SUMA 27
febrero 1998, pp. 81-90

# Análisis probabilístico del sorteo de los excedentes de cupo. ¿Fue justo el proceso en su conjunto?

Roberto Marcellán Bueno

ARTÍCULOS

Este artículo tiene por objeto analizar desde un punto de vista probabilístico el sorteo de los excedentes de cupo celebrado el día 12 de noviembre de 1997. Para ello se utilizan conceptos que figuran entre los contendios del 2.° ciclo de la ESO y del Bachillerato, como son el modelo de asignación de probabilidades de Laplace, el concepto de probabilidad condicionada y el teorema de la probabilidad total. Así mismo se contrastan los resultados obtenidos con algunos análisis que aparecieron en la prensa en los días posteriores al sorteo, haciendo ver las numerosas incorrecciones cometidas en estos análisis.

EL PASADO miércoles 12 de noviembre se celebró el sorteo para determinar quiénes de entre los 165.342 jóvenes llamados a realizar el servicio militar en el año 1998 resultaban excedentes de cupo.

En total 16.442 mozos serían los afortunados que conseguirían dar esquinazo a los preceptivos nueve meses de disciplina castrense.

He de confesar que yo era uno de los que aguardaban, expectante, a que el azar dictaminase su veredicto. Lo que en ese momento desconocía era que el número que me representaba en ese sorteo me otorgaba casi la mitad de posibilidades de resultar excedente de cupo respecto a otros compañeros con números más afortunados.

Es mi propósito con este artículo justificar el porqué de la anterior afirmación. Para ello me voy a servir de conceptos probabilísticos que aparecen entre los contenidos de la ESO como son el modelo de asignación de probabilidades de Laplace y el concepto de probabilidad condicionada, así como del teorema de la probabilidad total que aparece sólo en el Bachillerato.

Asimismo me planteo dar un repaso al tratamiento del que fue objeto este evento por parte de los medios de comunicación. Por ello he seleccionado algunas de las informaciones que aparecieron en la prensa en los días posteriores al sorteo. Espero poder hacer ver al amable lector o lectora que la mayor parte de los contenidos de esos artículos no se ajusta a la verdad, ni en lo referente a los resultados que se aportan ni a los razonamientos empleados. Pretendo con ello, como única intención, fomentar una de las actitudes que vienen prescritas en el DCB de Matemáticas, concretamente en el bloque correspondiente al tratamiento del azar. Tal actitud es la de «valorar de forma crítica las informaciones probabilísticas en los medios de comunicación, rechazando los abusos y usos incorrectos de las mismas».

Previamente a iniciar el análisis, quizás conviniera recordar cuál fue el proceso global que se siguió para elegir a los excedentes de cupo. Días antes de la celebración del sorteo, cada uno de los llamados a filas recibimos un número entre el 1 y el 165.342. Dicho número nos fue asignado de forma equiprobable mediante un programa informático. En el sorteo, se eligió un número entre el 1 y el 165.342 mediante la extracción de seis bolas de seis bombos, la primera correspondiente a las centenas de millar y la última a las unidades. El primero de los bombos, el de las centenas de millar, contenía diez bolas, cinco con el número 0 y cinco con el número 1. El resto de los bombos contenían diez bolas numeradas del 0 al 9. Resultarían excedentes de cupo todos aquellos mozos cuyo número coincidiese o bien con el extraído, o bien con cualquiera de los 16.441 siguientes, de forma que si se rebasaba el 165.342 se empezaba a contar de nuevo por el 1.

Centrémonos ahora en glosar el procedimiento que dirigió el desarrollo del sorteo con los bombos. Éste consistió en extraer las bolas de forma secuencial, empezando por las centenas de millar y. acabando por las unidades, de forma que, si una determinada extracción daba lugar a una cifra que no permitiera que el número resultante estuviese comprendido entre el 1 y el 165.342, se repetía la extracción hasta dar con una cifra permisible.

Así ocurrió en realidad. La primera cifra resultó ser un 1. La segunda fue un 8. Como no tenían sentido números de la forma 18#. ###, esa deficiencia se subsanó volviendo a introducir la bola en el bombo y repitiendo la extracción, resultando elegida en esta segunda tentativa la cifra 5.

## El modelo probabilístico

Analicemos en profundidad el modelo probabilístico (al que llamaré *modelo A*) subyacente a este forma de elegir un número entre el 1 y el 165.342. Para ello se consideran las siguientes familias de sucesos asociados a la extracción de un número de seis cifras mediante los seis bombos:

$(A_i)_{i=0}^{1}$ donde $A_i$: «El bombo de las centenas de millar determina la cifra *i*», $0 \le i \le 1$

$(B_i)_{i=0}^{9}$ donde $B_i$: «El bombo de las decenas de millar determina la cifra *i*», $0 \le i \le 9$

$(C_i)_{i=0}^{9}$ donde $C_i$: «El bombo de las unidades de millar determina la cifra *i*», $0 \le i \le 9$

$(D_i)_{i=0}^{9}$ donde $D_i$: «El bombo de las centenas determina la cifra *i*», $0 \le i \le 9$

$(E_i)_{i=0}^{9}$ donde $E_i$: «El bombo de las decenas determina la cifra *i*», $0 \le i \le 9$

$(F_i)_{i=0}^{9}$ donde $F_i$: «El bombo de las unidades determina la cifra *i*», $0 \le i \le 9$

Planteémonos, por ejemplo, cuál es la probabilidad de que el número elegido hubiera sido el 1:

$$P(1) = P(A_0 \cap B_0 \cap C_0 \cap D_0 \cap E_0 \cap F_1) =$$
$$P(A_0)\cdot P(B_0 \mid A_0)\cdot P(C_0 \mid A_0 \cap B_0)\cdot P(D_0 \mid A_0 \cap B_0 \cap C_0)\cdot P(E_0 \mid A_0 \cap B_0 \cap C_0 \cap D_0)$$
$$\cdot P(F_1 \mid A_0 \cap B_0 \cap C_0 \cap D_0 \cap E_0)$$

donde en la segunda igualdad se ha usado, de forma reiterada, el concepto de probabilidad condicionada.

Evidentemente, $P(A_0) = \frac{5}{10} = \frac{1}{2}$

Para esta asignación de probabilidad, así como para las que siguen, se ha hecho uso de la regla de Laplace.

Si la primera cifra ha sido 0, en el segundo bombo puede salir cualquier cifra entre 0 y 9, y todas tienen sentido, luego:

$$P(B_0 \mid A_0) = \frac{1}{10}$$

Análogamente se deducen:

$$P(C_0 \mid A_0 \cap B_0) = \frac{1}{10}$$
$$P(D_0 \mid A_0 \cap B_0 \cap C_0) = \frac{1}{10}$$
$$P(E_0 \mid A_0 \cap B_0 \cap C_0 \cap D_0) = \frac{1}{10}$$

Ahora, si las cinco primeras cifras han sido 0, tendrán sentido como unidades cualquier cifra entre 1 y 9, pero no el 0, pues el menor número posible es el 1, luego:

$$P(F_1 \mid A_0 \cap B_0 \cap C_0 \cap D_0 \cap E_0) = \frac{1}{9}$$

Así:

$$P(1) = \frac{1}{2}\cdot\frac{1}{10}\cdot\frac{1}{10}\cdot\frac{1}{10}\cdot\frac{1}{10}\cdot\frac{1}{9} = \frac{1}{180.000}$$

Calculemos ahora la probabilidad de que el número elegido hubiese sido el último posible, es decir, el 165.342.

$$P(165.342) = P(A_1 \cap B_6 \cap C_5 \cap D_3 \cap E_4 \cap F_2) =$$
$$P(A_1)\cdot P(B_6 \mid A_1)\cdot P(C_5 \mid A_1 \cap B_6)\cdot P(D_3 \mid A_1 \cap B_6 \cap C_5)\cdot P(E_4 \mid A_1 \cap B_6 \cap C_5 \cap D_3)$$
$$\cdot P(F_2 \mid A_1 \cap B_6 \cap C_5 \cap D_3 \cap E_4)$$

Evidentemente $P(A_1) = \frac{5}{10} = \frac{1}{2}$

Ahora, si la primera cifra ha sido 1, la segunda tendrá que ser un número entre 0 y 6, pues el máximo número posible es de la forma 16#. ###. Así:

$$P(B_6 | A_1) = \frac{1}{7}$$

Si la primera cifra ha sido 1 y la segunda 6, entonces la tercera tendrá que ser un número entre 0 y 5, pues el máximo número posible es de la forma 165. ###. Así :

$$P(C_5 | A_1 \cap B_6) = \frac{1}{6}$$

Si la primera cifra ha sido 1, la segunda 6 y la tercera 5, entonces la cuarta tendrá que ser un número entre 0 y 3, pues el máximo número posible es de la forma 165.3##. Así:

$$P(D_3 | A_1 \cap B_6 \cap C_5) = \frac{1}{4}$$

Análogamente:

$$P(E_4 | A_1 \cap B_6 \cap C_5 \cap D_3) = \frac{1}{5}$$

$$P(F_2 | A_1 \cap B_6 \cap C_5 \cap D_3 \cap E_4) = \frac{1}{3}$$

Por tanto:

$$P(165.342) = \frac{1}{2} \cdot \frac{1}{7} \cdot \frac{1}{6} \cdot \frac{1}{4} \cdot \frac{1}{5} \cdot \frac{1}{3} = \frac{1}{5.040}$$

Adoptando la misma estrategia que nos ha permitido calcular $P(1)$ y $P(165.342)$ podemos calcular la probabilidad de cualquier otro número. Así, se obtiene:

| Si | Probabilidad |
|---|---|
| $1 \leq n \leq 9$ | $\frac{1}{2} \cdot \frac{1}{10} \cdot \frac{1}{10} \cdot \frac{1}{10} \cdot \frac{1}{10} \cdot \frac{1}{9} = \frac{1}{180.000} \approx 5{,}56 \cdot 10^{-6}$ |
| $10 \leq n \leq 99.999$ | $\frac{1}{2} \cdot \frac{1}{10} \cdot \frac{1}{10} \cdot \frac{1}{10} \cdot \frac{1}{10} \cdot \frac{1}{10} = \frac{1}{200.000} \approx 5 \cdot 10^{-6}$ |
| $100.000 \leq n \leq 159.999$ | $\frac{1}{2} \cdot \frac{1}{7} \cdot \frac{1}{10} \cdot \frac{1}{10} \cdot \frac{1}{10} \cdot \frac{1}{10} = \frac{1}{140.000} \approx 7{,}14 \cdot 10^{-6}$ |
| $160.000 \leq n \leq 164.999$ | $\frac{1}{2} \cdot \frac{1}{7} \cdot \frac{1}{6} \cdot \frac{1}{10} \cdot \frac{1}{10} \cdot \frac{1}{10} = \frac{1}{84.000} \approx 11{,}91 \cdot 10^{-6}$ |
| $165.000 \leq n \leq 165.299$ | $\frac{1}{2} \cdot \frac{1}{7} \cdot \frac{1}{6} \cdot \frac{1}{4} \cdot \frac{1}{10} \cdot \frac{1}{10} = \frac{1}{33.600} \approx 29{,}76 \cdot 10^{-6}$ |
| $165.300 \leq n \leq 165.339$ | $\frac{1}{2} \cdot \frac{1}{7} \cdot \frac{1}{6} \cdot \frac{1}{4} \cdot \frac{1}{5} \cdot \frac{1}{10} = \frac{1}{16.800} \approx 59{,}52 \cdot 10^{-6}$ |
| $165.340 \leq n \leq 165.342$ | $\frac{1}{2} \cdot \frac{1}{7} \cdot \frac{1}{6} \cdot \frac{1}{4} \cdot \frac{1}{5} \cdot \frac{1}{3} = \frac{1}{5.040} \approx 198{,}41 \cdot 10^{-6}$ |

Tabla 1

El siguiente gráfico nos puede dar una idea de la descompensación existente:

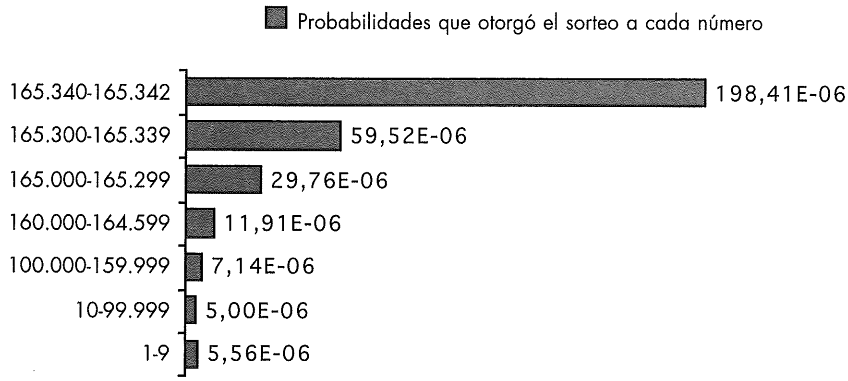



Gráfico 1

Téngase en cuenta que si el sorteo hubiese sido equiprobable, todos los números hubieran tenido como probabilidad:

$$P(n) = \frac{1}{165.342} \approx 6{,}05 \cdot 10^{-6}$$

Podemos observar que un número n ∈ {165.340, 165.341, 165.342} era 40 veces más probable que otro cualquiera entre 10 y 99.999. Asimismo, el número que resultó elegido, 155.611, fue uno de los «privilegiados» por el sorteo, aunque no de los que más.

Contrastemos la distribución de probabilidades recogida en la tabla 1 con la que se puede ver en la información (cuadro 3) según la cual las probabilidades que el sorteo asignó a cada número fueron:

| Si | Probabilidad |
|---|---|
| $1 \leq n \leq 99.999$ | $\frac{1}{2} \cdot \frac{1}{99.999} = \frac{1}{199.998}$ |
| $100.000 \leq n \leq 165.342$ | $\frac{1}{2} \cdot \frac{1}{65.342} = \frac{1}{130.684}$ |

Tabla 2

Obsérvese como este modelo de asignación de probabilidades no se ajusta al mecanismo del sorteo. De hecho dicha asignación es doblemente incorrecta, pues, aun adoptando el punto de vista (erróneo) del autor, es evidente que:

$$P(n) = \frac{1}{2} \cdot \frac{1}{65.343}, \text{si } 100.000 \leq n \leq 165.342$$

Volviendo a nuestro estudio merece la pena resaltar que, *aun cuando el primer bombo hubiera tenido la proporción justa de ceros y unos* —en orden a representar correctamente el distinto número de guarismos de la forma 0##.### respecto de los de la forma 1##.###—, *el* *sorteo hubiera sido sesgado*, pues el único cambio en este nuevo supuesto consistiría en sustituir 1/2, o bien por 99.999/165.342 (si se trata de calcular la probabilidad de un número que comience por 0), o bien por 65.343/165.342 (si se trata de calcular la probabilidad de un número que comience por 1), manteniéndose invariantes el resto de factores (véase la tabla 1). Dicho de otra forma, el motivo por el que el sorteo no fue equitativo no fue la inadecuada proporción de ceros y unos en el primer bombo, sino el hecho de que si en una determinada extracción resultaba elegida una cifra que no diera lugar a un número entre el 1 y el 165.342, se optara por reponer la bola extraída y repetir la extracción. Ésta, y no otra, fue la causa de la no equiprobabilidad del sorteo.

Sin embargo, muchas informaciones parecen reflejar como motivo de la no equiprobabilidad del sorteo la inadecuada proporción de ceros y unos en el primer bombo. Véase si no el texto extraido de *El Periódico de Aragón*, del 14 de noviembre, que aparece bajo estas líneas. Obsérvese como se dan proporciones alternativas para el primer bombo, como si la raíz del problema estuviera ahí. Ninguno de los dos cambios que se proponen hubiera hecho equiprobable el sorteo, como es fácil comprobar razonando de la misma forma que se ha hecho para calcular las probabilidades de la tabla 1.

## La probabilidad de cada mozo

Retomemos nuestro análisis para ver cómo repercutió el desequilibrio probabilístico del sorteo con los bombos en la probabilidad de que el mozo representado con el número $n$, $1 \le n \le 165.342$, resultara elegido excedente de cupo.

Considero la siguiente familia de sucesos:

$$\left(EXC\_n\right)_{n=1}^{165342}$$

donde $EXC\_n$ es el suceso «El mozo con número $n$ resulta excedente de cupo».

Comencemos calculando $P(EXC\_1)$. Para que el mozo con el número 1 resultara excedente de cupo, tal y como se ha explicado al comienzo, en el sorteo debería salir un número entre el 148.902 y el 165.342 (ambos inclusive), o bien el 1. Por tanto, a partir de las probabilidades calculadas en la tabla 1, deducimos que:

$$P(EXC\_1) = \frac{11.098}{140.000} + \frac{5.000}{84.000} + \frac{300}{33.600} + \frac{40}{16.800} + \frac{3}{5.040} + \frac{1}{180.000} = \frac{27.127}{180.000} \approx 0{,}150706$$

Cuadro 1

# El bombo catea las matemáticas

## Los expertos descalifican el sistema de Defensa porque provocó desigualdades

EL PERIÓDICO
**Madrid**

Los matemáticos y expertos en estadística coinciden en descalificar el sorteo de excedentes de cupo. El gran problema es, señalan, el empleo de los bombos, que no garantiza la igualdad de oportunidades. El sistema llegó a ser calificado de **"chapuza"** desde el punto de vista matemático por José Luis Sagredo, profesor de Matemáticas de la Escuela de Ingenieros Aeronáuticos de Madrid.

Enrique González Arangüena, profesor titular de la Escuela Universitaria de Estadística, agregó que existen motivos para que los mozos acudan a un tribunal para recurrir el sistema de elección empleado. **"Yo les daría la razón"**, afirmó tajantemente.

El método utilizado por el Ministerio de Defensa **"no es el más apropiado"**, apuntó. En opinión de González Arangüena, hay muchos procedimientos que se hubieran ajustado más a la igualdad para todos los jóvenes. **"Cualquiera que tenga una mínima idea de la teoría de la probabilidad podía haber pensado en una generación de números aleatorios"**, dijo.

El **"delicado"** procedimiento de bombos hubiera sido **"menos erróneo"** si en el primer bombo, hubieran puesto cinco 0 y tres 1, **"de tal modo que las cifras superiores a 100.000 no tuvieran un plus de probabilidad"**. Otro método barajado por Arangüena hubiera sido poder marcar en el primer bombo los dos primeros dígitos. En este caso, habría bastado con colocar bolas con los números de 0 a 16 para marcar las decenas de miles de los 16.442 jóvenes eximidos de hacer la mili.

El profesor de matemáticas de la Escuela de Ingenieros Aeronáuticos de Madrid, José Luis Sagredo, consideró que el sistema empleado por Defensa **"es una chapuza"** en términos matemáticos, ya que la franja que va desde el número 116.442 hasta el 165.342 cuenta con un 35% más de posibilidades de quedar exentos que el grupo comprendido entre el 16.000 y el 99.999. Sagredo advirtió que además hay un grupo intermedio –del 1 al 16.4419–, en el que cada mozo tiene una **"probabilidad distinta a otro dentro de su misma franja"**. **"Es curioso que el sorteo se haya decantado en el grupo donde más posibilidades había"**, concluyó.

En el sorteo se utilizaron seis bombos, uno para cada dígito. En la primera esfera, un soldado profesional introdujo 10 bolas: cinco con el número 0 y otras cinco con el número 1. De ahí nació el primer desequilibrio en el cálculo de probabilidades, ya que los mozos con número a partir de 100.000 eran 65.342 y, al ser menos que los 99.999 que les antecedían, tenían más posibilidades de quedar exentos.

Para corregir esta desigualdad se debió introducir en ese primer bombo de las centenas de millar tres bolas con el número 1, en vez de cinco.

Otra posibilidad hubiera sido la utilización de un solo bombo para determinar los dos primeros dígitos –correspondientes a las centenas y a las decenas de millar– con una secuencia de bolas del 0 al 16.

| Si | P(EXC_n) |
|---|---|
| $1 \le n \le 9$ | $\left(\frac{11.098}{140.000}+\frac{5.000}{84.000}+\frac{300}{33.600}+\frac{40}{16.800}+\frac{3}{5.040}+\frac{1}{180.000}\right)-(n-1)\cdot\left(\frac{1}{140.000}-\frac{1}{180.000}\right)=\frac{27.127}{180.000}-(n-1)\cdot\frac{1}{630.000}$ |
| $10 \le n \le 11.098$ | $\left(\frac{11.089}{140.000}+\frac{5.000}{84.000}+\frac{300}{33.600}+\frac{40}{16.800}+\frac{3}{5.040}+\frac{9}{180.000}+\frac{1}{200.000}\right)-(n-10)\cdot\left(\frac{1}{140.000}-\frac{1}{200.000}\right)=$ $=\frac{210.967}{1.400.000}-(n-10)\cdot\frac{3}{1.400.000}$ |
| $11.099 \le n \le 16.098$ | $\left(\frac{5.000}{84.000}+\frac{300}{33.600}+\frac{40}{16.800}+\frac{3}{5.040}+\frac{9}{180.000}+\frac{11.089}{200.000}+\frac{1}{200.000}\right)-(n-11.099)\cdot\left(\frac{1}{84.000}-\frac{1}{200.000}\right)=$ $=\frac{1.777}{14.000}-(n-11.099)\cdot\frac{29}{4.200.000}$ |
| $16.099 \le n \le 16.398$ | $\left(\frac{300}{33.600}+\frac{40}{16.800}+\frac{3}{5.040}+\frac{9}{180.000}+\frac{16.089}{200.000}+\frac{1}{200.000}\right)-(n-16.099)\cdot\left(\frac{1}{33.600}-\frac{1}{200.000}\right)=$ $=\frac{3.881}{42.000}-(n-16.099)\cdot\frac{13}{525.000}$ |
| $16.399 \le n \le 16.438$ | $\left(\frac{40}{16.800}+\frac{3}{5.040}+\frac{9}{180.000}+\frac{16.389}{200.000}+\frac{1}{200.000}\right)-(n-16.399)\cdot\left(\frac{1}{16.800}-\frac{1}{200.000}\right)=\frac{3.569}{42.000}-(n-16.399)\cdot\frac{229}{4.200.000}$ |
| $16.439 \le n \le 16.441$ | $\left(\frac{3}{5.040}+\frac{9}{180.000}+\frac{16.429}{200.000}+\frac{1}{200.000}\right)-(n-16.439)\cdot\left(\frac{1}{5.040}-\frac{1}{200.000}\right)=\frac{17.387}{210.000}-(n-16.439)\cdot\frac{2.437}{12.600.000}$ |
| $16.442 \le n \le 16.450$ | $\left(\frac{9}{180.000}+\frac{16.432}{200.000}+\frac{1}{200.000}\right)-(n-16.442)\cdot\left(\frac{1}{180.000}-\frac{1}{200.000}\right)=\frac{16.443}{200.000}-(n-16.442)\cdot\frac{1}{1.800.000}$ |
| $16.451 \le n \le 99.999$ | $\frac{16.442}{200.000}$ |
| $16.451 \le n \le 99.999$ | $\left(\frac{16.441}{200.000}+\frac{1}{140.000}\right)-(n-100.000)\cdot\left(\frac{1}{140.000}-\frac{1}{200.000}\right)=\frac{115.097}{1.400.000}-(n-100.000)\cdot\frac{3}{1.400.000}$ |
| $16.451 \le n \le 99.999$ | $\frac{16.442}{140.000}$ |
| $16.451 \le n \le 99.999$ | $\left(\frac{16.441}{140.000}+\frac{1}{84.000}\right)-(n-160.000)\cdot\left(\frac{1}{84.000}-\frac{1}{140.000}\right)=\frac{3.083}{26.250}-(n-160.000)\cdot\frac{1}{210.000}$ |
| $16.451 \le n \le 99.999$ | $\left(\frac{5.000}{84.000}+\frac{11.441}{140.000}+\frac{1}{33.600}\right)-(n-165.000)\cdot\left(\frac{1}{33.600}-\frac{1}{140.000}\right)=\frac{5.651}{40.000}-(n-165.000)\cdot\frac{19}{840.000}$ |
| $16.451 \le n \le 99.999$ | $\left(\frac{300}{33.600}+\frac{5.000}{84.000}+\frac{11.141}{140.000}+\frac{1}{16.800}\right)-(n-165.300)\cdot\left(\frac{1}{16.800}-\frac{1}{140.000}\right)=\frac{31.099}{210.000}-(n-165.300)\cdot\frac{11}{210.000}$ |
| $16.451 \le n \le 99.999$ | $\left(\frac{40}{16.800}+\frac{300}{33.600}+\frac{5.000}{84.000}+\frac{11.101}{140.000}+\frac{1}{5.040}\right)-(n-165.340)\cdot\left(\frac{1}{5.040}-\frac{1}{140.000}\right)=\frac{189.409}{1.260.000}-(n-165.340)\cdot\frac{241}{1.260.000}$ |

Tabla 3

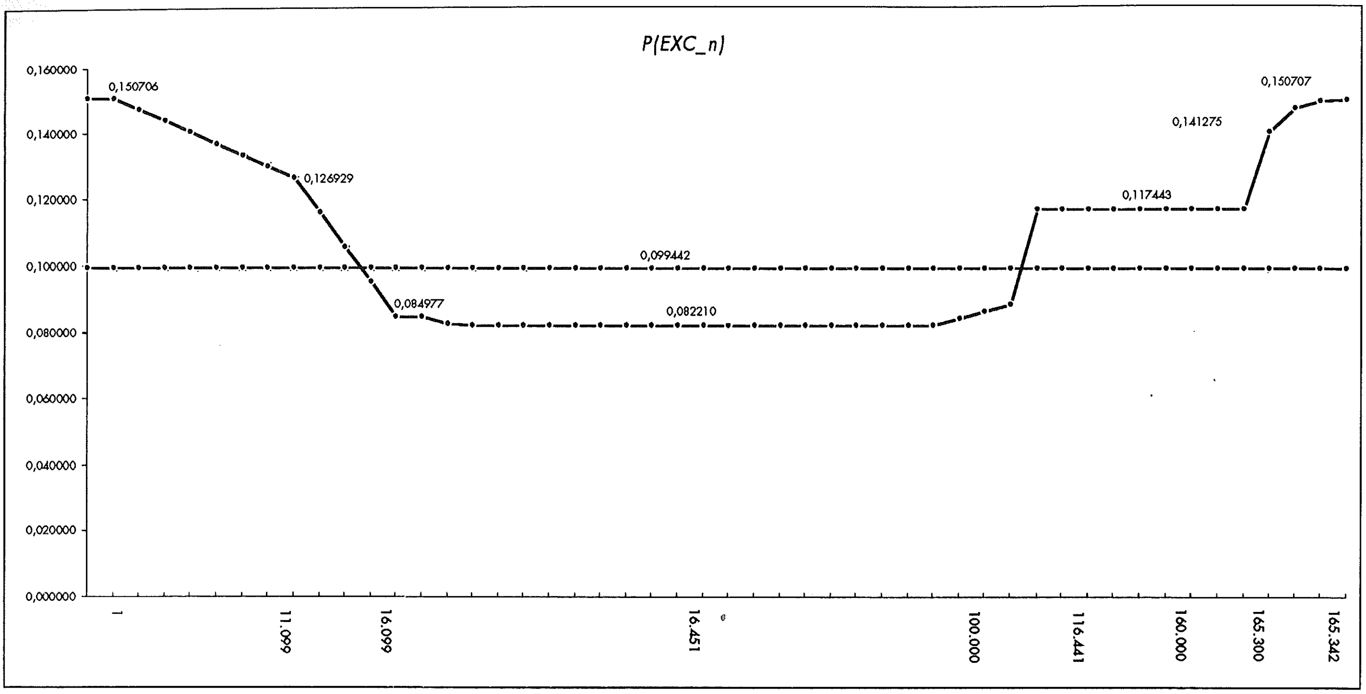



Gráfico 2

Procediendo con un poco de paciencia, es posible determinar las probabilidades de cada uno de los sucesos *EXC_n*. En la tabla 3 podemos seguir el cálculo detallado de dichas probabilidades.

Para hacernos una idea de cómo varía *P(EXC_n)* según *n* podemos obrservar el gráfico 2.

En dicho gráfico, la línea horizontal determina la situación que se hubiera dado si el sorteo de los excedentes de cupo hubiese sido equiprobable, es decir, si todos los números otorgasen a su poseedor la misma probabilidad de salir excedente de cupo:

$$P(EXC\_n)=\frac{16.442}{165.342}\approx 0{,}099442$$

El número que más probabilidades proporcionó a su poseedor para ser elegido excedente de cupo fue el 165.342, con:

$$P(EXC\_165.342)=\frac{21.099}{140.000}\approx 0{,}150707$$

Por el contrario, los números que menos probabilidades otorgaron a sus dueños —entre los que se encontraba, por desgracia, el que me representaba a mí—, fueron los comprendidos entre el 16.451 y el 99.999, con:

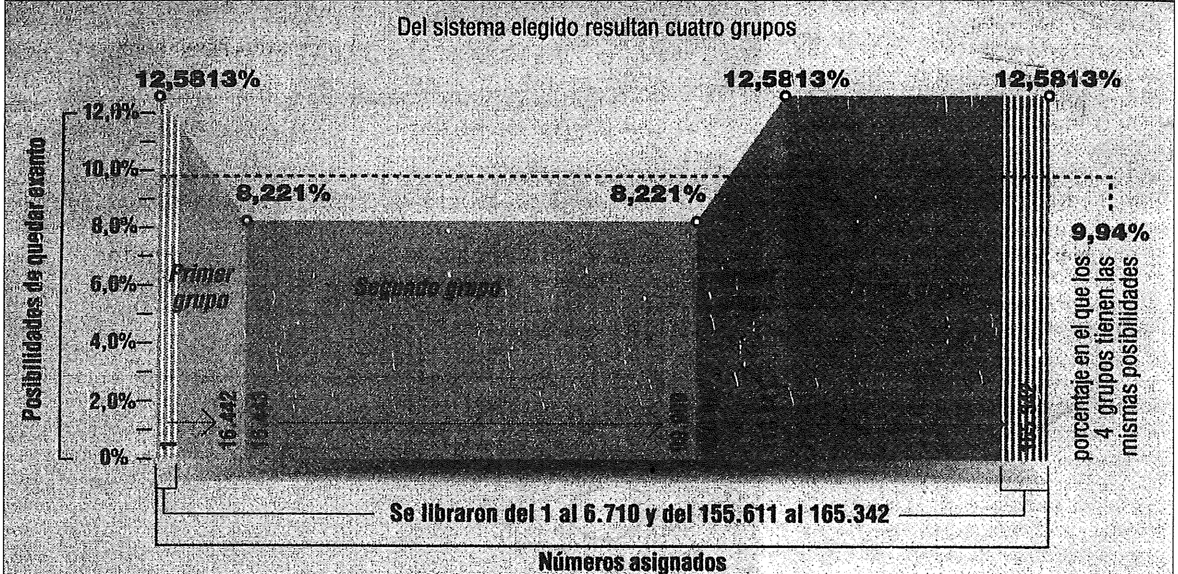
$$P(EXC\_j)=\frac{16.442}{200.000}; \qquad \text{donde } 16.451\leq j\leq 99.999$$

*De hecho, el resultado del sorteo vino a refrendar este desequilibrio probabilístico...*

Puede comprobarse que los números comprendidos entre el 15.080 y el 108.040 (ambos inclusive) tuvieron menos probabilidades que las que les hubiera asignado un sorteo justo. De hecho, el resultado del sorteo vino a refrendar este desequilibrio probabilístico pues, a la postre, resultaron elegidos los números comprendidos entre el 155.611 y el 165.342, o bien entre el 1 y el 6.710, todos ellos beneficiados por la no equiprobabilidad del sorteo.

Retornemos de nuevo al análisis de algunas noticias aparecidas en la prensa. Merece resaltarse el contraste entre el gráfico 2 y los que aparecen en las informaciones de *El Mundo* y de *El País* del día 14 de noviembre (ver cuadros 2 y 3). Frente a las 14 variaciones de pendiente que podemos observar en la tabla 3 y que manifiesta la línea que une los puntos representando las probabilidades que cada número otorgó a su poseedor para ser elegido excedente de cupo, ambos gráficos únicamente muestran 4. Evidentemente el error en dichos gráficos proviene de la inadecuada asignación de probabilidades ya comentada (véase la tabla 2).

## Ventajas y desventajas en el sorteo




El segundo grupo es el que tiene menos posibilidades.

$$\frac{1}{2} \times \frac{16.442}{99.999} = 0{,}08221 \rightarrow 8{,}221\%$$

El cuarto grupo es el que tiene más posibilidades.

$$\frac{1}{2} \times \frac{16{,}442}{65{,}343} = 0{,}125813 \rightarrow 12{,}5813\%$$

La razón aritmética de estas progresiones decreciente la primera y creciente la tercera es el 0.0002651

Los 48.901 jóvenes del 4º grupo tienen un 53% más de probabilidades de salir excedentes de cupo que los 83.557 jóvenes del 2º grupo.

Documentación: Leandro Alvarez Villalón

EL MUNDO

# El sorteo aleatorio previo no elimina la injusticia

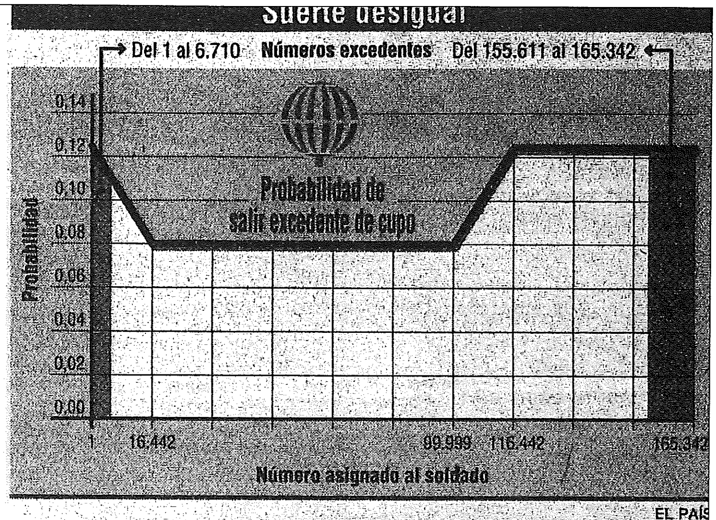
El sorteo para decidir los excedentes de cupo en el servicio militar fue claramente irregular, de manera que unos jóvenes fueron beneficiados frente a otros.

◆ Defensa se acoge al sorteo previo de los mozos para sostener que la lotería posterior dio las mismas oportunidades a todos. Destaca que realizó anteriormente un reparto aleatorio de un *décimo* a cada recluta con el que concursaba después y que este hecho igualaba oportunidades.
Sin embargo, su justificación no es matemáticamente cierta. Pese a que el orden de reparto de números sí que fue puro (un ordenador secuenció todas las combinaciones posibles con los 165.000 reclutas y se escogió una de las mismas al azar) el hecho de que los boletos no tuvieran todos la misma oportunidad sigue siendo insoslayable: en un ejemplo extremo, para que a uno le toque la lotería, su décimo debe entrar en el sorteo; si esto no ocurre, o sus oportunidades de salir son menores, se produce un desequilibrio, aunque el nombre y apellido del damnificado sea escogido al azar.

◆ El gráfico demuestra que los mozos que sortearon podían encontrarse en cuatro grupos distintos de posibilidades. El segmento más perjudicado por las cinco bolas con un 0 y cinco con un 1 es el que se encuentra entre los números 16.443 y 99.999, mientras que el parcial más beneficiado se encuentra entre el 116.442 y el 165.342.

◆ Como resultado, a consecuencia de su colocación aleatoria en la lista, el grupo cuarto tiene un 53% más de probabilidades de salir excedente de cupo que los integrados en el segundo apartado, mientras que el primero y el tercero se hallan en un estadio intermedio, con una desventaja variable en el propio segmento.

Cuadro 2




## Un problema con fácil solución

JAVIER PORTELA

Para realizar correctamente el sorteo era necesario tomar un número entre el 000001 y el 165.342, asignando idéntica probabilidad a cada uno de ellos. Sin embargo, el sorteo se hizo tomando de un bombo el número 0 o 1, con igual probabilidad, escogiendo **después**, en un segundo bombo, un número del 0 al 9, de forma que el primer número escogido fue un 1 y el segundo un 8, por lo que se volvió a tomar una bola **del segundo bombo,** hasta que salió un número inferior o igual a seis. El error matemático radica en mantener fijo el resultado del primer bombo.

Como consecuencia, no se asignó la misma probabilidad de salir a todos los números entre el 1 y el 165.342. La probabilidad de que tocara el número concreto A, situado entre el 1 y el 99.999, era de un medio por 1 dividido 99.999, mientras que la probabilidad de que tocara el número concreto B, situado entre el 100.000 y el 165.342, era de un medio por 1 dividido 65.342.

Esto trae consecuencias sobre la probabilidad de salir excedente de cupo **en el caso particular de cada recluta**. Esta probabilidad varía según el número asignado al recluta, como se refleja en el gráfico. Es decir, los reclutas con número entre el 16.442 y el 99.999 tenían una probabilidad de salir excedente de cupo aproximadamente de 0,082, mientras que los situados en el tramo que va de 116.442 a 165.342 tenían una probabilidad aproximada de 0,1258, cerca del 50% más.

Para los reclutas situados entre el 1 y el 16.442, la probabilidad dependía linealmente del número asignado, de manera que el número 3 tenía más probabilidades que el 4. Lo mismo ocurría con el tramo de 100.000 a 116.342, pero inversamente: el recluta número 110.000 tenía menos probabilidades que el 110.101.

El sorteo estaría bien hecho si se hubieran tirado los 6 bombos seguidos, repitiendo todo el proceso de sorteo, incluido el primer bombo, si saliera un número superior a 165.342. La probabilidad de tener que repetir una vez el sorteo por esta causa es relativamente baja, de 0,173.

Si la asignación inicial de los números a los reclutas fue realmente aleatoria, no se ve la necesidad de un sorteo adicional: bastaba con declarar excedente de cupo a los 16.442 primeros.

---

**Javier Portela** es profesor de la Escuela de Estadística de la Universidad Complutense de Madrid.

Cuadro 3

### ¿Sorteo equitativo?

Llegados a este punto, estamos en situación de plantearnos la pregunta que a mí, como parte afectada, más me interesa. ¿Fue el proceso en su conjunto equitativo? Es decir. ¿Cualquier mozo tenía al inicio del mismo la misma probabilidad de ser excedente de cupo? Si logramos zafarnos del influjo de los cálculos anteriores, vamos a ver que sí.

*¿Cualquier mozo tenía al inicio del mismo la misma probabilidad de ser excedente de cupo?*

Recordemos que el proceso había comenzado con la asignación equiprobable a cada mozo de un número entre el 1 y el 165.342, y que, posteriormente, el sorteo determinó 16.442 números cuyos poseedores fueron los excedentes de cupo.

Ya hemos visto que hubo unos números que dotaron a sus dueños de más probabilidad que otros. Pero advirtamos que los procesos de asignación de un número aleatorio a cada mozo y de elección de los 16.442 números determinando los excedentes de cupo fueron procesos independientes, es decir, el resultado del primer proceso no tuvo ninguna influencia en el desarrollo del segundo. Por tanto, aun cuando el sorteo para la elección de los excedentes no otorgó las mismas probabilidades a todos los números, la probabilidad que tenía cada mozo de que le correspondiera uno de los 16.442 números posteriormente determinados por el sorteo era igual para todos:

$$\frac{16.442}{165.342}$$

pues la asignación de números a cada mozo sí fue equiprobable.

Podemos llegar a la misma conclusión razonando desde otro punto de vista, más formal. Para ello consideremos un mozo cualquiera, $X$, de los 165.342 posibles. Se trata de determinar la probabilidad del siguiente suceso:

$X\_exc$ := «X resulta excedente de cupo»

Consideremos ahora la siguiente familia de sucesos:

$$\left(N_i^X\right)_{i=1}^{165.342}$$

donde:

$N_i^X$ : «A $X$ le corresponde el número $i$ en la asignación aleatoria de números»

Aplicando el teorema de la probabilidad total obtenemos que:

$$P[X\_exc] = \sum_{i=1}^{165.342} P[X\_exc | N_i^X] \cdot P[N_i^X] =$$

$$= \frac{1}{165.342} \cdot \sum_{i=1}^{165.342} P[X\_exc | N_i^X]$$

Ahora nótese que el suceso $X\_exc | N_i^X$ coincide con el que antes hemos denotado por *EXC_i*.

Sustituyendo los valores de *P[EXC_i]* dados en la tabla 3, se comprueba que:

$$\sum_{i=1}^{165.342} P[X\_exc | N_i^X] = \sum_{i=1}^{165.342} P[EXC\_i] = 16.442$$

*De todo este estudio cabe concluir que ninguno de los que fuimos parte interesada en todo este asunto debe sentirse perjudicado por el proceso de elección de los excedentes de cupo.*

Así: $$P[X\_exc] = \frac{16.442}{165.342}$$

y esto es válido para cualquier mozo $X$.

De todo este estudio cabe concluir que ninguno de los que fuimos parte interesada en todo este asunto debe sentirse perjudicado por el proceso de elección de los excedentes de cupo. Puede hablarse de que unos números fueron más afortunados que otros pero no que hubiera, a priori, mozos con más probabilidad que otros de resultar excedentes de cupo.

No parece opinar lo mismo el autor de la información aparecida en *El Mundo* (ver cuadro 2), para quien la argumentación anterior «no es matemáticamente cierta».

## Disgresión final

Pese a que, juzgado desde el principio, el proceso de elección de los excedentes de cupo puede considerarse justo, no deja de ser lamentable el modo en que se realizó el sorteo con los bombos, de acuerdo a los fines que se perseguían: la elección de un número entre el 1 y el 165.342 de forma equiprobable. Tal propósito se hubiera conseguido, en la situación del sorteo, con el siguiente procedimiento: si en alguna de las seis extracciones, la cifra determinada diera lugar a un número no comprendido entre el 1 y el 165.342, entonces toda la secuencia de extracciones debe repetirse desde el principio, es decir desde el bombo de las centenas de millar.

¿Por qué este procedimiento garantiza la equiprobabilidad en la elección de un número entre 1 y 165.342? Por el siguiente motivo. Recuérdese que partimos de seis bombos, uno para cada cifra. El primero con diez bolas, cinco con el 0 y cinco con el 1. El resto con diez bolas numeradas del 0 al 9. En tal situación, lo que sí es evidente es que podemos elegir un número entre el 0 y 199.999 de forma equiprobable. En términos más técnicos diríamos que estamos en condiciones de simular una variable aleatoria uniforme discreta con rango {0, 199.999}. Ahora conviene hacer notar que la restricción de una variable aleatoria uniforme a cualquier subdominio de su rango vuelve a ser una variable aleatoria uniforme con rango dicho subdominio.

Así, con el procedimiento consistente en desechar, de todos los números que podían determinarse con los seis bombos ({0, 199.999}), aquéllos no comprendidos entre 1 y 165.342, lo que estamos haciendo es simular una variable uniforme discreta con rango {1, 165.342}, o lo que es lo mismo, estamos garantizando que todos los números entre 1 y 165.342 tienen la misma probabilidad de ser elegidos.

Denotemos al modelo probabilístico subyacente a este modo de proceder como *modelo B*.

A continuación me gustaría plantear la siguiente reflexión. Supóngase que el número elegido, 155.611, lo hubiera sido sin necesidad de repetir ninguna extracción (recuérdese que tuvo que repetirse la extracción de la segunda cifra, pues inicialmente salió un 8). En tal supuesto, hubiese sido imposible discernir si el modelo probabilístico que regía el sorteo era el *modelo A*, que como hemos visto no asignó las mismas probabilidades a todos los números, o bien era el *modelo B* que sí asigna igual probabilidad a todos los números. Es decir, no hubiésemos tenido evidencias para poder afirmar que el sorteo no era equiprobable. Por tanto, ante el aluvión de críticas que igualmente se hubiera producido, un responsable del Ministerio de Defensa habilidoso en cuestiones probabilísticas hubiese podido alegar que el modelo que regía el sorteo era el B (equiprobable) y no el A (no equiprobable), y nadie, a partir de un número elegido sin repetir ninguna extracción, hubiera podido demostrar lo contrario.

Lamentablemente para la reputación del Ministerio de Defensa, se repitió una de las extracciones con lo que quedó de manifiesto que el modelo probabilístico que gobernó el sorteo fue el *A* (no equiprobable) y no el *B* (equiprobable).

En este contexto, podemos preguntarnos por la probabilidad de que en el proceso de elección de las seis cifras de los seis bombos (según el *modelo A*), tuviera que repetirse al menos una de las extracciones. Uno de los procedimientos posibles para su cálculo es el siguiente. Debemos hallar la probabilidad del siguiente suceso:

*S*: «Se repite al menos una de las seis extracciones en la elección de un número de seis cifras con los seis bombos»

Recuérdense las familias de sucesos:

$$(A_i)_{i=0}^{1}, (B_i)_{i=0}^{9}, (C_i)_{i=0}^{9}, (D_i)_{i=0}^{9}, (E_i)_{i=0}^{9}, (F_i)_{i=0}^{9}$$

ya utilizadas en los cálculos iniciales.

A partir de teorema de la probabilidad total, se obtiene:

$$P[S] = P[S|A_0] \cdot P[A_0] + P[S|A_1] \cdot P[A_1] =$$

$$= 0 \cdot \frac{1}{2} + P[S|A_1] \cdot \frac{1}{2}$$

De nuevo por el teorema de la probabilidad total:

$$P[S|A_1] = P\left[S\middle|A_1 \cap \left(\bigcup_{i=1}^{5} B_i\right)\right] \cdot P\left[\bigcup_{i=1}^{5} B_i\right] + P[S|A_1 \cap B_6] \cdot P[B_6]$$

$$+ P\left[S\middle|A_1 \cap \left(\bigcup_{i=7}^{9} B_i\right)\right] \cdot P\left[\bigcup_{i=7}^{9} B_i\right] = 0 + P[S|A_1 \cap B_6] \cdot \frac{1}{10} + 1 \cdot \frac{3}{10}$$

Análogamente:

$$P[S|A_1 \cap B_6] = 0 + P[S|A_1 \cap B_6 \cap C_5] \cdot \frac{1}{10} + 1 \cdot \frac{4}{10}$$

$$P[S|A_1 \cap B_6 \cap C_5] = 0 + P[S|A_1 \cap B_6 \cap C_5 \cap D_3] \cdot \frac{1}{10} + 1 \cdot \frac{6}{10}$$

$$P[S|A_1 \cap B_6 \cap C_5 \cap D_3] = 0 + P[S|A_1 \cap B_6 \cap C_5 \cap D_3 \cap E_4] \cdot \frac{1}{10} + 1 \cdot \frac{5}{10}$$

$$= 0 + \frac{7}{10} \cdot \frac{1}{10} + 1 \cdot \frac{5}{10} = \frac{57}{100}$$

Luego sustituyendo, se tiene:

$$P[S|A_1 \cap B_6 \cap C_5] = \frac{57}{100} \cdot \frac{1}{10} + \frac{6}{10} = \frac{657}{1.000}$$

$$P[S|A_1 \cap B_6] = \frac{657}{1.000} \cdot \frac{1}{10} + \frac{4}{10} = \frac{4.657}{10.000}$$

$$P[S|A_1] = \frac{4.657}{10.000} \cdot \frac{1}{10} + \frac{3}{10} = \frac{34.657}{100.000}$$

$$P[S|A_1] = \frac{34.657}{100.000} \cdot \frac{1}{2} = \frac{34.657}{200.000} \approx 0{,}1733$$

Así, puede considerarse que el Ministerio de Defensa tuvo «mala suerte» al producirse la repetición de una extracción.

Vuelvo a insistir en que, de no haberse producido dicha repetición, el proceso seguido se hubiera podido interpretar como la simulación de una variable aleatoria uniforme discreta con rango {1, 165.342} (*modelo B*). O dicho de otra forma más clara: no hubiésemos tenido indicios para pensar que el sorteo era injusto.

Finalizo manifestando mi deseo de que este artículo haya servido para mostrar las múltiples caras de un problema que, a la vista de las opiniones vertidas en los medios de comunicación, no parece haber sido objeto de un examen excesivamente riguroso.

**Roberto Marcellán**
Alumno de CAP
ICE de Zaragoza.
Sociedad Aragonesa de Profesores de Matemáticas «Pedro Sánchez Ciruelo»