

DESARROLLO Y EXPLICACIONES ADICIONALES DEL ARTÍCULO DE LA VANGUARDIA “RESTAR NO SUMA” Por Fernando Trias de Bes

La conveniencia de erradicar de nuestro sistema educativo las preguntas que restan en los exámenes tipo test

1. Cálculo de los puntos aleatorios de un chimpancé en un test

Texto del artículo de La Vanguardia:

Supongamos un examen test (a, b, c, d) de 40 preguntas. Cada acierto son 0,25 puntos (40 aciertos es un 10 de nota).

Si un chimpancé contestase al azar ese examen obtendría, en promedio, 10 respuestas correctas de las 40 preguntas (omito cálculos de probabilidades, pero es así). Traducido a puntos, el mono sacaría un 2,5 sobre 10.

EXPLICACIÓN

Al tener 4 posibilidades, la probabilidad de acertar al azar es 1 dividido entre 4, esto es, un 25%. Como hay 40 preguntas, el chimpancé obtendría (esperanza matemática) el 25% de las 40 preguntas correctas = 10 preguntas.

Como cada una puntúa 0,25 puntos, la nota del chimpancé sería de $0,25 \times 10$ preguntas = 2,5 puntos sobre 10, que equivale, lógicamente, a un 25 sobre 100, exactamente su probabilidad aleatoria de aciertos (25%).

El cálculo de los puntos aleatorios que obtiene un alumno (llamado en el artículo “puntos de salida”) para cualquier tipo de test es uno dividido por el número de alternativas que se ofrece en las preguntas, multiplicado por el número de preguntas y multiplicado por el número de puntos de cada pregunta.

2. Cálculo de la nota de aprobado en caso de no restar

Texto del artículo de La Vanguardia:

Propongo situar el aprobado en una cifra superior al 5. No sería 7,5 (el 5 de aprobar más los 2,5 que la probabilidad regala al chimpancé), sino 6,25 (encontrará este cálculo en www.triasdebes.net). Las notas se establecerían de acuerdo a esta nueva escala: “Suspenso”, por debajo de 6,25 puntos; “Aprobado”, a partir de 6,25 puntos; “Bien”, a partir de 7 puntos; “Notable”, a partir de 7,75 puntos; “Excelente”, a partir de 8,875 puntos – “Matrícula de honor”, 10 puntos.

Recordemos que se trata de un examen test (a, b, c, d) de 40 preguntas. Cada acierto son 0,25 puntos (40 aciertos es un 10 de nota).

EXPLICACIÓN

Hay tres formas de demostrar este cálculo. Dos sencillas y una más compleja.

a) La primera es que, si asumimos que el chimpancé obtiene 10 de las 40 preguntas al azar, podemos considerar que el resto de aciertos por encima de 10, no son ya azarosos, y que corresponden (en términos de esperanza matemática) a aciertos reales. Eso significa que si las 30 preguntas restantes fueran una prueba aparte, el aprobado se situaría en 15 preguntas acertadas. Es decir, la mitad.

Considerando ahora la totalidad de 40 preguntas de nuevo, el aprobado obligaría a 10 preguntas (que se considera que el alumno no ha hecho nada para ganarse, aunque pueda haberlas acertado gracias a haber estudiado) y a 15 preguntas más que, ahora sí, consideramos desprovistas ya de aleatoriedad (toda la aleatoriedad está incluida en los 10 primeros aciertos).

Eso significa que el aprobado estaría en $10+15=25$ preguntas sobre 40. A 0,25 puntos cada una, eso significa $25 \text{ aciertos} \times 0,25 \text{ puntos/acierto} = 6,25 \text{ puntos}$.

b) Otra forma (un poco más complicada) de llegar a la misma conclusión es imaginar que un alumno pudiera escoger voluntariamente la modalidad al entregar el examen (que le resten puntos debiendo obtener un 5 para aprobar o que no le resten debiendo obtener 6,25 puntos para aprobar). Debemos demostrar que en ambos casos su probabilidad de aprobar es la misma.

Veámoslo. Ese alumno contestaría primero 20 preguntas que sabe a ciencia cierta que son correctas. "Ya tengo un aprobado", se dirá el alumno. "Pero ahora quiero ir a por nota. Me planteo pasar de un sistema que resta a otro que no resta por errores. Pero entonces, necesito 5 aciertos más para aprobar (25 respuestas correctas=6,25 puntos). ¿Me la juego?..."

Al alumno le quedan 20 cuestiones por responder. Con 5 más, aprueba. Como pasa a un sistema de no restas, puede tranquilamente contestar las 20 sin miedo a perder puntos de las primeras 20 que sabe correctas. Si de las segundas 20, no tuviese ni idea, aleatoriamente, con (a, b, c, y d) como posibilidades, obtendría un 25% de aciertos. $25\% \times 20 \text{ preguntas más} = 5 \text{ aciertos aleatorios}$ (insisto, como esperanza matemática).

Por tanto, podría jugársela y pasar de un sistema de restas a otro de no restas porque por probabilidades, es lo mismo: contestar 20 seguras y acertarlas todas equivale a contestar 40, acertando esas mismas 20 por conocimientos y obteniendo 5 más por azar. Ahí se demuestra que el 5 restando = 6,25 sin restar (en casos de 4 opciones para responder).

c) Un último modo (este bien simple) de demostrar que el 6,25 sin restar equivale al 5 restando es aceptar que, en todas y cada una de las preguntas del test (de 4 posibles respuestas cada una) yo asumo que hay un 25% de posibilidades de que el alumno acierte por azar. Por tanto, le voy a pedir un 25% más de puntos de lo que es un aprobado normal. 5 puntos incrementados un 25% equivalen a 6,25 puntos.

3. Cálculo de la equivalencia de puntos y calificaciones en ambos casos

Demostrado que un 5 en el Sistema Tradicional (te restan) equivale a un 6,25 en el Sistema Trias de Bes (no te restan), para el resto de calificaciones tan solo debe calcularse proporcionalmente las equivalencias de tramos entre el 5 de aprobado y el 10 de nota máxima y el de la escala que va de 6,25 de aprobado a 10 de nota máxima.

Por ejemplo, el "Bien", que es un 6, en el Sistema Tradicional, es el 20% del tramo (1 punto dividido entre 5 puntos, los que van del 5 al 10). El "Notable" (7) es el 40% del intervalo. Un "Excelente" (8,5) es el 70% del intervalo. Y la "Matrícula" es el 100% de ese intervalo. Veámoslo:

Correspondencia entre calificaciones y notas / Sistema Tradicional

Aprobado = 5 ptos

Bien = $20\% \times (10 \text{ ptos} - 5 \text{ ptos}) = 1 \text{ pto} + 5 \text{ ptos de aprobado} = 6 \text{ ptos}$

Notable = $40\% \times (10 \text{ ptos} - 5 \text{ ptos}) = 2 \text{ ptos} + 5 \text{ ptos de aprobado} = 7 \text{ ptos}$

Excelente = $70\% \times (10 \text{ ptos} - 5 \text{ ptos}) = 3,5 \text{ ptos} + 5 \text{ ptos de aprobado} = 8,5 \text{ ptos}$

Matrícula de Honor = $100\% \times (10 \text{ ptos} - 5 \text{ ptos}) = 5 \text{ ptos} + 5 \text{ ptos de aprobado} = 10 \text{ ptos}$

Veamos lo propio en el Sistema Trias de Bes

Correspondencia entre calificaciones y notas / Sistema Trias de Bes

Mantenemos los porcentajes y los aplicamos al nuevo tramo que va del 6,25 al 10.

Aprobado = 6,25 ptos

Bien = $20\% \times (10 \text{ ptos} - 6,25 \text{ ptos}) = 0,75 \text{ ptos} + 6,25 \text{ ptos de aprobar} = 7 \text{ ptos}$

Notable = $40\% \times (10 \text{ ptos} - 6,25 \text{ ptos}) = 1,5 \text{ ptos} + 6,25 \text{ ptos de aprobar} = 7,75 \text{ ptos}$

Excelente = $70\% \times (10 \text{ ptos} - 6,25 \text{ ptos}) = 2,625 \text{ ptos} + 6,25 \text{ ptos de aprobar} = 8,875 \text{ ptos}$

Matrícula de Honor = $100\% \times (10 \text{ ptos} - 6,25 \text{ ptos}) = 3,75 \text{ ptos} + 6,25 \text{ ptos de aprobar} = 10 \text{ ptos}$

4. Otros comentarios

Más allá de las dos alternativas aquí propuestas, existen un sinfín de opciones que enriquecerían los exámenes tipo test y abren nuevas posibilidades. Me gustaría destacar tres:

a) La primera es que, en lugar de cuatro opciones (a, b, c, d) se ofrezcan hasta cinco, seis o incluso siete posibles respuestas. Con siete opciones a marcar, a pesar de ser más trabajo para el examinador el preparar la prueba y más tiempo para el alumno responderla, la probabilidad de obtención de puntos aleatorios se reduce muchísimo.

En el caso de siete posibles respuestas (a, b, c, d, e, f, g), la probabilidad del chimpancé baja al 14,3%, con lo que ya parte con 1,43 puntos de salida en

lugar de 2,5. Su esperanza matemática es acertar ya solamente entre cinco y seis preguntas de las 40.

Sin embargo, el problema de que hay puntos aleatorios al alcance de la mano del alumno persiste. Podría igualmente aplicarse el Sistema Trias de Bes, elevando el aprobado. Si seguimos los mismos cálculos, el aprobado pasaría de un 5 a un 5,7. En tal caso, siendo escasa la diferencia, algunos profesores desistirían ya de tan nimia resta y considerarían equiparables ambos sistemas.

b) La segunda opción es la de que en un sistema de 4 respuestas posibles, la correcta sea una de las cuatro, dos de las cuatro, tres de las cuatro o incluso todas sean correctas (omito el que todas sean falsas, cosa que se suele aplicar añadiendo una opción que especifique “ninguna de las anteriores”). En tal caso, la probabilidad de que el chimpancé acierte una pregunta baja en picado.

El cálculo es el siguiente:

Hay 14 posibles respuestas correctas:

Solo la a
Solo la b
Solo la c
Solo la d
La a y la b
La a y la c
La a y la d
La b y la c
La b y la d
La c y la d
La a, la b y la c
La a, la c y la d
La b, la c y la d
La a, la b, la c y la d

Por lo que la probabilidad es ya solo del 7,1% (1 posibilidad entre las 14). El chimpancé acertaría solo entre 2 y 3 de las 40 preguntas y el alumno partiría ya solo con 0,07 puntos de ventaja. El Sistema Trias de Bes establecería situar el aprobado en el 5,36. En definitiva, podría ya dejarse el aprobado en el 5 sin necesidad de restar.

El inconveniente de este tipo de exámenes es que alumnos que hayan estudiado suelen cometer muchos errores porque cada cuestión se multiplica en realidad por cuatro. Todo puede ser cierto. Es un sistema que premia a los que se saben la lección al dedillo.

c) La última opción, que ha compartido conmigo un médico que estudió en el extranjero, mientras realizaba el trabajo de campo e investigación previa a la escritura de este artículo, es la de que hayan muchas preguntas (unas 500, por ejemplo), y que no todas puntúen igual. Las fáciles producen menos puntos que las difíciles, que proporcionan muchos más puntos. En las fáciles hay

cuatro opciones a escoger y en las difíciles pueden darse respuestas combinadas como las del epígrafe anterior.

Eso requiere mucho tiempo de preparación del test por parte del examinador pero permite diseños muy precisos en que se puede discernir muy bien en qué grado de conocimientos se sitúa el "aprobado", el "bien", "notable" o "excelente".

En estos casos, la resta no se considera necesaria porque el sistema de puntuación y dificultad de las preguntas ya se establece para erradicar la aleatoriedad de sacar buena nota sin haber estudiado.

5. Conclusiones

De todos modos, estas tres alternativas no invalidan el debate sobre "restar" o "no restar". Porque en cualquiera de ellos, en mayor o menor medida, existe aleatoriedad.

Por tanto, se mantiene vigente la propuesta de mi artículo: no restar y elevar el aprobado en la misma medida que la aleatoriedad (esperanza matemática) produce puntos al alumno. Se solventa la ventaja que el sistema de preguntas cerradas ofrece y dejamos de desincentivar la audacia.

Finalmente, para quienes piensen que la solución sea: "que no se hagan exámenes tipos test", aclarar que, en algunos casos, especialmente cuando se trata de medir el conocimiento del alumno en temarios muy extensos (por ejemplo, el examen del MIR), el número de personas que se examinan es muy elevado o se desea la máxima objetividad, el sistema tipo test es mejor que el de preguntas abiertas.

Se trata, únicamente, de establecer un sistema de evaluación justo y que no castigue los errores.

Recuerde contestar a la consulta popular sobre este asunto en este link:

https://docs.google.com/forms/d/13p60XB3XwZJhEHvQxxNwEEu9yBFE2D_3n9pCk-UMB30/viewform

que, a su vez, puede enviar a otras personas para que participen.

Puede encontrar el artículo original de La Vanguardia en www.triasdebes.net o en twitter opinar y participar a través de @ftriasdebes #restarnosuma

Fernando Trías de Bes