Snap

The game of SNAP is played with 2 standard decks of 52 cards (well shuffled!). Cards are dealt simultaneously from the top of each deck and SNAP is called if the two dealt cards are identical (value and suit!). Here's the problem: what is the probability of getting at least one SNAP in a deal of the entire deck?

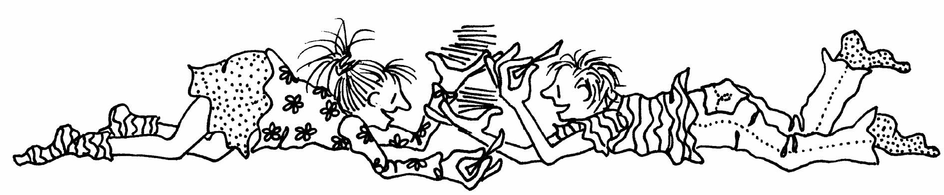
Let me describe one of the more interesting sessions I had with this problem I started with the question of whether the probability of getting a SNAP should be greater than or less than 1/2. That is, if you play the game again and again and each time have to bet on the outcome, how do you bet? After a small amount of arguing, they voted almost unanimously in favour of no SNAP. That is, they figured that you should get a SNAP less than half the time.

We then generated some data. I put the 50 students into 25 pairs and had them play 10 games of SNAP each, of course stopping if ever they got a SNAP, reshuffling, and starting again. Thus we had a total of 250 games of which 164 produced a SNAP, giving an estimate of the SNAP probability of 164/250 = 0.656. This was noticeably greater than ½—they were quite surprised.

I then posed the problem of how we might calculate the theoretical probability of getting a SNAP. I gave them a few minutes to think about this, but nobody did very much. Then all of a sudden Jenine arose from her seat, came up and wrote a formula on the board.

This is an old problem, but a good one. It has a number of classic variations, for example if I have a number of letters and corresponding addressed envelopes, and I put the letters at random into the envelopes, what is the probability that at least one letter will get into the correct envelope? It's especially good with cards, because the students have some intuition about what to expect, and they can gather lots of data quite easily. There are a number of elegant solutions around, but this problem demonstrates, I think, the value in not trying to force your own favorite technical trick on the class. (At least it did for me!) It also gives, at the end, a nice glimpse of that mysterious number e.

If you don't have a huge number of decks of cards, a 13-card deck works well—give each student a complete suit, and snap is called if the cards have the same value. Interestingly enough, the answer for the 13-card deck agrees with that for the 52-card deck to at least 10 decimal places! If you want to know why—read on!



Jenine's formula. The SNAP probability is

$$P = 1 - \left(\frac{51}{52}\right)^{52} = 0.636.$$

Interesting—and very close to the experimental estimate too. The class looked at her with a mixture of puzzlement and awe. I asked for an explanation.

"Well, every time I put down a card the chances are 51/52 that my partner's card <u>won't</u> be a match. Since this happens 52 times, the probability of having <u>no</u> match on every card is $(51/52)^{52}$. So the probability of having a SNAP is 1 minus this."

I asked the class what they thought; there was much uncertain nodding of heads, and a lot of close studying of <u>me</u> to discern what <u>I</u> thought.

How might we evaluate such an argument? Could we try it out in a simpler situation where we might know ahead of time what the answer was? Someone suggested we check it out with a smaller deck. For example, what would the formula give us for a deck of size 2? It would give

$$P = 1 - \left(\frac{1}{2}\right)^{2} = \frac{3}{4}.$$

But what is the SNAP probability for a deck with 2 cards? It should be easy to work this out simply by looking at all the possibilities. Indeed it is. With 2 cards, either the two decks are in the same order, or they're reversed—and each of these is equally likely. In the first case there's a SNAP and in the second there isn't, so *P* = 1/2.

Jenine’s formula doesn't seem to work. Hmm.

In order to focus the discussion, I asked whether anyone could give me a 2-card snap situation in which Jenine’s calculation *would* be correct—with an answer of 3/4. Someone came up with the idea of using a coin instead of cards. If two players each had a coin, and tossed them simultaneously, and called SNAP if they matched, then what’s the probability of getting a SNAP in two repetitions of this?

While we're at it, can you see how to modify the 52-card game to make Jenine’s formula correct? Somehow, you have to make successive plays (of a pair of cards) "independent", and have the same probability of SNAP at each play.

The answer is that after each play, you should gather up the two cards that were played and shuffle both decks. Then, at each play, the probability of no SNAP is 51/52, and the above calculation holds if you do this 52 times. An interesting and important digression!

Well, with each pair of flips, the probability of getting no SNAP is 1/2, and so the probability of no SNAP on both trials is $(1/2)^2$ = 1/4 , and so the probability of at least one SNAP is 1–(1/4) = 3/4. And that's Jenine’s formula.

Now what makes this different from the card problem? Well, it's that successive tosses of a coin are "independent" in the sense that what happens on the first toss has no effect on the possibilities for the next.

But now back to the original problem. The idea was now out there that we might try using smaller decks, and we pursued this—hoping maybe a pattern might be found as we increased the deck size. Suppose we denote by $P(n)$ the probability of a SNAP with a deck size of n. Let's try to calculate the first few $P(n)$. We already have $P(1) = 1$ and $P(2) = 1/2$, as noted above.

A deck size of 3.

A few moments were needed for $P(3)$, and there were a couple of different approaches. One, which was quite instructive, was to simply write down all possible pairs of decks and count up the number of SNAPs. If we call the cards of the first deck A, B and C, in order, then the 6 (= 3!) possibilities for the second deck are displayed as the six columns at the right.

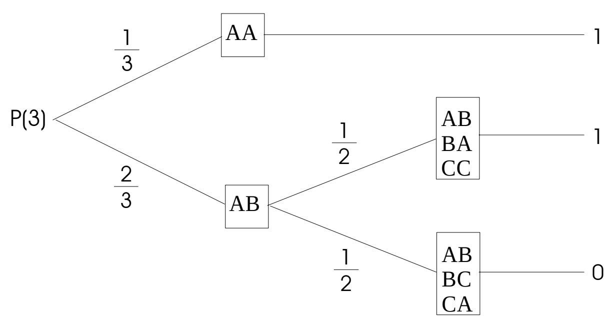
first	*second*					
A	A	A	B	B	C	C
B	B	C	A	C	A	B
C	C	B	C	A	B	A
snap?	✓	✓	✓			✓

The 4 checked cases give SNAPs, for a SNAP probability of 4/6 = 2/3. The nice thing about this approach is that it clearly displays the *probability* as the *proportion* of successes among all the equally likely possibilities.

Any drawbacks to this approach?—well it might get unwieldy for large decks. For example, for a deck of size 4, there are 4! = 24 columns—which you could still list in a couple of minutes—but for a 52-card deck, there are 52! and that's about 10^{68}. That's a lot when you consider that a computer which looked at a million columns per second running since the beginning of the universe would have covered less than 10^{24} columns.

The other kind of argument I got was an analysis of cases. Either there's a SNAP right away ($p = 1/3$) or not ($p = 2/3$). In the second case, the top cards are different, and if you look at how the remaining two pairs of cards can sit, there will be a SNAP with probability 1/2. This was an important approach for the purposes of generalizing to bigger decks, and so we worked a while on the presentation. The scheme I encouraged was the use of a tree.

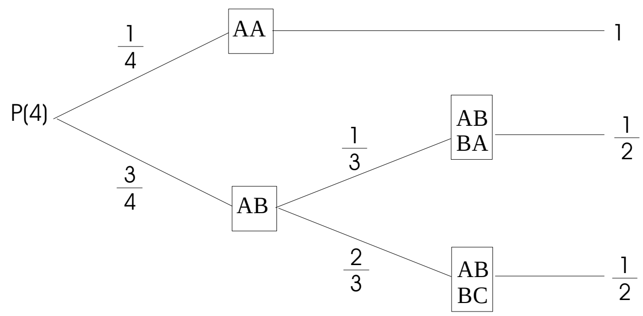
$$P(3) = \frac{1}{3}(1) + \frac{2}{3}\left(\frac{1}{2}\cdot 1 + \frac{1}{2}\cdot 0\right) = \frac{2}{3}$$



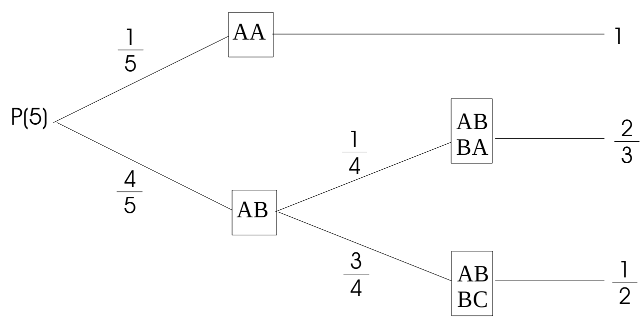
In this diagram, branches diverging from a common point represent different possibilities, and they are labeled with the probability of that alternative. The numbers at the extreme right are the SNAP probabilities of the different routes. In the first pair, we argue that either the first two cards are the same, *in which case we call them AA*, or they are different, *in which case we call them AB*. In the latter case, we then consider the two possibilities for how the rest of the deck might look, and record the SNAP probability in each case. Notice that we have tried to be economical in the numbers of branches, with the result that the way we use the symbols A, B and C is different from before, but more powerful. With this notational scheme, we don't keep track of the order of the cards in a deck (that's not what's important), but just what's opposite what. So B designates the card in deck-2 opposite A in deck-1, whatever that happens to be (different from A). And then the two boxes on the right describe the cases in which the deck-2 A is or is not opposite the deck-1 B.

Decks of size 4 and 5.
It took a few minutes to get P(4), and then a few more to get P(5).

$$P(4) = \frac{1}{4}(1) + \frac{3}{4}\left(\frac{1}{3}\cdot\frac{1}{2} + \frac{2}{3}\cdot\frac{1}{2}\right) = \frac{5}{8}$$



$$P(5) = \frac{1}{5}(1) + \frac{4}{5}\left(\frac{1}{4}\cdot\frac{2}{3} + \frac{3}{4}\cdot\frac{1}{2}\right) = \frac{19}{30}$$



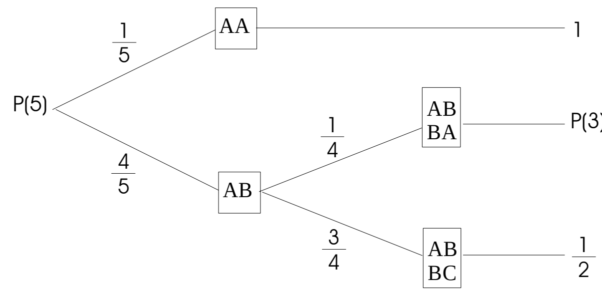
In both cases the argument has the same structure. Let the top card of deck-1 be A. Then maybe the card opposite A is A (top path). Otherwise, let the card opposite A be called B (bottom path). Now we ask what's opposite B? There are two cases: either A is opposite B (top path) or something else called C (bottom path). In each case, ask what the rest of the deck can look like. This last part of the argument is not detailed in the diagram, but only the answer is given. (We certainly can't continue to record all the branches!) Actually, a few students were having a bit of trouble following the argument, and we took a few moments to expand the lower branches on the P(5) tree just to see that the probabilities on the right were actually proportions of successes among the different possibilities.

How far are we going to go like this? Should we work out P(6), P(7), etc.? In fact, it was clear to everyone that this approach was going to get pretty tough pretty soon, and so we'd better find some patterns. Is there a pattern emerging in the numbers? Well, no one was able to see one. [*Except, another time I did this, a student did find quite a remarkable pattern. See problem 2.*]

n	*P*(*n*)
1	1
2	1/2
3	2/3
4	5/8
5	19/30

However, there did seem to be some patterns in the arguments themselves. At any rate, a number of students remarked that when investigating, for example, the case n=5, some of the arguments used were familiar—they had already occurred earlier. Is there some way to get a partial "recursion" out of this?

One observation that was made immediately is that, for the P(5) argument, when we get to $\begin{bmatrix} AB \\ BA \end{bmatrix}$, the rest of the deck is really a 3-card deck, and the SNAP probability is P(3). So the 2/3 in the P(5) diagram is really P(3). That's rather nice and we inserted that into the P(5) tree.



Can we do something similar with the lower part of the diagram? In this case we have $\begin{bmatrix} AB \\ BC \end{bmatrix}$. What can the rest of the deck look like? We went over the argument again, sensing real echos of previous arguments, but not quite able to put our finger on anything. And then a hand went up in the back of the room (that wonderful hand that shows you where the phrase "lending a hand" really came from). *If you forget about B, you really have a 4-deck, with an AC on top. That is:*

$$\boxed{\begin{array}{l} AB \\ BC \\ C? \\ D? \\ E? \end{array}} \quad \text{is equivalent to} \quad \boxed{\begin{array}{l} AC \\ C? \\ D? \\ E? \end{array}}$$

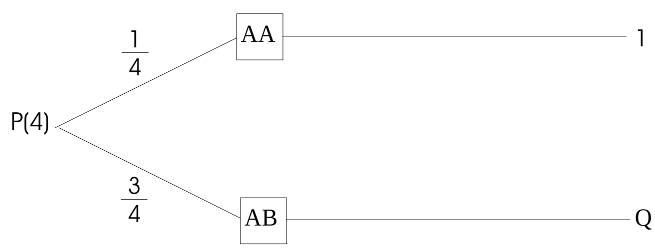
And the set-up on the right has occurred before—it's right up in P(4).

This is a wonderful observation. It identifies an isomorphism *(a word that literally means "same structure") between a branch of the P(5) tree and a branch of the P(4) tree. And that allows us to complete our recursion.*

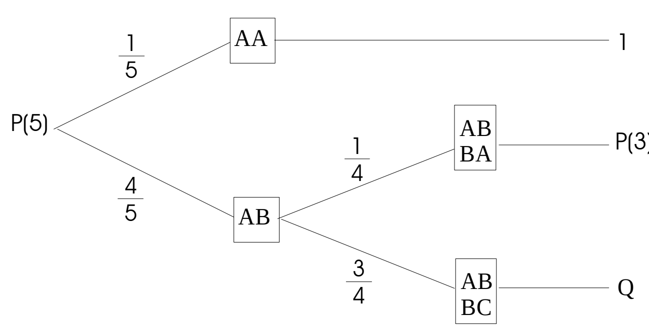
Even so—finishing the argument off, that is, making the Q substitution, is a sophisticated technical manoeuvre.

Indeed it has: it's the SNAP probability for the lower arm. That is, if we call this probability Q, then Q occurs both in P(4) and in P(5), as follows.

$$P(4) = \frac{1}{4}(1) + \frac{3}{4}(Q)$$



$$P(5) = \frac{1}{5}(1) + \frac{4}{5}\left(\frac{1}{4}P(3) + \frac{3}{4}Q\right)$$



Of course we know that Q is equal to 1/2, but that's not the point—the point is that we have a "connection" between the P(4) tree and the P(5) tree, and that's important because our strategy is to find some way of working up from small decks to large decks. If we substitute the quantity ☟Q from the P(4) equation into the P(5) equation, and simplify things a bit, we get

This is an enormous discovery and it has come Wham! right out of the blue. Not only is it a perfectly wonderful formula, its compelling form convinces us that if there's any justice in this universe, it should be generally true—that is, the corresponding formula should hold for P(6), P(7), etc. giving us a way of easily generating higher P values.

$$P(5) = \frac{1}{5}(1) + \frac{4}{5}\left(\frac{1}{4}P(3) + P(4) - \frac{1}{4}\right)$$

$$P(5) = \frac{1}{5}P(3) + \frac{4}{5}P(4) \ .$$

This remarkable formula displays P(5) as a "weighted average" of P(3) and P(4) with weights 1/5 and 4/5. Geometrically, if we think of the real line, this means that P(5) is between P(3) and P(4), and divides the distance between them in the ratio 🗐 to 😐, being closer to P(4) (since it has the bigger weight).

In fact, the same argument works quite generally and gives us the recursion:

$$\boxed{P(n) = \frac{1}{n}P(n-2) + \frac{n-1}{n}P(n-1)}$$

Again, P(*n*) is a weighted average of the two previous P values, *but the weights depend on n*—as *n* increases, P(*n*) gets relatively closer to P(*n*–1). If we track a few successive P values on the real line, we see that the fact that each P(*n*) is between the two previous values forces the P(*n*) to oscillate. They also seem to have to approach some limiting value. I wonder what *that* is.

What's next? Well, we could use the above formula to calculate quite easily a lot more values of P(*n*). In particular, we could get a computer to calculate P(52) in no time. In fact, one student had a programmable calculator, and programmed the routine. His program calculated and displayed each P(*n*) using the previous two P-values and the value of *n* which he stored each time. When he ran it, he noticed an interesting thing—around $n = 12$, his display stopped changing. It appeared that from $n = 12$ up, all the P(*n*) were the same to 9 decimal digits and all equal to:

$$P(n) = 0.632120559 \quad (n > 12) \ .$$

A useful corollary of this is that the game of SNAP with a 12-card deck is no different from the 52-card game, provided all you care about is whether or not you get a SNAP, and provided you're not sensitive to the 10th decimal place!

This recursion allowed us to easily extend our table of values for P(*n*) and we added a few entries, and stared at the table for awhile. Any patterns?

n	*P*(*n*)
1	1
2	1/2
3	2/3
4	5/8
5	19/30
6	91/144
7	531/840

At last someone tried successive differences P(*n*) – P(*n*–1) (one of the standard tools for any "what's the next term in this sequence" question). We got:

$$-\frac{1}{2},\ \frac{1}{6},\ -\frac{1}{24},\ \frac{1}{120},\ -\frac{1}{720},\ \frac{1}{5040}$$

Seen this before? It's the sequence 1/*n*! (*Eureka!*) with alternating signs.

$$-\frac{1}{2!},\ \frac{1}{3!},\ -\frac{1}{4!},\ \frac{1}{5!},\ -\frac{1}{6!},\ \frac{1}{7!},\dots$$

Now this allows us to actually write a formula for P(*n*). We start with P(1) = 1, and then get each P(*n*) from the previous one by adding the appropriate difference.

$$P(1) = 1$$
$$P(2) = 1-\frac{1}{2!}$$
$$P(3) = 1-\frac{1}{2!}+\frac{1}{3!}$$
$$P(4) = 1-\frac{1}{2!}+\frac{1}{3!}-\frac{1}{4!}$$

and in general

$$P(n) = 1-\frac{1}{2!}+\frac{1}{3!}-\frac{1}{4!}\dots\pm\frac{1}{n!}$$

.

It can be verified directly that this formula for P(n) solves the recursion. [A nice thing about solving equations: we may not know how to do it, but any time we manage, somehow, to get a candidate, we can always find out whether it's a solution by plugging it in!] In particular the formula gives, for n = 52, an exact expression for our standard-deck SNAP probability P(52).

That was quite a journey. In fact, this is a hard problem, but I find that even students in grade 9 and 10 can gain a lot from it. It is not acknowledged nearly enough that at the early stages in mathematical learning, it's important just to see structures "appear" in all their power and mystery, and produce answers to your problems, even when you don't quite understand what's happening.

Problems

1.(a) Consider a deck of 52 cards, numbered 1 to 51 and J (for joker). Now play SNAP with two such decks, with the rule that if the two jokers go down together this does not count as a snap. Calculate the snap probability P*(52).

(b) Consider a deck of 52 cards, numbered 1 to 50 and two identical jokers J and J. Now play SNAP with two such decks, with the rule that if any two jokers go down together this does not count as a snap. Calculate the snap probability P**(52).

(c) Consider a deck of 52 cards, numbered 1 to 52 and two identical jokers J and J. Now play SNAP with two such decks, with the rule that if any two jokers go down together this will count as a snap. Calculate the snap probability R(n).

2. Last time I did this problem, one of the students found a remarkable pattern. For each deck size *n*, he tabulated the number M(*n*) of cases (out of *n*!) in which there was no SNAP. For example, for *n*=3, out of the 6 possible orderings of deck 2, 4 are SNAPs, giving M(3) = 6–4 = 2. He then noticed that if you add two successive entries in the M column, and multiply by the higher *n* value, you get the next entry in the M column. For example:

$$(0+1)\times 2 = 2$$
$$(1+2)\times 3 = 9$$
$$(2+9)\times 4 = 44, \text{ etc.}$$

He also observed that this was not such a strange "rule". It also worked for factorials. This seems to me like a good starting point for an analysis of the problem.

n	M(*n*)
1	0
2	1
3	2
4	9
5	44
6	265

n	*n*!
1	1
2	2
3	6
4	24
5	120
6	720

3. Suppose that we are interested not in the probability of getting a SNAP, but in the *number* of SNAPs. That is, if we get a SNAP, we don't pick up the deck and start again, but we continue and keep track of the number of SNAPs we get by going to the end of the deck. If we do this again and again, what is the *average* number of SNAPs per deck? [There's an elegant way to think about this problem. A useful idea is found in Darts.]

4. A crucial stage in our development above was the discovery of the wonderful recursion:

$$\mathrm{P}(n) = \frac{1}{n}\mathrm{P}(n-2) + \frac{n-1}{n}\mathrm{P}(n-1) .$$

We could have finished up at this point by directly “solving” this recursion, but we didn’t know how. It’s worth noting that we have encountered second-order recursions before and we found a way to solve them. For example, in **Binet** Ex. 1, we solved the recursion

$$t_n = t_{n-1} + 2t_{n-2} .$$

Here the coefficients (1 and 2) are constants, independent of n, but in the P-recursion this is not the case, and our **Binet** method won’t work. What we did do, is look at the differences

$$\mathrm{Q}(n) = \mathrm{P}(n) - \mathrm{P}(n-1)$$

and we found a simple arithmetic pattern in these. Well, that gives me an idea for how we might have solved the P-recursion in the first place: define the differences Q(n) and try to obtain a recursion for them. I guess there’s a chance it might turn out to be a bit simpler than the P-recursion. Follow this idea up. What was it about the P-recursion that made this approach work? Can you formulate a general result about how to solve certain kinds of recursions?

5.(a) The formula

$$\mathrm{P}(n) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \ldots \pm \frac{1}{n!} .$$

might ring a bell for anyone who has seen the famous infinite series expansion for the exponential function:

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots$$

Now the series for P(n) is finite, while that for e^x goes on forever, but use the fact that for $n > 12$ the P(n) are all the same to 9 decimal places to find a simple formula for P(52) in terms of the number e accurate to 9 decimal places. Use your calculator to check your answer.

(b) (For those who know a bit about infinite series) In fact the accuracy of the approximation of (a) is considerably better than 9 decimal places. How good is it? [Hint: the series for P(n) is alternating.]

(c) Let’s go back to Jenine’s (invalid) argument that

$$\mathrm{P}(52) = 1 - (51/52)^{52} = 0.636.$$

Her answer is exceedingly close to the correct answer P(52) = 0.632 (everything to 3 decimal places). [In a sense that suggests that with a large deck successive cards are almost independent.] Can we understand why these should be so close?

[Hint. The trick here is to use another famous limit formula for e: $$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

Actually, we need the more general version: $$e^r = \lim_{n\to\infty}\left(1+\frac{r}{n}\right)^n$$

What does this give us for $1/e$?]