Mathcad Community Challenge January 2024
Statistics - Shuffling a deck of cards

This month's challenge is based around statistics, specifically permutations and combinations around shuffling a deck of cards. It's inspired by this article:

https://www.mcgill.ca/oss/article/did-you-know-infographics/there-are-more-ways-arrange-deck-cards-there-are-atoms-earth

1) Calculate the number of ways to shuffle a deck of cards. (Standard 52 card deck, no Jokers).

Hmm, "the number of ways to shuffle a deck of cards" - really?? Lets count: 1. riffle shuffle, 2. overhand shuffle, 3. ....
That's not what you mean? OK, guess what you really are interested in is the number of different ways the 52 different cards could be ordered. So the question is after the number of arrangements, or we may also call it sequences.

Mathematically we are talking here about "permutations without repetitions". The associated mathematical term is "enumerating combinatorics" rather than "statistics", as stated in the challenge description.
Formulas to be used can be found here for example: *Entry in the German Wikipedia*
The "formula" to be used is rather basic and should be known from school.
n different objects can be arranged in n! ways, where the ! denotes factorials.

The number of ways that 52 different cards can be arranged is $52! = 8.066 \cdot 10^{67}$. Yes, thats quite a lot!

2) Calculate the number of ways to shuffle a deck of cards if the suit (heart, diamond, spade, club) doesn't matter

So now we are talking about "permutations with repetitions".
We have 52 objects and 13 groups of four all equal objects
Applying the appropriate formula gives us

The number of sequences in a 52-card deck ignoring the suits is $\dfrac{52!}{4!^{13}} = 9.202 \cdot 10^{49}$

3) Calculate the number of ways to shuffle a deck of cards consisting of just the face cards (jacks, queens, and kings). If you could shuffle the deck one time a second, how long would it take to shuffle each possible way? (Express the result in hours, days, or years, whichever is easiest for people to grasp.

So now we are just dealing with a deck of card consisting of 12 different cards.
Mathematically again a permutation without repetiton as suits once again matter.

Number of ways to arrange 12 different cards: $arrangements := 12!$ $\quad arrangements = 4.79 \cdot 10^{8}$
Let $\boldsymbol{arrangement} := 1.$
At a rate of $rate := 1 \dfrac{\boldsymbol{arrangement}}{\boldsymbol{s}}$

it would take $\dfrac{arrangements}{rate} = 15.179 \ \boldsymbol{yr}$ to create each of the possible sequences.
Thats a bit **more than 15 years** and two months!

4) How long would it take to shuffle the face cards until there's a 0.5 probability you've shuffled the same way twice? If you add the aces, how long would it take for that same 0.5 probability that you've shuffled the same way twice?

Lets be more precise and assume the we should find **at least two** identical sequences.
We would also accept triples or more than one twin-pair, correct?
Furthermore I guess that the problem as given does not have a solution! We will not end up exactly at a 50% probability. So let's change the text to *"... until there's **at least** a 0.5 probability ..."*
And we should also assume a real random shuffle. Because if we would use a Faro shuffle (perfectly interleaving the cards of two 26-cards halves) where the outer cards remain in place, the original arrangement would be arrived at again after just 8 shuffles ;-)
After shuffling n seconds, we have created n sequences and we ask for the probability of at least two equal ones.
Thats similar to the well known Birthday Paradoxon.
See here Birthday problem in English Wikipedia or here Geburtstagsparadoxon in German Wikipedia

Or problem is similar but we now have (see 3)) $12! \rightarrow 479001600$ different possibilities and not just 365.
We can solve the given task using the probability of the complementary event, that is the probability for all sequences being different within n arangements.
The first one sure is different from all precedings as there are none -> probability is 1
The second one can be any of the 12!-1 arrangements differently from the first -> probability is (12!-1)/12!
The k-th sequence can be any of the 12!-k+1 remaining ones -> probability is (12!-k+1)/12!
So for n sequences (n not larger than 12!) we have to multiply all these probabilites for k=1 up to n and then subtract the result from 1 to get the probability for at least one matching pair.

$$P(n) := 1 - \prod_{k=1}^{n} \left(\frac{12! - k + 1}{12!} \right) \qquad \text{or} \qquad P(n) := 1 - \prod_{k=1}^{n-1} \left(1 - \frac{k}{12!} \right)$$

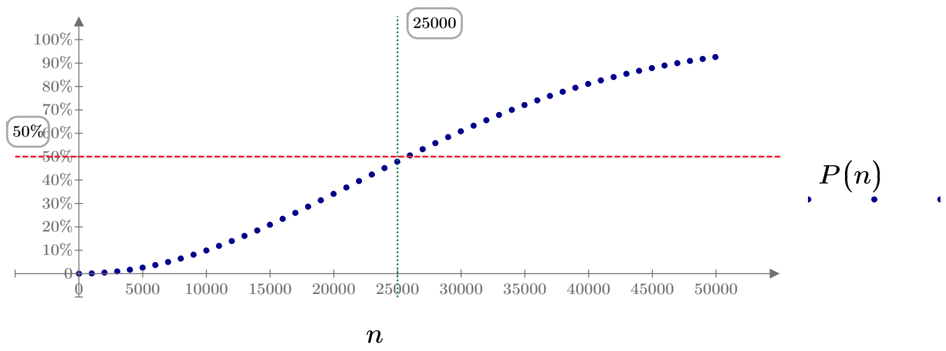
Could also be written as $P_2(n) := 1 - \operatorname{combin}(12!, n) \cdot \frac{n!}{12!^n}$ or $P_3(n) := 1 - \frac{n!}{(16! - n + 1)! \cdot 16!^{n-1}}$

but thats less useful for numerical evalation.

We now need to solve the inequality $P(n) \geq 0.5$ for n to get the first number n with probability at least 50%.
But neither Primes numerical methods (root function, solve block) nor the built-in symbolic engine (solve) are able to solve this equation because P(n) is not a continuous but rather a discrete function.

So lets have a look at the graph of the function P(n):

$n := 1, 1000 .. 50000$



just to make sure the definition would not interfere with subsequent calculations: $\operatorname{clear}(n)$

As we can clearly see, the value we are looking for is slightly above 25000.
So let's use a brute force attack, trying all values n from 25000 up and stop when we first have P(n)>=0.5

$$n50\% := \text{for } n \in 25000 .. 12! \;\; \| \; \text{if } P(n) \geq 0.5 \;\; \| \; \text{return } n$$

$$n50\% = 25770$$

Given that one shuffle is done in 1 second according to task 3), this means that it takes
more than seven hours: $n50\% \; s = 7.158 \; hr$

Adding the aces just increases the number of cards from 12 up to 16.
So we have $all := 16! = 20922789888000$ different possible sequences.

Thats $\frac{16!}{12!} = 43680$ times more than we had before!

We again define $P(n) := 1 - \prod_{k=1}^{n-1} \left(1 - \frac{k}{all}\right)$ and can find the 50% threshold similar as before using brute force.

It was a bit more tricky and difficult to find a suitable starting value for the attack and it would have been helpful if Primes native plots would provide facilities like trace and zoom as used from real Mathcad.

```
n50% := for n ∈ 5385640 .. 6000000
          if P(n) ≥ 0.5
            return n
```

$n50\% = 5385643$

So it would take $n50\%\ s = 62.334\ day$ which is **more than two months**
instead of just a few hours as before. Thats the effect of just four cards added!
I assumed that one shuffle of a 16-card deck takes the same time (1 sec.) as one shuffle of a 12-card deck.

The numeric engine has to multiply more than 5 millions values and its very likely that we run into some numerical inaccuracies adding up. So I won't fully trust the calculated value, even though the numeric evaluation seems to be OK.

$P(n50\% - 1) = 49.9999895\%$
$P(n50\%) = 50.00000237\%$

We could check using the symbolic engine

$$\begin{bmatrix} P(5385642) > \frac{1}{2} \\ P(5385643) > \frac{1}{2} \end{bmatrix} \rightarrow ?$$

but the calculation takes much too long (maybe it even fails - don't know as i cancelled the calculations) so I had to disable this region. If the calculated value would be correct, the result should have been the vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

We can check our result for plausibility using the well known inequality $1 - x < e^{x}$, valid for $x > 0$

Using this inequality we get $P(n) = 1 - \prod_{k=1}^{n-1} \left(1 - \frac{k}{all}\right) > 1 - \prod_{k=1}^{n-1} e^{\frac{-k}{all}} = P_{approx}(n)$

So lets define $P_{approx}(n) := 1 - \prod_{k=1}^{n-1} e^{\frac{-k}{all}} \xrightarrow{simplify} -e^{\frac{-n^2+n}{41845579776000}} + 1$

Thats a function which could be used for the symbolic solver

$$P_{approx}(n) > \frac{1}{2} \xrightarrow[]{\substack{assume,\, n>0 \\ solve,\, n}} n > \frac{\sqrt{167382319104000 \cdot \ln(2) + 1} + 1}{2} \xrightarrow{float} n > 5385643.0466819572239$$

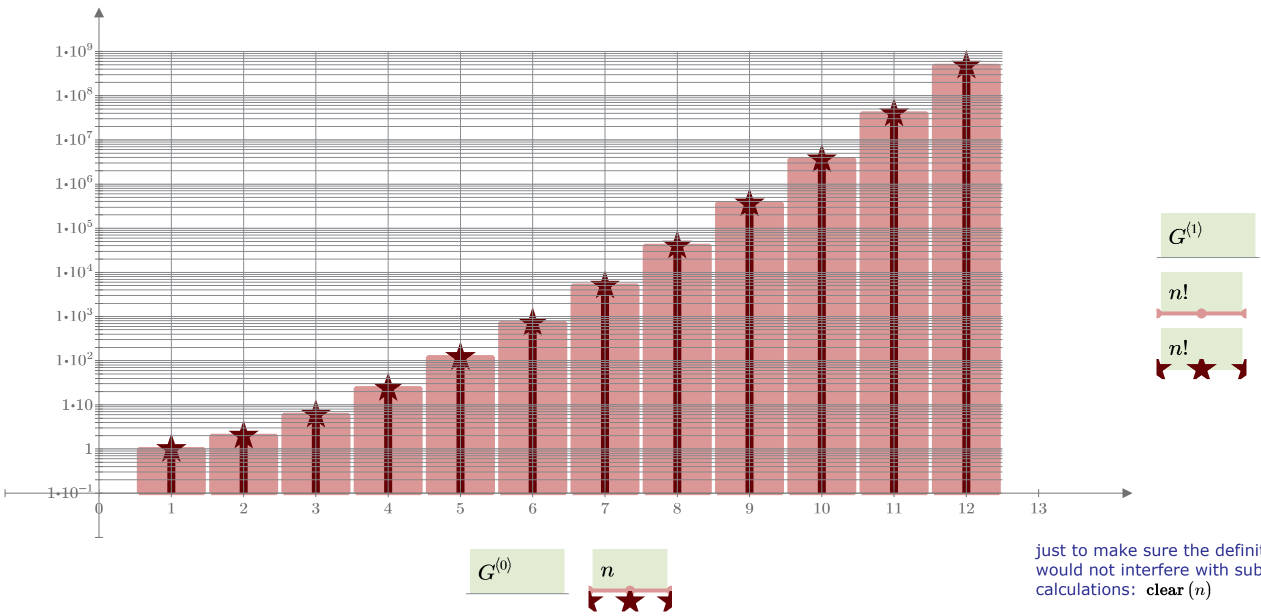
This shows that at least starting with $n = 5385644$ we have a probability above 50% for sure.
But because its just an approximation and $P(n) > P_{approx}(n)$ and also because the result of the symbolic solve is just slightly larger than 5385643, I now tend to trust the result of the numeric calculation.

BTW, we also could use the numeric "root" function to solve the equation (not the inequation):
$\mathrm{root}\left(P_{approx}(n) - 0.5, n, 1, 16!\right) = 5385643.047$

5) *Create a Chart Component to depict the number of ways to shuffle a deck from 1 to 12 cards. (Hint: use the Chart Component's logarithmic scale capability.)*

Because of a severe display bug in the Chart Component which is effective with my hardware configuration and also because the native plots are far faster and easier to use (albeit unfortunately much less capable) I am doing this simple plot using Primes native plot facility.

Number of different arrangements of a deck with *n* cards ($n := 1..12$)



just to make sure the definition would not interfere with subsequent calculations: $\mathbf{clear}\,(n)$

6) *How many cards would be in a deck in which the number of ways to shuffle would be on an order of magnitude with the number of atoms in the Earth? What about the universe? Assume 10 to the 50th power for number of atoms in the Earth and 10 to the 80th for the universe.*

We would have to solve the 'simple' equation

$n! = 10^{50} \xrightarrow{solve,\,n} ?$ or using the numeric root $\mathbf{root}\left(n! - 10^{50}, n, 10, 50\right) = ?$

Given that n! is defined for integers only its understandable that no solution could be provided.
So lets use the Gamma function instead of the factorials:

$\Gamma(n+1) = 10^{50} \xrightarrow{solve,\,n} ?$ resp. $\mathbf{root}\left(\Gamma(n+1) - 10^{50}, n, 10, 50\right) = ?$

Thats disappointing! I would have expected at least the symbolics to find the solution.The symbolics in Mathcad 15 has no problem doing so: $\Gamma(n+1) = 10^{50}$ solve, n $\rightarrow 41.293636460956417421$

So at least using legacy Mathcad we get the solution and can say that with a deck of **41 to 42 cards** (remove the picture cards and add one or two jokers) the number of ways to arrange the cards is in the order of magnitude with the number of atoms in the Earth.

Similarily we get the answer for 10^80: $\Gamma(n+1) = 10^{80}$ solve, n $\rightarrow 58.919951971368015923$

So just add about seven cards to a normal deck to get the **59 cards** necessary to obtain a number of possible arrangements of similar magnitude as the number of atoms in the universe!

But how could we come up with the solution for 6) using just Prime?
One way could be to plot the factorials similar to 5) to see in which range to look for the solution and then use trial and error. This is unsatisfactory, but you get the result relatively quickly and it doesn't have to be that precise for this task anyway, as nobody has actually ever counted the atoms ;-)

$41! = 3.345 \cdot 10^{49}$ $42! = 1.405 \cdot 10^{51}$ $58! = 2.351 \cdot 10^{78}$ $59! = 1.387 \cdot 10^{80}$

One additional idea is to use an approximation, the stirling formula for factorials:

$$fact(n) := \sqrt{2\,\pi \cdot n} \cdot \left(\frac{n}{e}\right)^n$$

$$\boxed{\mathbf{root}\left(fact(n) - 10^{50}, n, 10, 50\right)} = ?$$

$$\boxed{fact(n) = 10^{50}} \xrightarrow{solve, n} ?$$

Disappointing again.

Last try - let's use the logarithm:

$$\mathbf{root}\left(\ln(fact(n)) - \ln\left(10^{50}\right), n, 10, 50\right) = 41.294$$

$$\ln(fact(n)) = 50 \cdot \ln(10) \xrightarrow{solve, n} 41.294177077965907509$$

Heureka! It should not have been necessary to use the logarithm but at least its a workaround so that Prime at least gives us a solution using the Stirling approximation.

Sure it also works with 10^80: $\mathbf{root}\left(\ln(fact(n)) - \ln\left(10^{80}\right), n, 10, 70\right) = 58.92$

Fazit: Primes (version 9) symbolics and numerics are not capable enough to provide exact solutions like he symbolics in Mathcad 15 is able to to.
But by using the Stirling formula as an approximation and the trick of logarithmizing the equation, Prime ultimately managed to get a solution that was at least good enough for this task.

7) Add a header and footer to your worksheet.

Done! :-)