

Cert: un modelo matemático y tecnológico de evaluación

Antonio Ariza García, Ricardo Barroso Campos, Jose M^a Gavilán Izquierdo y
Angel Sánchez Sotelo

Universidad de Sevilla

En este artículo, los autores, profesores del Área de Didáctica de las Matemáticas de la Facultad de Ciencias de la Educación de la Universidad de Sevilla, realizan un estudio de las componentes matemáticas del modelo de evaluación propuesto en el CERT (Cooperative Project of Evaluation Results Training). Este trabajo se basa fundamentalmente en modelos estadísticos básicos.

El proyecto CERT, es una innovación en el uso de las nuevas tecnologías (en este caso el ordenador) aplicadas al proceso de evaluación.

In this article, the authors, teachers of the Area of Didactic of Mathematics of the Faculty of Science of Education of the University of Seville, accomplish a study of the mathematics components of the evaluation model proposed in the CERT (Cooperative Project of Evaluations Results Training). This work is based fundamentally on basic statistic models.

The project CERT, it is an innovation in the use of the new technologies (in this case of the computer) applied to the evaluation process.

DESCRIPTORES: Evaluación, Nuevas tecnologías, Modelos matemáticos, Identificación del Modelo Estadístico.

1. ¿Qué es el CERT?

Se ha comprobado, que la utilización de pruebas objetivas para la evaluación, sobre todo en el caso de grandes grupos, es muy adecuada, aunque una buena preparación de estas es muy compleja, permiten una "fácil" corrección siendo esta una ventaja entre otras como pueden ser la posibilidad de evaluar sistemáticamente niveles superiores de actividad mental o la objetividad de la corrección. La introducción del ordenador ha dotado de un instrumento rápido y preciso con un amplio campo de posibilidades a las pruebas objetivas.

Debido a la bondad de este tipo de pruebas se han desarrollado proyectos de investigación como "the Cooperative project on Evaluation of Results of Training" (CERT) (Alonso 1992). Como indica De Pablos (1993) en el seno de este proyecto se desarrolló un programa informático que apoya y facilita el uso de pruebas objetivas en la evaluación educativa, posibilitando gestionar gran cantidad de información recogida a través de pruebas objetivas de opción múltiple y analizarla, agilizando el proceso de manipulación y cálculos manuales de los formadores para evaluar aprendizajes. Una de

las características principales de este programa es la inclusión de los denominados niveles de confianza que son una estimación que el alumno realiza en base a una escala de 0-5 del grado de confianza o seguridad con que responde a cada una de las cuestiones de la prueba y que sustituye a la corrección del factor de respuesta por azar en la puntuación de la prueba.

Así a la hora de baremar las respuestas de los alumnos tendremos en cuenta dos criterios: la exactitud de la respuesta y el grado de confianza con que se responde. Una respuesta correcta con un grado de confianza elevado tendrá una puntuación más elevada que la misma respuesta con un grado de confianza más bajo. Por el contrario una respuesta incorrecta tendrá una penalización mayor al aumentar el grado de confianza en la respuesta. La puntuación final tiene que ser calculada de forma que el alumno tenga interés en decir la verdad a la hora de responder.

El proyecto CERT propone el siguiente baremo:

| | Puntuación en caso de respuesta | |
|---------------------------|---------------------------------|------------|
| | Correcta | Incorrecta |
| entre 0 y 25% (nivel 0) | 13 | 4 |
| entre 25 y 50% (nivel 1) | 16 | 3 |
| entre 50 y 70% (nivel 2) | 17 | 2 |
| entre 70 y 85% (nivel 3) | 18 | 0 |
| entre 85 y 95% (nivel 4) | 19 | -6 |
| entre 95 y 100% (nivel 5) | 20 | -20 |

El proyecto CERT, en base al número de respuestas correctas y dependiendo el nivel de confianza empleado en las mismas, calcula la calificación y diversos índices matemáticos entre ellos, destaca el relativo a la autoevaluación del alumno denominado "**índice de realismo**" para su cálculo, se emplean la noción de **tasas de exactitud (TE)**, es decir el porcentaje de aciertos de un alumno en cada nivel de confianza, el número de estas variarán de 1 a 6 en función de cuantos niveles de confianza haya empleado el alumno al responder a la prueba.

En este índice se indica en que medida la confianza que el alumno tiene en sí mismo corresponde a la realidad. La probabilidad de acierto predicha por él mismo ¿es cercana o lejana en relación a los aciertos obtenidos en la realidad (TE)?

Todo el proyecto CERT se basa en las posibilidades del ordenador y su programación aplicadas a procesos de evaluación de aprendizajes.

2. ¿Qué es un modelo matemático?.

Hasta el siglo XI, era comúnmente aceptado que la naturaleza estaba regida por leyes matemáticas. A principios del XX la tesis es, que las leyes matemáticas son útiles para describir los fenómenos de la naturaleza. Este cambio nos lleva a plantear que las matemáticas nos van a servir para modelar la realidad, bajo determinadas condiciones (restricciones que impone el observador).

La propia idea de modelización matemática de la realidad lleva aparejada las ideas de incertidumbre y de falibilidad de la resolución matemática de un fenómeno natural. El proceso de inducción que se produce en la búsqueda del modelo es básico para una buena descripción de la realidad. Para Perero (1994) la búsqueda de regularidades y propiedades comunes es básica.

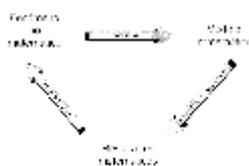
Como define Dreyfus (1991),

"Como encontrar una representación matemática para un objeto o proceso no matemático, es construir una estructura o teoría matemática que incorpora las características esenciales del objeto o proceso. La estructura o teoría, el modelo, puede ser usado en orden a estudiar la conducta del proceso u objeto modelado."

Además hay que tener presente, como resalta Perero (1994), las conclusiones a las que se llega en el modelo son solo posibles (de ahí la incertidumbre) y pueden ser refutadas (de ahí la falibilidad) con otros datos.

Con estos referentes interpretamos que un modelo matemático es el resultado de labúsqueda de regularidades que subyacen en una situación no necesariamente matemática, a partir de unas conjeturas o suposiciones iniciales. El modelo da forma matemática a la situación lo que permite aplicar todas las reglas, procedimientos y procesos de las matemáticas (el razonamiento matemático), no crea una identificación entre ambos (no se debe confundir el modelo con la situación real), pero si una correspondencia que permitirá interpretar los resultados del modelo en términos de la situación no matemática estudiada.

La interpretación gráfica sería:



3. Modelo Subyacente.

Un análisis de las técnicas cuantitativas aplicadas por el programa de evaluación permite identificar una variable bidimensional (gráfico 2). Todos los cálculos serán referentes a ella. Como toda variable bidimensional, permite un

tratamiento dicotómico, es decir, la consideración de las variables por separado, que son denominadas distribuciones marginales, lo cual nos permite aplicarle todas las técnicas y resultados estadísticos sobre variables unidimensionales; o bien la consideración conjunta de las variables que permitirá aplicarle todos los parámetros bidimensionales.



La variable x, es una variable continua en una escala de cero a cien o bien, normalizada de cero a uno, representando la probabilidad de "acertar". No es equiespaciada, sino que los intervalos van decreciendo conforme nos acercamos a cien. Desde el punto de vista estadístico esto puede obligar de algún modo a introducir unos "pesos" que ponderen de manera adecuada las frecuencias para cada nivel de confianza en el momento de calcular cualquier estadístico por ejemplo, la Media Aritmética o la Varianza. Se sigue un proceso clásico de discretización de esta variable continua, que significa que se consideraran únicamente los valores centrales (marca de clase) como posible de la variable.

Este proceso de discretización, ha sido una necesidad con el objeto de facilitar los cálculos. Con la potencia actual de los ordenadores y de software este paso podría suprimirse o bien la discretización podría ser más "fina" (más número de intervalos).

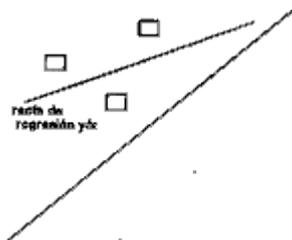
La variable y, representa la tasa de exactitud de las respuestas. A diferencia de la variable anterior esta cuantifica una característica objetiva, medida desde el exterior y a posteriori, mientras que la variable x, tiene una fuerte carga subjetiva medida simultáneamente a la propia realización de la prueba (de ahí el aspecto positivo de autoevaluación que presenta el proyecto CERT). Desde una perspectiva estadística esta variable podemos considerarla discreta, ya que cada experimento concreto, viene determinada a priori su dominio.

Algunos autores (De Pablos, García, etc.,1993) señalan que un aspecto importante es la obtención o representación de datos de forma visual, esta visualización se denomina estadísticamente "nube de puntos o diagrama de dispersión" que además permite estudiar la relación de dependencia aleatoria lineal sin necesidad de recurrir al cálculo de parámetros, (Coeficiente de correlación de Pearson, rectas de regresión, pendientes de dichas rectas).

Para que dicha dependencia pueda ser cuantificada, y permita hacer comparaciones entre individuos sí es necesario el cálculo de algunos parámetros, entre ellos el que el proyecto CERT denomina índice de coherencia.

Pensamos que el coeficiente dado por la pendiente de la recta de regresión de la variable Y sobre X, que es la covarianza de X, Y dividida por la varianza de X, puede medir en alguna forma la coherencia entre lo que opina el alumno de

su aprendizaje y la realidad del mismo. Para una mejor "fotografía" del aprendizaje (la pendiente de la recta de regresión no la determina unívocamente, pues existen infinitas rectas todas ellas paralelas entre sí con la misma pendiente) se puede utilizar la "ordenada en el origen". Si además queremos medir la adecuación de los datos de la nube de puntos a la recta de regresión usamos el coeficiente de correlación lineal (gráfico 3).



Las salidas o resultados del CERT incluyen un parámetro que se denomina índice de realismo, que viene a proporcionar una medida de la confianza que el alumno tiene sobre su propio conocimiento y de la calidad de su aprendizaje en relación a los resultados reales. Este índice es lo que se conoce en estadística como varianza. Debido a que en su cálculo aparecen las desviaciones al cuadrado, se pierde la información sobre el sentido (signo positivo o negativo) de dichas desviaciones, lo que puede inducir a error en la interpretación, tanto en la sobrevaloración como en la infravaloración, lo que nos obliga a recurrir a la información gráfica que proporciona el diagrama de dispersión.

Con este análisis lo que se ha pretendido es modelizar matemáticamente el proyecto CERT identificando sus distintas componentes desde la perspectiva estadística. Queremos señalar que se podrían haber tomado otros estadísticos para dicha cuantificación.

Referencias bibliográficas.

ARIZA A. , SANCHO M., Y TORREBLANCA J. (1996): La utilización de base de datos como herramientas de evaluación. **Pixel bit. Revista de medios y educación**, 7. 73-82

DE KETELE, J. M. (1984): **Observar para educar. Observación y evaluación en la práctica educativa**. Madrid. Visor libros.

DE PABLOS J. y OTROS (1993): La evaluación del alumno en la Universidad: el proyecto CERT. **Revista de enseñanza universitaria**, 6. 49-71

DREYFUS, T (1991): **Advanced Mathematical Thinking Processes**, en Tall D. Editor.

PERERO, M. (1994): **Historia e Historias de Matemáticas**, Grupo editorial Iberoamericano S.A., México D.F.

