Probability of Dice Sequences

Jonathan Stokes

February 16, 2024

In rolling 6 dice what is the probability of rolling $1, 2, 3, 4$? Or more generally in rolling n dice what is the probability for rolling a length m sequence starting with 1? It is straight forward to estimate the probability of rolling these sequences using $100,000$ monte carlo simulations per dice count, see fig. 2.

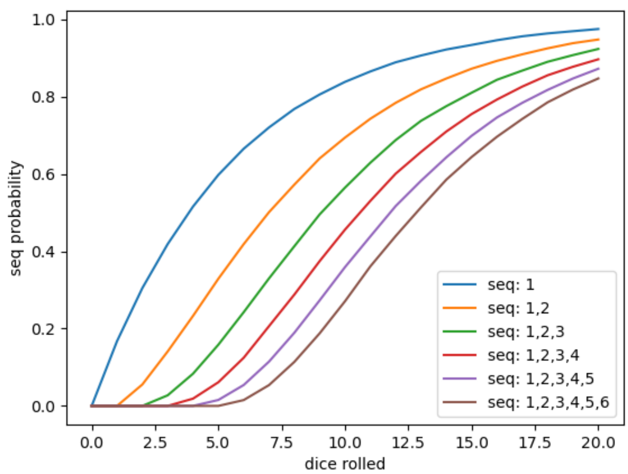

Figure 1

At first this looks similar the probability of drawing a straight in 5 card draw poker. If one draws a straight in 5 card draw poker including a royal flush there are

1. $\binom{10}{1}$ ways to choose the low card in the straight, recalling that the ace may be 1 or follow the king in a straight.

2. $\binom{4}{1}^5$ ways to select the card suits.

and generally there are $\binom{52}{5}$ ways of drawing 5 cards. So the probability of drawing the straight in 5 card draw poker is

$$\frac{\binom{10}{1}\binom{4}{1}^5}{\binom{52}{5}} \approx 0.00394 \tag{1}$$

So generalizing the probability of drawing a length n straight for $n \in \{2 \dots 13\}$ is

$$\frac{\binom{15-n}{1}\binom{4}{1}^n}{\binom{52}{n}} \tag{2}$$

Now lets try applying this same logic to the probability of rolling a length m sequence in a roll of n dice for $n \geq m$ dropping the requirement that the sequence start at 1.

1. $\binom{7-m}{1}$ ways to choose the low number in the sequence.
2. $\binom{6}{1}^{n-m}$ ways to roll the remaining dice.

and generally there are $\binom{6}{1}^n$ ways of rolling n dice. Therefore the probability of rolling a length m sequence in a roll of n dice is,

$$\frac{\binom{7-m}{1}\binom{6}{1}^{n-m}}{\binom{6}{1}^n} = \frac{\binom{7-m}{1}}{\binom{6}{1}^m} = \frac{7-m}{6^m} \tag{3}$$

which looking at fig. 2 is clearly incorrect for $m = 6$ and sequence $1, 2, 3, 4, 5, 6$ as eq. (3) is constant in n the number of dice rolled.

What did we do wrong? The term $\binom{7-m}{1}$ implicitly assumes that on any roll of the dice there are always 6 different numbers to choose from. In the case of cards this type of assumption is correct since every deck of has 13 distinct values (14 if the ace can be low and high). However in the case of a roll of 6 dice we may not roll a 1 leaving only 5 numbers to choose from in which case there are not $\binom{7-m}{1}$ ways to choose the lowest number in the sequence.

So what is the solution? Well the short answer is that I have not worked out a closed for solution for the general case of a sequence of length m starting at 1. However below I give a solution for the case where $m = 2$.

What is the probability of rolling a 1 and a 2 when rolling 3 dice? First lets define the following events.

$$A\text{: 1 does appear in a roll of 3 dice.} \tag{4}$$

$$B\text{: 2 does appear in a roll of 3 dice.} \tag{5}$$

To find the probability of events A and B we need to find the probability of $\neg A$ and $\neg B$ which is the probability of rolling any of the other 5 dice faces of the 3 independent dice. Therefore

$$P(\neg A) = P(\neg B) = \left(\frac{5}{6}\right)^3 \tag{6}$$

and by analogy the probability of not rolling a 1 and a 2 is

$$P(\neg A \cap \neg B) = \left(\frac{4}{6}\right)^3 \tag{7}$$

It follows from eq. (6) that the probability of rolling either a 1 or a 2 is

$$P(A) = P(B) = 1 - P(A) = 1 - \left(\frac{5}{6}\right)^3 = \frac{91}{216} \tag{8}$$

To answer our initial question above we need to find $P(A \cap B)$.

Given eq. (7) the conditional probability of not rolling a 1 given not rolling a 2 is

$$P(\neg A|\neg B) = \frac{P(\neg A \cap \neg B)}{P(\neg B)} = \frac{\left(\frac{4}{6}\right)^3}{\left(\frac{5}{6}\right)^3} = \left(\frac{4}{5}\right)^3 \tag{9}$$

and given the law of total probability

$$P(A) = P(A|B)P(B) + P(A|\neg B)P(\neg B) \tag{10}$$

Therefore from eq. (9) and since $P(A|\neg B) + P(\neg A|\neg B) = 1$,

$$P(A|B) = \frac{P(A) - P(A|\neg B)P(\neg B)}{P(B)} = \frac{\left(1 - \left(\frac{5}{6}\right)^3\right) - \left(1 - \left(\frac{4}{5}\right)^3\right)\left(\frac{5}{6}\right)^3}{1 - \left(\frac{5}{6}\right)^3} = \frac{30}{91} \tag{11}$$

and it follows that

$$P(A \cap B) = P(A|B)P(B) = \left(\frac{30}{91}\right)\left(\frac{91}{216}\right) = \frac{5}{36} \approx 0.138 \tag{12}$$

Looking at fig. 2 this looks about right.

Now does this generalize for $n \geq 2$ dice rolled? Lets first define the following events

$$A_n\text{: 1 does appear in a roll of } n \text{ dice.} \tag{13}$$

$$B_n\text{: 2 does appear in a roll of } n \text{ dice.} \tag{14}$$

$$\begin{aligned} P(A_n \cap B_n) &= \left(\frac{\left(1 - \left(\frac{5}{6}\right)^n\right) - \left(1 - \left(\frac{4}{5}\right)^n\right)\left(\frac{5}{6}\right)^n}{1 - \left(\frac{5}{6}\right)^n}\right)\left(1 - \left(\frac{5}{6}\right)^n\right) \\ &= \left(1 - \left(\frac{5}{6}\right)^n\right) - \left(1 - \left(\frac{4}{5}\right)^n\right)\left(\frac{5}{6}\right)^n \\ &= 1 - 2\left(\frac{5}{6}\right)^n + \left(\frac{2}{3}\right)^n \end{aligned}$$

and as $n \to \infty$, $P(A_n \cap B_n) \to 1$ as expected.

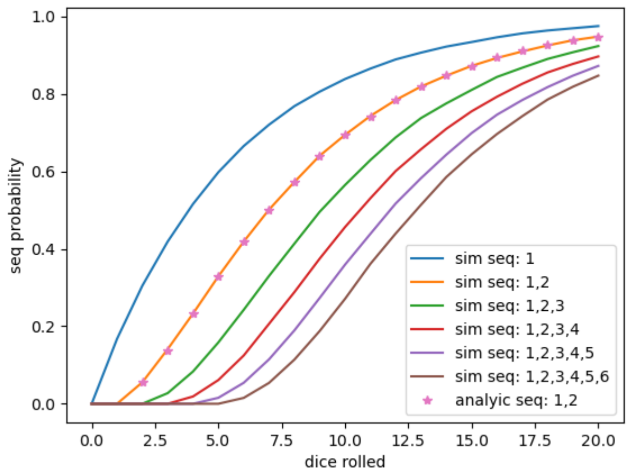

Figure 2

?? compares the analytic and simulated results.
ADD plot for the analytical solution

Below seems overly complicated and probably wrong. Perhaps try making the recursion more explicit. Show how m=3 is build from m=2? Also check that m=3 solution is correct in simulations.

Now say we want to know the probability of rolling the sequence of dice 1 to m where $m \in \{2,3,4,5\}$. Given $n > m$ (DOES THIS WORK FOR n=m????, what happens if m=6???????) lets define the events

$$A_{n,i}\text{: } i \text{ does appear in a roll of } n \text{ dice.} \tag{15}$$

for $i \in \{1 \ldots m\}$.

$$P(\neg A_{n,i}) = \left(\frac{5}{6}\right)^n \tag{16}$$

$$P\left(\bigcap_{i=1}^{m} \neg A_{n,i}\right) = \left(\frac{6-m}{6}\right)^n \tag{17}$$

Now expanding the conditional probability **(that is not the right term)** we know that.

$$P\left(\bigcap_{i=1}^{m} A_{n,i}\right) = P\left(A_{n,1}|\bigcap_{i=2}^{m} A_{n,i}\right) P\left(\bigcap_{i=2}^{m} A_{n,i}\right) \tag{18}$$

$$= \prod_{j-1}^{m} P\left(A_{n,j}| \bigcap_{i=j+1}^{m} A_{n,j}\right) \tag{19}$$

and we also know from above that $P\left(A_{n,j}|\bigcap_{i=j+1}^{m} A_{n,j}\right)$ has the form

$$P\left(A_{n,j}| \bigcap_{i=j+1}^{m} A_{n,j}\right) = \frac{P(A_{n,j}) - P(A_{n,j}|\bigcap_{i=j+1}^{m} \neg A_{n,i})\left(1 - P\left(\bigcap_{i=j+1}^{m} \neg A_{n,j}\right)\right)}{\left(1 - P\left(\bigcap_{i=j+1}^{m} A_{n,j}\right)\right)} \tag{20}$$

Now from eq. (17) we can see that

$$\left(1 - P\left(\bigcap_{i=j+1}^{m} \neg A_{n,j}\right)\right) = \left(\frac{6-m+i}{6}\right)^n \tag{21}$$

and from eq. (16)

$$P(A_{n,j}) = 1 - \left(\frac{5}{6}\right)^n \tag{22}$$

also

$$1 - P\left(\bigcap_{i=j+1}^{m} A_{n,j}\right) = 1 - P\left(\bigcap_{i=1}^{c=m-j} A_{n,i}\right) \tag{23}$$

setting up a recursion by substitution into eq. (18) with $c = 2$ as the base case which we solved for above. This leaves

$$P(A_{n,j}| \bigcap_{i=j+1}^{m} \neg A_{n,i}) = 1 - P(\neg A_{n,j}| \bigcap_{i=j+1}^{m} \neg A_{n,i}) \tag{24}$$

$$= 1 - \frac{P(\bigcap_{i=j}^{m} \neg A_{n,i})}{P(\neg A_{n,j})} \tag{25}$$

$$= 1 - \frac{\left(\frac{6-m+j}{6}\right)^n}{\left(\frac{5}{6}\right)^n} \tag{26}$$

$$= 1 - \left(\frac{6-m+j}{5}\right)^n \tag{27}$$

substitute all this back into eq. (20) and the result of that into eq. (18).