Si el estudiante pertenece a la categoría de puntos 13/37 ($p=30\%$) o menor en HDI (considerando puntajes 2006), se sugiere:

- a. no cursar todas las asignaturas del primer semestre simultáneamente
- b. optar por cursar los cuatro primeros semestres de la carrera en seis semestres
- c. cursar las asignaturas GAL1 y CAL1 en modalidad Trayectos Diferenciados

- B- Si el estudiante pertenece a la categoría de puntos 13/37 ($p=30\%$) o menor en HDI y además su procedencia geográfica o de instituto refiere a un factor adicional de riesgo de fracaso académico (Interior y Público/UTU respectivamente), se sugiere
- a. no cursar todas las asignaturas del primer semestre simultáneamente
 - b. optar por cursar los cuatro primeros semestres de la carrera en seis semestres
 - c. cursar las asignaturas GAL1 y CAL1 en modalidad Trayectos Diferenciados

Sería recomendable pensar nuevas opciones adecuadas para cada grupo, lo cual supondrá necesariamente el desarrollo y evaluación de nuevas propuestas de enseñanza para brindar el apoyo pertinente a la población de estudiantes que más lo necesita.

Se entiende que, de tomarse alguna resolución a este respecto, debería difundirse ampliamente de modo que los ingresantes puedan prepararse como consideren adecuado para la HDI.

Este análisis forma parte de un estudio más amplio que realiza la UEFI acerca del avance y rendimiento estudiantil; considerado globalmente se evidencia que es necesario pensar en propuestas de innovación que atiendan a la problemática del reiterado fracaso académico estudiantil que se relaciona con la problemática del aprendizaje, con las problemáticas de enseñanza así como con otras más generales que trascienden a esta Facultad (Informe de avance, exp. N° 060020-000234-07). Se sugiere se considere la posibilidad de poner a disposición de las autoridades de la UdelaR estos informes de modo que puedan ser insumos, por ejemplo, para el grupo de trabajo "Generación 2007".

4. Referencias

McCullagh, P. y Nelder, J. A. 1989. *Generalized Linear Models*. Chapman & Hall.

Mesa, A. y Basterrech, S. 2006. *Modelos Lineales Generalizados* Apuntes del curso del mismo nombre dictado en la Facultad de Medicina de la Universidad de la República en septiembre-octubre de 2006.

Informes HDI 2005, 2006 y 2007. Unidad de Enseñanza de la Facultad de Ingeniería.