

Un Método de Evaluación Aplicable al Examen Escrito(*)

Eusebio Cleto del Rey (**)

-----0-----

1. Introducción

Si alguna originalidad tiene el presente trabajo, ella se encuentra en los / detalles. La idea central refleja sólo el resultado de nuestra larga experien-// cia, como estudiante sometido a exámenes escritos en la Universidad Nacional de Tucumán, la Universidad Nacional de Cuyo y The University of Chicago, y como profesor que tomó ese tipo de exámenes en la Universidad Nacional de Tucumán y la / Universidad Nacional de Salta.

La experiencia arriba mencionada tuvo lugar en las disciplinas Economía y Es tadística, pero estamos convenidos de que ella puede resultar de utilidad para / los docentes de otras áreas del conocimiento.

El método de valuación que describimos parte de la siguiente idea básica: De be calificarse a cada alumno de acuerdo a su situación relativa dentro del grupo de alumnos que rindieron el mismo examen. Creemos que tal principio rige a todo sistema de clasificación, pero el método que aquí consideramos lo toma explícita mente en cuenta. Esa idea está presente en lo que decimos en las Secciones 2 a 7 de este trabajo, o sea en el análisis de todo el proceso que va desde la prepara ción del enunciado hasta la asignación de las calificaciones.

La Secc. 8 tiene la intención de "vender" el método: esto es, de convencer / al lector de sus ventajas y de la conveniencia de su aplicación.

2. Preparación del Enunciado

Resulta conveniente tener especial cuidado en la elaboración del enunciado / del examen, cuando se va a aplicar el método que se describe más abajo. De ello depende no solo la ulterior facilidad en la asignación de puntajes a cada estu-// diante, sino también el grado de justicia de los resultados de la evaluación.

Debe tenerse en cuenta que la que hemos llamado "idea básica" supone que la suerte de cada alumno se decide mediante su comparación con los demás. El enun-// ciado debe ser pensando de modo que facilite, en lo posible, la posterior compa- ración de las respuestas obtenidas. No resultan apropiadas para ello las "pregun-// tas ensayo", o sea aquellas en las que se pide al examinando que escriba una lar- ga exposición sobre determinado tema. Así, resulta inconveniente pedir al alum-// no: "Desarrolle el Tema tal del Programa", o "Explique el proceso por el cual...", o "escriba la evolución de ...", etc. Deben utilizarse, mas bien, los ejerci-// cios prácticos, los "falso, cierto, incierto", las preguntas de elección múlti-// ple, el pedido de definiciones o conceptos cortos y precisos, etc.

Algunas de las técnicas que facilitan la comparación entre examinandos, o // que tienden a eliminar la subjetividad del examinador en la corrección, facili-// tan, a la vez, el copiado entre los estudiantes durante la prueba. Es necesario,

(*) Trabajo presentado a las III Jornadas Nacionales Universitarias de Docen-// tes de Contabilidad General, que se realizaron en nuestra Universidad entre el 7 / 24 y el 26 de Septiembre de 1981.

El autor agradece al Cr. Mario Guido García la sugerencia de realizar es-// te trabajo y sus valiosos comentarios al borrador del mismo.

(**) Profesor Titular de Economía, Dto. de Cs. Económicas, Jurídicas y Socia-// les, UNSa.

en consecuencia, tomar precauciones especiales en tales casos. Como ejemplos de preguntas que pueden presentar este problema mencionaremos las de elección múltiple, y los "falso, cierto, incierto" (cuando no se exige explicación para la respuesta).

Dos defectos que es deseable erradicar de los enunciados son: La repetición y el "encadenamiento" de las preguntas y ejercicios.

El material utilizado en cada examen debe ser creado especialmente para él. Los ejercicios o preguntas tomados de exámenes anteriores o, peor aún, del material bibliográfico disponible benefician a quienes, por casualidad, vieron antes esos ejercicios o preguntas, y, por comparación, perjudican a los otros estudiantes. Lo que no resulta fácil evitar es la utilización de material parecido al de enunciados de exámenes anteriores de la misma asignatura. Ello, por otra parte, puede resultar beneficioso para la formación de los alumnos, pues los incentiva a ejercitarse con el material preexistente. El equilibrio entre originalidad e incentivos a resolver exámenes viejos es una cuestión de arte de preparar enunciados, y sólo depende de la habilidad y experiencia de cada docente.

Si el haber respondido bien a una pregunta o ejercicio anterior es condición necesaria para que el alumno pueda contestar correctamente a otro punto del mismo examen, decimos que existe "encadenamiento" perfecto. Sin llegar a la perfección, un cierto grado de "encadenamiento" está presente cuando la calidad con // que se responde a un punto del examen depende de la calidad de otra respuesta. Un alto grado de "encadenamiento" puede producir serias distorsiones en el resultado de la valuación. Así, puede ocurrir que un alumno que yerra en la respuesta a la primera pregunta o al primer problema obtenga un mas bajo puntaje, debido a / que ese error se traslada al resto del desarrollo, que otro que sabe menos, pero acertó al comienzo de su examen. Dentro de un determinado problema o ejercicio, es muchas veces necesario "encadenar" los puntos que lo componen. En esos casos resulta conveniente poner varios problemas, que sean independientes entre sí, para evitar el efecto del "encadenamiento" interno de cada uno de ellos.

Es imprescindible mantener el enunciado en el más absoluto secreto hasta el momento mismo de iniciar la prueba, aún a costa de molestas precauciones. El conocimiento previo de parte o de todo el contenido del examen beneficia de un modo incorrecto, inmoral, a quién obtiene esa información, y perjudica al resto de los examinandos, por efectos de la comparación entre sus respuestas.

Al margen del sistema oficial de calificación (notas de 0 a 10, por ejemplo), asignase a cada pregunta o ejercicio un puntaje máximo alcanzable, o sea aquel / que logrará el alumno que dé una respuesta perfecta. Generalmente se busca que la suma de los puntajes máximos correspondientes a todos los items del examen sea / un número redondo, tal como 100, o como 90 si la prueba durará una hora y media, o como 60 si durará una hora, etc. La distribución del total entre las preguntas y ejercicios debe hacerse al preparar el enunciado, teniendo en cuenta la dificultad, laboriosidad, importancia, etc. de cada uno de esos items. Una práctica conveniente es informar a los examinandos, al iniciarse la prueba o en el mismo enunciado, los puntajes máximos parciales, para que les sirvan de guía en la administración de su tiempo.

3. La Prueba

La prueba puede ser "a libro abierto" o sin material de consulta. Es conveniente que el alumno conozca previamente cual de esas formas será la utilizada, para que los resultados no sean fruto de la sorpresa. El examen "a libro abierto" permite al profesor relajar la vigilancia durante la prueba, pero le obliga a poner mas ingenio en la preparación del enunciado, a fin de que él resulte discriminatorio entre examinandos mejores y peores. Aclaremos esto con un ejemplo, que corresponde a un examen no discriminatorio: El enunciado pide la transcripción de determinados artículos de cierta ley, y los alumnos tienen en su poder, durante la prueba, el texto de la ley. ¡Todas las respuestas resultarán iguales! Tal resultado es equivalente a no tomar el examen, pues no aporta información alguna respecto al conocimiento de los estudiantes. El examen "a libro abierto" // tiende a premiar el razonamiento, mientras que el examen "a libro cerrado" tiende a premiar la memoria de los alumnos.

Lo que merece especial cuidado, durante la realización de la prueba, es que el trabajo de los examinados sea estrictamente individual. Un alumno que copia se beneficia, pero, por efectos de la comparación de puntajes que realiza nuestro método, perjudica a todos los otros, incluido el que lo dejó copiar. Logra mejorar su posición en el grupo, a costa de la de los demás, mediante un medio inmoral.

Es también fundamental el control del tiempo que dura el examen. Lo corriente es establecer un máximo de tiempo para realizar la prueba, pudiendo el alumno renunciar a parte de ese tiempo, entregando su respuesta antes de que él termine. Llegado el instante final del máximo mencionado, el profesor debe exigir la entrega inmediata de las respuestas al examen, tratando de evitar el robo de minutos por parte de algunos examinados. El efecto de esta picardía es, nuevamente, un beneficio para los que la hacen y un perjuicio para los otros.

4. Sistema de Corrección

Recordemos, que nuestra "idea básica" es la comparación entre los estudiantes examinados, y tengamos esto presente mientras procedemos a corregir las pruebas.

Una primera recomendación en tal sentido es que se debe corregir pregunta // por pregunta, o ejercicio por ejercicio, y no alumno por alumno. Esto quiere decir que el docente tomará la pregunta o problema número 1, evaluará las respuestas dadas a él por todos los examinados, y recién pasará a la pregunta número 2, etc. El sistema alternativo, consistente en revisar todas las respuestas del alumno Juan Pérez, para pasar luego al examinado siguiente, no es tan eficiente para realizar la comparación deseada.

El evaluador conoce el puntaje máximo que le corresponde a la pregunta que, en cada momento, está corrigiendo. Debe graduar los puntos que en esa pregunta o problema (o ítem de un problema) da a cada alumno teniendo en cuenta ese máximo y la calidad de la respuesta. Tal calidad se refleja en atributos como corrección, precisión, claridad, elegancia, ingenio, etc., y debe valuar, en lo posible, por comparación con los mismos atributos de las respuestas de los otros estudiantes al mismo punto del examen. Ello no siempre resulta tarea fácil, especialmente si los examinados son numerosos. Pero debe tenerse en cuenta que el esfuerzo que se realice para lograr una buena comparación se verá compensado por un alto grado de justicia en las calificaciones resultantes.

Una tendencia que debe evitarse es la de asignar el total del puntaje máximo a las respuestas consideradas aceptables, y ningún punto a las otras, en cada ítem del examen. Por el contrario, es conveniente graduar los puntos de modo que corresponda el máximo a las respuestas "perfectas", menos puntaje a las que se consideren como algo inferiores, y así, paulatinamente, hasta llegar a cero punto para las no respuestas o los grandes errores $\frac{1}{5}$. Así, si determinada pregunta tiene asignado un puntaje máximo de 5 puntos sobre cien del total del examen, es inconveniente asignar a unos alumnos 5 y a otros 0, según que la respuesta esté "bien" o "mal"; es preferible utilizar los seis primeros alternativos: 0, 1, 2, 3, 4 ó 5 (y aún, si se quiere, utilizar fracciones de punto), para asignarlos a las distintas aproximaciones a una "perfecta" respuesta.

Si, en razón de ser muy grande el número de alumnos examinados, la tarea de corrección debe ser realizada por varios docentes, se debe asignar, a cada uno de ellos, uno o varios puntos del examen, que serán corregidos por él en las respuestas de todos los estudiantes que rindieron juntos. Esto facilita la asignación comparativa del puntaje. Si la distribución de tareas consistiera, por el contrario, en que cada docente reciba las respuestas de un grupo de estudiantes, a los que corregirá el examen completo, el puntaje total obtenido por cada examinado depende no solo de lo que él fué capaz de escribir, sino también del criterio del corrector, que le tocó la suerte. Las posibilidades de comparación, en este //

Una posibilidad interesante es asignar puntaje negativo a los errores, en especial a los muy fundamentales. Tal sistema debe ser anunciado a los estudiantes antes de la prueba, para que puedan optar por no contestar, ante la duda y el riesgo de pérdida de puntaje si yerran.

último caso, son mínimas, y se puede llegar a notables injusticias, salvo que las características del examen permitan la corrección mecánica, carente de subjetividad por parte del docente.

Mientras se realiza el proceso de valuación de cada pregunta o problema del examen, debe escribirse, en forma clara y visible, el puntaje que se le adjudicó, en la hoja en que respondió cada uno de los alumnos. Esto facilita la primera tarea que describiremos en la Secc. 5.

5. Tratamiento del Puntaje

El siguiente paso consiste en sumar los puntos obtenidos por cada estudiante en todas y cada una de sus respuestas a los diferentes ítems del examen. Es necesario tener sumo cuidado en no omitir en la suma el puntaje de ninguna respuesta, pues ello tendría graves consecuencias en el proceso posterior. El total de puntos obtenidos por cada individuo debe ser escrito claramente en la propia hoja / que él empleó, o en una planilla especialmente preparada a tal fin. Nótese que / ese total está expresado en puntos "sobre 100", o "sobre 90", o "sobre 60", etc., según sea el máximo puntaje asignado a todo el examen.

Hecho esto, se procede a promediar el puntaje obtenido por los estudiantes, empleando para ello la fórmula de la media aritmética simple. Esto es: Si X_i es el puntaje obtenido por el estudiante número i ; n es el número de alumnos examinados; \bar{X} es la media que queremos obtener; entonces:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Necesitamos también la desviación estándar del puntaje, a la que calculamos con la conocida fórmula estadística:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}$$

Donde: σ simboliza desviación estándar.

Nótese que el valor de σ se obtiene dividiendo la suma de los desvíos cuadráticos respecto a la media por n , no por $(n-1)$. Ello se debe a que nuestro proceso es descriptivo, no inferencial.

En este lugar conviene recordar un par de conceptos, que ayudarán a entender la asignación de notas propuesta en la Secc. 6. La media es una medida de posición. Nos indica la posición del grupo en la escala de puntaje elegida. Así, por ejemplo, siendo el total máximo alcanzable 100 puntos, una $\bar{X} = 75$ nos indica que el grupo examinado está ubicado en la parte superior de la escala; mientras que $\bar{X} = 25$ nos muestra, por el contrario, a un grupo ubicado en la parte inferior de ella. La desviación estándar es una medida de dispersión. Nos indica si los puntajes tienden a encontrarse alejados o próximos a la media. Así $\sigma = 20$ indica // que los puntajes tienden a estar más dispersos, alrededor de \bar{X} , que si $\sigma = 5$.

Nuestra idea básica impone que el alumno sea calificado de acuerdo a su situación relativa en el grupo. La media nos da un punto de referencia para establecer esa situación: ¿Está el puntaje de determinado estudiante por arriba o // por abajo de \bar{X} ? En otras palabras: ¿El estudiante considerado es mejor o peor // que el promedio del grupo? Una vez que sabemos que el examinado observado es // "peor", digamos, que el promedio, nos interesará saber cuán "peor" es. No merece la misma nota quién tiene un puntaje inmediatamente inferior a \bar{X} , que quien lo // tiene inferior y muy alejado de ella. La desviación estándar nos permite medir // la distancia existente entre la media y el puntaje obtenido por cada estudiante. El método de la sección siguiente hace uso, indirectamente, de estos conceptos.

6. Asignación de las Notas

Una vez conocido el puntaje, X_i , correspondiente a cada estudiante, y calculado el promedio, \bar{X} , y la desviación estándar, σ , de todos esos puntajes, resta el problema de asignar, en base a esos elementos, la nota en escala entre 0 y 10 (u otra que corresponda), a cada alumno examinado. Para ello construimos la siguiente tabla:

TABLA N° 1

INTERVALOS		NOTA	NOTA
(En puntaje sobre 100, ó sobre 90, ó sobre 60, etc.)		(Criterio más exigente)	(Criterio menos exigente)
1 -	Entre $\bar{X} - 2,5 \sigma$ y $\bar{X} - 2 \sigma$		1
2 -	" $\bar{X} - 2 \sigma$ y $\bar{X} - 1,5 \sigma$	1	2
3 -	" $\bar{X} - 1,5 \sigma$ y $\bar{X} - \sigma$	2	3
4 -	" $\bar{X} - \sigma$ y $\bar{X} - 0,5 \sigma$	3	4
5 -	" $\bar{X} - 0,5 \sigma$ y \bar{X}	4	5
6 -	" \bar{X} y $\bar{X} + 0,5 \sigma$	5	6
7 -	" $\bar{X} + 0,5 \sigma$ y $\bar{X} + \sigma$	6	7
8 -	" $\bar{X} + \sigma$ y $\bar{X} + 1,5 \sigma$	7	8
9 -	" $\bar{X} + 1,5 \sigma$ y $\bar{X} + 2 \sigma$	8	9
10 -	" $\bar{X} + 2 \sigma$ y $\bar{X} + 2,5 \sigma$	9	

Construida una relación de notas con puntajes con uno de los dos criterios propuestos en la tabla anterior, estamos en condiciones de clasificar a cada uno de los alumnos. Elijamos el "criterio más exigentes", a título de ejemplo, y supongamos que el quinto alumno tiene X_5 puntos sobre 100, y que ese puntaje es tal que:

$$\bar{X} + 0,5 \sigma < X_5 \leq \bar{X} + \sigma$$

Dicho en otras palabras: El puntaje del quinto alumno se encuentra dentro del séptimo intervalo de la tabla. En tal caso, a ese alumno le corresponde la nota 6 // (seis). Resulta, pues, un trabajo fácil y mecánico el asignar las notas.

Quedan por aclarar algunos detalles del procedimiento. Tomemos el "criterio más exigente": ¿Qué ocurre si el puntaje de un alumno es mayor que $\bar{X} + 2,5 \sigma$? // Simplemente le corresponde un 10 (diez). ¿Y si fuera menor que $\bar{X} - 2 \sigma$? En principio le correspondería 0 (cero), pero ello puede parecer demasiado fuerte, pues // esa nota es más que un aplazo: Significa un castigo. Nuestro criterio es el siguiente: Si se trata de un examen parcial, cuya nota no va a los registros y será promediada con la de los otros parciales, ponemos la nota exacta que surja de la tabla, incluido el cero; si se trata de un examen final, cuya nota quedará registrada, reemplazamos los valores 0, 1, 2, y 3 que surjan de la tabla por la nota / 2 (dos). Cada profesor debe elegir su modo de operar en esta cuestión. Lo dicho / en este párrafo es también válido para el "criterio menos exigente", pero debe tenerse en cuenta que en tal caso los límites máximo y mínimo establecidos en la tabla son $\bar{X} + 2$ y $\bar{X} - 2,5 \sigma$.

Merece especial atención la diferencia entre el "criterio más exigente" y el "criterio menos exigente". Puesto que nuestro método se basa en la comparación de cada examinando con el grupo de ellos, representado por la media de los puntajes obtenidos, \bar{X} , parece razonable y objetivo que la nota 5 (que es el promedio de / las notas en la escala de 0 a 10) sea asignada a quienes obtuvieron puntajes cer-

canos a \bar{X} . Pero "cercanos" puede significar próximos desde abajo o próximos desde arriba de la media. Si ponemos un 5 a los que están cercanos a la media, pero arriba de ella, estamos aplicando el "criterio más exigente". Si el cinco corresponde a los que están inmediatamente bajo de la media, nuestro criterio será "menos exigente". En la tabla se puede observar que la elección de uno u otro criterio afecta toda la escala de notas, en relación con los intervalos.

La elección del criterio a aplicar, de entre los arriba aplicados, queda al arbitrio del profesor; pero consideramos que debe ser elegido antes de tomar el examen, a fin de eliminar una fuente de influencia subjetiva del docente sobre los resultados. Ayuda en esta decisión el saber que, con el "criterio más exigente", se puede esperar que aproximadamente el 30,85% de los alumnos resultarán aplazados, y que ese porcentaje se reduce al 15,87% con el "criterio menos exigente". Estos dos números surgen de sumar las frecuencias relativas teóricas, correspondientes a las notas 0, 1, 2 y 3 y a cada uno de los criterios referidos, en la Tabla N° 2 de la Sección 7.

Nosotros hemos usado generalmente uno de esos dos criterios (el "más exigente" para exámenes finales, y el otro para exámenes parciales), pero se puede trabajar con un "criterio intermedio", que resulta aún más objetivo. Consiste en asignar la nota 5 al intervalo de puntajes que tenga en su punto medio a \bar{X} , o sea a:

$$\text{"Entre } \bar{X} - 0,25 \sigma \text{ y } \bar{X} + 0,25 \sigma \text{"}$$

Los límites de los otros intervalos se obtienen restando o sumando sucesivamente $0,5 \sigma$, a partir de estos límites. El porcentaje de alumnos aplazados será, en tal caso, próximo al 22,66%, cifra que se puede obtener de la Tabla N° 2.

Generalmente los límites de los intervalos que proponemos resultan con varios decimales. Los puntajes, por su parte, suelen ser números enteros. Tales circunstancias hacen imposible que el total de puntos obtenidos por un alumno coincida con el límite de un intervalo. Si, a pesar de ello la mencionada coincidencia pudiera ocurrir (porque los límites resultan enteros, o porque los puntajes tienen decimales), sería necesario especificar mejor los intervalos, por ejemplo del siguiente modo:

$$\text{"} \bar{X} - \sigma \text{ a menos de } \bar{X} - 0,5 \sigma \text{"}$$

Esto significa que un puntaje de exactamente $\bar{X} - \sigma$ está en este intervalo, pero / uno de exactamente $\bar{X} - 0,5 \sigma$ está en el siguiente.

7. Consideraciones Estadísticas

Parece aceptable esperar que sean más numerosos los estudiantes con puntajes cercanos al promedio que aquellos con un número de puntos excepcionalmente grande o excepcionalmente pequeño. La experiencia nos dice que ello ocurre en la generalidad de los casos. No parece fuera de razón, suponer, por lo tanto, que el puntaje obtenido por los alumnos en un examen se distribuye en forma aproximadamente normal. Sin embargo, pueden presentarse circunstancias que hagan que la forma de la distribución observada de los puntajes sea muy diferente a la distribución normal o de Gauss. Ello ocurriría, por ejemplo, si hubieran en el grupo dos niveles bien definidos de estudiantes, pues en tal caso podemos esperar que se presente una distribución bimodal, que nada tendría que ver con la teórica que proponemos. Un test de la bondad del ajustamiento nos permite saber si el supuesto de normalidad es aceptable o no.

Si la distribución de los puntajes fuera exactamente normal o gaussiana, y si el número de estudiantes examinados fuera infinito, la tabla que se da a continuación se cumpliría exactamente. La primera columna de la tabla contiene las notas, mientras que en las tres siguientes aparecen los porcentajes de alumnos que obtendrían cada nota, si ella se asigna de acuerdo a nuestro método y con el criterio que encabeza cada columna.

TABLA N°
Frecuencias Relativas Teóricas
(En Porcentajes)

N O T A	Criterio más exigente	Criterio intermedio	Criterio menos exigente
0	2,28	1,22	0,62
1	4,40	2,79	1,66
2	9,19	6,55	4,40
3	14,98	12,10	9,19
4	19,15	17,47	14,98
5	19,15	19,74	19,15
6	14,94	17,47	19,15
7	9,19	12,10	14,98
8	4,40	6,55	9,19
9	1,66	2,79	4,40
10	0,62	1,22	2,28
T o t a l	100,00	100,00	100,00

Para interpretar esta tabla debe tenerse en cuenta que los porcentajes que ella presenta se cumplen exactamente solo bajo los supuestos que arriba especificamos. Ellos son frecuencias teóricas, que solo por casualidad se cumplen en la realidad con exactitud, pero que sirven de guía para saber lo que, aproximadamente, se puede esperar.

Si el ajustamiento de la curva normal a nuestros datos fuera muy malo, las frecuencias teóricas de la Tabla N°2 no nos servirían como punto de referencia, / pero nuestro método de evaluación puede ser utilizado, a pesar de ello, pues no / requiere normalidad.

Es evidente que el supuesto de infinito número de examinados jamás se cumple. Pero sirve para llamarnos la atención hacia un punto muy importante: Nuestro método funciona mejor cuando mayor es el número de alumnos examinados. Esta afirmación no solo es correcta porque si ese número, al que hemos llamado n , crece, / nuestros porcentajes de notas tenderán a aproximarse a los de la Tabla N° 2, bajo el supuesto de normalidad, sino porque, si n es muy pequeño, nuestro método puede conducir a resultados ridículos o muy arbitrarios.

Ejemplifiquemos lo dicho en el párrafo anterior con dos casos extremos, utilizando el "criterio más exigente" (ver Tabla N° 1). Si $n = 1$, tenemos un solo // puntaje y , por lo tanto, la media es igual a ese puntaje y $\sigma = 0$. Nuestro único / examinando merece, con nuestro método, un 4 ó un 5 (está justo en el límite). ¡Lo ridículo es que tal nota le corresponde tanto si no obtuvo ningún punto como si / su puntaje fué el máximo posible. Si $n = 2$ y ambos alumnos obtienen el mismo puntaje estamos en el caso anterior. Si sus puntajes difieren, uno de ellos tendrá $X_1 = \bar{X} + \sigma$ y obtendrá un 6 ó un 7. El otro tendrá: $X_2 = \bar{X} - \sigma$, y su nota será un 2 ó un 3. ¡Ello es siempre cierto, independientemente de que los puntajes (sobre 100) sean: $X_1 = 99$ y $X_2 = 98$, o sean: $X_1 = 2$ y $X_2 = 1$, o cualquiera otra combinación posible.

La moraleja que se desprende de lo anterior es que el número de examinados no debe ser pequeño. Pero ¿Qué número es el mínimo para confiar en nuestro método? En nuestros exámenes imponamos que n sea igual a 20 o más, y hasta ahora obtuvimos resultados aceptables. Cuando a un examen se presentan menos de 20 estudiantes (suele ocurrir en algunos exámenes finales) tomamos prueba oral; pero ella podría ser escrita, utilizando otro método de valuación.

Una condición más para que el método funcione es que exista independencia en la conducta de los estudiantes durante la prueba. Una pequeña desviación respecto a esta norma (un poco de copia, por ejemplo) no invalidaría del todo los resultados. Pero no podríamos aplicar nuestro sistema si todos los alumnos se pusieran de acuerdo y entregaran la hoja firmada en blanco. Clara está que, en este / último caso, no haría falta método alguno para ponerles cero a todos.

8. Ventajas del Método

Todo sistema de evaluación es, en el fondo, un modo de comparar exámenes. En el examen oral, por ejemplo, implícitamente comparamos la exposición de cada alumno con una ideal, que está en nuestra mente, a la que debemos mantener constante a lo largo de todo el tiempo que permanezcamos integrando el tribunal, si queremos ser justos. Es esa una forma de comparar entre ellos a los exámenes de todos los examinandos.

Nuestro método se basa explícitamente en la comparación de cada individuo // con el grupo, representado por la media y la desviación estándar. Trata de que // esa comparación sea lo más matemática, lo más mecánica posible, minimizando la // subjetividad del profesor en el proceso de calificación. Esta es, sin duda, su // principal ventaja.

El proceso de elaboración de los puntajes para llegar a las notas puede parecer una pesada carga. Nosotros utilizamos para ello un calculador científico de bolsillo, preparado para cálculos estadísticos. Basta con introducir los puntajes de todos los examinados para obtener la media y la desviación estándar. Queda luego la tarea de establecer los intervalos y asignar la nota a cada alumno. El trabajo resulta así aceptable, aún para gran número de alumnos. La disponibilidad de una computadora y la automaticidad del método permiten, por otra parte, que el profesor tenga menos trabajo que con otros métodos de valuación. La tarea se reduciría a la preparación de un buen enunciado, a un estricto control durante la prueba, y a la corrección y asignación de puntaje a las respuestas. El resto es trabajo de un operador y de la computadora, que debe estar convenientemente programada. Es de hacer notar que quién opera la computadora debe ser un docente de la cátedra, que merezca confianza del profesor, pues tendrá la oportunidad de alterar los resultados. El operador informa a la computadora el puntaje // que cada alumno mereció en cada pregunta o problema, la máquina procesa esos datos y obtiene la nota que corresponde a cada estudiante. Todo el proceso de elaboración que hemos descrito más arriba se produce dentro de ella.

Una gran ventaja del método que proponemos es que amortigua los efectos de // los exámenes "fáciles" o "difíciles". No es tarea sencilla para un profesor preparar el enunciado de un examen de un modo tal que resulte bien graduado, en relación al nivel del material utilizado en el curso, a la capacidad de comprensión // de ese material por parte de los alumnos, a la capacidad de hacerse entender del propio profesor, etc. Resulta así que un examen es a veces fácil, y lo // aprueban aún quienes no lo merecen, o, por el contrario, es difícil, y resultan aplazados quienes debieron aprobar. Nuestro método tiene un sistema automático // de corrección para ese fenómeno. Si el examen fué fácil, los puntajes obtenidos por todos los examinados tenderán a ser altos, y, por lo tanto, también lo será nuestro punto de comparación: La media \bar{X} . Así, puede ocurrir que un estudiante // obtenga más de 50 puntos sobre cien y resulte aplazado, si la media estuvo en 70 y el desvío estándar fué de 9. Si, por el contrario, el examen fué difícil, todos los alumnos tenderán a obtener bajo puntaje, y \bar{X} resultará también baja. Un alumno con 25 puntos sobre cien tendrá una buena nota si la media fué de 20. No hay, pues, ninguna razón para que un examen difícil traiga como consecuencia un mayor porcentaje de aplazos que un examen fácil.

De un modo similar, el método disminuye los efectos de un error en el tiempo asignado para escribir la prueba. Si fué escaso lo fué para todos, y la media // tenderá a ser menos. Lo contrario ocurre cuando fué abundante. Ello corrige el // efecto que la disponibilidad de tiempo tuvo sobre los realizado por cada individuo, al asignarse las notas por comparación con el promedio.

A pesar de lo dicho en los dos últimos párrafos, el profesor debe cuidar que el examen sea aceptablemente equilibrado en cuanto a dificultad y tiempo. Un de-

sequilibrio muy grande en cualquiera de estas características puede traer serias consecuencias, aún cuando apliquemos nuestro método. En efecto, si el examen fue demasiado fácil, o el tiempo muy generoso, o ambos, nuestra distribución se "pugaría contra el techo". La media estaría muy cercana al máximo de puntos posibles, digamos 100, y no dejaría lugar para la existencia de notas altas. Así, si 100 está a solo una distancia s de la media, no existirían el 8, el 9, ni el 10 entre las notas, cualquiera sea el criterio aplicado (ver Tabla N° 1). Lo contrario ocurre con un examen difícil y/o con poco tiempo. La distribución se "aplanta contra el piso". La media estaría demasiado próxima a cero, y no dejaría lugar para los aplazos. Si, por ejemplo, el puntaje cero estuviera a una distancia de solo 0,5 s por abajo de \bar{X} , no habría ningún aplazado, aún con el "criterio // más exigente". Nótese que el efecto es el inverso al que se podría esperar intuitivamente: ¡Los exámenes fáciles tienden a eliminar las notas altas y los difíciles a eliminar los aplazos. En definitiva, nuestro método corrige los efectos de la dificultad del examen o del tiempo que se le asignó, salvo cuando el profesor cometió un error muy grave en alguno de estos aspectos.

Cuando un alumno reclama por la nota que obtuvo, se lo puede convencer de // que ella fue justa explicándole el método, haciéndole conocer la media y la desviación estándar del examen que él rindió, y mostrándole la posición que le cupo en la distribución de puntajes. Nuestra experiencia enseña que tales argumentos son lo suficientemente fuertes como para convencer al estudiante. La razón es, // posiblemente, la siguiente: Si el reclamante sabe Estadística, entenderá lo sustancial del método y se quedará sin argumentos para defender sus pretensiones de mayor clasificación: si no sabe Estadística, aunque las cosas no le quedan muy // en claro, captará al menos la idea de que se trata de un procedimiento mecánico, objetivo, y terminará por rendirse ante la evidencia. Es ésta otra ventaja del // método. Pero esta ventaja requiere que el procedimiento sea cuidadosa y lealmente aplicado. Especial cuidado debe ponerse en la operación más subjetiva, y por lo tanto más vulnerable, del proceso: La asignación de los puntajes a cada res-// puesta. La falta de lealtad en el manejo del método, además de ser inmoral, puede traer serias consecuencias desde el punto de vista que nos ocupa. Así, yendo a un ejemplo extremo, un profesor puede escudarse tras de medias y desviaciones estándares para cometer arbitrariedades: pero a riesgo de tutar la credibilidad de los estudiantes no solo con respecto a su persona, sino también con respecto al método.

Lo dicho en esta Sección nos permite afirmar que nuestro método tiene ventajas sobre otros más frecuentemente aplicados, como ser el examen oral, el examen escrito con relación rígida entre puntaje y notas, etc., en cuanto a la equidad en la calificación. Esperamos que quienes investigan sobre los métodos de evaluación nos señalen los defectos y limitaciones que tuviere.

UNIVERSIDAD NACIONAL DE SALTA

DEPARTAMENTO DE CIENCIAS ECONOMICAS, JURIDICAS Y SOCIALES

AREA DE ECONOMIA

Reuniones de Discusión

<u>N°</u>	<u>Fecha</u>	<u>Autor</u>	<u>Título</u>
1	8-9-76	Eusebio Cleto del Rey	"Distribución de los Porcentajes de Ingresos - Egresos de Depósitos".
2	19-9-80	Eduardo D. Antonelli	"Una Nota sobre las Teorías Neoclásica y Keynesiana de la Determinación del Empleo y la Renta".
3	28-11-80	Ricardo Jiménez	"Metodología para el Cálculo de la Prima en los Seguros Elementales".
4	12-12-80	Eusebio Cleto del Rey	"Observaciones al Método de Actualización en la Evaluación de Proyectos".
5	28-5-81	Eduardo D. Antonelli	"Una Nota sobre las Teorías Neoclásica y Keynesiana de la Determinación del Empleo y la Renta" - II.
6	7-7-81	Eusebio Cleto del Rey	"Cálculo de la Duración de una Carrera Universitaria".
7	3-8-81	Guillermo Lloret	"El Costo Directo de Estudiar una Carrera Universitaria".
8	1-10-81	Rita Lavin Figueroa	"El Costo de la vida: un concepto económico y otro estadístico".
9	29-10-81	Eusebio Cleto del Rey	"Un Método de Evaluación Aplicable al Examen Escrito".
