

DETECCIÓN DE SESGOS POR SEXO EN LA PRUEBA DE APTITUD ACADÉMICA APLICADA EN EL ITCR

Edgar Guillén*

I. Justificación

Vives y Blando (1998) en un estudio que describe el comportamiento por sexo en el proceso de atracción, selección y admisión al Instituto Tecnológico de Costa Rica (ITCR) en los últimos cinco años, encontraron que el porcentaje de hombres que “ganan” el examen de admisión, respecto del total de hombres que se presentan a realizarlo, es superior al porcentaje de mujeres que “ganan” el examen de admisión, respecto del total de mujeres que se presentan a realizarlo.

El *examen de admisión* en el ITCR es el resultado de un promedio ponderado de tres componentes: las *calificaciones de cuarto ciclo* de la secundaria y las *áreas verbal y matemática* de la *Prueba de Aptitud Académica*. Cuando se realiza el cálculo del *puntaje de admisión* se le asigna mayor peso al área matemática y se le da menor ponderación al área verbal.

“Las mujeres presentan una tendencia a tener mejores promedios de *cuarto ciclo* que los hombres” (Vives y Blando: p.11). No obstante, la proporción de hombres que “ganan” el *examen de admisión* es superior a la proporción de mujeres que lo “ganan”, según se anotó anteriormente. Pareciera entonces que en las *áreas verbal y/o matemática* de la *Prueba de Aptitud Académica* los hombres obtienen puntajes superiores a las mujeres. Esto a su vez podría tener un efecto sinérgico con la forma de ponderar los componentes del examen de admisión. Si los hombres obtienen puntajes más altos en el *área matemática* y ésta a su vez tiene mayor peso en la determinación del *puntaje de admisión*,

las consecuencias diferenciales para uno y otro sexo son evidentes.

La literatura especializada, por su parte, es prolija en señalar que tanto en el SAT (Scholastic Aptitude Test) como en el ACT (American College Testing Program), utilizados como uno de los criterios para definir la admisión a universidades estadounidenses, se presentan diferencias por sexo en los puntajes obtenidos por hombres y mujeres (a favor de los hombres) y en su poder predictivo sobre el rendimiento académico según sexo (Schrader:1971; Ramist:1984; Wainer and Steinberg:1992; Striker, Rock and Burton:1993).

Existe suficiente evidencia de que el SAT subpredice el rendimiento académico de las mujeres, en tanto sobrepredice el rendimiento académico de los hombres, lo cual es obviamente un sesgo a favor de los hombres. “Los resultados de diversos estudios indican que la extensión de ésta subpredicción es muy pequeña, aunque altamente consistente” (Wainer and Steinberg: ob.cit.: p.710) cuando se usan el SAT y el ACT para predecir el promedio de las calificaciones obtenidas por los estudiantes en la universidad (típicamente durante el primer año). Además en la universidad las mujeres obtienen notas, en promedio, iguales o superiores a las obtenidas por los hombres (Striker, Rock and Burton: ob.cit. p.323).

II. Definición del problema

El presente estudio se propone responder las siguientes preguntas relacionadas con el proceso de admisión del ITCR:

- ¿Se presentan diferencias

* Lic. en Psicología, Departamento de Orientación y Psicología, ITCR. eguillen@itcr.ac.cr

estadísticamente significativas, por sexo, en los resultados del *puntaje de admisión* y de cada uno de sus componentes: *cuarto ciclo, área verbal, área matemática*?

- ¿La fórmula usada para ponderar los componentes del *puntaje de admisión* crea sesgos a favor de uno u otro sexo?
- ¿La fórmula de ponderación empleada se fundamenta técnicamente a la luz de los resultados empíricos obtenidos?

III. Definición de términos

PUNTAJE DE ADMISIÓN. Como ya se señaló, el examen de admisión al ITCR, que da como resultado el puntaje de admisión o puntaje normalizado, es un promedio ponderado de sus tres componentes: el promedio de cuarto ciclo de secundaria y los dos componentes de la prueba de aptitud académica, a saber, el área matemática y el área verbal. Para evitar confusiones dejaremos de lado el término examen de admisión y emplearemos el término más preciso de *puntaje de admisión*.

PRUEBA DE APTITUD ACADÉMICA. Es la prueba que elabora, administra y procesa el Comité de Examen de Admisión. Consta de un *área matemática* con ítems (normalmente entre 50 y 55) de contenido matemático (operaciones básicas, propiedades de los números, ecuaciones y expresiones, porcentajes, geometría) y un *área verbal* con ítems (entre 45 y 50) de contenido verbal (antónimos, analogías, comprensión de lectura).

ÉXITO ACADÉMICO. Definimos éxito académico como el rendimiento académico obtenido por el estudiante durante su primer semestre de permanencia en el Instituto.

Como indicador de rendimiento académico se emplea el promedio ponderado (según número de créditos) obtenido por el estudiante en su primer semestre de permanencia en el Instituto. Para su cálculo se promedian todas aquellas asignaturas cuyas calificaciones son numéricas y que aparecen como cursadas en los archivos del Departamento de Admisión y Registro. Se dejan por fuera las asignaturas retiradas o congeladas y aquellas que solamente se reportan como aprobadas o reprobadas.

IV. Metodología

1. Procedimiento

Para determinar si se dan diferencias significativas, por sexo, en los puntajes obtenidos en los componentes y en el *puntaje de admisión* como un todo, se utilizaron los archivos que para tal efecto lleva el Departamento de Admisión y Registro. Se utilizó la información de todas las personas que realizaron la Prueba de Aptitud Académica.

Para determinar si la ponderación que se hace actualmente de las partes del examen de admisión se justifica técnicamente, se calcularon las correlaciones de cada uno de los componentes del *puntaje de admisión* con el *rendimiento académico*. Los análisis se basaron en los y las estudiantes que se matricularon y cursaron al menos una materia en el Instituto.

2. Población y muestra

Tanto para la detección de diferencias por sexo en los resultados de los componentes del *puntaje de admisión*, como para la evaluación técnica de la pertinencia de los pesos utilizados en el cálculo del promedio ponderado del *puntaje de admisión*, se trabajó con los archivos del Departamento de Admisión y Registro, con todas las personas que realizaron la Prueba de Aptitud Académica en los años 1994 a 1998 (identificamos el año por el número de carné que corresponde a los estudiantes que ingresan).

3. Técnicas de análisis de datos

Para someter a prueba las hipótesis de que no hay diferencias significativas según sexo en el *puntaje de admisión, cuarto ciclo, área verbal y área matemática* de la Prueba de Aptitud Académica, se practicaron *análisis de varianza* de un factor. Debe advertirse que en realidad se trata de la prueba t de Student que, realizada mediante un análisis de varianza para dos grupos, ofrece los mismos resultados (Kerlinger: 1975: p.156); se recurre al análisis de varianza porque la “salida” en el análisis por computadora ofrece las sumas de cuadrados, lo que no ofrece la t de Student.

Se practicó además la prueba U de Mann-Whitney, que es la mejor alternativa no paramétrica a la t de Student (Siegel:1985: p.143), con el propósito de corroborar los resultados obtenidos con el *análisis de varianza* toda vez que, en términos estrictos, los datos con que se trabajó violan supuestos de ese análisis.

Para verificar la pertinencia técnica de la ponderación que se da a los componentes del *puntaje de admisión* se calcularon las correlaciones r de Pearson entre los componentes de ese puntaje y el promedio ponderado obtenido por los y las estudiantes matriculadas en el ITCR, durante su primer semestre de permanencia en el Instituto.

V. Resultados

1. Detección de sesgos en el puntaje de admisión y sus componentes

1.1 Puntaje de Admisión

El *análisis de varianza* provee la sumatoria de los cuadrados atribuible a las diferencias entre los grupos, a la variabilidad que se da al interior de los grupos mismos y la sumatoria total de cuadrados. El porcentaje que representa la suma de cuadrados entre grupos respecto de la suma total de cuadrados indica qué porcentaje de la variabilidad de los datos es atribuible a las diferencias entre grupos. Esto es conocido como el *porcentaje de varianza* explicada por las diferencias entre grupos.

En los análisis que hemos practicado incluimos la Prueba de Levene para la homogeneidad de las varianzas, por cuanto es uno de los supuestos del *análisis de varianza*. Al igual que con la probabilidad

de F, un valor inferior a 0,05 generalmente se toma como indicador de que las varianzas son distintas (no provienen de una misma población).

En el Cuadro 1 se presenta un resumen de los datos más relevantes de las diferencias por sexo en el *puntaje de admisión*. En la primera columna aparece el año analizado. En la segunda columna se presenta el *porcentaje de varianza* explicada por la pertenencia a uno u otro sexo. En la tercera columna se presenta la probabilidad de F, esto es, la probabilidad de que ambos grupos pertenezcan a una misma población. En la cuarta columna se presenta la probabilidad del test de Levene. En la quinta columna se presenta el promedio obtenido por los hombres en el *puntaje de admisión*. En la sexta columna se presenta el promedio obtenido por las mujeres en el *puntaje de admisión*. En la última columna se presenta la diferencia entre los promedios de los hombres y las mujeres.

El *análisis de varianza* muestra que se presentan diferencias significativas en los promedios obtenidos por hombres y mujeres en el *puntaje de admisión*, en los cinco años analizados.

Los hombres obtienen puntajes más altos que las mujeres. Las diferencias en los puntajes van desde 10,66 puntos hasta 16,26 puntos. Se trata, sin duda alguna, de diferencias importantes en términos de sus repercusiones en la admisión al Instituto y a la carrera.

La *varianza del puntaje de admisión* explicada por las diferencias de sexo es muy pequeña; va de 0,30% a 0,98%. Vale decir, de la variabilidad que muestra el *puntaje de admisión*, más del 99% no está asociada al

CUADRO 1. Puntaje de admisión por sexo y año.

Año	% Varianza explicada	Prob. F	Prob. Levene	Promedio Hombres	Promedio Mujeres	Diferencia Promedio
1998	0,38	0,0000	0,000	533,87	521,98	11,89
1997	0,30	0,0000	0,000	515,03	504,37	10,66
1996	0,49	0,0000	0,002	504,08	492,73	11,35
1995	0,98	0,0000	0,307	506,95	490,69	16,26
1994	0,98	0,0000	0,003	506,49	490,68	15,81

CUADRO 2. Cuarto ciclo por sexo y año.

Año	% Varianza explicada	Prob. F	Prob. Levene	Promedio Hombres	Promedio Mujeres	Diferencia Promedio
1998	1,26	0,0000	0,000	80,71	82,33	-1,62
1997	1,63	0,0000	0,004	79,87	81,68	-1,81
1996	1,53	0,0000	0,069	79,78	81,61	-1,83
1995	0,39	0,0000	0,000	77,80	78,74	-0,94
1994	1,16	0,0000	0,000	77,73	79,42	-1,69

sexo como variable aislada.

En tres de los cinco años analizados aparecen diferencias significativas en las varianzas de los grupos por sexo: uno de los supuestos del *análisis de varianza*. El otro supuesto, el de la normalidad de la distribución tampoco se cumple por cuanto el *puntaje de admisión* muestra una distribución claramente sesgada a la derecha. Este último supuesto, sin embargo, y el de que la variable en estudio ha sido medida al menos en escala de intervalo, son tratados con bastante flexibilidad en la literatura especializada.

Autores de reconocido prestigio relativizan el cumplimiento de los supuestos. Véase lo que plantean Cohen y Cohen:

“Afortunadamente, la evidencia disponible sugiere que aún una clara y sustancial separación de los supuestos frecuentemente resultará en errores de inferencia pequeños cuando los datos son tratados como si los supuestos fueran válidos” (Cohen and Cohen: 1975: p.48).

Kerlinger por su parte, citando a Lindquist anota: “la distribución F es sorprendentemente insensible a la forma de la distribución de las medidas de criterio en la población original... salvo que las varianzas sean tan heterogéneas que el investigador lo note a simple vista, esto es, que existan diferencias relativamente grandes, el efecto que ellas tengan en la prueba F probablemente será insignificante” (Kerlinger: 1979: p.205).

A la luz de la flexibilización de los supuestos que admite la prueba F podemos dar por buenos los resultados obtenidos con el *análisis de varianza*. No obstante, en aras del mayor rigor, procedimos a

aplicar la prueba U de Mann-Whitney, ésto con el objeto de confirmar los resultados encontrados con el análisis paramétrico de la varianza. Los resultados coinciden con los mostrados en el Cuadro 1.

1.1.1 Cuarto ciclo

En el Cuadro 2 se presenta un resumen de los datos más relevantes de las diferencias por sexo en las calificaciones de *cuarto ciclo*. El formato de presentación es el mismo usado para el Cuadro 1.

El *análisis de varianza* muestra que se presentan diferencias significativas en los promedios obtenidos por hombres y mujeres en *cuarto ciclo*, en los cinco años analizados.

Las mujeres obtienen puntajes más altos que los hombres. Las diferencias en los puntajes van desde 0,94 puntos hasta 1,83 puntos.

La varianza de *cuarto ciclo* explicada por las diferencias de sexo es muy pequeña; va de 0,39% a 1,63%.

Los resultados obtenidos con la prueba U de Mann-Whitney coinciden con los mostrados en el Cuadro 2.

1.1.2 Area Matemática

En el Cuadro 3 se presenta un resumen de los datos más relevantes de las diferencias por sexo en los promedios del *área matemática*. El formato de presentación es el mismo usado para el Cuadro 1.

El *análisis de varianza* muestra que se presentan diferencias significativas en los promedios obtenidos por hombres y mujeres en el *área matemática* de la Prueba de Aptitud Académica, en los cinco años analizados.

Las hombres obtienen puntajes más altos que las mujeres. Las diferencias en los

CUADRO 3. Área matemática por sexo y año.

Año	% Varianza explicada	Prob. F	Prob. Levene	Promedio Hombres	Promedio Mujeres	Diferencia Promedio
1998	3,51	0,0000	0,169	40,12	35,54	4,58
1997	4,11	0,0000	0,403	39,55	34,32	5,23
1996	2,69	0,0000	0,811	36,30	32,05	4,25
1995	4,07	0,0000	0,000	35,74	30,34	5,40
1994	3,97	0,0000	0,007	36,42	30,88	5,54

CUADRO 3. Área matemática por sexo y año.

Año	% Varianza explicada	Prob. F	Prob. Levene	Promedio Hombres	Promedio Mujeres	Diferencia Promedio
1998	1,30	0,0000	0,001	16,60	15,58	1,02
1997	0,71	0,0000	0,160	15,04	14,26	0,78
1996	1,37	0,0000	0,025	16,39	15,24	1,15
1995	1,77	0,0000	0,113	14,48	13,28	1,20
1994	2,06	0,0000	0,022	15,97	14,54	1,43

puntajes van desde 4,25 puntos hasta 5,54 puntos.

La varianza del *área matemática* explicada por las diferencias de sexo es muy pequeña; va de 2,69% a 4,11%.

Los resultados obtenidos con la prueba U de Mann-Whitney coinciden con los mostrados en el Cuadro 3.

1.1.3 Área Verbal

En el Cuadro 4 se presenta un resumen de los datos más relevantes de las diferencias por sexo en los promedios del *área verbal* de la Prueba de Aptitud Académica. El formato de presentación es el mismo usado para el Cuadro 1.

El *análisis de varianza* muestra que se presentan diferencias significativas en los promedios obtenidos por hombres y mujeres en el *área verbal* de la Prueba de Aptitud Académica, en los cinco años analizados.

Las hombres obtienen puntajes más altos que las mujeres. Las diferencias en los puntajes van desde 0,78 puntos hasta 1,43 puntos.

La varianza del *área verbal* explicada por las diferencias de sexo es muy pequeña; va de 0,71% a 2,06%.

Los resultados obtenidos con la prueba U de Mann-Whitney coinciden con los mostrados en el Cuadro 4.

Los resultados presentados hasta aquí coinciden parcialmente con los que reporta la literatura especializada. En una revisión de estudios de validez del SAT que analizan separadamente el comportamiento por sexo y que incluye a más de quinientas universidades, Ramist encuentra que mientras los grupos de hombres muestran un puntaje claramente superior en el *área matemática*, las mujeres muestran un puntaje levemente superior en el *área verbal* (Ramist: 1984: p.154).

1.2 Pertinencia de la fórmula de ponderación empleada.

El *puntaje de admisión* es un promedio ponderado de los tres elementos que lo componen. Al *área matemática* se le da un peso de 50%, a *cuarto ciclo* un peso de 30% y al *área verbal* un peso de 20%.*

Como se ha demostrado con los análisis practicados hasta aquí, los tres componentes del *puntaje de admisión* brindan resultados diferentes para uno y otro sexo. En *cuarto ciclo* las mujeres muestran puntajes

promedio superiores a los hombres, en tanto en las *áreas matemática y verbal* de la Prueba de Aptitud Académica los hombres obtienen puntajes superiores a las mujeres. En estas circunstancias la ponderación de los componentes del *puntaje de admisión* se vuelve crucial por cuanto puede tanto contrarrestar como agudizar los resultados diferenciales obtenidos por hombres y mujeres.

En principio la ponderación que se dé a los distintos elementos del *puntaje de admisión*, debería depender de su contribución diferencial a explicar la varianza del *rendimiento académico*. Con vistas a analizar esta problemática, en el Cuadro 5 se presentan las correlaciones entre los componentes del *puntaje de admisión* y el *promedio ponderado* obtenido por los estudiantes en su primer semestre en el Instituto, por año. En la primera columna aparece el año de que se trata. En la segunda columna se presenta la correlación (r de Pearson) entre *rendimiento académico* y *cuarto ciclo*. En la tercera columna se presenta la correlación entre *rendimiento académico* y el *área matemática* de la Prueba de Aptitud Académica. En la cuarta columna se presenta la correlación entre *rendimiento académico* y el *área verbal* de la Prueba de Aptitud Académica. En la última columna se presenta el número de estudiantes sobre los cuales se calcularon las correlaciones para el año respectivo.

El cuadro muestra claramente que para cuatro de los cinco años bajo análisis, *cuarto ciclo* es mejor predictor de *rendimiento*

académico que las *áreas matemática y verbal* de la Prueba de Aptitud Académica. El año en el cual la correlación entre el *área matemática* y *rendimiento académico* es superior a la correlación entre *cuarto ciclo* y *rendimiento académico*, la diferencia entre las correlaciones es ínfima.

A la vista de estos resultados no se justifica la mayor ponderación que se le da al *área matemática* en el cálculo del *puntaje de admisión*. Los resultados más bien señalan que *cuarto ciclo* debería tener el mayor peso. En todo caso, si los resultados de los dos últimos años marcaran una nueva tendencia, se aconsejaría dar un peso equitativo a *cuarto ciclo* y al *área matemática*. El *área verbal* prácticamente no contribuye a explicar *rendimiento académico*, por lo que su peso debería ser mínimo, si es que se le quiere dar algún peso (se calcularon las respectivas correlaciones parciales para sustentar la anterior afirmación; no se incluyen para no hacer este informe innecesariamente largo).

La ponderación que se emplea actualmente está creando un sesgo injustificado a favor de los hombres, toda vez que ellos obtienen un mayor puntaje en el *área matemática* y ésta tiene un poder explicativo sobre *rendimiento académico*, inferior al mostrado por *cuarto ciclo*.

La fórmula de ponderación empleada no sólo no se justifica técnicamente a la luz de los resultados obtenidos en el presente estudio, sino que además se separa drásticamente de la práctica que se sigue con el SAT. Ramist, de una revisión de estudios de validez del SAT en 685 universidades, encuentra que la ponderación promedio más adecuada es un 54% para las calificaciones de secundaria, un 26% para el *área verbal* y un 20% para el *área matemática*; aunque advierte que se dan variaciones considerables entre universidades. (Ramist: 1984: p.142)

Referencias bibliográficas

Cohen, J. and Cohen, P. (1975). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. New Jersey: Lawrence Erlbaum Associates Publishers.

* Comunicación personal de la MSc. Sharay Meneses, coordinadora del Comité de Examen

CUADRO 5. Correlaciones entre el promedio ponderado en el primer semestre y los componentes del puntaje de admisión por año.

Año	Cuarto ciclo	Área matemática	Área verbal	n
1998	0,1323	0,1347	0,1105	1015
1997	0,1165	0,1067	0,0195	885
1996	0,2372	0,1564	-0,0072	873
1995	0,2193	0,1342	-0,0009	881
1994	0,1923	0,1572	0,0402	1078

Todas las correlaciones son significativas más allá de ,001, excepto para el Área Verbal, para la cual sólo es significativa en el año 1998.

- Kerlinger, F.N. (1979). *Investigación del comportamiento*. México D.F.: Nueva Editorial Interamericana.
- Norusis, M.J./SPSS Inc. (1993) *SPSS for Windows Base System User's Guide*. Rel. 6.0
- Ramist, L. (1984). Predictive validity of the ATP Tests. En: Donlon, T.F., (Editor). *The College Board Technical Handbook of the Scholastic Aptitude Test and Achievement Test*. N.Y.: College Entrance Examination Board.
- Schrader, W.B. (1971). The predictive validity of College Board Admissions Test. En: Angoff, W.H., (Editor) *The College Board Admissions Testing Program*. N.Y.: College Examination Board.
- Siegel, S. (1985). *Estadística no paramétrica, aplicada a las ciencias de la conducta*. México D.F.: Trillas.
- Stricker, L.J.; Rock, D.A. & Burton, N.W. (1993). Sex differences in predictions of college grades from Scholastic Aptitude Test scores. *Journal of Educational Psychology*, vol 85, N° 4, 710-718.
- Vives Brenes, W. y Blando Valverde, M. (1998). *Estudio sobre el comportamiento de la demanda estudiantil en el ITCR*. Cartago: Instituto Tecnológico de Costa Rica, Programa Equidad de Género.
- Wainer, H. & Steinberg, L.S. (1992). *Sex differences in performance on the mathematics section of the Scholastic Aptitude Test: a bidirectional validity study*. Boston: Harvard Educational Review, vol. 62, N° 3, 323-336.