Summations in Probability

Yongjun Lee

ylee2021@chadwickschool.org

Chadwick International School

1. Introduction

Summations are central to the calculation of probability in more complex binomial distributions. However, effectively using them to solve complex probability systems requires a significant level of critical thinking. Today, we will examine this concept through the lens of a sample problem:

What is the average number of six-sided die rolls in order to get a 6? [1]

At first glance, the problem seems trivial; since there is a $\frac{1}{6}$ probability that one rolls a 6 on each roll, it should take 6 rolls on average to reach a completely certain probability of

$\frac{1}{6} \times 6 = 1.$ This will be our initial conjecture:

The average number of six-sided die rolls needed to get a 6 is 6. (Conjecture)

However, this intuitive approach is not enough to provide a concrete answer to the question; this paper will explore how infinite summations come into play in proving the above conjecture and their wider application in the field of combinatorics.

2. Background

a) Binomial Distributions

A binomial distribution is formally defined as the "discrete probability distribution of obtaining exactly x successes out of n true-false trials where the result of each trial is true with probability p and false with probability $q = 1 - p$" [2]. The symbol for such a distribution is $B(n, p)$, with the probability that exactly x successes are obtained being

$$P(x) = nCx \times p^x(1-p)^{n-x}$$

We can note here that the core problem also indicates such a binomial distribution with $p = \frac{1}{6}$, as it calls for repeated trials of an outcome with probability $\frac{1}{6}$ (in this case, rolling a 6 on a six-sided die). Other binomial distributions in daily life include sampling with replacement, where the probability field stays constant due to the replacement of samples taken from the total set.

b) Expected Values

The expected value of a probability distribution is the average resulting value of over a large number of trials [3]. In practice, the expected value is calculated by multiplying each possible output value by the probability that it will be outputted, that is:

$$E(x) = \sum_{i=1}^{n} x_i p_i \quad (1)$$

Here, the system has n different output values x_1 through x_n with respective probabilities p_1 through p_n. In the case of our problem, since we are calculating the *average number of trials* needed to achieve our desired outcome of getting a 6, the values x_1 through x_n will be all of the possible numbers of trials that we might undergo in order to accomplish our outcome, whereas p_1 through p_n will be the respective probabilities that the outcome will be accomplished in said number of trials. We can also note here that since it is theoretically possible to roll a six-sided die an infinite number of times without ever getting a 6, n is infinite and therefore $E(x)$ constitutes an infinite summation.

c) Finite and Infinite Geometric Summations

An geometric summation is the addition of consecutive terms within an geometric sequence, that is, a sequence where each term is the product of the previous term and the common ratio r. Therefore, the summation S_n of the first n terms of the geometric sequence a_n with first term a and common ratio r can be expressed as

$$S_n = \sum_{k=1}^{n} a_k = a + ar + ar^2 + ar^3 + ar^4 + \cdots + ar^{n-1} \quad (2)$$

However, we can convert Equation 2 into a finite formula for S_n using a trick that will come in handy later during our proof stage:

$$S_n = a + ar + ar^2 + ar^3 + ar^4 + \cdots + ar^{n-1}$$

$$rS_n = ar + ar^2 + ar^3 + ar^4 + ar^5 + \cdots + ar^n$$

$$S_n - rS_n = a - ar^n$$

$$S_n(1 - r) = a(1 - r^n)$$

$$S_n = \frac{a(1 - r^n)}{(1 - r)} \ (3)$$

In order to calculate an *infinite* geometric sum, however, we must extend this equation further, namely by having n approach to infinity:

$$S_\infty = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{a(1 - r^n)}{(1 - r)}$$

Here, we can further simplify the equation by dividing it into cases:

$$\lim_{n\to\infty} r^n = \infty \ (r < -1 \ or \ r > 1)$$

$$\lim_{n\to\infty} r^n = 0 \ (-1 < r < 1)$$

Since the first case results in a divergent outcome with no meaningful result, we can focus on the second result with its unique limitations:

$$S_\infty = \frac{a}{1 - r} \ (-1 < r < 1) \ (4)$$

This equation is the official formula for the summation of an infinite geometric sequence with common ratio $-1 < r < 1$; it will be crucial to our approach in solving the original problem.

3. The Proof

The easiest way to solve infinite-sum problems is usually to list out a couple of terms before looking for a pattern.

For example, when the number of trials is 1, we can find that the probability of getting a 6 in one trial is $\frac{1}{6}$; when the number of trials is 2, we would need to roll a non-6 followed by rolling a 6, so the probability is $\frac{5}{6}\times\frac{1}{6}$, and so forth. Each subsequent trial adds the rolling of a non-6 and therefore a factor of $\frac{5}{6}$ to the result, as seen in this table:

Table 1: Probability values for number of die rolls ranging from 1 to 5.

Number of Trials	Probability
1	$\frac{1}{6}$
2	$\frac{5}{6}\times\frac{1}{6}$
3	$\left(\frac{5}{6}\right)^2\times\frac{1}{6}$
4	$\left(\frac{5}{6}\right)^3\times\frac{1}{6}$
5	$\left(\frac{5}{6}\right)^4\times\frac{1}{6}$

This allows us to express the expected value E as an infinite summation using Equation 2:

$$E = 1\times\frac{1}{6} + 2\times\frac{1}{6}\times\frac{5}{6} + 3\times\frac{1}{6}\times\left(\frac{5}{6}\right)^2 + \cdots$$
$$= \sum_{n=1}^{\infty} n\times\frac{1}{6}\times\left(\frac{5}{6}\right)^{n-1} \quad (5)$$

Here, we notice that we cannot immediately use the results of Equation 4 to come to a constant value due to the factor of n; this infinite sum in particular is known as an arithmetic-geometric sum, and it is solvable using the method we previously used to derive Equation 3, as follows:

$$E = 1\times\frac{1}{6} + 2\times\frac{1}{6}\times\frac{5}{6} + 3\times\frac{1}{6}\times\left(\frac{5}{6}\right)^2 + 4\times\frac{1}{6}\times\left(\frac{5}{6}\right)^3 + \cdots$$
$$\frac{5}{6}E = 1\times\frac{1}{6}\times\frac{5}{6} + 2\times\frac{1}{6}\times\left(\frac{5}{6}\right)^2 + 3\times\frac{1}{6}\times\left(\frac{5}{6}\right)^3 + \cdots$$

$$E - \frac{5}{6}E = 1 \times \frac{1}{6} + 1 \times \frac{1}{6} \times \frac{5}{6} + 1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^2 + 1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^3 + \cdots$$

Now, having arrived at an infinite geometric sum, we can use Equation 4 to reach our answer:

$$\frac{1}{6}E = \frac{1}{6} + \frac{1}{6} \times \frac{5}{6} + \frac{1}{6} \times \left(\frac{5}{6}\right)^2 + \frac{1}{6} \times \left(\frac{5}{6}\right)^3 + \cdots$$

$$\frac{1}{6}E = \frac{\frac{1}{6}}{1 - \frac{5}{6}} = \frac{\frac{1}{6}}{\frac{1}{6}} = 1$$

$E = 6$ thus proving our original conjecture.

4. Alternative Methods

a) The "Grid of Infinity"

Another, more creative method to solving this problem is by splitting Equation 5 into an infinite number of infinite summations. The way we go about this is by first artificially eliminating the factor of n in the front of the equation:

$$1 \times \frac{1}{6} + 1 \times \frac{1}{6} \times \frac{5}{6} + 1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^2 + 1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^3 + 1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^4 + \cdots = \frac{\frac{1}{6}}{1 - \frac{5}{6}} = 1$$

We can then go about rectifying our alterations to the equation, namely by adding another infinite geometric summation starting with the term $\frac{1}{6} \times \frac{5}{6}$ so that the coefficient for this term goes from 1 back to 2, as below:

$$1 \times \frac{1}{6} + 1 \times \frac{1}{6} \times \frac{5}{6} + 1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^2 + 1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^3 + 1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^4 + \cdots = \frac{\frac{1}{6}}{1 - \frac{5}{6}} = 1$$

$$1 \times \frac{1}{6} \times \frac{5}{6} + 1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^2 + 1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^3 + 1 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^4 + \cdots = \frac{\frac{1}{6} \times \frac{5}{6}}{1 - \frac{5}{6}} = \frac{5}{6}$$

Now, since the coefficient of the next term $\frac{1}{6} \times \left(\frac{5}{6}\right)^2$ is at 2, we can once again add another infinite summation starting with $\frac{1}{6} \times \left(\frac{5}{6}\right)^2$ to restore the original coefficient of 3, and so on, until an infinitely large triangular "grid" is formed:

$$1\times\frac{1}{6}+1\times\frac{1}{6}\times\frac{5}{6}+1\times\frac{1}{6}\times\left(\frac{5}{6}\right)^2+1\times\frac{1}{6}\times\left(\frac{5}{6}\right)^3+1\times\frac{1}{6}\times\left(\frac{5}{6}\right)^4+\cdots=\frac{\frac{1}{6}}{1-\frac{5}{6}}=1$$

$$1\times\frac{1}{6}\times\frac{5}{6}+1\times\frac{1}{6}\times\left(\frac{5}{6}\right)^2+1\times\frac{1}{6}\times\left(\frac{5}{6}\right)^3+1\times\frac{1}{6}\times\left(\frac{5}{6}\right)^4+\cdots=\frac{\frac{1}{6}\times\frac{5}{6}}{1-\frac{5}{6}}=\frac{5}{6}$$

$$1\times\frac{1}{6}\times\left(\frac{5}{6}\right)^2+1\times\frac{1}{6}\times\left(\frac{5}{6}\right)^3+1\times\frac{1}{6}\times\left(\frac{5}{6}\right)^4+\cdots=\frac{\frac{1}{6}\times\left(\frac{5}{6}\right)^2}{1-\frac{5}{6}}=\left(\frac{5}{6}\right)^2$$

$$1\times\frac{1}{6}\times\left(\frac{5}{6}\right)^3+1\times\frac{1}{6}\times\left(\frac{5}{6}\right)^4+\cdots=\frac{\frac{1}{6}\times\left(\frac{5}{6}\right)^3}{1-\frac{5}{6}}=\left(\frac{5}{6}\right)^3$$

By observing the results of these infinite summations, we can see that they themselves form an infinite geometric progression with starting term 1 and common ratio $\frac{5}{6}$, while the deconstructed left-hand sides of the equations add to our original E value defined in Equation 5. Therefore, we can now arrive at our answer using a straightforward application of Equation 4:

$$E=1+\frac{5}{6}+\left(\frac{5}{6}\right)^2+\left(\frac{5}{6}\right)^3+\cdots=\frac{1}{1-\frac{5}{6}}$$

$$E=6$$

b) Inductive Reasoning

However, the simplest way to solve this problem is by comparing the states before and after the first roll. Let us once again assume that the mean number of trials required is E. If the first toss is a success, then the resulting number of trials is clearly 1; if the first toss is a failure, given that the probability fields remain unchanged after our failure, there are still m trials remaining on average, and therefore the average number of trials is $E+1$.

This indicates that

$$E=\frac{1}{6}\times 1+\frac{5}{6}(E+1)$$

$$E-\frac{5}{6}E=\frac{1}{6}\times 1+\frac{5}{6}$$

$$\frac{1}{6}E = 1$$

$E = 6$ yielding the same result as above.

5. Generalization

We can also observe that all three lines of reasoning hold for any probability value; simply substituting p for $\frac{1}{6}$ and $1 - p$ for $\frac{5}{6}$ results in the following:

$$E = p \times 1 + (1 - p)(E + 1)$$

$$E - (1 - p)E = p + 1 - p$$

$$pE = 1$$

$$E = \frac{1}{p}$$

Therefore, we can generalize our original proof and synthesize our findings as follows:

Given repeated random trials of a binomial sample space where outcome A has probability $p > 0$, $E = \frac{1}{p}$ *where* E *is the expected number of trials required to yield the first instance of A. (Theorem)*

6. Extension: Deck of Cards

A similar expected-value problem worth our attention concerns the gambler's perennial struggle to calculate the odds of drawing an ace:

A standard deck of cards has 52 cards, 4 of which are aces. On average, how many cards must be drawn off the top of a standard deck of cards in order for the first ace to be drawn?

Once again, it seems prudent to start by listing some values. Obviously, the probability that an ace is drawn on the first card is $\frac{4}{52}$. For the ace to be drawn on the second card, we would need to draw a non-ace (with a probability of $\frac{48}{52}$) followed by drawing an ace with a probability of $\frac{4}{51}$. Note here the change in the denominator; as we are digging our way through the deck, we are sampling without replacement, hence the reduction in number of cards remaining. In similar fashion, we can identify the probability values for $1 \leq x \leq 5$ where x is the number of cards drawn, as in the table below.

Table 3: Probability values for number of cards drawn from 1 to 5.

Number of Cards	Probability
1	$\frac{4}{52}$
2	$\frac{48}{52} \times \frac{4}{51}$
3	$\frac{48}{52} \times \frac{47}{51} \times \frac{4}{50}$
4	$\frac{48}{52} \times \frac{47}{51} \times \frac{46}{50} \times \frac{4}{49}$
5	$\frac{48}{52} \times \frac{47}{51} \times \frac{46}{50} \times \frac{45}{49} \times \frac{4}{48}$
6	$\frac{48}{52} \times \frac{47}{51} \times \frac{46}{50} \times \frac{45}{49} \times \frac{44}{48} \times \frac{4}{47}$

Unfortunately, the changing denominators also mean that we no longer have a finite geometric sequence that allows us to simplify this system using Equation 3. However, we can notice that the 48's in both the numerator and the denominator of the result of $x = 5$ cancel out, as do both the 48's and 47's in the row of $x = 6$. This allows us to deduce that all of the other denominators less than 48 must also inevitably cancel each other out, leaving behind a denominator of $52 \times 51 \times 50 \times 49$.

Meanwhile, with regards to the numerator, we observe that it also consists of 4 times the product of three consecutive integers where the largest integer is $52 - x$; for example, when $x = 5$, the highest factor of the numerator is $52 - x = 47$ and the numerator itself is $47 \times 46 \times 45 \times 4$. Note that this rule applies even to the cases in which reduction does not occur; for example, when $x = 3$, artificially creating a denominator of $52 \times 51 \times 50 \times 49$ by multiplying $\frac{49}{49} = 1$ to the probability yields $\frac{48}{52} \times \frac{47}{51} \times \frac{4}{50} \times \frac{49}{49}$ and therefore a numerator of $49 \times 48 \times 47 \times 4$ as consistent with the pattern.

This allows us to express the probability $P(n)$ for $x = n$ as follows:

$$P(n) = \frac{4(52 - x)(51 - x)(50 - x)}{52 \times 51 \times 50 \times 49}$$

In turn, we can now use Equation 2 once again to calculate the expected value E for all x values from 1 (ace on the first card) to 49 (all of the non-ace cards plus one ace are drawn)

$$E = \sum_{x=1}^{49} \frac{4(52-x)(51-x)(50-x)}{52 \times 51 \times 50 \times 49}$$

Using the summation formulas (which lie outside the scope of this investigation), we can calculate that $E = 10.6$; that is, a gambler would need to draw 10.6 cards on average off the top of a random 52-card deck before getting his hands on an ace.

7. Conclusion

Through this investigation, we were able to explore various applications of finite and infinite summations in probability, particularly when pertaining to expected values and repeated trials. The problem-solving techniques and theorems developed throughout this inquiry are applicable to many other types of unique binomial distributions in a relatively lesser-known branch of combinatorics.

8. References

[1] Mosteller, Frederick. *Fifty Challenging Problems in Probability with Solutions.* Dover Publications Inc., 1965, https://mbapreponline.files.wordpress.com/2013/07/fifty_challenging_problems_in_2_.pdf

[2] Weisstein, Eric W. “Binomial Distribution.” From *MathWorld*—A Wolfram Web Resource. https://mathworld.wolfram.com/BinomialDistribution.html.

[3] Grinstead, Charles M., and J. Laurie Snell. “Expected Value and Variance.” *Introduction to Probability,* The American Mathematical Society, 1988, https://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/amsbook.mac.pdf.