LA EQUIPARACIÓN DEL EXPEDIENTE DE BACHILLERATO EN EL PROCESO DE SELECCIÓN DE ALUMNOS PARA EL ACCESO A LA UNIVERSIDAD

JOSÉ LUIS GAVIRIA SOTO*

RESUMEN. En el próximo futuro las universidades deben determinar el procedimiento que se ha de seguir en la selección de alumnos que quieran ingresar en primer curso de las carreras universitarias. Los cambios legales que se están produciendo así lo determinan.

El procedimiento utilizado hasta la fecha en las pruebas de acceso a la universidad parte del supuesto incontrastable y probablemente erróneo de que las notas del expediente de Bachillerato de todos los alumnos está en la misma escala, y por tanto, son comparables. En esta investigación se parte del supuesto contrario, más realista, de que las notas del Bachillerato son totalmente contextuales y no pueden ser comparadas sin antes realizar una transformación adecuada. Para ello se estudian procedimientos que permiten transformar las puntuaciones del Bachillerato poniéndolas todas en la misma escala, mejorando de esta forma la equidad del proceso de selección.

La investigación se subdivide en un estudio de viabilidad, para el que se realiza una simulación de Montecarlo, y en un estudio de plausibilidad, para el que se recogieron datos de selectividad correspondientes a 57.465 sujetos de centros públicos y privados, en cinco años, de las convocatorias de junio y de septiembre, de todas las opciones de tres universidades españolas, las universidades de Salamanca, Extremadura y Burgos.

ABSTRACT. In the near future, according to the new legislation, universities will have to determine the selection process for those students who wish to take 1st year university studies.

Until now, the process used in the university entrance exam was based on the uncontrastable and probably wrong assumption that all students' Baccalaureate record's marks are part of the same scale and therefore, comparable. This research project is based on the more realistic and opposite assumption that Baccalaureate marks completely fit a context and cannot be compared without previously undergoing an appropriate change. In order to do so, procedures to change Baccalaureate marks by placing them in the same scale are under study, which would, therefore, improve the equity of the selection process.

The research project consists of a viability survey, for which a Monte Carlo simulation has been carried out, and a plausibility survey, for which the *Selectividad* (university entrance exam) data of 57,465 students, from public and private educational establishments, over five years (June and September calls) were collected. The data refers to all the options of three Spanish universities: Salamanca, Extremadura and Burgos.

(*) Departamento de Métodos de investigación y Diagnóstico en Educación. Universidad Complutense de Madrid.

Revista de Educación, núm. 337 (2005), pp. 351-387.
Fecha de entrada: 15-04-2003 Fecha de aceptación: 04-11-2003

INTRODUCCIÓN

En un próximo futuro las universidades deben determinar el procedimiento que se ha de seguir en la selección de alumnos que quieran ingresar en primer curso de las carreras universitarias. Los cambios legales que se están produciendo así lo determinan.

En la actualidad se utiliza para este fin las notas del expediente de Bachillerato de los alumnos aspirantes y la nota obtenida en el examen de selectividad. Las dos notas mencionadas tienen una importancia indudable para proporcionar información acerca de la valía y capacidad académica de los aspirantes.

La nota del expediente en el bachiller tiene la ventaja de que refleja el rendimiento medio continuado del alumno a lo largo de los años de preparación para la entrada a la universidad. Sin embargo la comparación de los resultados de alumnos procedentes de distintos centros resulta ciertamente problemática. El nivel de exigencia de cada uno de los centros determina la nota media de los alumnos. Una puntuación de 7 en un centro puede representar mucho más nivel de trabajo y conocimientos que una puntuación de 8 ó 9 en otro centro.

Por su parte la nota de la Prueba de Aptitud para el Acceso a la Universidad (PAAU) tiene la ventaja de que supone el resultado de un mismo examen para todos los alumnos. Presenta sin embargo el inconveniente de que proporciona una información puntual en el tiempo, y el margen de error en la valoración de los conocimientos y destrezas individuales es demasiado alto como para confiar sólo en ese resultado.

El procedimiento empleado en la actualidad consistente en obtener una media ponderada de las puntuaciones mencionadas está justificado por las limitaciones que cada una de las informaciones por separado presenta. Sin embargo es posible mejorar substancialmente el uso de la información que ambas notas proporcionan.

En estas páginas se presenta una investigación cuyos resultados permiten poner en marcha un procedimiento que aprovecha la mejor información de las dos notas, eliminando sus inconvenientes.

ANTECEDENTES: LA INVESTIGACIÓN SOBRE LAS PRUEBAS DE ACCESO A LA UNIVERSIDAD

En el sistema educativo español y, al menos desde los años de la Guerra Civil, siempre ha existido alguna prueba situada a medio camino entre el bachillerato o sus últimos cursos y la Universidad. Unas veces como culminación de aquél, otras como puerta de acceso a ésta.

La Prueba de Aptitud para el Acceso a la Universidad o prueba de selectividad ha sido sometida a muchos estudios desde su creación. Se ha estudiado su capacidad predictiva: González y Valle (1990), Escudero (1987); diferencias entre tribunales y efectos sobre los alumnos en función de su nota en Bachillerato: Sanz (1992); estudios pormenorizados de resultados de algunas universidades: Miguel (1988); o la evolución de los resultados de la selectividad en el tiempo Miguel (1990), Miguel (1993).

Un tema que ha sido abordado en varias ocasiones es la posibilidad, muy directamente relacionada con el objetivo de este trabajo, de que algunos centros sobrevaloren la nota de bachiller. De hecho es un tema que es abordado en varias de las investigaciones sobre selectividad, incluso cuando ése no fuese el objetivo principal de la misma.

Para algunos parece evidente que los centros privados *no* sobrevaloran la nota, Aguirre de Cárcer (1986).

Con relación a este tema Muñoz-Repiso y otros (1997) resaltan que la función de la selectividad, tal y como se concibe en la LOGSE, es, entre otras cosas, garantizar homogeneidad de criterios de evaluación entre los centros.

> ... Es preciso garantizar que los alumnos que accedan a la Universidad han recibido una formación adecuada en sus respectivos centros de secundaria, *que ésta ha sido valorada con criterios similares en cada uno de esos centros*[1], que poseen la madurez adecuada y que, en caso necesario, sólo el rendimiento académico medido con equidad pueda decidir qué alumno tiene prioridad para realizar determinados estudios. Muñoz-Repiso y otros (1997).

En el mismo estudio Muñoz-Repiso y otros (1997) comprueban, en los datos de la PAAU en la UAM correspondiente a 1995, que existe una diferencia entre las notas medias de expediente de Bachillerato y prueba de acceso. Así informan de una nota media de 6,97 en el expediente de Bachillerato, y de 5,19 en la prueba de selectividad, con una desviación típica de 0,81 y de 1,37 respectivamente. La media de la nota final era de 6,37 con una desviación típica de 0,87. La diferencia media es por tanto de 1,78 puntos. La diferencia entre bachiller y PAAU es menor en los centros privados que en los públicos (1,69 frente a 1,85).

Las diferencias entre las notas de selectividad y las de Bachillerato ya habían sido comprobadas en ocasiones anteriores. Por ejemplo Muñoz-Repiso y otros (1988) habían comprobado que las de selectividad son como promedio casi dos puntos inferiores a las de bachillerato. En Muñoz-Repiso y otros (1991) vuelven a comprobarse diferencias de calificación entre los tribunales de cada universidad y según el tipo de centro en el que hayan estudiado. Los mismos autores encontraron una correlación entre el expediente de Bachillerato y la nota de selectividad de 0,63.

Es fácil imaginar por qué se producen esas diferencias. Las notas de los expedientes de Bachillerato están puestas por equipos de profesores concentrados en un solo centro, y que no tienen modo de comparar a sus alumnos con los de otros centros distintos. Las notas de Bachillerato son totalmente dependientes del contexto en el que se asignan y por tanto no admiten una interpretación libre del contexto ni comparaciones directas con las notas concedidas en otros centros.

En esta investigación se mantiene específicamente que las notas de los expedientes del Bachillerato de cada centro son puntuaciones que están expresadas en escalas distintas que no tienen ningún vínculo de unión, que para poder compararlas correctamente es preciso someterlas al proceso técnico conocido como *equiparación,* y que esa equiparación es posible llevarla a cabo solamente basándonos en los resultados de una prueba que sea común a todos los alumnos, como por ejemplo la nota de la prueba de selectividad, y que actuaría como *test de anclaje*.

LA SELECTIVIDAD COMO PROBLEMA TÉCNICO

La determinación del acceso a los estudios universitarios es un problema técnico que es perfectamente abordable desde la investigación educativa. Se trata en

(1) Cursiva añadida.

definitiva de evaluar y seleccionar de la manera más equitativa, y por tanto más justa.

La esencia de la cuestión que nos ocupa es un problema de equiparación, como ya hemos señalado. Se actúa como si las notas de los centros fuesen equiparables, cuando eso no es más que una hipótesis que debe ser sometida a prueba. En esta investigación se parte del punto contrario. Se parte del supuesto de que las puntuaciones no son equivalentes y deben ser equiparadas. Se considera al caso de puntuaciones equiparadas como un caso particular en el que las diferencias entre centros es cero.

Equiparar las puntuaciones de dos escalas consiste en establecer una función de transformación de las puntuaciones de una de ellas de forma que obtengamos las puntuaciones que les corresponderían en la otra.

En el problema que nos ocupa puede entenderse que una escala está constituida por las puntuaciones proporcionadas por la selectividad. En este mismo sentido las notas proporcionadas por un colegio constituyen otra escala.

Así nos encontramos con una situación en la que de todos los sujetos tenemos dos puntuaciones. Una de ellas proviene de la selectividad, que es una escala común para todos los sujetos, mientras que la otra es la nota proporcionada por el centro. Cada centro sin embargo establece una escala distinta. Nuestro objetivo consiste en poder comparar las puntuaciones producidas por distintos centros.

CONDICIONES DE EQUIPARACIÓN

Para que sea posible la equiparación tenemos que cumplir ciertas condiciones (Lord, 1980; Morris, 1982):

- *Igual habilidad*. Los dos instrumentos de medida están referidos a la misma característica de los sujetos, y las mismas variaciones en las características producirán variaciones en las medidas proporcionadas por los dos instrumentos. Naturalmente variaciones idénticas no producirán variaciones idénticas en las medidas, ya que en ese caso no sería necesaria la equiparación. Este supuesto se refiere al hecho de que los dos instrumentos son sensibles a la misma magnitud.
- *Equidad* (Distribuciones condicionales iguales). En su forma más general este supuesto implica que para cualquier alumno tiene que ser indiferente la forma del de la medida que ha de tomar. Esto supone que la fiabilidad de ambas medidas ha de ser la misma.
- *Invarianza* respecto de la población. Esta condición implica que la relación que se establece entre los dos instrumentos de medida no varía al cambiar la población sobre la que se calcula.
- *Simetría*. (Transformación reversible). Esta condición implica que la transformación, realizada desde el instrumento a hasta el b, pueda ser revertida al tomar la puntuación correspondiente desde el instrumento b y obteniendo como transformada la a. Es decir, tiene que ocurrir que si es la puntuación que en la escala del instrumento x corresponde a una puntuación en el instrumento y, e $y^*(x)$ es la puntuación que en la escala del instrumento y corresponde a una puntuación en el instrumento x, entonces $x = x^* (y^* (x))$ e $y = y^* (x^* (y))$. Esto tiene como consecuencia que si a un grupo de sujetos se le aplicó el instrumento x y a otro el instrumento y, tiene que inducirse el mismo orden si se rea-

liza la transformación de la escala de *x* en *y* que si realiza la transformación de la escala de *y* en *x*. En el fondo este requisito es otra forma distinta del principio de equidad. Efectivamente, si no se cumple el supuesto de simetría tendríamos que para un sujeto dado no es indiferente si toma la forma *x* o la forma *y*, puesto que cambiaría su posición en la ordenación final.

No todas las transformaciones cumplen con estas condiciones. El procedimiento que se ha venido utilizando para determinar el acceso a la universidad los asume implícitamente, sin que se haya hecho ningún intento de comprobación de los mismos. El solo hecho de combinar en una sola puntuación las notas de Bachillerato y de selectividad implican que están midiendo la misma cosa, pues de otro modo estaría poco justificada esta práctica. Asimismo la práctica de no transformar las puntuaciones de Bachillerato procedentes de distintos centros supone que se asume que son equiparables, y por tanto simétricas, invariantes y equitativas.

Tampoco todos los métodos de equiparación posibles tienen porqué cumplir en principio estos supuestos. Por ejemplo la equiparación basada en la regresión no cumple la condición de simetría, mientras que, aunque la transformación que se propone en esta investigación está basada en la regresión, dado que es regresión de todos los instrumentos (centros) sobre un elemento común, la selectividad, es una transformación reversible. Y además se cumple que la ordenación es la misma siempre, indistintamente de cuál sea la escala a la que se vierten los datos.

Por ejemplo en el caso de la regresión, si tenemos la transformación de equiparación de la forma uno a la dos dada por la expresión $x_2^* = \alpha_1 x_1 + \beta_1$ y de la forma dos a la uno por $x_1^* = \alpha_2 x_2 + \beta_2$, comprobamos que la transformación no es reversible, ya que

$$x_2^* = \alpha_1(\alpha_2 x_2 + \beta_2) + \beta_1 =$$

$$= \alpha_1\alpha_2 x_2 + \alpha_1\beta_2 + \beta_1 \neq x_2$$

Como consecuencia el orden que se induce es también distinto según se pasen los datos a una u otra escala.

Ninguna de estas dos cosas ocurre con el método que aquí se propone. En primer lugar, dada una puntuación, la antitransformada de la transformada de la puntuación es igual a la misma puntuación. Para el colegio *1* su cambio a la escala común se hace por medio de $x'_1 = \alpha_1 x_1 + \beta_1$ y para el colegio 2, por medio de $x'_2 = \alpha_2 x_2 + \beta_2$. Para cada centro los valores de los coeficientes α y β son los que corresponden a la regresión de *s*, la nota de selectividad, sobre *x*, la nota del expediente de Bachillerato, en ese centro. Las puntuaciones x_1 y x_2 son equiparables si sus transformadas a la escala común son iguales, es decir, si $x_2 = x_1$. Luego si deseo obtener la puntuación que en el colegio *2* correspondería a una puntuación en el colegio *1*, tendría $\alpha_1 x_1 + \beta_1 = \alpha_2 x_2 + \beta_2$ con lo que

$$x_2^* = \frac{1}{\alpha_2}(\alpha_1 x_1 + (\beta_1 - \beta_2))$$

y $$x_1^* = \frac{1}{\alpha_1}(\alpha_2 x_2 + (\beta_2 - \beta_1))$$

Para que sea simétrica se tiene que cumplir que $x_1^*(x_2^*(x_1)) = x_1$ y viceversa.

Así,

$$\begin{aligned} x_1^*(x_2^*) &= \frac{1}{\alpha_1}(\alpha_2 x_2^* + (\beta_2 - \beta_1)) = \frac{1}{\alpha_1}(\alpha_2 x_2^* + (\beta_2 - \beta_1)) = \\ &= \frac{1}{\alpha_1}(\alpha_2(\frac{1}{\alpha_2}(\alpha_1 x_1 + (\beta_1 - \beta_2)) + (\beta_2 - \beta_1)) = \\ &= \frac{1}{\alpha_1}(\frac{\alpha_2}{\alpha_2}(\alpha_1 x_1 + \beta_1 - \beta_2 + \beta_2 - \beta_1)) = \frac{1}{\alpha_1}\alpha_1 x_1 = x_1 \end{aligned}$$

Y a la inversa ocurre con la otra transformación. Es por lo tanto una transformación simétrica, y en consecuencia es indiferente la dirección en que se haga, y el orden que se induce es siempre el mismo.

Así, la condición para que tengamos una equiparación simétrica y, por ello, equitativa, es que se disponga de una nota procedente de una prueba común, en la que la posición de todos los centros y su dispersión pueda ser determinada. La prueba de selectividad debe hacer la función de *prueba común* y las notas de los distintos centros hacen la función de *pruebas variantes*.

DISEÑO DE LA INVESTIGACIÓN

OBJETIVO E HIPÓTESIS

El objetivo de la presente investigación consistía en demostrar que la equiparación de las notas de Bachillerato basándose en la nota de la prueba común de la selectividad proporciona una ordenación para el acceso a la Universidad, más adecuada justa y equitativa de los alumnos de los distintos centros.

No existe una sola forma de llevar a cabo la equiparación. Por eso en este trabajo se propuso comparar cuatro métodos distintos. Cada uno de estos métodos constituía un nivel de la variable independiente. El primero, denominado clásico era en realidad el correspondiente al grupo de control. En el apartado referido a las variables se explican cada uno de los métodos cuyos efectos han sido comparados.

HIPÓTESIS DE INVESTIGACIÓN

H_1: Cualquiera de los tres métodos alternativos producirá mejores resultados que el clásico.

H_2: Los métodos basados en regresión sobre la selectividad, OLS y Multinivel, dado que utilizan mayor cantidad de información, producirán mejores resultados que los otros, Clásico e IMS.

VARIABLES

Variable independiente: Como variable independiente tenemos el método de equiparación, con cuatro niveles que corresponden a cada uno de los métodos de equiparación que se han mencionado, y que se describen con más detalle en otro punto.

Variables dependientes:

La congruencia entre alumnos seleccionados utilizando la puntuación latente como criterio y alumnos seleccionados por cada uno de los métodos de la variable independiente.

La correlación entre la puntuación transformada y la puntuación de selección obtenida para cada alumno con cada uno de los cuatro métodos mencionados.

VARIABLE INDEPENDIENTE. MÉTODOS PROPUESTOS

Los métodos que se presentan en esta investigación tienen en cuenta la nota del expediente de bachiller del alumno, modificada en función de los resultados medios del centro de procedencia en la Prueba de Acceso a la Universidad.

Método clásico

Éste es en realidad el de control. Consiste en calcular la nota de acceso exactamente como se hace en todas las universidades siguiendo la normativa legal. Así la nota del expediente de Bachillerato de cada alumno se pondera por 0,6 y la nota de Selectividad por 0,4.

Método OLS (Mínimos cuadrado ordinarios)

La idea básica es muy simple. Si un alumno procede de un centro en el que se tiende a puntuar de forma excesivamente alta, su puntuación deberá corregirse a la baja, mientras que si la misma nota la hubiese obtenido el alumno en un centro en el que los profesores tienden a dar puntuaciones excesivamente bajas, debería corregirse al alza.

Imaginemos que fuese posible que las notas del expediente de todos los aspirantes estuviesen asignadas por el mismo claustro de profesores. Entonces no nos cabría la menor duda de la equivalencia de las notas. Sin embargo esto no es posible. Pero existe algo que es común a todos los alumnos. Se trata de las notas de la selectividad. Estas notas nos dan una referencia común para todos los alumnos. Ciertamente para un solo alumno puede ser arriesgado tomar esa nota como único referente. Pero es razonable pensar que el conjunto de errores que se pueda cometer con la nota de selectividad en la valoración de los conocimientos de los alumnos de un centro, se compensan unos con otros, es decir, la media en selectividad de los alumnos de un centro es un estimador insesgado de la verdadera nota media que les correspondería.

Se trata entonces de equiparar las notas de bachiller utilizando como información de anclaje las notas medias de cada centro en selectividad. Supongamos que llamamos T_{ij} a la nota del alumno i del centro j que nos servirá para realizar la selección, B_{ij} a la nota de bachiller, y S_{ij} a la nota de selectividad. Hasta ahora se calcula $T_{ij}=0.6B_{ij}+0.4S_{ij}$.

Se propone que la nota B_{ij} debe estar modificada de la forma a_jB+b_j. La nota S_{ij} es la única que está en una escala común a todos los alumnos, por tanto es lógico que sea la que sirva para establecer la métrica común. Las notas del expediente se someten a una transformación lineal, de modo que la media de B correspondiente al centro se equipare a la media

$$\sum_{i=1}^{n}(S_{ij}-(\alpha_j B_{ij}+\beta_j))^2=\text{mínimo}$$

del centro en S. Los valores de a_j y b_j se obtienen de forma que minimicemos la suma de los cuadrados de las diferencias entre la nota de la selectividad y la puntuación transformada de B.

Como resultado nuestro problema se reduce a hallar una ecuación de regresión lineal específica para cada centro de bachiller.

Método Multinivel (Regresión Multinivel)

Como una segunda posibilidad se avanza la idea de que una transformación basada en la regresión multinivel tendrá un efecto similar al anterior, con la ventaja de que la ecuación de transformación de los centros que presenten un número reducido de alumnos se vería beneficiada por la información aportada por los demás centros.

La idea es que si de cada centro tuviésemos un solo alumno, obtendríamos sólo una ecuación de regresión común a todos los centros. En el otro extremo, si tuviésemos muchos alumnos en cada centro, podríamos tener sin dudarlo una ecuación de regresión para cada centro. Por tanto es razonable contar con un método que cuando el número de sujetos de un centro sea muy grande, dé mucha importancia a esos datos, y cuando el número de alumnos sea muy pequeño, se apoye en los datos que disponemos acerca de la relación general entre las dos variables consideradas, la nota de la prueba común y la nota de bachiller.

Método IMS (Igualación de medias y desviaciones típicas)

Se trata de que en cada centro, la puntuación transformada de cada sujeto se obtiene asignándole en X* una puntuación que se aparte de la media tantas desviaciones típicas como la puntuación de selectividad del sujeto se aparte de la media de la nota de selectividad del centro. Es decir,

$$\frac{x^* - \overline{S}_c}{S_{sc}} = \frac{x - \overline{x}_c}{S_{xc}} \quad \text{, y}$$

$$x^* = \frac{S_{sc}}{S_{xc}}(x - \overline{x}_c) + \overline{S}_c$$

donde

- x^* es la puntuación de Bachillerato transformada
- x es la nota que se va a transformar
- $\overline{S}_c$ es la media del centro en la prueba de selectividad
- $\overline{x}_c$ es la media del centro en la nota de expediente de Bachillerato
- S_{sc} es la desviación típica del centro en la prueba de selectividad
- S_{xc} es la desviación típica del centro en el expediente de Bachillerato.

VARIABLES DEPENDIENTES

Para comparar los distintos métodos se han determinado dos variables dependientes. La primera es la variable COINCIDENCIAS. La justificación de esta variable es la siguiente: todo el proceso de selección tiene como principal finalidad, además de determinar qué alumnos pueden acceder a la universidad, el establecer el orden en que estos alumnos pueden elegir centro de estudios. Por simplicidad se seleccionan los cien primeros alumnos que resultan en la ordenación en la puntuación latente, que es la puntuación σ. Seguidamente, para cada uno de los métodos obtenemos las puntuaciones equiparadas, y basándonos en ellas, la puntuación final correspondiente. Así si x es la puntuación equiparada, la puntuación final se calcula como $0,60x+0,4S$. A continuación se ordenan los sujetos en función de la puntuación final de cada método, y se comprueba cuántos sujetos seleccionados con la puntuación latente son también seleccionados por cada método.

La segunda variable es la CORRELACIÓN que se obtiene entre la puntuación de la selectividad y la puntuación equiparada. Si el método es eficaz, la correlación entre la puntuación típica y la puntuación equiparada debe ser mayor que la correlación entre la puntuación de bachillerato sin equiparar y la puntuación de selectividad. Se obtiene por tanto la nueva correlación para cada muestra, y se constituye en variable dependiente.

Esta segunda variable dependiente es de gran utilidad, puesto que en los datos reales no es posible medir las coincidencias, ya que no se conoce la puntuación latente. Pero sin embargo sí es posible comprobar las variaciones en la correlación entre la selectividad y las puntuaciones equiparadas en comparación con la puntuación original del Bachillerato. Un patrón similar de comportamiento en esta variable en el caso de los datos reales respecto de los simulados permitiría afirmar la plausibilidad del método o métodos propuestos.

Los resultados finales se analizan utilizando una técnica tradicional de contraste de hipótesis como el análisis factorial de varianza con medidas repetidas.

METODOLOGÍA

Se han llevado a cabo dos estudios paralelos: uno de viabilidad y otro de plausibilidad.

En el *estudio de viabilidad*, se definió una estructura teórica para los datos, se generaron los datos a partir de esa estructura teórica, se aplicaron los métodos descritos a los datos así generados, y por último se compararon los resultados obtenidos con la estructura verdadera generada.

En el *estudio de plausibilidad* se trataba de determinar si los datos obtenidos de una situación real se asemejan en su estructura a los generados por medio de la simulación, y si los resultados obtenidos cuando son analizados con la metodología propuesta son congruentes con los producidos en aquélla.

ESTUDIO DE VIABILIDAD

En esta investigación en concreto se decidió generar datos simulando el proceso que se produce en la selectividad en la forma en que se describe a continuación.

En primer lugar se generó una variable con distribución Beta[2] que representa la puntuación latente, es decir, el verdadero nivel de conocimientos de los alumnos. Si esta variable fuese realmente conocida, sería la variable que se utilizaría como criterio para determinar qué alumnos accederán a cada carrera.

A partir de la puntuación latente se generaron dos puntuaciones distintas para cada sujeto. Una relacionada con la nota de expediente de Bachillerato y la otra con las notas de la prueba de selectividad.

La nota de la prueba de selectividad se generó como una medida con error de la puntuación latente, pero en la misma escala para cada sujeto, con valores comprendidos entre 0 y 10.

La nota del expediente de Bachillerato se generó como una medida con error de la puntuación latente, pero con una media y una desviación típica distinta para cada centro, y con la limitación de que sus valores estén comprendidos entre 5 y 10.

La nota del expediente de Bachillerato se generó con distintas distribuciones, en función de las condiciones experimentales de la simulación, como se explica más adelante.

Una vez generados los datos, se procedió a aplicar cada uno de los métodos de equiparación descritos.

Para el *método clásico* simplemente se multiplicó la nota de selectividad por 0,40 y la nota de bachillerato por 0,60 y se sumaron las dos cantidades.

(2) En el apartado *Generación de la puntuación latente* y siguientes se dan explicaciones detalladas del proceso de generación.

Para el *método OLS* se calculó para cada uno de los centros supuestos una ecuación de regresión distinta, con la nota de Bachillerato como predictor y la nota de selectividad como criterio. Con esa ecuación de regresión se obtuvieron las predicciones para cada sujeto. La predicción es la puntuación transformada. Seguidamente se multiplicó la transformada por 0,6 y se le suma a la nota de selectividad multiplicada por 0,4.

Con el *método multinivel*, se ajustó un modelo de dos niveles, con el centro como nivel 2, con la nota de Bachillerato como predictor y la nota de selectividad como criterio, con el intercepto y la pendiente variando aleatoriamente en el nivel 2. Con este modelo se obtuvo la predicción para cada sujeto, se multiplicó por 0,60 y se sumó a la nota de selectividad multiplicada por 0,40.

En el *método IMS*, para cada centro se calcularon los valores correspondientes de media y desviación típica en la nota del expediente del Bachillerato y en la de selectividad. A continuación se realiza la transformación de las puntuaciones de cada sujeto de forma que la puntuación que le corresponde en la escala transformada se desvía de la media el mismo número de desviaciones típicas de la escala de destino que las que su puntuación original se desvía de la media de la escala original.

Generación de los datos para la simulación

Supuestos

La forma en de generación de los datos simulados debía estar relacionada con la naturaleza de los datos reales. Se trataba de generar un banco de datos que se asimilase en su estructura y distribución lo más posible a los datos reales.

La primera pregunta que podríamos hacernos, es: ¿Qué es lo que realmente tratan de medir tanto la nota del expediente del bachillerato como la nota de la Prueba de Acceso a la Universidad? Responder con detalle a esta cuestión podría ser una tarea larga y compleja, pero de lo que no cabe ninguna duda es de que *tratan de medir la misma cosa*. De ninguna otra forma podría justificarse la práctica de combinar las dos puntuaciones en una sola nota que es la que determina la puntuación de acceso a la Universidad.

El que las dos puntuaciones reciban distinta ponderación puede responder a dos motivos. Primero, que las dos puntuaciones estén en distinta escala y esos coeficientes traten de equipararlos. Segundo, y más cierto en este caso, que los coeficientes reflejen la importancia relativa que por distintos motivos se desea asignar a cada puntuación.

Basándonos en esta idea, proponemos los siguientes supuestos:

- Tanto la prueba de selectividad como la nota de expediente de Bachillerato son estimaciones independientes de la misma capacidad latente.
- Los centros difieren en la media en la capacidad latente. No son por tanto muestras independientes de la misma población.
- La prueba de selectividad, por ser una prueba común a todos los alumnos, es una estimación independiente e insesgada de la puntuación latente y establece una escala que es común para todos los alumnos.
- La nota de expediente de Bachillerato es una estimación independiente de la puntuación latente, pero por las condiciones en que se obtiene, las puntuaciones de los sujetos no están en la misma escala. Su media está desplazada

respecto de la verdadera media, y lo mismo sucede con la desviación típica. Cada centro por tanto tiene su propia métrica, distinta de los demás.

- Sin embargo dentro de cada centro, la fiabilidad de la nota de bachillerato es superior a la fiabilidad de la nota de selectividad. Esto se justifica por el hecho de que la nota de Bachillerato refleja un proceso más largo y continuado de evaluación de los conocimientos, y varias y sucesivas aplicaciones de pruebas. Los posibles errores en una de ellas en cierta dirección quedarían compensados por errores en otra dirección distinta en otra de las evaluaciones. Una consecuencia de este supuesto es que *dentro de cada centro*, el orden de los sujetos queda mejor reflejado por el expediente de Bachillerato que por la nota de selectividad.
- La nota de la prueba de acceso tiene, *para cada sujeto*, mayor error que la nota de bachiller, ya que responde a una evaluación puntual sujeta a muchas más variables que afectan a la precisión de la medida. Sin embargo, la media *para cada centro*, refleja mejor la media de la puntuación latente del centro que la nota del expediente, puesto que está en la misma escala para todos los centros, cosa que no ocurre con la nota del expediente. Una consecuencia de este supuesto es que *entre los centros*, el orden de los colegios queda mejor reflejado por la media de selectividad que por la media del expediente de Bachillerato.

Fases de la generación de datos

Basándonos en estos supuestos se ha llevado a cabo la generación de los datos de la simulación. Para ello se han seguido varios pasos.

- Generación de la puntuación latente para cada sujeto.
  - Determinación de la media general.
  - Determinación de la media de cada centro.
  - Determinación del residuo de cada sujeto.
- Generación de la nota de selectividad a partir de la puntuación latente.
  - Determinación del residuo de cada sujeto respecto de σ.
- Generación de la nota de Bachillerato.
  - Determinación de los parámetros de desplazamiento de cada centro.
  - Determinación de los residuos de cada sujeto respecto de la nota de Bachillerato.

Generación de la puntuación latente

Así, para el alumno *i* del centro *j*, tendremos:

$$\sigma_{ij} = \gamma_0 + \delta_{0j} + \zeta_{ij}$$

Donde σ es la puntuación latente (y por tanto latente) que corresponde a dicho alumno

γ_o es la media general en la puntuación latente

δ_{oj} es lo que la media del centro *j* se aparta de la media general

ζ_{ij} es lo que el sujeto *i* del centro *j* se aparta de la media de su centro.

Empíricamente se ha comprobado que las medias de los centros están comprendidas en un intervalo de más menos dos puntos respecto a la media general.

Por ese motivo, la distribución condicional de la media de cada centro no será exactamente normal, puesto que ningún valor puede bajar de –2 puntos respecto de la media, ni pasar de +2 puntos por encima de la media. Por ese motivo se toma una aproximación con una distribución Beta con parámetros $v=w=5$. Así,

$$\delta_{oj} \sim 2B(5,5) + (-2)(1-B(5,5))$$

En cuanto a los residuos, sus valores son tales que los valores de la variable latente estén comprendida también entre 0 y 10. En esta caso los parámetros no pueden ser fijos. No se trata de la misma distribución para todos los centros, sino que ésta varía de un centro a otro.

$$\zeta ij \sim wB(v_\zeta,w_\zeta) + (-v(1 - B(v_\zeta,w_\zeta)))$$

Con

$$v_\zeta = \gamma_0 + \delta_{0j} \text{ y } w_\zeta = 10 - (\gamma_0 + \delta_{0j}).$$

Generación de la nota de la selectividad

La nota de la selectividad, s_{ij}, es una estimación insesgada de la capacidad latente de los sujetos. Sea σ_{ij} la verdadera capacidad del sujeto *ij* la puntuación s_{ij} obtenida es efecto del error de medida. Como los valores de s_{ij} están comprendidos entre 0 y 10, la distribución condicional de respecto de s_{ij} sólo será aproximadamente normal en el centro de la distribución. Por eso hacemos una aproximación con una distribución Beta reescalada. La media de la distribución condicional de s_{ij} es σ_{ij}. Como se explica en el apéndice,

$$v_s = \frac{k_s \sigma_{ij}}{10} \quad \text{y} \quad = \frac{(b-a)(\overline{x}-a)(1-\frac{\overline{x}-a}{b-a})}{\sigma_x^2} - 1$$

los valores de los parámetros de esta distribución Beta son (con $k_s = v_s + w_s$) y .

En nuestro caso $\overline{x} = \sigma_{ij}$, (b–a) = 10, mientras que se comprobó empíricamente que el valor más verosímil para la varianza de los residuos con los datos de las Universidades de Burgos, Extremadura y Salamanca era 0.6875.

Por tanto para el sujeto *i* del grupo *j*, su nota en la selectividad tiene una distribución linealmente dependiente de una beta de la cual la media es σ_{ij} , y cuyos parámetros vienen dados por las expresiones anteriores. La relación lineal viene determinada por los límites del recorrido de la variable. Y como éstos son $a=0$ y $b=10$, tendremos en definitiva $s_{ij} = a + (b - a)\, B(v_s, w_s) = 10B(v_s, w_s)$. De esta forma la distribución Beta tiene una forma que depende de los parámetros, v_s y w_s, que a su vez son función de la puntuación latente, de la amplitud del recorrido de la variable, y de la fiabilidad, esto último a través de la varianza de épsilon.

Generación de la nota de expediente de Bachillerato

Esta parte de la simulación es probablemente la más compleja. En primer lugar tenemos el hecho de que x debe estar comprendido entre *5* y *10* puntos. Esto se debe naturalmente a que ningún sujeto puede llegar a tener nota de la selectividad si antes no tiene aprobado el Bachillerato.

Debemos asumir relaciones no lineales producidas por el hecho de que se trata de variables acotadas, con unas varianzas que son relativamente grandes respecto a las cotas.

Por estas razones se adoptó una distribución Beta.

Así,

$x_{ij} = f(\beta_j,\alpha_j\sigma_{ij},\omega_{ij}$

β_j, α_j representan el desplazamiento y el cambio de unidad de la escala de la media del centro *j*.

σ_{ij} es la puntuación en la variable latente del sujeto *i* del centro *j*.

ω_{ij} es el residuo asociado al sujeto *i* del centro *j*.

Generación de β_j

Este parámetro representa el desplazamiento de la media de las puntuaciones de expediente de Bachillerato respecto de la media de la puntuación latente en cada centro. Ese parámetro tiene una distribución Beta tal que

$$\beta_j \sim a + (b-a)B(v_\beta, w_\beta)$$

con parámetros $v_\beta = \dfrac{k_\beta(\mu_{\beta c} - a)}{(b-a)}$,

$$k_\beta = \frac{(b-a)(\mu_{\beta c} - a)(1 - \frac{(\mu_{\beta c} - a)}{b-a})}{\sigma^2_{\sigma\, entre}} - 1$$

y $k_\beta = v_\beta + w_\beta$[3] .

En estas expresiones $\mu_{\beta c}$ es la media de la distribución condicional de beta. Este valor depende para cada centro de $\gamma_0 + \delta_{0j}$. Pero mientras que ese valor puede ir de *0* a *10*, el valor de beta para cada centro sólo puede tomar valores entre *5* y *10*. Por ese motivo la media condicional tiene una relación de función logística, como aparece en la siguiente expresión:

$$\mu_{\beta c} = a + (b-a)/(1 + \exp(-((\gamma_0 + \delta_{0j}) - P)))$$

donde *a*=5, *b*= 10, y *P* es un parámetro de posición, al que en este caso se le asignó el valor de γ_0 + 1.5.

(3) Véase apéndice.

Generación del resto de la parte fija de x_{ij}

La parte fija de x_{ij} tiene una relación no lineal con x_{ij}, ya que depende de s_{ij} que toma valores entre *0* y *10*. Otra vez la relación es una función logística. Pero esa relación tiene que ser de tal forma que cuando σ_{ij} tome el valor de la media del centro en *s* entonces *x* tome el valor de la media del expediente.

Esto quiere decir que si Δ_j es el parámetro de localización de la función, la relación tiene que ser tal que

$\beta_j = f(\alpha_j\ (\bar{\sigma}_j - \Delta_j))$.

Así pues, como

$$\beta_j = a + \frac{b-a}{1 + \exp(-\alpha_j((\gamma_0 + \delta_{0j}) - \Delta_j))}$$

entonces

$$\Delta_j = \frac{\ln\left[\frac{b-a}{\beta_j - a} - 1\right] + \alpha_j(\gamma_0 + \delta_{0j})}{\alpha_j}$$

Así la fórmula para x_{ij} será

$$x_{ij} = f(a + \frac{b-a}{1 + \exp(-\alpha_j(\sigma_{ij} - \Delta_j))}, \omega_{ij})$$

donde Δ_j toma el valor indicado. ω_{ij} es la variable aleatoria correspondiente al residuo, cuya generación será explicada en el siguiente apartado.

En esta fórmula α_j representa la mayor o menor pendiente de la función (de hecho es proporcional a la pendiente) y por tanto la mayor o menor relación entre σ y *x*. Dado que α_j es proporcional a la tangente de la función en el punto de inflexión, puede tomar valores entre menos y más infinito. Pero vamos a considerar en este caso que sólo son aceptables valores positivos.

Este parámetro, como beta, tiene una distribución aleatoria. Su valor determina la varianza sistemática en *x* dentro del centro *j*. A mayor valor de alfa, valores más extremos, y por tanto mayor varianza. A valores más bajos, valores más homogéneos, y por tanto menos varianza sistemática.

Así que se trata de una variable aleatoria que toma valores entre cero e infinito, y con una media de 1. Si además tenemos en cuenta que está relacionada con la varianza de *x*, es fácil aceptar que tiene una distribución χ^2 con un grado de libertad.

$$\alpha_j \sim \chi^2(v = 1)$$

α_j afecta también a la fiabilidad. En efecto, valores muy grandes de este parámetro determinan valores muy extremos de *x*, y dado que la variable está acotada, la varianza máxima también lo está. De esta forma la varianza de error será menor. Y por el contrario, valores bajos de α_j dan valores bajos de la varianza, con lo que la varianza de error será mayor.

Generación de los residuos ω_{ij}

Por último también, y por idénticas razones que las variables anteriores, w_{ij} tiene una distribución Beta con límites *b* y *a* iguales a *10* y *5* respectivamente.

En este caso la media de la distribución Beta vendrá dada por la parte fija generada en los puntos anteriores, y que llamaremos x'_{ij}.

$$x'_{ij} = a + \frac{b-a}{1+\exp(-\alpha_j(\sigma_{ij} - \Delta_j))}$$

y $\quad x_{ij} \sim a + bB(v_\omega, w_\omega)$

Los valores de los parámetros son

$$v_\omega = \frac{k_\omega(x'_{ij} - a)}{b-a} \quad \text{y}$$

$$k_\omega = \frac{(b-a)(x'_{ij} - a)(1 - \frac{x'_{ij}-a}{b-a})}{s^2_{x\text{intra}} - 1}$$

con $\quad w_\omega = k_\omega - v_\omega, b = 10 \text{ y } a = 5$

En este caso el valor de la varianza intra de *x* ha sido determinado empíricamente (a partir de los datos disponibles de las universidades de Burgos, Salamanca y Extremadura) en 0.75.

Se trata, como hemos podido comprobar, de una estructura recursiva bastante compleja, en la que las relaciones no lineales de un nivel paramétrico se anidan en las relaciones no lineales del siguiente nivel paramétrico.

Programa

El método descrito fue programado en Visual Basic VBA para EXCEL.

Condiciones experimentales

Por la naturaleza del método se entiende fácilmente que el mayor o menor efecto que puede tener la equiparación en el establecimiento de una escala común, y en cómo esa escala común hace aumentar el número de coincidencias o mejorar las correlaciones entre las notas de Bachillerato y las de selectividad, depende de al menos tres variables importantes.

M_b: La primera es la medida en que las medias de los centros en el expediente de Bachillerato difieren entre sí. Se definió una variable llamada M_b que multiplica al numerador de la expresión que define el valor del parámetro ***k***. M_b hace que ***k***

sea más grande o más pequeño. Cuanto más grande k, más pequeña la varianza y viceversa. Si la varianza es mayor, los centros tienen mayor dispersión respecto a la escala original, y la correlación entre s y x será mayor, y viceversa.

De hecho, cuanto más parecidas entre sí sean esas medias, mayor será la correlación entre la nota de selectividad y la del Bachillerato.

$\mathbf{M_x}$: La segunda es la fiabilidad interna de la nota de Bachillerato, es decir, la relación entre esta puntuación y la puntuación latente a partir de la cual se ha generado. Así, ***céteris páribus,*** a mayor fiabilidad mayor será la correlación entre la nota de selectividad y la del Bachillerato. Se definió por tanto una variable, M_x, que está multiplicando al numerador de la expresión del parámetro k de la distribución de w. M_x hace que dentro de los centros la varianza sea mayor o menor, y por tanto la fiabilidad será menor o mayor, haciendo que la correlación entre s y x sea menor o mayor.

$\mathbf{G_{00}}$: En tercer y último lugar el valor de la media general de la puntuación latente. Dado que las variables están acotadas, valores de este parámetro más próximos a las cotas harán que haya menos influencia del efecto principal y más de los términos de error, con lo que las coincidencias y las correlaciones entre la nota de selectividad y la del bachillerato serán menores.

Método: La variable 'Método', con los 4 niveles que se han determinado anteriormente se conforma como la variable experimental principal.

Valores de los condiciones experimentales y efectos teóricos sobre los parámetros

Valores de M_b:	Efecto sobre k	Efecto sobre la varianza de $Beta_j$	Efecto sobre la correlación de s y x.
0.25	Menor	Mayor	Menor
0.5	Medio	Medio	Medio
1	Mayor	Menor	Mayor

Valores de M_x:	Efecto sobre k	Efecto sobre la varianza de $Omega_{ij}$	Efecto sobre la correlación de s y x.
1	Menor	Mayor	Menor
2	Mayor	Menor	Mayor

Valores de G_{00}:	Efecto sobre la correlación de s y x.
1	Menor
2	Mayor

Con todo lo anterior se obtiene un diseño experimental de medidas repetidas de tres factores interunidades de 2x2x3 casillas con cuatro medidas en cada casilla.

M_x=2	M_b=0.25	M_b=0.5	M_b=1
G_{00}=5.5	Muestras 1 11	Muestras 1 12	Muestras 1 13
G_{00}=7	Muestras 1 21	Muestras 1 22	Muestras1 23

M_x=1	M_b=0.25	M_b=0.5	M_b=1
G_{00}=5.5	Muestras 2 11	Muestras 2 12	Muestras 2 13
G_{00}=7	Muestras 2 21	Muestras 2 22	Muestras2 23

Muestras

Para cada una de las casillas anteriores se replicaron 10 muestras distintas, cada una de ellas con 70 centros simulados, cada centro con un número aleatorio de alumnos, extraído de una distribución uniforme entre 1 y 100.

En consecuencia se generaron 120 muestras con 70 centros cada una, y un número de unidades de alrededor de 3.500 por muestra.

Equiparación de las puntuaciones de Bachillerato con los datos simulados

El método clásico implicaba el simple cálculo de la puntuación final ponderada.

El método IMS exigía obtener en primer lugar la media y la desviación típica del expediente y de la selectividad de cada centro. Seguidamente se realizaba la transformación descrita en un punto anterior.

La obtención de los predictores para los métodos OLS y Multinivel exige la estimación de los parámetros de los correspondientes modelos.

En el caso del método OLS se procedió a calcular para cada centro una ecuación de regresión. A partir de los parámetros de esa ecuación de regresión se obtuvo para cada unidad de nivel *1* la predicción correspondiente.

En el caso del método Multinivel hubo que ajustar primero el modelo general y obtener las predicciones correspondientes.

Todos estos cálculos se realizaron programando varias macros en el programa MLWIN, que fue el que se utilizó para calcular todas las puntuaciones equiparadas.

ESTUDIO DE PLAUSIBILIDAD

En el análisis de adecuación a los datos reales se procede de manera similar, excepto que en este caso de cada sujeto no conocemos su puntuación latente. Pero se obtiene para cada sujeto la puntuación equiparada con cada uno de los tres métodos alternativos, además del control denominado «Método Clásico».

Seguidamente se calcula la correlación entre la nota de selectividad de los sujetos con las puntuaciones equiparadas con cada método y con la puntuación de control.

Por medio de un análisis de varianza se determina su hay diferencias significativas entre las correlaciones obtenidas por los distintos métodos.

Muestras

Para el estudio de plausibilidad se obtuvieron datos de tres universidades españolas: la Universidad de Burgos, la Universidad de Salamanca y la Universidad de Extremadura.

En cada una de estas universidades se obtuvieron datos de los últimos cinco años, de un número suficientemente grande de centros públicos y privados, y de la convocatoria de junio y de la septiembre.

En todos los casos había datos pertenecientes a todas las opciones.

En las universidades de Burgos y Salamanca los datos correspondían exclusivamente a la opción de COU.

En la Universidad de Extremadura se dispuso de datos tanto de alumnos procedentes de COU como de alumnos procedentes del bachillerato LOGSE.

En las tablas I a IV que aparecen en el Anexo podemos ver el número de centros y alumnos por opciones y por convocatorias en las tres universidades.

La Universidad de Salamanca es de las tres la de mayor número de alumnos. En el curso 2000-01 contaba con 32.880 alumnos, frente a 28.186 de la Universidad de Extremadura y 10.318 de la Universidad de Burgos. En total se tomaron datos correspondientes a 57.465 alumnos.

Diseño del estudio de plausibilidad

Análisis preliminares dejaron ver que el comportamiento de los datos de las convocatorias de junio y septiembre eran claramente distintos. Como puede verse en la distribución de los datos de las tres universidades, el número de alumnos que se presentan en septiembre es siempre mucho menor, y no es necesario insistir mucho en que se trata de alumnos que no pertenecen a la misma población estadística que los alumnos de junio.

Otro tanto ocurre con las opciones. Dado que se trata de exámenes distintos para cada una de las opciones, y que las notas del expediente se refieren a conjuntos de asignaturas diferentes para cada una de las opciones, tampoco parecía conveniente combinar esos datos.

Por este motivo el estudio de equiparación se realizó de manera independiente con las muestras definidas por cada una de las opciones, en cada una de las convocatorias, de cada una de las universidades, excepto en la Universidad de Extremadura en que se diferenció además entre alumnos de COU y alumnos de bachillerato LOGSE.

Esto hace un total de 38 muestras distintas, cada una de ellas incluyendo los datos de al menos 5 años de selectividad, y con el número de centros y de alumnos que se especifica en las tablas I a IV, hasta alcanzar un total de 57.465 alumnos correspondientes a al menos cinco años en las convocatorias de junio y septiembre en tres universidades.

En cada una de las muestras se ajustó un modelo multinivel que era una versión multinivel del modelo OLS del método denominado OLS. En cada una de ellas se eliminaron los parámetros que no resultaban significativos, hasta obtener un modelo satisfactorio. Obtenido éste, se procedió a calcular los valores de los predictores para el modelo OLS, en este caso ajustado al nivel dos, el de los años anidados en los centros, y el predictor correspondiente al método multinivel. Seguidamente se obtuvieron los valores del método clásico, ponderando las notas de bachiller y selectividad por 0.6 y 0.40 respectivamente. A continuación, para cada año dentro de cada centro se obtuvieron los valores de las medias y desviaciones típicas de las notas de bachiller y de selectividad, y con ellos se calculó el valor de la equiparación por el método IMS.

Con los datos reales no se podía calcular la coincidencia entre las selecciones por la puntuación latente y cada una de las equiparaciones, puesto que la primera es desconocida. En su lugar se obtuvo la correlación de la puntuación equiparada de cada método con la nota de selectividad. El objetivo era comprobar si el aumento en la correlación en los métodos alternativos respecto del clásico correspondía a lo obtenido en la simulación.

Para ello, con las correlaciones calculadas en cada muestra se formó un archivo de datos que fue analizado utilizando SPSS, con dos diseños distintos. Dado que a diferencia de la simulación teníamos sólo un valor por cada casilla, se llevaron a cabo dos análisis distintos.

El primer análisis correspondía a un diseño en el que las variables independientes eran «Método», «Opción» y «Convocatoria», por tanto con 4x4x2 casillas.

En el segundo análisis se sustituyó la variable «Opción» por la variable «Universidad», conformando un diseño de 4x2x4 casillas.

RESULTADOS

ANÁLISIS DE DATOS

ESTUDIO DE VIABILIDAD

Análisis de las coincidencias

En el primer caso tenemos un diseño de medidas repetidas, con tres factores intersujetos con 2x2x3 casillas, y un factor intrasujetos con cuatro niveles, que es el método de equiparación.

En la tabla V tenemos los contrastes multivariados correspondientes, realizados con el estadístico Lambda de Wilks. Este estadístico está relacionado con la función de verosimilitud, y toma valores entre *0* y *1*. Valores próximos a *0* implican que las medias son distintas, mientras que valores próximos a *1* suponen que las medias son iguales.

En esta tabla tenemos el contraste de los efectos principales y las interacciones, junto con los valores de F asociados, los correspondientes grados de libertad y la probabilidad del contraste.

En este caso vemos que los efectos principales de la variable «Método» son estadísticamente significativos, además de la interacción con «Mx», «G0» y «Mb».

En la tabla VI tenemos el contraste univariado de los efectos intrasujetos con la corrección correspondiente al límite inferior[4]. Vemos que en ese caso hay diferencias significativas asociadas a los efectos principales de la variable «Método» (etiquetado como «COINCIDE»), su interacción con «Mx», y con «Mb».

En la tabla VII aparece el contraste de los efectos intersujetos. En este caso vemos que aparecen como significativos los efectos correspondientes a las variables «Mx» y «G0».

En la tabla VIII tenemos los valores de las medias marginales de las coincidencias por cada uno de los métodos, junto con su error típico y los límites del intervalo de confianza del 95%.

Podemos comprobar que los métodos OLS y Multinivel tienen los valores medios de coincidencias más altos, seguidos del método IMS, y a mayor distancia del método clásico, es decir, de la ausencia de equiparación.

Es muy importante destacar, no sólo la ventaja de las puntuaciones

(4) Se realizó el test de Mauchly y se rechazó la hipótesis nula. Esto supone que debemos modificar los grados de libertad. En nuestro caso hemos seleccionado la corrección correspondiente al límite inferior, por tratarse de la más conservadora.

transformadas, sino el valor tan bajo de coincidencias que en cualquiera de los métodos se produce, y especialmente en el método clásico. Efectivamente, en este caso sólo el 32% de los alumnos que debieran entrar en el centro elegido tienen la posibilidad de hacerlo. Hay por tanto un 68% de alumnos que no debieran haber sido seleccionados y, lo que es peor, un número absoluto igual de alumnos que debiendo haber ingresado no lo han hecho.

Se trata desde luego de datos simulados, y como veremos en el análisis de las correlaciones, en algunas universidades la congruencia entre las notas de selectividad y las de Bachillerato es algo mayor, lo que supone mayor fiabilidad de todo el proceso. Sin embargo los valores utilizados son perfectamente compatibles con los datos de muchas universidades. En esos casos sería necesario replantear muy seriamente todo el proceso.

No debe olvidarse que las coincidencias aumentan cuando aumenta la fiabilidad, como veremos seguidamente. Si esto es así, el aumento de la fiabilidad de las pruebas de acceso es no sólo una exigencia técnica, sino un requisito de estricta justicia para muchos alumnos.

En la tabla IX tenemos los contrastes posteriores para los 4 niveles de la variable experimental. Podemos ver que la diferencia más pequeña se da entre los métodos Multinivel y OLS, con 1,317 puntos a favor del segundo, mientras que la mayor diferencia se da entre el método Clásico, que es realmente el control, y el método OLS, con 12,458 puntos de diferencia media.

Estos resultados son muy importantes. Si utilizamos el método OLS o el método Multinivel, tendremos 12 alumnos más de los que tenían que estar, pero también 12 menos de los que no tenían que estar. En total son 24 alumnos implicados en la modificación del acceso. Y esto en un solo centro universitario. Podemos darnos cuenta de que de hecho supone un cambio muy importante en la configuración de los grupos de alumnos de toda una universidad.

Las ilustraciones I a VI nos presentan las medias marginales para los efectos principales y las interacciones significativos. En la I se ve cómo los valores mayores aparecen asociados a los dos métodos mencionados, Multinivel y OLS. Les sigue IMS, y por último el control o método Clásico.

El análisis de la ilustración V nos permite comprobar cómo se produce la interacción correspondiente. Tanto en los métodos Clásico como IMS, cuando «Mx» toma el valor 1, es decir, cuando la fiabilidad es menor, las coincidencias de esos métodos son todavía menores, mientras que las diferencias entre los otros dos métodos son constantes. Este resultado es importante. Cuando disminuye la fiabilidad en las pruebas de acceso de un universidad, tanto en el método Multinivel como en el método OLS la disminución del porcentaje de coincidencias no es mayor que el debido al efecto principal de la fiabilidad. Se trata según esto de dos métodos no sólo más eficaces, sino también más robustos ante variaciones de las demás condiciones experimentales.

La ilustración VI nos proporciona una aclaración muy interesante. Vemos que cuando la media general es mayor, se produce una disminución de la eficacia de los métodos IMS y OLS, mientras que la disminución en el método Multinivel es menor. Resulta así que en las universidades donde la nota media es más alta, prácticamente desaparecen las diferencias entre los métodos Multinivel y OLS. Este dato tiene suma importancia puesto que de hecho la media de la nota de selectividad y de Bachillerato de algunas universidades españolas están más próximas al valor 7 que al 5,5. Esto unido al hecho de que el método Multinivel es más estable en los casos en los que los centros tienen grandes variaciones en el número de sujetos parece aconsejar el uso de dicho método cuando se den tales circunstancias.

Análisis de las correlaciones

Comprobar qué ocurre con las correlaciones entre las notas de selectividad y las notas de Bachillerato con los datos simulados nos permite interpretar los resultados obtenidos en el análisis de los datos reales de las universidades analizadas.

El diseño es el mismo que en el caso del análisis de las coincidencias.

En la tabla X tenemos los contrastes multivariados con el estadístico Lambda de Wilks. En la tabla X el efecto etiquetado como «Correlac» corresponde a la variable «Métodos», con los cuatro niveles correspondientes. Podemos comprobar que en el contraste multivariado son significativos los efectos del factor intrasujetos, y todas sus interacciones binarias con los factores entre sujetos.

En la tabla XI tenemos que los efectos principales de los métodos son significativos, mientras que sólo la interacción entre el método de equiparación y «Mb» tiene efectos estadísticamente significativos[5].

En la tabla XII podemos ver que en el diseño intersujetos tanto los efectos principales de la variable «G0» como los de la variable «Mx» resultan estadísticamente significativos.

En la tabla XIII vemos los valores medios de las correlaciones por cada uno de los métodos de equiparación. Como era de esperar, los métodos se ordenan del mismo modo que con las coincidencias, aunque las diferencias son algo menores.

Cuando analizamos las diferencias en los contrastes posteriores, tabla XIV, de nuevo ocurre que el método de control o Clásico queda significativamente por debajo de los otros tres. La más pequeña diferencia se da entre el método OLS y Multinivel, y la mayor entre el método de control y el método OLS.

En las ilustraciones VII a XIV podemos observar que la configuración de las medias propias de cada método es muy parecida a la observada para la variable dependiente Coincidencias.

ESTUDIO DE PLAUSIBILIDAD

Análisis de las correlaciones

Como se ha indicado en un punto anterior, en el estudio de plausibilidad se han contrastado dos diseños distintos. El primero de ellos es un diseño factorial con tres factores intersujetos, con 4, 4 y 2 niveles respectivamente. Se trata de las variables «Método», con los mismos niveles anteriores, es decir, «Clásico», «Multinivel», «OLS» e «IMS»; el factor «Opción», con los niveles «Ciencias y Tecnología», «Ciencias Biosanitarias», «Ciencias sociales» y «Humanidades»; y «Convocatoria», con los niveles «Junio» y «Septiembre».

En la tabla XV vemos que resultan estadísticamente significativos los efectos asociados a la variable «Método», a la variable «convocatoria», y a la interacción de ambos factores.

En la tabla XVI tenemos los contrastes posteriores entre los niveles del primer factor. Existen diferencias significativas del método Clásico con los otros tres. El método Multinivel presenta diferencias con el método «Clásico», pero no con los otros dos. El método OLS se diferencia

(5) Por los mismos motivos que se informaban en la nota 3, se modificaron los grados de libertad en este contraste.

significativamente del método Clásico y de IMS, pero no de Multinivel. Por último el método IMS no se diferencia de Multinivel pero sí de Clásico, al que supera, y de OLS, que le supera.

Las ilustraciones XVI a XVII nos permiten comprobar que las diferencias entre los métodos son muy similares a los valores que hemos obtenido en la simulación. Cabe destacar especialmente la gran diferencia entre las convocatorias de junio y septiembre. Así, en la ilustración XVII comprobamos cómo en la convocatoria de septiembre el método clásico produce una correlación media que no supera el valor de 0,35, mientras que llega a 0,70 de valor medio en la convocatoria de junio. En ésta última la diferencia entre el valor medio de la correlación que reproduce el método clásico, 0,7, y el método OLS es de 0,12, que supone, para esos valores de correlación, un incremento muy sustantivo. Pero la diferencia es todavía más llamativa en la convocatoria de septiembre, llegando a los 0,40 de diferencia.

El segundo diseño sustituye la variable «Opción» por la variable «Universidad». Este último factor consta de cuatro niveles, puesto que los datos de los alumnos que han accedido a las pruebas de acceso desde el COU o desde el bachillerato LOGSE han sido diferenciados.

En la tabla XVII vemos que resultan estadísticamente significativos los efectos asociados a los factores «Método», «Convocatoria», «Universidad», y las interacciones entre «Método» y «Convocatoria», «Método» y «Universidad» y «Convocatoria» por «Universidad».

Al realizar las comparaciones múltiples entre los niveles del principal factor, vemos que la introducción del factor 'Universidad' la disminución de la correspondiente suma cuadrática hace que sean estadísticamente significativas todas las diferencias (tabla XVIII).

La ilustración XVIII nos permite comprobar que las diferencias observadas se deben fundamentalmente a que la Universidad de Burgos presenta una correlación media entre las notas Bachillerato y las de selectividad excepcionalmente alta. Esto puede deberse muy probablemente a que se trata de una Universidad muy joven, pequeña, y con un ámbito territorial muy reducido, lo que asegura una mayor homogeneidad entre los centros que presentan alumnos a la selectividad.

La ilustración XIX corrobora lo anteriormente dicho respecto a los métodos y las convocatorias.

Resulta muy interesante la inspección de la ilustración XX, ya que nos permite comprobar que incluso en una universidad con una alta correlación entre las variables estudiadas, los métodos de equiparación producen un aumento muy considerable en la correlación media. Efectivamente, podemos comprobar que, universidad por universidad, la misma ordenación por eficacia de los métodos se mantiene.

En la ilustración XXI vemos que la diferencia entre las convocatorias de junio y septiembre es una constante en todas las universidades, auque, probablemente por las razones que ya hemos apuntado, esa diferencia es menor en la de Burgos.

CONCLUSIONES

Basándonos en el estudio de simulación que hemos diseñado, hemos demostrado la viabilidad de la utilización de métodos de equiparación basados en la utilización de un examen común a todos los alumnos, en este caso la Prueba de Selectividad, como elemento de anclaje para situar en la misma escala las puntuaciones de los expedientes de Bachillerato procedentes de distintos centros.

Hemos comprobado que los métodos basados en la regresión, OLS y Multinivel, producen mejores resultados cualquiera que sea la variable dependiente utilizada, tanto «Coincidencias» como «Correlaciones».

El método IMS, aunque menos eficaz que los anteriores, resulta sin embargo también mejor que la práctica habitual, denominada Clásica.

En el caso en que las notas medias son altas, las diferencias entre el método OLS y el método Multinivel se reducen hasta hacerse no significativas. Dadas las características del segundo método, conviene utilizarlo cuando se dan las circunstancias de notas medias altas y dispar número de sujetos por centro.

Con los datos reales hemos comprobado la plausibilidad de los supuestos utilizados para generar los datos, y hemos comprobado que se obtienen resultados muy similares a los producidos con los datos simulados.

Con los datos reales hemos comprobado que también resultan superiores los métodos OLS y Multinivel.

Incluso en los casos en los que la correlación entre la nota de Bachillerato y la de selectividad es un valor muy alto, como ocurre con la Universidad de Burgos, cualquiera de los métodos de equiparación es capaz de elevar la correlación obtenida.

Hemos comprobado que la fiabilidad afecta mucho al grado de coincidencia entre los alumnos que deberían ser seleccionados y los que de hecho lo son. Este dato, junto con las demás conclusiones a que nos ha llevado esta investigación, implica una seria exigencia de mayor esfuerzo en la investigación de estos procesos, siempre con el fin de mejorar el proceso de selección para lograr resultados más precisos y, lo que es consecuencia de ello, más justos y equitativos para todos los alumnos, tal como nuestras leyes y nuestro sentido común nos imponen.

Como resultado de la investigación se comprueba que el método propuesto es viable, plausible, y altamente recomendable, puesto que con los datos simulados aumenta en más de un 50% la precisión y eficacia de la selección, y con los datos reales de las universidades se comprueba que aumenta la correlación usada como criterio también en más de un 50% respecto del método utilizado hasta la fecha. Eso supone que es técnicamente posible mejorar la justicia y equidad de los procedimientos de acceso a la universidad.

Los resultados de esta investigación pueden ser muy importantes para orientar los procedimientos que las universidades y comunidades autónomas adopten en el próximo futuro para determinar el acceso a los centros universitarios que más demanda presentan.

Las universidades y comunidades autónomas deben plantearse de forma inmediata el procedimiento por el que se regulará el acceso a los distintos centros universitarios. Hemos demostrado que es erróneo considerar que las notas del expediente de Bachillerato por sí solas sean criterio suficiente para regular el acceso a los estudios universitarios. El proceso de equiparación de las puntuaciones de Bachillerato que aquí se ha propuesto aumenta muy significativamente la justicia de la selección. Este es probablemente el mejor momento para plantearse la mejora de las condiciones técnicas en que la selección se organiza, garantizando de este modo una mayor equidad del proceso.

BIBLIOGRAFÍA

AGUIRRE DE CÁRCER, I: *Validez concurrente de las calificaciones otorgadas en el COU*. Madrid, CIDE, Memoria de investigación inédita, 1986.

AGUIRRE DE CÁRCER, I. y otros: *Las pruebas de selectividad en la Universidad Autónoma de Madrid*. Madrid, CIDE, Memoria de Investigación inédita, 1984.

ESCUDERO ESCORZA, T.: *Seguimiento a la selectividad universitaria*. Zaragoza, ICE de la Universidad, 1987.

GONZÁLEZ Y VALLE, G.: «El acceso a la universidad en la C.E», en *Actas de las jornadas: la investigación educativa sobre la universidad,* Madrid, CIDE, (1990), pp. 159-177.

LORD, F. M.: *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, N. J., Lawrence Erlbaum, 1980.

MIGUEL, M. DE: *Las calificaciones en las pruebas de acceso en la Universidad de Oviedo*. Oviedo, Universidad de Oviedo. 1988. Inédito. Citado en Muñoz-Repiso (1997).

MIGUEL, M. DE: «Cambios generacionales y acceso a la enseñanza superior», en *Actas de las jornadas: La investigación educativa sobre la Universidad*, Madrid, CIDE. 1990

— *El acceso a los estudios universitarios. Análisis y seguimiento de la demanda en Asturias*. Madrid, CIDE, 1993.

MORRIS, C. N.: «On the foundations of Test Equating». En Paul W. Holland y Donald B. Rubin. *Test Equating*. Academic Press, New York, 1982.

MUÑOZ-REPISO, M. y otros: *Las calificaciones en las pruebas de aptitud para el acceso a la Universidad*. Madrid, MEC, CIDE. 1991.

MUÑOZ-REPISO, M., y otros: *Las calificaciones en las pruebas de acceso a la Universidad*. Madrid, CIDE. Memoria de Investigación inédita, 1988.

SANZ PAZ, I.: *Análisis por asignaturas de las pruebas de acceso a la universidad*. Madrid, CIDE, Inédito. 1992.

ANEXO

TABLAS

TABLA I

Datos de alumnos y centros por opciones en la Universidad de Burgos

		A	B	C	D
JUNIO	Centros	27	27	27	27
	Alumnos	2.400	1.670	1.734	570
SEPTIEMBRE	Centros	27	27	27	27
	Alumnos	695	496	696	232

TABLA II

Datos de alumnos y centros por opciones en la Universidad de Salamanca

		A	B	C	D
JUNIO	Centros	48	48	48	48
	Alumnos	4.204	3.734	3.492	1.728
SEPTIEMBRE	Centros	48	47	48	46
	Alumnos	1.474	1.310	1.674	729

TABLA III

Datos de alumnos y centros por opciones en la Universidad de Extremadura

		A	B	C	D
JUNIO	Centros	65	67	63	63
	Alumnos	3.795	4.025	3.538	2.136
SEPTIEMBRE	Centros	63	64	56	62
	Alumnos	1.126	1.262	1.288	751

TABLA IV

Datos de alumnos y centros por opciones en la Universidad de Extremadura

		A	B	C	D
JUNIO	Centros	79	75	74	77
	Alumnos	1.894	2.835	2.101	2.913
SEPTIEMBRE	Centros	62	61	60	64
	Alumnos	451	789	568	975

TABLA V
Contrastes multivariados. Lambda de Wilks

Efecto	Valor	F	Gl de la hipótesis	Gl del error	Significación
COINCIDE	,127	243,461	3	106	,000
COINCIDE * MX	,644	19,492	3	106	,000
COINCIDE * G0	,869	5,314	3	106	,002
COINCIDE * MB	,828	3,502	6	212	,003
COINCIDE * MX * G0	,993	,255	3	106	,857
COINCIDE * MX * MB	,949	,935	6	212	,471
COINCIDE * G0 * MB	,956	,805	6	212	,567
COINCIDE * MX * G0 * MB	,952	,885	6	212	,507

c Diseño: Intercept+MX+G0+MB+MX * G0+MX * MB+G0 * MB+MX * G0 * MB Diseño intra sujetos: COINCIDE

TABLA VI
Pruebas de efectos intra-sujetos. Límite inferior

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
COINCIDE	11.787,123	1	11.787,123	248,984	,000
COINCIDE * MX	779,940	1	779,940	16,475	,000
COINCIDE * G0	333,123	1	333,123	7,037	,009
COINCIDE * MB	193,071	2	96,535	2,039	,135
COINCIDE * MX * G0	7,106	1	7,106	,150	,699
COINCIDE * MX * MB	59,729	2	29,865	,631	,534
COINCIDE * G0 * MB	69,171	2	34,585	,731	,484
COINCIDE * MX * G0 * MB	90,163	2	45,081	,952	,389
Error (COINCIDE)	5.112,825	108	47,341		

TABLA VII

Pruebas de los efectos inter-sujetos. Variable transformada: Promedio

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Intercept	743.006,719	1	743.006,719	7.504,925	,000
MX	2.655,502	1	2.655,502	26,823	,000
G0	535,519	1	535,519	5,409	,022
MB	142,212	2	71,106	,718	,490
MX * G0	26,602	1	26,602	,269	,605
MX * MB	89,554	2	44,777	,452	,637
G0 * MB	246,912	2	123,456	1,247	,291
MX * G0 * MB	33,454	2	16,727	,169	,845
Error	10.692,275	108	99,003		

TABLA VIII

Estimaciones

	Media	Error típ.	Intervalo de confianza al 95%.	
COINCIDE			Límite inferior	Límite superior
CLÁSICO	32,092	,616	30,872	33,312
MULTINIVEL	43,233	,580	42,085	44,382
OLS	44,550	,481	43,597	45,503
IMS	37,500	,523	36,463	38,537

TABLA IX

Comparaciones por pares

		Diferencia entre medias (I-J)	Error típ.	Significación	Intervalo de confianza al 95% para diferencia	
(I) COINCIDE	(J) COINCIDE				Límite inferior	Límite superior
CLÁSICO	MULTINIVEL	-11,142	,592	,000	-12,315	-9,968
	OLS	-12,458	,519	,000	-13,487	-11,429
	IMS	-5,408	,581	,000	-6,561	-4,256
MULTINIVEL	OLS	-1,317	,396	,001	-2,102	-,531
	IMS	5,733	,553	,000	4,637	6,830
OLS	IMS	7,050	,396	,000	6,264	7,836

Basadas en las medias marginales estimadas.

* La diferencia de las medias es significativa al nivel ,05.

a Ajuste para comparaciones múltiples: Diferencia menos significativa (equivalente a la ausencia de ajuste).

TABLA X

Contrastes multivariados. Lambda de Wilks

Efecto	Valor	F	Gl de la hipótesis	Gl del error	Significación
CORRELAC	,015	2.372,305	3	106,000	,000
CORRELAC * MX	,773	10,405	3	106,000	,000
CORRELAC * G0	,851	6,176	3	106,000	,001
CORRELAC * MB	,264	33,490	6	212,000	,000
CORRELAC * MX * G0	,945	2,070	3	106,000	,109
CORRELAC * MX * MB	,942	1,062	6	212,000	,386
CORRELAC * G0 * MB	,982	,314	6	212,000	,929
CORRELAC * MX * G0 * MB	,925	1,396	6	212,000	,217

a Estadístico exacto
c Diseño: Intercept+MX+G0+MB+MX * G0+MX * MB+G0 * MB+MX * G0 * MB Diseño intra sujetos: CORRELAC

TABLA XI

Pruebas de efectos intra-sujetos

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
CORRELAC	3,138	1	3,138	1.006,807	,000
CORRELAC * MX	1,156E-02	1	1,156E-02	3,708	,057
CORRELAC * G0	9,163E-03	1	9,163E-03	2,940	,089
CORRELAC * MB	,335	2	,167	53,705	,000
CORRELAC * MX * G0	5,039E-03	1	5,039E-03	1,617	,206
CORRELAC * MX * MB	6,485E-03	2	3,243E-03	1,040	,357
CORRELAC * G0 * MB	1,981E-03	2	9,903E-04	,318	,728
CORRELAC * MX * G0 * MB	1,615E-02	2	8,075E-03	2,591	,080
Error(CORRELAC)	,337	108	3,117E-03		

TABLA XII

Pruebas de los efectos inter-sujetos. Variable transformada: Promedio

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Intercept	32,568	1	32,568	29.133,087	,000
MX	2,710E-02	1	2,710E-02	24,240	,000
G0	2,560E-02	1	2,560E-02	22,902	,000
MB	3,347E-03	2	1,674E-03	1,497	,228
MX * G0	2,848E-03	1	2,848E-03	2,547	,113
MX * MB	8,456E-03	2	4,228E-03	3,782	,026
G0 * MB	3,780E-03	2	1,890E-03	1,691	,189
MX * G0 * MB	1,418E-03	2	7,092E-04	,634	,532
Error	,121	108	1,118E-03		

TABLA XIII

Correlación media según el método

	Media	Error típ.	Intervalo de confianza al 95%.	
MÉTODO			Límite inferior	Límite superior
1	,404	,004	,396	,412
2	,577	,005	,567	,587
3	,612	,002	,607	,617
4	,491	,004	,483	,498

TABLA XIV

Comparaciones por pares

		Diferencia entre medias (I-J)	Error típ.	Significación	Intervalo de confianza al 95% para diferencia	
(I) CORRELAC	(J) CORRELAC				Límite inferior	Límite superior
CLÁSICO	MULTINIVEL	-11,142	,592	,000	-12,315	-9,968
	OLS	-12,458	,519	,000	-13,487	-11,429
	IMS	-5,408	,581	,000	-6,561	-4,256
MULTINIVEL	OLS	-1,317	,396	,001	-2,102	-,531
	IMS	5,733	,553	,000	4,637	6,830
OLS	IMS	7,050	,396	,000	6,264	7,836

Basadas en las medias marginales estimadas.
* La diferencia de las medias es significativa al nivel ,05.
a Ajuste para comparaciones múltiples: Diferencia menos significativa (equivalente a la ausencia de ajuste).

TABLA XV

Pruebas de los efectos inter-sujetos. Variable dependiente: CORRELAC

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo corregido	3,223	31	,104	8,414	,000
Intersección	58,035	1	58,035	4.696,681	,000
MÉTODO	1,261	3	,420	34,027	,000
OPCIÓN	7,292E-02	3	2,431E-02	1,967	,124
CONVOCAT	1,482	1	1,482	119,898	,000
MÉTODO * OPCIÓN	3,284E-02	9	3,649E-03	,295	,974
MÉTODO * CONVOCAT	,336	3	,112	9,059	,000
OPCIÓN * CONVOCAT	1,344E-02	3	4,481E-03	,363	,780
MÉTODO * OPCIÓN * CONVOCAT	2,498E-02	9	2,775E-03	,225	,990
Error	1,186	96	1,236E-02		
Total	62,444	128			
Total corregida	4,409	127			

a R cuadrado = ,731 (R cuadrado corregida = ,644)
Comparaciones múltiples
Variable dependiente: CORRELAC

TABLA XVI

Contrastes posteriores. Prueba de Scheffé

		Diferencia entre medias (I-J)	Error típ.	Significación	Intervalo de confianza al 95%	
(I) MÉTODO	(J) MÉTODO				Límite inferior	Límite superior
Clásico	Multinivel	-,1985547	2,78E-02	,000	-,278	-,1194715
	OLS	-,2689909	2,78E-02	,000	-,348	-,1899077
	IMS	-,1299109	2,78E-02	,000	-,209	-5,0827739E-02
Multinivel	Clásico	,1985547	2,78E-02	,000	,119	,2776379
	OLS	-7,0436250E-02	2,78E-02	,100	-,149	8,646948E-03
	IMS	6,864375E-02	2,78E-02	,114	-1,044E-02	,1477269
OLS	Clásico	,2689909	2,78E-02	,000	,190	,3480741
	Multinivel	7,043625E-02	2,78E-02	,100	-8,647E-03	,1495194
	IMS	,1390800	2,78E-02	,000	5,999E-02	,2181632
IMS	Clásico	,1299109	2,78E-02	,000	5,083E-02	,2089941
	Multinivel	-6,8643750E-02	2,78E-02	,114	-,148	1,043945E-02
	OLS	-,1390800	2,78E-02	,000	-,218	-5,9996802E-02

Basado en las medias observadas.
* La diferencia de medias es significativa al nivel ,05.

TABLA XVII

Pruebas de los efectos inter-sujetos. Variable dependiente: CORRELAC

Fuente	Suma de cuadrados tipo III	gl	Media cuadrática	F	Significación
Modelo corregido	4,062	31	,131	36,230	,000
Intersección	58,035	1	58,035	16.046,589	,000
MÉTODO	1,261	3	,420	116,257	,000
CONVOCAT	1,482	1	1,482	409,641	,000
UNIVERSI	,700	3	,233	64,472	,000
MÉTODO * CONVOCAT	,336	3	,112	30,950	,000
MÉTODO * UNIVERSI	8,060E-02	9	8,955E-03	2,476	,014
CONVOCAT * UNIVERSI	,156	3	5,215E-02	14,419	,000
MÉTODO * CONVOCAT * UNIVERSI	4,666E-02	9	5,185E-03	1,434	,185
Error	,347	96	3,617E-03		
Total	62,444	128			
Total corregida	4,409	127			

a R cuadrado = ,921 (R cuadrado corregida = ,896)

TABLA XVIII

Comparaciones múltiples. Variable dependiente: CORRELAC. Scheffé

		Diferencia entre medias (I-J)	Error típ.	Significación	Intervalo de confianza al 95%	
(I) MÉTODO	(J) MÉTODO				Límite inferior	Límite superior
Clásico	Multinivel	-,1985547	1,50E-02	,000	-,2413394	-,1557700
	OLS	-,2689909	1,50E-02	,000	-,3117756	-,2262062
	IMS	-,1299109	1,50E-02	,000	-,1726956	-8,7126247E-02
Multinivel	Clásico	,1985547	1,50E-02	,000	,1557700	,2413394
	OLS	-7,0436250E-02	1,50E-02	,000	-,1132209	-2,7651559E-02
	IMS	6,864375E-02	1,50E-02	,000	2,585906E-02	,1114284
OLS	Clásico	,2689909	1,50E-02	,000	,2262062	,3117756
	Multinivel	7,043625E-02	1,50E-02	,000	2,765156E-02	,1132209
	IMS	,1390800	1,50E-02	,000	9,629531E-02	,1818647
IMS	Clásico	,1299109	1,50E-02	,000	8,712625E-02	,1726956
	Multinivel	-6,8643750E-02	1,50E-02	,000	-,1114284	-2,5859059E-02
	OLS	-,1390800	1,50E-02	,000	-,1818647	-9,6295309E-02

Basado en las medias observadas.

- La diferencia de medias es significativa al nivel ,05

ILUSTRACIONES

Ilustración I

Medias marg. de Coincidencia con Sigma

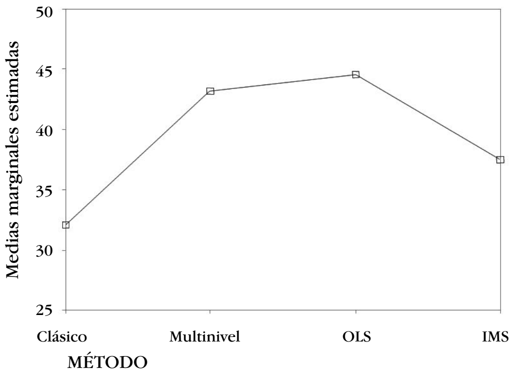


Ilustración IV

Medias marg. de Coincidencia con Sigma

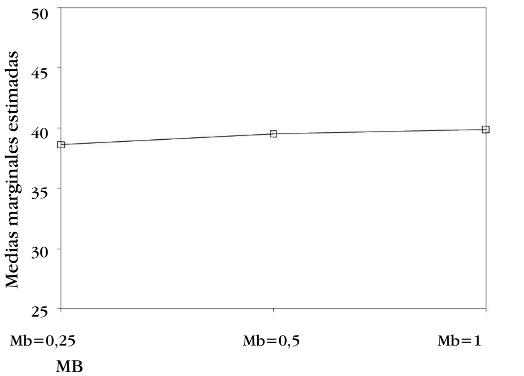


Ilustración II

Medias marginales estimadas de Coincidencia

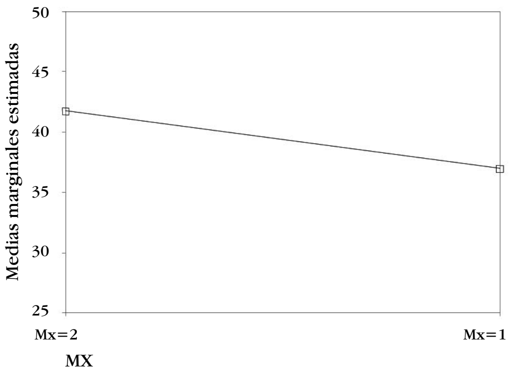


Ilustración V

Medias marg. de Coincidencia con Sigma

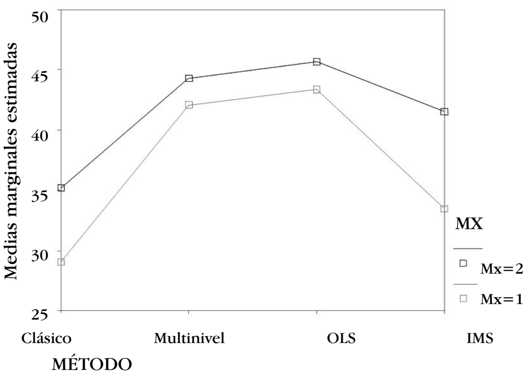


Ilustración III

Medias marg. de Coincidencia con Sigma

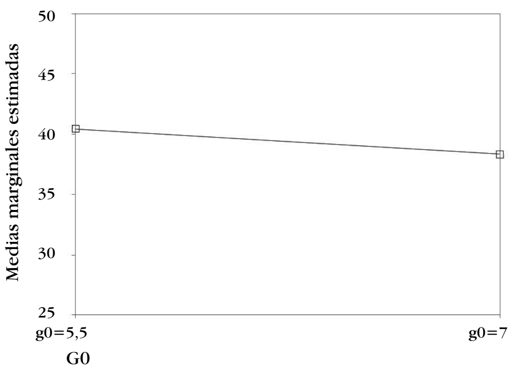


Ilustración VI

Medias marg. de Coincidencia con Sigma

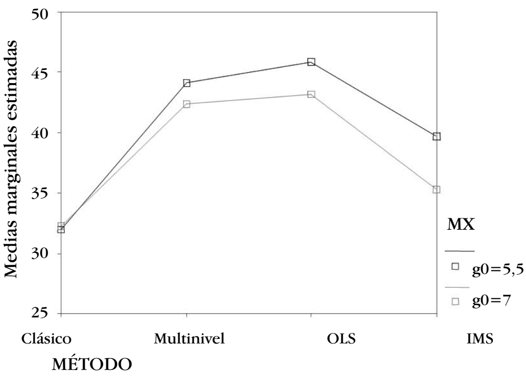


Ilustración VII

Medias marginales estimadas de CORR

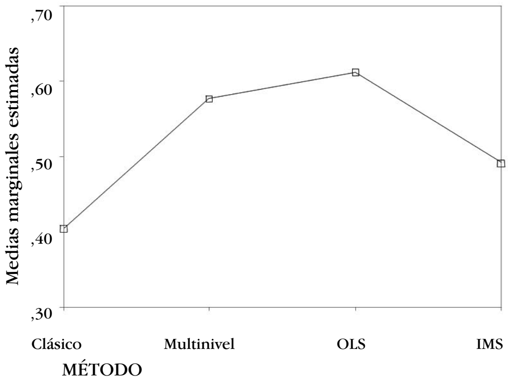


Ilustración X

Medias marginales estimadas de CORR

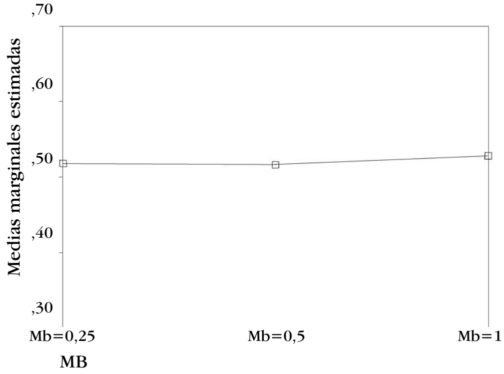


Ilustración VIII

Medias marginales estimadas de CORR

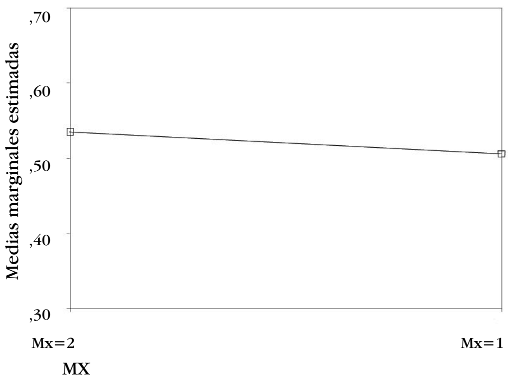


Ilustración XI

Medias marginales estimadas de CORR

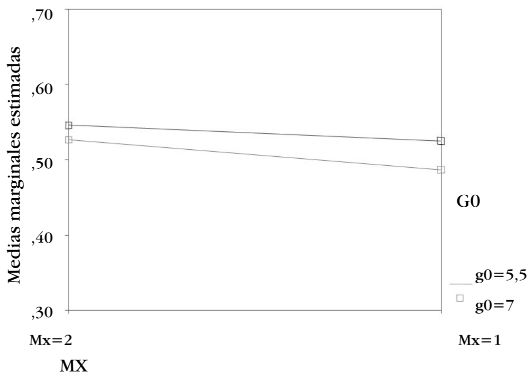


Ilustración IX

Medias marginales estimadas de CORR

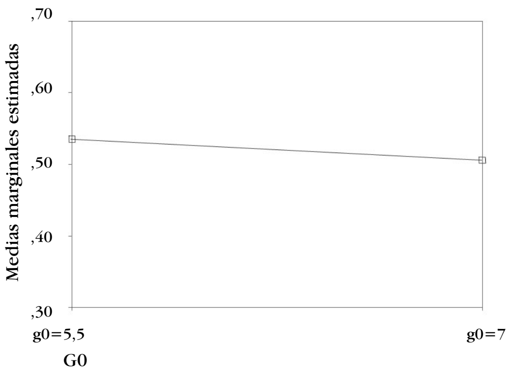


Ilustración XII

Medias marginales estimadas de CORR

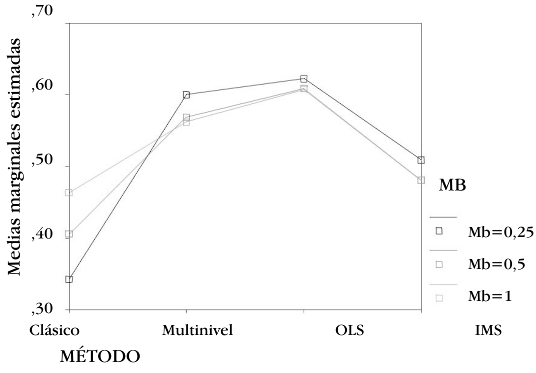


Ilustración XIII

Medias marginales estimadas de CORR

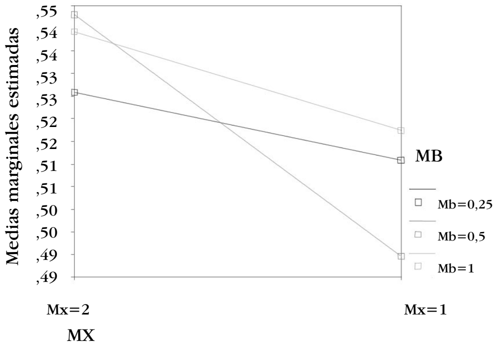


Ilustración XVI

Medias marg. estimadas de Correlación

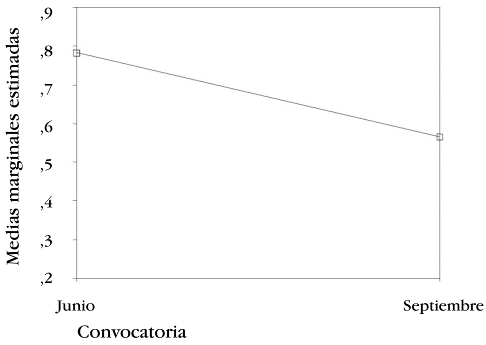


Ilustración XIV

Medias marginales estimadas de CORR

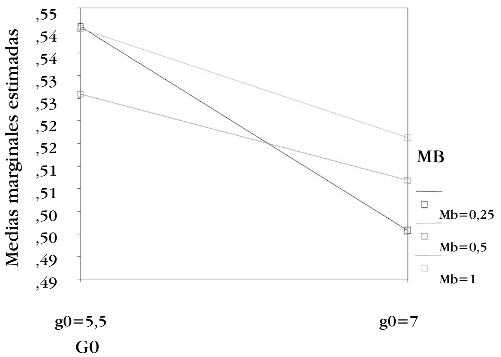


Ilustración XVII

Medias marg. estimadas de Correlación

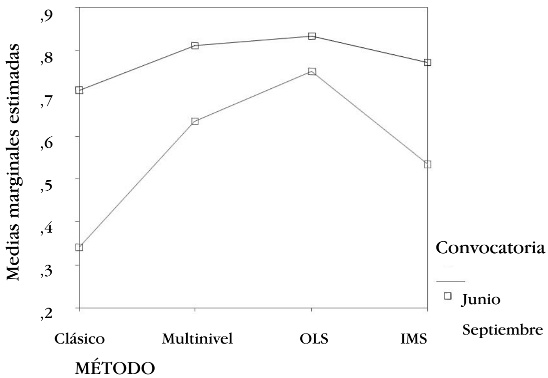


Ilustración XV

Medias marg. estimadas de Correlación

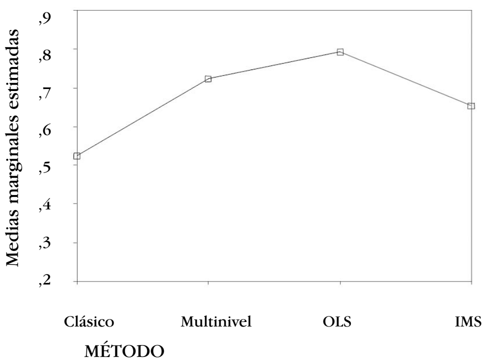


Ilustración XVIII

Medias marg. estimadas de Correlación

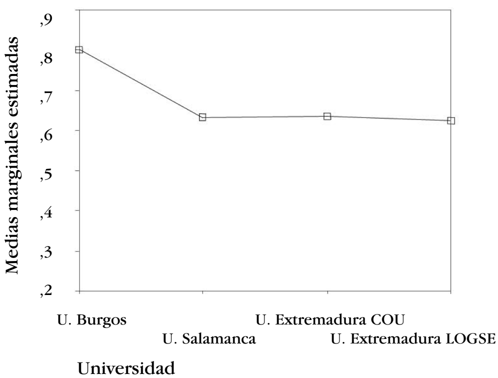


Ilustración XIX

Medias marginales estimadas de Correlación

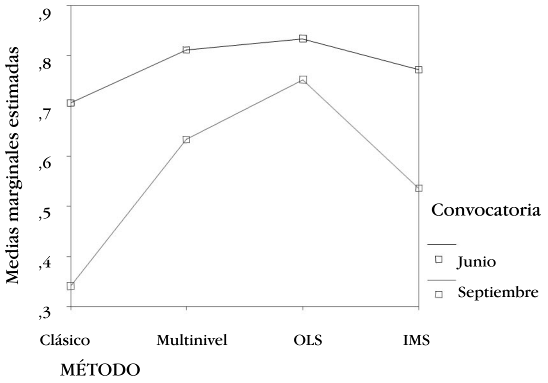


Ilustración XX

Medias marginales estimadas de Correlación

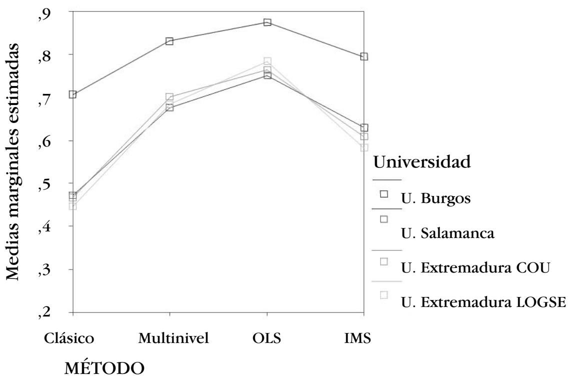
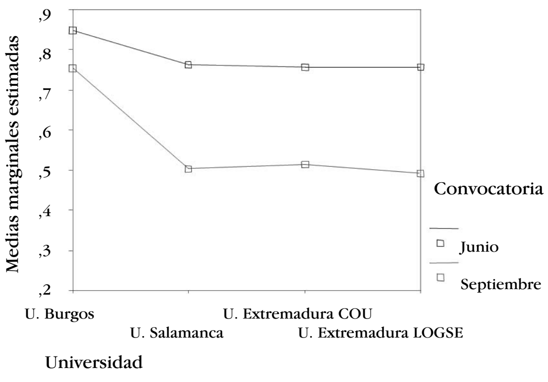


Ilustración XXI

Medias marginales estimadas de Correlación



APÉNDICE: PARÁMETROS DE LA DISTRIBUCIÓN BETA

Para determinar los parámetros de una distribución Beta que queremos utilizar para modelizar el comportamiento de una variable x que está acotada entre a y b, y de la que conocemos su media y su varianza, nos basamos en la relación entre medias y varianzas.

Para la media es muy sencillo:

$$\frac{v}{v+w}=\frac{\overline{x}-a}{b-a}$$

y haciendo

$k=v+w$, resulta

$$v=\frac{k(\overline{x}-a)}{b-a} \ ,$$

ya que del mismo modo que

$0\leq \mu_\beta \leq 1$, también

$$0\leq \frac{(\overline{x}-a)}{b-a}\leq 1 \ .$$

Para obtener el otro parámetro utilizamos la varianza.

$$\sigma^2_\beta=\frac{vw}{(v+w)^2(v+w+1)}$$

En este caso caben dos soluciones. Podemos poner en relación la varianza con la varianza máxima, o alternativamente con el recorrido de la variable.

En el primer caso tenemos en cuenta que $a \leq x \leq b$. Por eso su varianza máxima se producirá siempre que los valores se alejen lo máximo de la media. Es decir

$$\frac{n_a(a-\frac{a+b}{2})^2+n_b(b-\frac{a+b}{2})^2}{n_a+n_b}=$$

$$=\frac{n_a(a-\frac{a}{2}+\frac{b}{2})^2+n_b(b-\frac{a}{2}+\frac{b}{2})^2}{n_a+n_b}=$$

$$=\frac{1}{4}(b-a)^2$$

Y como ambas varianzas son proporcionales a sus máximos, tendremos que

$$\frac{\sigma^2_\beta}{\sigma^2_x}=\frac{0.25}{\frac{1}{4}(b-a)^2}$$

y por tanto $\quad \sigma^2_\beta=\dfrac{\sigma^2_x}{(b-a)^2}$

Si $\quad k=v+w,$

$$\frac{vw}{(v+w)^2(v+w+1)}=\frac{v(k-v)}{k^2(k+1)}=\frac{\sigma^2_x}{(b-a)^2}$$

y en consecuencia

$$\frac{\frac{k(\overline{x}-a)}{b-a}(k-\frac{k(\overline{x}-a)}{b-a})}{k^2(k+1)}=\frac{\sigma^2_x}{(b-a)^2}$$

$$\frac{\frac{(\overline{x}-a)}{b-a}(1-\frac{(\overline{x}-a)}{b-a})}{k+1}=\frac{\sigma^2_x}{(b-a)^2}$$

$$k=\frac{\frac{(\overline{x}-a)}{b-a}(1-\frac{(\overline{x}-a)}{b-a})(b-a)^2}{\sigma^2_x}-1$$

En el segundo caso la varianza es proporcional al cuadrado del recorrido, es decir, de $(b-a)^2$.