Miscelánea Matemática **52** (2011) 53–61 SMM

Revolviendo cartas suficientes veces

Leonardo Martínez

Lic. en Matemáticas, FC-UNAM
ssbmplayer@gmail.com

¿Cuántas veces hay que barajar un mazo de cartas para que quede *bien revuelto*? [1] La probabilidad se ha usado desde sus inicios para modelar juegos de azar y encontrar estrategias que maximicen ganancias, por ejemplo, para predecir que al tirar dos dados justos el número que tiene mayor probabilidad de obtenerse es el 7. En 1986, Persi Diaconis y David Aldous, matemáticos estadounidenses, le dieron un nuevo enfoque al uso de la probabilidad en juegos de azar. Al estudiar los paseos aleatorios que se dan en distintos grupos, pasaron por el grupo de permutaciones y su relación con barajar las cartas. Su trabajo con respecto a este tema se centra en estudiar la pregunta que mencionamos al inicio del artículo.

Si tuvieramos una máquina que revolviera las cartas perfectamente, no habría necesidad de hacernos tal pregunta. Si pudieramos obtener cualquiera de los posibles acomodos de un mazo de cartas con la misma probabilidad, entonces bastaría aplicar sólo una vez este proceso para tener cartas revueltas justamente. Sin embargo, al barajar las cartas en la vida real seguimos ciertas costumbres, por ejemplo, la de partir el mazo cerca de la mitad y luego dejar caer las cartas de ambas mitades haciendo lo posible por alternarlas. Al hacer esto solamente una vez, aún quedan cartas de la ronda anterior juntas y en este sentido podemos decir que las cartas aún no están bien revueltas. Por otro lado, no podemos realizar demasiadas veces este proceso pues tomaría mucho tiempo, y también queremos agilizar el juego. Formalizando matemáticamente estos conceptos, veremos cómo se puede resolver el problema de revolver cartas con herramientas de probabilidad.

[1] Este es un trabajo de divulgación basado en el capítulo 24 “Shuffling Cards” de [4]. El lector interesado puede encontrar una exposición más larga de este tema en dicha referencia o en otras fuentes que se enlistan en la bibliografía.

1. Modelos para barajar

Vamos a trabajar con mazos de cartas de tamaño arbitrario, y luego haremos el caso particular para las 52 cartas de la baraja inglesa. Tomemos un mazo de n cartas, y numerémoslas de 1 a n para pensarlas como números enteros. Una *permutación* del conjunto $\{1, 2, \ldots, n\}$ es una función biyectiva del conjunto en sí mismo, es decir, podemos pensar en una permutación como en acomodar a los números de 1 a n en línea recta en algún orden. Las permutaciones de $\{1, 2, \ldots, n\}$ bajo la composición forman un grupo de $n!$ elementos, al cual denotaremos por S_n. Así, aplicar una permutación en S_n a nuestro mazo hace que éste cambie de orden, de modo que podemos pensar en *barajar las cartas* como la aplicación de una de estas permutaciones π, que lleva la carta en la posición i a la posición $\pi(i)$.

Sin embargo, al revolver las cartas de un mazo no estamos seleccionando cualquiera de estas permutaciones con la misma probabilidad. Por ejemplo, una forma de barajar el mazo es tomar la carta superior del mazo e insertarla en un lugar aleatorio. Al aplicar una vez esta forma de revolver las cartas, no estamos obteniendo cualquier permutación con la misma probabilidad. Es más, hay algunas permutaciones, como la que invierte el orden de todo el mazo, que no se pueden obtener barajando reiteradamente de este modo. Pero ésta es únicamente una de las formas de barajar, podemos pensar en algunas otras como tomar dos cartas aleatoriamente e intercambiarlas de lugar o "partir el mazo[2]".

En cada una de estas formas de barajar las cartas estamos asignándole una probabilidad $P(\pi)$ a cada una de las permutaciones π de S_n. La forma de barajar que elijamos le dará cierta preferencia a unas permutaciones sobre otras, es decir, les dará mayor probabilidad de ser elegidas. Así, definimos un *modelo para barajar* como una función de distribución con dominio S_n.

El modelo de la máquina perfecta que selecciona con la misma probabilidad cualquier permutación, es decir, que desde la primera vez que revolvemos el mazo de cartas obtenemos un orden totalmente aleatorio se llama *modelo uniforme*. A este modelo lo denotaremos por U. Como tenemos $n!$ permutaciones y es equiprobable elegir cualquiera de ellas, tenemos que $U : S_n \to [0, 1]$ está dada por:

$$U(\pi) = \frac{1}{n!} \qquad \text{para toda } \pi \in S_n$$

[2]Hacer un corte en un punto aleatorio e intercambiar de lugar las mitades del mazo que se forman.

El segundo modelo para barajar que nos interesa es el *modelo tradicional*, el cual estamos acostumbrados a ver en los juegos de cartas. Consiste en dividir el mazo aproximadamente a la mitad, tomar cada mitad con una mano y dejar caer las cartas intentando que se alternen las de una mitad y otra. En 1955, Gilbert y Shannon propusieron un modelo sencillo y que se apega a esta forma de barajar las cartas. Su función de distribución es la siguiente:

$$Trad : S_n \to [0,1]$$

dada por:

$$Trad(\pi) := \begin{cases} \dfrac{n+1}{2^n} & \text{si } \pi \text{ es la identidad,} \\ \dfrac{1}{2^n} & \text{si } \pi \text{ consiste de dos sucesiones crecientes,} \\ 0 & \text{en otro caso.} \end{cases}$$

Lo que considera este modelo es que la probabilidad de que la mano derecha tome k cartas es $\binom{n}{k}$, con lo cual se asigna una pequeña probabilidad a que la mano derecha tome muy pocas o muchas cartas. Además, considera que la probabilidad de que caiga una carta de cierta mano es proporcional a la cantidad de cartas que tiene.

Esta forma de revolver las cartas es invertible. Para regresar el procedimiento que hicimos, lo que hay que hacer es tomar un subconjunto de cartas, sacarlo sin alterar su orden relativo y colocarlo encima del resto de las cartas. Al realizar este proceso seleccionando un conjunto de cartas aleatoriamente, cada uno con la misma probabilidad, obtenemos un modelo para barajar. A este modelo correspondiente a barajar al revés lo llamamos el *modelo tradicional inverso* y lo denotaremos por $Trad^{-1}$.

Así, los modelos para barajar que nos interesan son el uniforme (U), el tradicional ($Trad$) y su inverso ($Trad^{-1}$). En la siguiente figura tenemos un ejemplo de lo que le hacen cada uno de estos modelos para barajar a un mazo con 10 cartas.

Veremos qué quiere decir que algo esté *bien revuelto*.

2. Mazos *bien revueltos*

Para poder comparar las distribuciones entre sí, definimos una métrica en el espacio de distribuciones. Dados dos modelos para barajar P y Q definimos la *distancia variacional* entre ellos como:

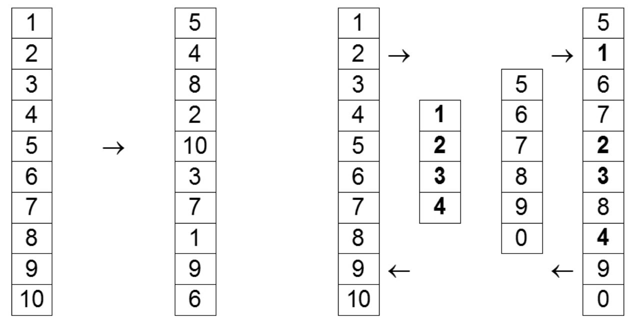

Figura 1: Ejemplo de posibles barajeos con U (izquierda), $Trad$ y $Trad^{-1}$ (derecha).

$$||P - Q|| := \frac{1}{2} \sum_{\pi \in S_n} |P(\pi) - Q(\pi)|$$

Éste es el punto clave para entender a qué nos referimos con *bien revuelto*. La distancia variacional nos dice qué tan cerca[3] están dos modelos para barajar con respecto a las probabilidades que le asignan a las permutaciones de S_n. Si dos modelos para barajar se parecen mucho en las permutaciones a las cuales les dan preferencia, entonces los términos $|P(\pi) - Q(\pi)|$ de la suma serán pequeños y por tanto su suma también lo será. Si un modelo para barajar tiene una distancia pequeña a la distribución uniforme, entonces aproximadamente será equiprobable elegir cualquier permutación. Es con respecto a esto que diremos que un modelo para barajar *revuelve bien* las cartas.

Hasta el momento sólo hemos hablado de barajar *una* vez con nuestros modelos. Sin embargo, dado un modelo para barajar P, podemos considerar el modelo obtenido tras aplicar k veces P, al cual denotaremos por P^k. Notemos que esto nos representa cómo se comporta la permutación tras repetir k veces la mezcla de cartas correspondiente a P.

Con el desarrollo de estas ideas podemos hacernos finalmente una pregunta más significativa: dado un modelo para barajar P, ¿cómo se comporta $d(k) := ||P^k - U||$ conforme k crece? Notemos que esto ya se acerca a saber cuántas veces debemos barajar un mazo para convencernos de que está bien revuelto, pues entender el comportamiento de esta función es entender cómo se acerca P^k al modelo uniforme. En [1]

[3]Recordemos que $P(\pi)$ y $Q(\pi)$ son números entre 0 y 1 pues P y Q son distribuciones. Usando la desigualdad del triángulo podemos ver que $||P - Q|| \leq 2$, de modo que el factor $\frac{1}{2}$ hace que la distancia variacional sea un número entre 0 y 1.

podemos encontrar una definición de "suficientemente aleatorio[4]" con la cual d se acerca a cero exponencialmente conforme k crece.

Más aún, en ciertas formas de barajar las cartas resulta que a partir de cierta k' la distancia decrece repentinamente, (aproximadamente como se muestra cerca de 15 en la figura 1). Esta k' es de gran importancia, pues es un punto en el cual barajar hasta ahí nos ayuda a reducir considerablemente la distancia entre las distribuciones P^k y U en pocas repeticiones. Por supuesto, tomar valores más grandes que k' hará que esta distancia se acorte más, pero la mejora requerirá un mayor esfuerzo[5]. De esta forma, la meta para responder la pregunta original es encontrar dicha k' para el modelo $Trad$.

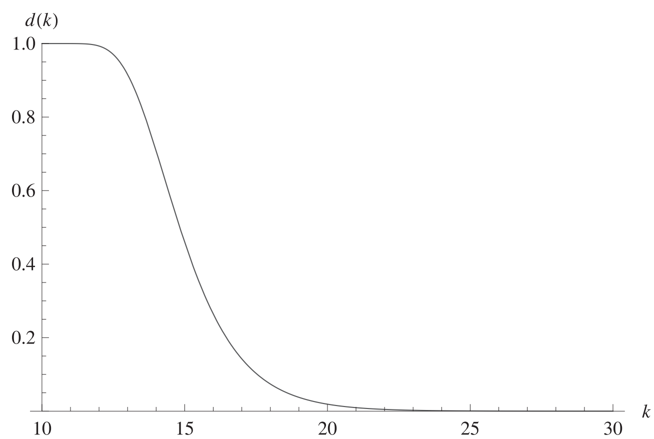

Figura 2: $d(k)$ decrece exponencialmente conforme k crece.

3. Reglas de alto y resultados

Ahora pasamos a una idea espectacular desarrollada por David Aldous y Persi Diaconis [2]: las reglas de alto fuertes uniformes. Decimos que modelo para barajar tiene una regla de alto fuerte uniforme si existe un criterio que nos diga cuándo dejar de barajar para que la distribución que quede en ese momento sea exactamente la distribución uniforme.

[4]Dentro de nuestras distribuciones aleatorias también están incluidas algunas formas de barajar deterministas. Por ejemplo, si le asignamos probabilidad 1 a la permutación que invierte el orden del mazo, entonces sabemos exactamente cómo quedará el mazo tras revolver las cartas. Evitando este tipo de situaciones es mediante lo cual obtenemos modelos para barajar "suficientemente aleatorios".

[5]Formalmente esta k' es un número *crítico* para el cual $d(k' + o(k')) \approx 1$ pero $d(k' - o(k')) \approx 0$. Esta condición es la que causa el fenómeno de corte.

Un ejemplo es el siguiente. Supongamos que para barajar tomamos la carta superior del mazo y la colocamos en una posición aleatoria, haciendo este proceso repetidas veces. Observemos la carta inferior del mazo. Cuando lleguen dos cartas debajo de ella, estarán ordenadas en cualquier orden con la misma probabilidad. Cuando llegue la tercera, ahora las 6 permutaciones de estas cartas serán equiprobables. Siguiendo este argumento, cuando la carta que estaba hasta abajo llegue arriba del mazo, las otras $n-1$ cartas tendrán cualquier orden con la misma probabilidad, de modo que en el momento en que coloquemos la carta inferior (ahora hasta arriba) en medio del mazo, todas las cartas estarán revueltas uniformemente. Así, la regla "deja de revolver cuando la última carta pase a ser la primera y la coloques en el mazo" es una regla de alto fuerte uniforme para el modelo.

Lo sorprendente es que existe una relación entre el tiempo que se tarda un modelo para barajar en alcanzar su regla de alto fuerte uniforme y la distancia $||Q^k - U||$ entre barajar varias veces y el modelo uniforme. Más concretamente, se tiene que si Q es un modelo para barajar y T la variable aleatoria que cuenta la cantidad de pasos necesarios para que Q alcance su regla de alto uniforme entonces para toda $k \geq 0$ se tiene que:

$$||Q^k - U|| \leq Pr(T > k) \tag{1}$$

La demostración de este resultado se puede encontrar en [4]. Ahora sí, ¿cuántas veces debemos barajar tradicionalmente un mazo de cartas para que quede bien revuelto?

Teorema. *Tras haber barajado k veces tradicionalmente en un mazo de n cartas, la distancia variacional a la distribución uniforme satisface:*

$$||Trad^k - U|| \leq 1 - \prod_{i=1}^{n-1}(1 - \frac{i}{2^k}).$$

Demostración. Para demostrar este resultado podemos analizar a $Trad^{-1}$ en vez de $Trad$ pues como la forma de barajar de $Trad^{-1}$ es la inversa de $Trad$, podemos concluir que $Trad^{-1}(\pi) = Trad(\pi^{-1})$. La ventaja de trabajar con $Trad^{-1}$ es que tiene una regla de alto fuerte uniforme muy clara, como veremos a continuación.

Recordemos que cada vez que barajamos con $Trad^{-1}$ una carta puede ser elegida o no. Así, a cada carta le podemos asignar una cadena de dígitos $a_k a_{k-1} \ldots a_1$, donde a_i toma el valor 1 si la carta fue movida al barajar la i-ésima vez y 0 en caso contrario. Afirmamos que una regla de alto fuerte uniforme es detenerse en cuanto todas las cartas tengan una cadena distinta de dígitos. Esto se debe a que justo cuando esto suceda, las cartas van a estar ordenadas de acuerdo al orden lexicográfico que llevan sus cadenas de dígitos. Como estas cadenas fueron elegidas aleatoriamente, entonces el orden de las cartas será uniformemente aleatorio.

Cuando barajamos la k-ésima vez quisieramos que n cartas tengan todas cadenas binarias distintas. En total hay 2^k cadenas binarias que podemos hacer con k dígitos. Así, la probabilidad de que tengamos que hacer $T > k$ pasos es el complemento de este evento, a saber, que haya al menos dos cartas que tengan la misma cadena de dígitos.

Este es un problema conocido de probabilidad [6], de donde sabemos que:

$$P(T > k) = 1 - \prod_{i=1}^{n-1}(1 - \frac{i}{2^k})$$

Pero de acuerdo a (1) sabemos que este número acota a $||(Trad^{-1})^k - U||$. Como cada permutación tiene un único inverso, tenemos entonces que $||(Trad^{-1})^k - U|| = ||Trad^k - U||$, de donde se sigue el resultado deseado. □

Ya con este resultado a la mano podemos tomar $n = 52$ cartas y observar cómo se comporta $d(k)$ conforme varía k en la siguiente tabla: Así,

k	$d(k)$
7	1,00
8	1,00
9	0,94
10	0,75
11	0,49
12	0,29
13	0,16
14	0,08

por medio de este análisis podemos ver que aproximadamente necesitamos barajar 12 o 13 veces con el modelo tradicional para revolver bien

[6] El problema general es: si se colocan a objetos en b cajas, con $0 \leq a \leq b$ ¿cuál es la probabilidad de que alguna caja tenga dos o más? Se responde calculando la probabilidad del complemento: que los objetos queden en cajas distintas. Colocando los objetos en orden, tras colocar i objetos en cajas distintas, la probabilidad de que el siguiente no quede en una caja ya ocupada es $1 - \frac{i}{b}$. Así, obtenemos $\prod_{i=1}^{a-1}(1 - \frac{i}{b})$. De forma que la respuesta al problema original es $1 - \prod_{i=1}^{a-1}(1 - \frac{i}{b})$.

un mazo de 52 cartas, ya que este es el punto en el cual la distancia a la distribución uniforme comienza a hacerse pequeña.

Al ser un poco más flexibles con la distancia al modelo uniforme, la demostración del teorema principal puede darnos una cota más pequeña. En [5] se puede consultar un análisis en el cual se muestra que barajar 7 veces son suficientes para el modelo tradicional.

Este método de usar reglas de alto fuertes uniformes y la distancia variacional admite aplicaciones más generales. La más directa es usarlo para otros modelos para barajar. En [3] se puede encontrar el estudio del modelo en el cual colocamos la primer carta del mazo en un lugar aleatorio. Para este modelo, se necesitan barajar aproximadamente $n \log(n)$ veces para asegurar aleatoriedad. Para $n = 52$ esto nos da cerca de 205 veces.

En general, podemos utilizar el método para explorar caminatas aleatorias en un grupo. Persi Diaconis y Davis Aldous muestran en [2] ejemplos de esto, encontrando reglas de alto fuertes uniformes para la caminata en Z_n y en ciertas clases de grupos.

La distancia variacional es tan sólo una de las muchas distancias que se pueden dar entre dos formas de barajar. También en [2] podemos encontrar definiciones equivalentes de dicha distancia y una pequeña discusión que justifica por qué es un criterio adecuado para *bien revuelto*.

La pregunta de obtener un mazo de cartas bien revuelto barajándolo pocas veces podría parecer inocente, pero implica problemas de optimalidad con importantes complicaciones. El planteamiento mismo del problema nos lleva a reflexionar en los métodos que usamos para obtener objetos aleatorios. Es llevar el juego al terreno exacto de las matemáticas para encontrar estrategias, proponer juegos justos y descubrir realmente que tanto azar está involucrado. Persi Diaconis, antes mago y ahora matemático, menciona:

"Si dices que eres un profesor en Stanford, la gente te respeta. Si dices que inventas trucos de magia... evitan presentarte a su hija".

Agradecimientos

Agradezco los valiosos comentarios recibidos por los árbitros anónimos, pues ayudaron a mejorar la presentación y el contenido de este trabajo. También agradezco a la Dra. Ana Meda Guardiola quien fue la que me animó a iniciar este artículo, me ayudó en el proceso y me apoyó en concluirlo.

Referencias

[1] E. Behrends. *Introduction to Markov Chains.* Vieweg, 2000.

[2] Persi Diaconis. Shuffling cards and stopping times. *American Mathematical Monthly*, 333–348, 1986.

[3] Persi Diaconis. *Group Representations in Probability and Statistics.* Institute of Mathematical Statistics, California, 1988.

[4] Martin Aigner y Günter M. Ziegler. *Proofs from the Book*. Springer, 3^a ed. 1998.

[5] Dave Bayer y Persi Diaconis. Trailing the dovetail shuffle to its lair. *Annals Applied Probability*, 294–313, 1992.