

Un problema sobre repetidas apuestas al azar

Eleonora Catsigeras ¹

10 de marzo de 2003.

Resumen

En estas notas se da el enunciado y una demostración de un conocido resultado sobre la probabilidad de éxito o fracaso en sucesivas apuestas al azar. La exposición utiliza técnicas elementales de cálculo de probabilidades y está dirigida a un lector con conocimientos de matemática básica a nivel preuniversitario.

1. Introducción

La solución sobre el problema de repetidas apuestas al azar, que es bastante conocido, se resume en lo siguiente:

Si la banca juega con algo de ventaja frente al apostador, por menor que sea esa ventaja, es preferible para el apostador, si está indefectiblemente decidido a jugar, que juegue una sola vez empleando el dinero que dispone para jugar, todo junto, en una única apuesta.

Si el apostador, en cambio, divide el dinero en N partes iguales para emplearlo en realizar muchas apuestas sucesivas, entonces la probabilidad de que él pierda todo el monto que dispone antes de llegar a duplicar el monto inicial que tenía, se acerca al 100 %, cuando N aumenta.

Dicho de otra forma: cuanto más veces juegue el apostador, si tiene una cantidad acotada de dinero para jugar y si la banca tiene alguna ventaja sobre él (por más leve que sea), entonces más cercana al 100 % es la probabilidad de que lo pierda todo antes de llegar a duplicar lo que tenía al principio.

En estas notas daré un enunciado matemático preciso de esa propiedad y su demostración detallada usando solo cálculos elementales de probabilidades.

El problema me fue planteado por un amigo ingeniero, al que tuve el placer de reencontrar después de muchos años. Su nombre es Luis Eduardo Gaggero. A él respondo con estas notas y agradezco que haya formulado las preguntas que las motivaron.

Planteo del problema.

En un juego de azar cada moneda apostada se dobla con probabilidad p y se pierde con probabilidad $q = 1 - p$. El número real p es dado, cualquiera tal que $0 < p < 1$.

Se disponen de N monedas de igual valor que se apuestan de a una, repetidas veces, tantas veces como sea necesario. El juego termina o bien cuando se perdieron las N monedas, o bien cuando se alcanzó por primera vez duplicar las N monedas del principio.

El número natural $N \geq 1$ es dado cualquiera. Todas las repeticiones del experimento tienen igual probabilidad p de ganar, y son independientes entre sí.

Calcular la probabilidad e (éxito) de terminar el juego ganando, es decir duplicando las N monedas. Calcular la probabilidad f (fracaso) de terminar el juego perdiendo las N monedas.

Estudiar e en función de N con p fijo y discutir según el valor de p .

Respuesta Probaremos el siguiente resultado:

¹Instituto de Matemática y Estadística Prof. Rafael Laguardia (IMERL), Facultad de Ingeniería, Universidad de la República, J. Herrera y Reissig 565, Montevideo, Uruguay. Correo electrónico: eleonora@fing.edu.uy

Proposición 1 Las probabilidades e y f de éxito y fracaso respectivamente son:

$$e = \frac{1}{1 + (q/p)^N}; \quad f = \frac{1}{1 + (p/q)^N}; \quad e + f = 1$$

Si $p < 1/2$ la probabilidad e de éxito es decreciente con N y tiende a 0 cuando N tiende a infinito, mientras que la probabilidad f de fracaso es creciente con N y tiende a 1 cuando N tiende a infinito.

En cambio, si $p > 1/2$ la probabilidad e de éxito es estrictamente creciente con N y tiende a 1 cuando N tiende a infinito, mientras que la probabilidad f de fracaso es estrictamente decreciente con N y tiende a 0 cuando N tiende a infinito.

Finalmente, si $p = q = 1/2$ entonces $e = f = 1/2$, constante para todo N .

En todo caso, para cada N fijo, la probabilidad e es mayor y la probabilidad f es menor, cuanto mayor es p .

Observación: No me parece evidente, antes de demostrar la Proposición 1, que $e + f = 1$. Sea g la probabilidad de que el juego no termine nunca, es decir, de que se mantenga indefinidamente repitiendo el experimento sin nunca llegar a duplicar las N monedas ni llegar a perderlas todas. En efecto, en la última sección probaré el siguiente resultado:

Proposición 2 Cualesquiera sean N , p y q dados:

Las probabilidades e y f de éxito y fracaso respectivamente, verifican:

$$e + f = 1$$

La probabilidad h de que el juego termine en éxito o fracaso al cabo de alguna cantidad finita de repeticiones es:

$$h = 1 - g = 1$$

Inmediatamente a partir de las definiciones de e , f , g y h se tienen las igualdades:

$$e + f + g = 1, \quad h = f + g$$

El objetivo general es calcular e , f , h y g , demostrando que $h = 1$.

2. Cálculo de las probabilidades de éxito y de fracaso condicionadas a que el juego termine en tiempo finito.

Sea $h = 1 - g$ la probabilidad de que el juego termine al cabo de alguna cantidad finita de repeticiones.

Afirmamos que

$$0 < h \leq 1$$

En efecto, si se diera el caso particular de N pérdidas consecutivas, o si se diera el caso particular de N ganancias consecutivas, el juego terminaría en exactamente N repeticiones, que es un tiempo finito. La unión de esos dos sucesos particulares tiene probabilidad $q^N + p^N > 0$. Entonces $h \geq q^N + p^N > 0$.

Sea e^c la probabilidad condicionada de éxito, es decir de haber duplicado las N monedas iniciales dado que el juego se acaba en alguna cantidad finita de repeticiones. Sea f^c la probabilidad condicionada de fracaso, es decir de perder las N monedas iniciales dado que el juego se acaba en alguna cantidad finita de repeticiones. Debido a la fórmula de cálculo de probabilidades condicionadas se obtiene que:

$$e^c = e/h, \quad f^c = f/h$$

En la próxima sección veremos que $h = 1$ cualesquiera sean los datos N, p y q . Por lo tanto los resultados del cálculo de las probabilidades condicionadas serán los mismos que el de las probabilidades sin condicionar.

Por ahora se sabe que $e + f = h$ y por lo tanto, dividiendo entre h se obtiene

$$e^c + f^c = 1$$

Un primer objetivo es calcular e^c y graficarlo en función de $N \geq 1$, discutiendo según el valor de p dado.

Lema 1 a) Si el juego se acaba en exactamente n repeticiones, entonces $n \geq N$.

b) Si se acaba el juego en exactamente n repeticiones y se tuvo éxito, entonces $n = N + 2k$, donde k es la cantidad de pérdidas y $n - k = N + k$ es la cantidad de ganancias.

c) Si se acaba el juego en exactamente n repeticiones y se tuvo fracaso, entonces $n = N + 2k$, donde k es la cantidad de ganancias y $n - k = N + k$ es la cantidad de pérdidas.

d) En cualquiera de ambos casos $n - N$ es par.

f) Por lo tanto, si $n - N$ es negativo o impar, es imposible que el juego acabe exactamente en n repeticiones. Luego, para esos n , es nula la probabilidad de que el juego acabe en exactamente n repeticiones.

Prueba:

Si se tuvo éxito en exactamente n repeticiones y llamamos $k \geq 0$ a la cantidad de pérdidas que hubo en esas n repeticiones, entonces $n - k$ fue la cantidad de ganancias, y se cumple $n - k - k = N$, o sea $n = N + 2k$ y $n - k = N + k$, lo que prueba la afirmación b).

Análogamente, cambiando las palabras pérdida y ganancia entre sí, se prueba la afirmación c)

Las afirmaciones a) y d) de esta proposición se deducen inmediatamente del cálculo anterior, y la afirmación e) es consecuencia inmediata de la d). $\square$

Lema 2 Sea $n \geq N$ tal que $n - N$ es par. Sea $k \geq 0$ tal que $n = N + 2k$. Sean e_n y f_n las probabilidades de que el juego termine con éxito, y con fracaso respectivamente, al cabo de exactamente n repeticiones. Se cumple que:

$$\begin{aligned} e_n &= p^N A(N, k) (pq)^k \\ f_n &= q^N A(N, k) (pq)^k \end{aligned}$$

donde $A(N, k) \geq 0$ es un número natural que depende solo de N y k , (o sea de N y n), pero no depende de p ni de q .

Prueba:

Usando el Lema 1, en caso de éxito (o de fracaso) al cabo de exactamente n repeticiones, k es la cantidad de pérdidas (ganancias respectivamente) y $n - k = N + k$ es la cantidad de ganancias (pérdidas respectivamente).

Denotemos con las variables reales x e y a los números p y q del siguiente modo: en caso de éxito $x = p$, $y = q$; y en caso de fracaso $x = q$, $y = p$. Para indicar las pérdidas y ganancias obtenidas consecutivamente durante el juego, en caso de éxito representamos con el símbolo X a cada ganancia y con el símbolo Y a cada pérdida. Y en el caso de fracaso representamos con el símbolo X a cada pérdida y con el símbolo Y a cada ganancia.

Al acabar el juego (ya sea con éxito o con fracaso) al cabo de exactamente $n = N + 2k$ repeticiones, se generó una palabra de n letras que cumple las siguientes dos condiciones:

a) Cada letra es o bien una X o bien una Y .

b) Aparece exactamente $N+k$ veces la letra X , y exactamente k veces la letra Y . Por ejemplo, pudo haberse generado la palabra:

$$XYXYXXYXYYYYXXXXXXXXX$$

para $N = 5$, $k = 7$ $n = N + 2k = 19$.

Como el juego NO terminó antes de la repetición n -ésima, cada palabra que se pueda generar, debe cumplir además la siguiente tercera condición:

c) Para todo natural $m \geq 1$ tal que $m \leq n$, la sub-palabra formada con las m primeras letras de la palabra completa, tiene una cantidad de veces la letra X , menos una cantidad de veces la letra Y , diferente de N y de $-N$.

Sea $A(N, k) \geq 0$, con N y k fijos (o sea con N y n fijos), la cantidad de palabras diferentes de $n = N + 2k$ letras, que cumplen las condiciones a) b) y c) a la vez. El número natural $A(N, k)$ (que no necesitaremos calcular, y ni siquiera probar que es diferente de cero) es, por construcción, independiente de los valores que tengan p y q . Además, como está definido con los símbolos X e Y (no importa lo que estos representen), el número $A(N, k)$ es el mismo en el caso de fracaso que de éxito.

Cada una de las $A(N, k)$ palabras diferentes posibles tiene probabilidad igual a $x^{N+k}y^k$, por lo tanto, la probabilidad de éxito y la probabilidad de fracaso son iguales a

$$A(N, k) x^{N+k} y^k$$

donde, como se definió al principio, $x = p$ e $y = q$ en el caso de éxito, y $x = q$ e $y = p$ en el caso de fracaso. $\square$

De los lemas anteriores se deduce que:

$$e = \sum_{n=N}^{+\infty} e_n = p^N \sum_{k=0}^{+\infty} A(N, k) (pq)^k$$

$$f = \sum_{n=N}^{+\infty} f_n = q^N \sum_{k=0}^{+\infty} A(N, k) (pq)^k$$

La probabilidad h de que el juego se acabe en algún tiempo finito es

$$h = e + f = (p^N + q^N) \sum_{k=0}^{+\infty} A(N, k) (pq)^k$$

Por lo tanto, sustituyendo las expresiones anteriores en el cociente $e^c = e/h$ se obtiene la probabilidad condicionada de éxito:

$$e^c = \frac{e}{h} = \frac{p^N}{p^N + q^N} = \frac{1}{1 + (q/p)^N} \quad \text{si } p = 1 - q \neq 0; \quad e^c = 0 \quad \text{si } p = 1 - q = 0$$

Análogamente se obtiene la probabilidad condicionada de fracaso:

$$f^c = \frac{f}{h} = \frac{q^N}{p^N + q^N} = \frac{1}{1 + (p/q)^N} \quad \text{si } q = 1 - p \neq 0; \quad e^c = 0 \quad \text{si } q = 1 - p = 0$$

Es fácil ver que estas probabilidades e^c y f^c cumplen todas las propiedades enunciadas para e y f en la Proposición 1. Por lo tanto para terminar de demostrar las Proposiciones 1 y 2, basta probar que $h = 1$.

3. El juego casi seguramente termina en tiempo finito.

En Matemática la expresión “casi seguramente” significa que el suceso se produce con probabilidad 1, o sea que es 100 % probable aunque quizás no ocurra en el 100 % de los casos posibles.

Probaremos que es 1 la probabilidad h de que el juego termine al cabo de alguna cantidad finita de repeticiones, terminando así de demostrar las Proposiciones 1 y 2.

Sea $g = 1 - h$ la probabilidad de que el juego no termine nunca, es decir que continúe indefinidamente. Hay que probar que $g = 0$.

Observemos que aunque queremos demostrar que es nula la probabilidad de que el juego se mantenga indefinidamente, no es imposible que esto suceda: por ejemplo, si $N \geq 2$, y si se obtiene la sucesión periódica:

$$XYXYXYXYXYXY \dots XY \dots$$

donde X representa a cada ganancia e Y a cada pérdida, el juego se mantendría indefinidamente sin llegar al éxito ni al fracaso nunca.

Sin embargo, probaremos que es nula la probabilidad g de aparición, por ejemplo de esa sucesión, más la de todas las otras que hacen posible que el juego no termine nunca.

Lema 3 *Sea $\mathcal{S}$ una sucesión de dos símbolos X e Y , donde X representa a la ganancia en cada repetición del juego, e Y representa a la pérdida. Sea $k \geq 1$ un número natural cualquiera.*

Si $\mathcal{S}$ es cualquier sucesión que hace posible que el juego siga indefinidamente, entonces sus primeras $2kN$ letras forman k palabras consecutivas, de $2N$ letras cada una, y ninguna de esas k palabras está formada por una única letra repetida $2N$ veces.

Nota: el recíproco del lema 3 es falso.

Prueba:

La cantidad acumulada de veces que aparece la letra X , menos la cantidad acumulada de veces que aparece la letra Y , en la palabra formada por las primeras $2kN$ letras de la sucesión $\mathcal{S}$, es un número natural $d_k \geq 0$. Por conveniencia denotamos $d_0 = 0$.

Afirmamos que $|d_k| < N$ para todo $k \geq 1$. En efecto: a cada repetición del juego se aumenta o disminuye en 1 la diferencia acumulada entre la cantidad de X y la cantidad de Y . Por absurdo, si fuera $|d_k| \geq N$ para cierto $k \geq 1$, entonces, como como $d_0 = 0$, se tendría antes de o al llegar a la repetición $2kN$ del experimento, una diferencia entre la cantidad de X y la cantidad de Y igual a N o a $-N$. Entonces se habría acabado el juego en ese momento contradiciendo la hipótesis del lema.

Ahora probemos la tesis del lema por absurdo. Si para algún $k \geq 1$ la palabra k -ésima estuviera formada por la misma letra (X o Y) repetida $2N$ veces, entonces

$$d_k = d_{k-1} \pm 2N; \quad \text{o sea: } 2N = |d_k - d_{k-1}|$$

Por la propiedad triangular del valor absoluto y usando la afirmación anterior, se obtiene:

$$2N = |d_k - d_{k-1}| \leq |d_k| + |d_{k-1}| < N + N = 2N$$

por lo que $2N < 2N$, absurdo. $\square$

Consideremos una sucesión cualquiera de dos símbolos X e Y , (no necesariamente representativa de algún caso en que el juego sigue indefinidamente), donde la aparición de cada letra es independiente de las anteriores e igualmente distribuida. Sea p la probabilidad de aparición de X en un lugar cualquiera determinado, y sea $q = 1 - p$ la probabilidad de aparición de Y en ese lugar.

Separemos la sucesión en palabras de $2N$ letras: la primera palabra está formada por las primeras $2N$ letras de la sucesión; la segunda palabra por las $2N$ letras de la sucesión, que ocupan los lugares $2N + 1, 2N + 2, \dots, 4N$; la j -ésima palabra $\mathcal{P}_j$, para todo natural $j \geq 1$, está formada por las $2N$ letras de la sucesión que ocupan los lugares $2N(j - 1) + 1, 2N(j - 1) + 2, \dots, 2Nj$.

Lema 4 *La probabilidad a_k de que ninguna de las primeras k palabras consecutivas $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_j, \dots, \mathcal{P}_k$ de $2N$ letras cada una, esté formada por una misma letra repetida $2N$ veces, es:*

$$a_k = (1 - p^{2N} - q^{2N})^k$$

Prueba:

Fijado $j \geq 0$, la probabilidad de que la palabra $\mathcal{P}_j$ esté exactamente formada por $2N$ veces consecutivas el símbolo X , es p^{2N} . La probabilidad de que la palabra $\mathcal{P}_j$ esté exactamente formada por $2N$ veces consecutivas el símbolo Y , es q^{2N} . Así, la probabilidad b_j de que la palabra $\mathcal{P}_j$ no sea ninguna de esas dos, es

$$b_j = 1 - p^{2N} - q^{2N}$$

Las sucesivas repeticiones del juego son independientes y están igualmente distribuidas. Por lo tanto, si consideramos las primeras $2kN$ letras de la sucesión generada, estas forman k palabras consecutivas, de longitud $2N$ cada una, que son independientes entre sí e igualmente distribuidas.

Entonces la probabilidad a_k de que ninguna de esas k palabras $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_j, \dots, \mathcal{P}_k$ esté formada por $2N$ letras consecutivas iguales es:

$$a_k = b_1 b_2 \dots b_k = (1 - p^{2N} - q^{2N})^k$$

como se quería probar. $\square$

La probabilidad g de que el juego siga indefinidamente es, por construcción, la probabilidad de que la sucesión infinita $\mathcal{S}$ generada por las letras X e Y sea alguna de las sucesiones que hacen posible que el juego siga indefinidamente.

De los lemas 3 y 4 se deduce entonces que

$$g \leq a_k, \quad \text{para todo } k \geq 1$$

de donde

$$g \leq (1 - p^{2N} - q^{2N})^k, \quad \text{para todo } k \geq 1$$

y por lo tanto

$$g \leq \lim_{k \rightarrow \infty} (1 - p^{2N} - q^{2N})^k = 0$$

terminando de demostrar las proposiciones 1 y 2.