International Journal of Pure and Applied Mathematics
Volume 85 No. 5 2013, 849-857
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
doi: http://dx.doi.org/10.12732/ijpam.v85i5.3

ijpam.eu

THE PROBABILITY OF CARDS MEETING AFTER A SHUFFLE

Yutaka Nishiyama

Department of Business Information
Faculty of Information Management
Osaka University of Economics
2, Osumi Higashiyodogawa Osaka, 533-8533, JAPAN

Abstract: Even after giving a standard 52-card deck a good shuffle, there will likely be instances where two cards with the same number end up next to each other. This article concerns my investigations into the probability of such an occurrence. In the end I found that this is by no means an uncommon event, there being a 21.7% probability of finding two cards of a given number adjacent, and a 95.4% chance of finding at least one pair of adjacent same-numbered cards. I made these calculations using combinatorics for up to 20 cards, and a random number-based simulation for the cases of 24 to 52 cards.

AMS Subject Classification: 60C05, 97K50, 00A08
Key Words: probability, combinatorics, counting, complimentary events, random numbers, simulation

1. Kings and Queens

I received an interesting e-mail from my colleague Steve Humble in the UK. Given a well-shuffled deck, he asked, what would be the probability of finding a King and a Queen either next to each other, or separated by only one card? Would it exceed 70%? I gave it a shot using an actual deck of cards, and sure

Received: January 17, 2013

enough it was quite common, in excess of 70%, it seemed, to find either an adjacent pair (KQ, QK) or the pair separated by one card (K | Q, Q | K). I wondered if there might be a mathematical explanation for this. While I was looking for clues, I received a second e-mail from him. He had found a website [1] claiming that the probability of finding in a well-shuffled deck two adjacent cards with the same number was 96.35%, which seemed awfully high.

An assumption of randomness is inherent in most card games. A common way of visually checking to see if a deck has been sufficiently shuffled is to scan through it, looking for pairs of matched numbers. Upon finding such a pair, one might even split it up by moving one of the cards to another location within the deck, though doing so does not guarantee that a new pair is not formed. The website that Steve had found did not show how this probability was calculated. I therefore decided to do a little investigation into the accuracy of this 96.35% value.

2. The Probability of a Pair of 8s

Our final goal is to determine the probability of finding any pair of adjacent numbers, but a bit of preparation is required.

We will start by looking for a pair of a specific number, say, 8. There are 52 cards altogether, and the laws of permutations tell us that there are 52! ways of arranging them. 52! is a big number.

$$52! \approx 8.066 \times 10^{67}$$

There are 4 rank 8 cards, and ${}_{52}C_4$ tells us the number of ways that these 4 cards can be positioned within a 52-card deck. This is calculated as follows.

$${}_{52}C_4 = \frac{52!}{(52-4)!4!} = 270725$$

After these four cards have been removed from the deck, there are 48 cards left. We'll represent those cards using a | character (Figure 1).

| | | | | | | | | | | | | | | | | · · · · · · · · · · · · | |

| | | | 8 | | | 8 | | 8 | | | · · · · · · · · · · · · | 8 | | | | |

Figure 1. The probability of no adjacent 8s

The number 8 cards will not be adjacent to each other if we insert them one by one between the |s. When we include the right and left ends, there are $48 + 1 = 49$ locations where we can perform these insertions. The number of ways to choose 4 out of these 49 locations is ${}_{49}C_4$, which is calculated as follows.

$${}_{49}C_4 = \frac{49!}{(49-4)!4!} = 211876$$

From the above, we can see that the probability of none of the four 8s appearing next to each other is as follows.

$$\frac{{}_{49}C_4}{{}_{52}C_4} = \frac{211876}{270725} = 0.782624$$

It therefore follows that the probability that two of the 8s *are* next to each other is the complement of this,

$$P_1 = 1 - 0.782624 = 0.217376,$$

which gives a final probability of about 22%.

This doesn't feel like high odds as long as we are focusing on a particular number, but a full 52-card deck has thirteen ranks in four suits each. We found that the probability of none of the four cards of a given rank being adjacent was 0.782624, which means that the probability of none of the four cards of each of the thirteen ranks being adjacent is, roughly speaking, the 13th power of this number,

$$0.782624^{13} = 0.041323,$$

or about 4.13%. The likelihood of at least one pair is therefore the complement of this,

$$P_2 = 1 - 0.782624^{13} = 0.958677,$$

or about 95.87%.

3. Writing it All Up

The 95.87% that we found is close to the 96.35% claim of the website - within an order of magnitude, at least - but could use some further refining. We need to take into account the fact that the thirteen ranks are not wholly independent of each other, but rather have a relationship among the 52 cards they lie within.

Namely, we cannot treat the probability that no cards of a given rank end up adjacent as a separate event from same probability with regards to another rank. As mentioned above, there are

$$52! \approx 8.066 \times 10^{67}$$

ways to order a deck of cards, and we need to look for cases where two adjacent cards have the same rank. I was a bit daunted at first, as 10^{67} is an astronomically large number, one not easily handled on a personal computer, despite the advancements of recent years.

I therefore decided to start not with 52 cards, but just four. Taking the Ace as our example, there is only one way to arrange 4 Aces - AAAA - and the probability of having two adjacent cards of the same rank is 1. So what happens when we add 2s to the mix? Then we have 8 cards, within which there are ${}_8C_4 = 70$ ways to arrange the Aces. Of these, there are only two patterns in which no Aces are adjacent, 2A2A2A2A and A2A2A2A2 (and therefore 68 in which they are), so the probability of adjacent Aces becomes

$$68/70 = 0.9714.$$

Now we add in the 3s, for a total of 12 cards. We now want to place our 4 Aces among 12 cards, and there are ${}_{12}C_4 = 495$ ways of doing that. After that there are ${}_8C_4 = 70$ ways of adding the four 2s, so the total number of ways of ordering the 12 cards becomes

$${}_{12}C_4 \times {}_8C_4 = 495 \times 70 = 34650$$

When we remove the 1092 cases where none of the three ranks are adjacent, the probability becomes as follows:

$$(34650 - 1092)/34650 = 0.9685$$

To find that there are 1092 cases of no adjacent ranks, I used a certain freely available program that creates ”combination lists” to find all ${}_{12}C_4 = 495$ combinations and all ${}_8C_4 = 70$ combinations. (This program, by the way, uses a wonderfully elegant algorithm to do this, but I will save discussion of that for another time.)

I first used a spreadsheet program to list all the ${}_{12}C_4 = 495$ positions for the Aces. I then did the same for the ${}_8C_4 = 70$ positions for the rank 2 card. There are 12 cards altogether, so I created 12 arrangements, first filling in the Aces, and then the 2s among the remaining open spaces, and finally the 3s in what was left. Doing this gave me 34650 patterns.

I next created a macro to sequentially search for patterns with no adjacent pairs of ranks, which resulted in 1092 hits. That means that there are 34650 − 1092 = 33558 patterns that do have adjacent pairs, which accounts for 96.85 % of the total.

$$33558/34650 = 0.9685$$

So for just 12 cards in three ranks there are 34650 patterns. Actually, it is not necessary to look for every one - just 5775, or 1/6 of the total, would suffice.

$$\frac{{}_{12}C_4}{3} \times \frac{{}_{8}C_4}{2} = 165 \times 35 = 5775$$

Nonetheless, this is still not an operation that one would want to perform by hand.

In a similar manner, I found that for 16 cards in four suits there are a total of 63063000 patterns, of which 60797976 have adjacent pairs, for a probability of 96.41%. For 20 cards, the probability is 96.14% (see Table 1).

The number of combinations increases exponentially as the number of cards increases. For a full 52-card deck the number is extremely large.

$$\frac{{}_{52}C_4}{13} \times \frac{{}_{48}C_4}{12} \times \cdots \times \frac{{}_{8}C_4}{2} = 20825 \times 16215 \times \cdots \times 35 = 1.478 \times 10^{40}$$

Not quite so bad as $52! \approx 8 \times 10^{67}$ perhaps, but each of them must be examined. This is just not possible given the memory capacity and calculation speeds of a present-day personal computer. Indeed, 20 cards is about the limit.

In that case I found a 96.14% probability of adjacent cards, but this is slightly less than the 96.35% figure given on the website. From looking at the trend, I predicted a lower number still as the number of cards increased to 52.

Cards	Total patterns	Adjacent cases	Prob.
4	1	1	100%
8	70	68	97.14%
12	34650	33558	96.85%
16	63063000	60797976	96.41%
20	305540235000	293752962960	96.14%

Table 1: Manual counting method

4. Random Simulation

Learning that I wasn't likely to get beyond about 20 cards (4 each of five ranks) using my method of listing up each case and searching them for adjacent pairs, I began looking for an alternative approach for handling a full deck. From the beginning I had intended to avoid an investigation by simulation, but circumstances seemed to be pushing me in this direction.

All modern computers are capable of generating uniform pseudorandom numbers in the form of a real (floating point) number in the range (0, 1), so I wondered if that wouldn't be a help. Quite some time back I created a bingo game simulation to calculate probabilities using random numbers [2], so I decided to try repurposing that.

In bingo, each player has a card with the numbers 1 through 25 randomly placed in a 5×5 grid. Numbered balls are then randomly drawn, and each player marks the square containing the drawn number. The player who first creates a horizontal, vertical, or diagonal row of filled-in numbers wins the game.

When investigating this before, I found that the probability of completing a row, column, or diagonal after r balls are drawn is P_r, which is defined as follows.

$$P_r = \frac{12 \times {}_{20}C_{r-5}}{{}_{25}C_r}$$

In the case of $r = 8$, for example, the probability is $P_8 = 0.1265 \times 10^{-1}$. My simulator had given a predicted value of 0.1258×10^{-1}, which is relatively close to the actual value.

The bingo game simulation generated random balls numbered 1 through 25 using random numbers in the range (0, 1). Since there are 52 cards in a standard deck, I tried mapping the open range (0, 1) of real numbers to as a closed range [1, 52] of integers. We can do this as follows.

We start by choosing one of the 52 numbers. Since we can't reuse selected numbers, the second choice will be from among 51 numbers, and the third choice from among 50 numbers. Letting the selected numbers be I_1, I_2, I_3, we can use a random number x to generate them as follows.

$$I_1(integer) = x(real) \times 52 + 1$$

$$I_2(integer) = x(real) \times 51 + 1$$

$$I_3(integer) = x(real) \times 50 + 1$$

Using real numbers to generate integers is not difficult: simply discard the digits following the decimal. For example, if 0.55817, 0.08649, and 0.91929 are chosen as the x values, the calculations are as follows.

$$0.55817 \times 52 + 1 = 30.02470$$

$$0.08649 \times 51 + 1 = 5.41087$$

$$0.91929 \times 50 + 1 = 46.96441$$

Dropping the digits after the decimal result in the integers 30, 5, and 46.

We next want to use the random numbers obtained to represent card numbers from 1 through 52. For example, if the random integers were $I_1 = 30, I_2 = 5, I_3 = 46$, we would have the following.

$$1 \leq 30 \leq 52$$

$$1 \leq 5 \leq 51$$

$$1 \leq 46 \leq 50$$

We prepare 52 boxes, and assign to them the numbers 1 through 52, as well as cards, also numbered 1 through 52. Our first random number was 30, so we place the first card in the 30th box and put a lid on it. We now have 51 remaining open boxes that can accept cards. Our next random number is 5, so we place the second card in the 5th open box, and close that one. We now have 50 available boxes. Our next random number is 46, so we count to the 46th open box. Note that since we've already closed the 5th and 30th boxes those will be skipped, so the number of the box that the next card is placed into will be the box in the $46 + 2 = 48$th position. After placing the card and closing its box, we repeat this process until all 52 cards are placed into the 52 boxes. We then open each box in turn, take out its card, and use the cardfs number to generate a string of random integers.

We still have to perform a second step to randomize the cards. While there are 52 cards in a deck, there are 4 of each rank $(A, 2, 3, \cdots, K)$, so the random numbers need to be associated with the following sequence.

$$(A, A, A, A, 2, 2, 2, 2, 3, 3, 3, 3, \cdots, K, K, K, K)$$

Thankfully, this is easily done using the computer.

I next wanted to verify that the generated random numbers were suitable for use in an actual simulation. I had already found that the probability in the case of 8 cards was

$$P_1 = 1 - \frac{{}_{49}C_4}{{}_{52}C_4} = 1 - 0.782624 = 0.217376,$$

so I tested the simulation against this value. Table 2 shows the number of simulations run along with the resulting test value, and the number of significant values to which the results agreed with the theoretical value. After 10 million trials, the values met to 4 significant figures.

Trials	Simulated probability	Significant values
100000(=10^5)	0.21629	2
1000000(=10^6)	0.217102	3
10000000(=10^7)	0.2173721	4
Theoretical value	0.217375566 $\cdots$	

Table 2: Verification of random simulation

One might predict that increasing the number of trials will result in increasingly close agreement with the theoretical value, but in actual practice the cycle of the random number generator ($2^{31} \approx 2.1 \times 10^9$) cannot be exceeded, and there are limits to how far a personal computer can be pushed, so 10 million ($= 10^7$) trials seems like a sufficiently valid number. So setting the number of trials to 10 million, I compared the results of the counting method and the simulation for the cases of 8 to 20 cards, finding that the theoretical and simulated values met to within 3 - 4 significant values. Having confirmed that the random number method gave more or less valid results, I reran the simulation for the cases with 24 - 52 cards (Table 3).

The value that I found for 52 cards using the rough method was 95.87%, and that of the random simulation was a very close 95.44%. This still seemed off from the website's given value of 96.35%, so I tried contacting the author. I have yet to receive a reply, although this is perhaps not surprising since the article was written in 1996, now over 15 years ago. In any case, we can at least say that the probability of at least one pair of adjacent same-ranked cards is probably around 95 - 96%.

While flipping through a deck to look for adjacent cards as a measure of how well the deck was shuffled might be common, it is likely not very effective. As we have calculated here, that should happen around 95% of the time. So if you're shuffling and notice a pair or two in the deck, don't sweat it.

Cards	Counting	Random
4	100%	
8	97.14%	97.12%
12	96.85%	96.86%
16	96.41%	96.40%
20	96.14%	96.13%
24		95.95%
28		95.82%
32		95.70%
36		95.64%
40		95.56%
44		95.51%
48		95.49%
52		95.44%

Table 3: Comparison of probabilities from counting and random numbers

References

[1] Random Card Shuffling Probabilities, *The Math Forum @Drexel*, http://mathforum.org/library/drmath/view/52153.html

[2] Y. Nishiyama, A Bingo Game, *Mathematics Seminar*, **22**, No. 4 (1983), 61-66, In Japanese.