Probability Part I - Experiments with Random Variables

Yan Tao

Advanced 1 - Fall 2023

1 Review of Definitions

This week, we'll perform some experiments related to random chance. This section will serve as a refresher on the topic of probability, with relevant definitions and examples.

Definition 1 *A finite* ***probability space*** *consists of the following:*

- *A* ***sample space*** Ω*, which is a finite set.*
- ***Outcomes****, which are the elements of the sample space* Ω*.*
- ***Events****, which are subsets of the sample space* Ω*. In other words, events are sets of outcomes.*[1]
- *A* ***probability function*** $\mathbb{P}$*, which assigns a real number between* 0 *and* 1 *to every event, such that*
  - $\mathbb{P}(\emptyset) = 0$ *and* $\mathbb{P}(\Omega) = 1$ *(that is, "nothing" has probability* 0 *and "everything" has probability* 1*).*
  - *For disjoint events* E_1, E_2*,* $\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$ *(that is, if* E_1 *and* E_2 *are disjoint, the probability of* E_1 *or* E_2 *is the sum of the probabilities of* E_1 *and* E_2*).*

The probability space as a whole is denoted $(\Omega, \mathbb{P})$*.*

While we can abstractly define probability spaces, they largely arise as a model of real-life chance-based "outcomes" and "events" (hence the names). In these cases, we will need to figure out what Ω and $\mathbb{P}$ are. The sample space is the set of outcomes, so we should write down all possible outcomes, and then the chance of each outcome.

Problem 1 *Consider an experiment where we flip a coin three times.Find the sample space* Ω*, by listing all possible outcomes.*

Solution: $\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$

[1] For an infinite sample space, it may be the case that it's not possible (or desireable) to measure the probability of every possible subset. But because we will assume Ω is finite, every subset will in fact be an event.

Events are often described by words. For instance, when flipping a fair coin three times, the event $\{o \in \Omega \mid o \text{ has exactly two heads}\}$is the same as the subset $\{HHT, HTH, THH\}$ of Ω.

Problem 2 *Consider the same experiment where we flip a fair coin three times.*

- *Find* $\mathbb{P}(\{o \in \Omega \mid o \textit{ has at least two heads}\})$.

  Solution: Adding together the four outcomes gives $1/8 + 1/8 + 1/8 + 1/8 = 1/2$.

- *Find* $\mathbb{P}(\{o \in \Omega \mid o \textit{ has an odd number of heads}\})$.

  Solution: Adding together the four outcomes gives $1/8 + 1/8 + 1/8 + 1/8 = 1/2$.

- *Find* $\mathbb{P}(\{o \in \Omega \mid \textit{Not all three flips of } o \textit{ are heads}\})$ *two different ways, using both rules for* $\mathbb{P}$.

  Solution: Adding together the seven outcomes gives $7/8$. *Alternatively, we notice that this event with the disjoint event* $\{HHH\}$ *together make up the whole sample space, so its probability plus* $1/8$ *is* 1.

Problem 3 *Consider the same experiment where we flip a fair coin three times. Let's try to describe the whole probability function* $\mathbb{P}$.

- *How many events are there? In order to determine* $\mathbb{P}$, *do we need to find the probability of every event, or is there a smaller list of events whose probabilities will determine the entire function?*

  Solution: There are $2^8 = 256$ *subsets of an 8-element set. However, by the additivity of* $\mathbb{P}$, *we only need to know the probability of each event containing exactly one outcome, since every event can be written as the union of these.*

- *Assuming that the coin is fair (that is, each flip is equally likely to land heads or tails), find* $\mathbb{P}$.

  Solution: $\mathbb{P}(\{HHH\}) = \mathbb{P}(\{HHT\}) = \mathbb{P}(\{HTH\}) = \mathbb{P}(\{THH\}) = \mathbb{P}(\{HTT\}) = \mathbb{P}(\{THT\}) = \mathbb{P}(\{TTH\}) = \mathbb{P}(\{TTT\}) = 1/8$.

2 Random Variables

When running chance-based experiments, we'd often like to measure some quality of the outcomes. For instance, when flipping a fair coin three times, we might want to count the number of heads. Previously, we sorted the outcomes into events based on the number of heads. The act of assigning a number to each outcome is exactly the definition of a function, so we have a sometimes confusingly-named definition:

Definition 2 *Given a finite probability space* $(\Omega, \mathbb{P})$*, a (discrete)* ***random variable*** X *on it is a function* $X : \Omega \to \mathbb{R}$.

Despite the name, a random variable is a function (which can appear to be a variable in a real-life experiment, hence the name). Though all of our examples give real number values to the outcomes, it is possible to change the definition to give other kinds of numbers, such as complex numbers. As with events, random variables can be described in words.

Just like any function, another way to describe a random variable is by writing down its values on each outcome. Finally, for each real number x, there is a (possibly empty) set of outcomes where $X = x$. Therefore $X = x$ is an event, so we can measure its probability.

Problem 4 *Consider the experiment of flipping a fair coin three times, and let* X *be the random variable measuring the number of heads. Find the event* $X = 1$*, and its probability.*

Solution: The event $X = 1$ is $\{HTT, THT, TTH\}$, with probability 3/8.

Definition 3 *Given a probability space* $(\Omega, \mathbb{P})$ *and a random variable* X *on it, the* ***probability mass function*** *(PMF for short)* $p_X(x)$ *of* X *is the function*

$$p_X(x) := \mathbb{P}(X = x)$$

Since there are a finite set of outcomes, there can only be a finite set of possible values of X, so there is a finite set of real numbers x where $p_X(x)$ is nonzero.

Problem 5 *Consider the experiment of flipping a fair coin three times, and let* X *be the random variable measuring the number of heads. Find the PMF of* X.

Solution:

$$p_X(x) = \begin{cases} \frac{1}{8} & x = 0 \\ \frac{3}{8} & x = 1 \\ \frac{3}{8} & x = 2 \\ \frac{1}{8} & x = 3 \end{cases}$$

Problem 6 *Consider the experiment of rolling a fair six-sided die once, and let X be the random variable measuring the number on its face. Find the PMF of X.*

Solution:

$$p_X(x) = \begin{cases} \frac{1}{6} & x = 1 \\ \frac{1}{6} & x = 2 \\ \frac{1}{6} & x = 3 \\ \frac{1}{6} & x = 4 \\ \frac{1}{6} & x = 5 \\ \frac{1}{6} & x = 6 \end{cases}$$

Problem 7 *Consider the experiment of rolling a fair six-sided die twice, and let X be the random variable measuring the sum of the numbers on its faces. Find the PMF of X.*

Solution:

$$p_X(x) = \begin{cases} \frac{1}{36} & x = 2 \\ \frac{1}{18} & x = 3 \\ \frac{1}{12} & x = 4 \\ \frac{1}{9} & x = 5 \\ \frac{1}{6} & x = 6 \\ \frac{7}{36} & x = 7 \\ \frac{1}{6} & x = 8 \\ \frac{1}{9} & x = 9 \\ \frac{1}{12} & x = 10 \\ \frac{1}{18} & x = 11 \\ \frac{1}{36} & x = 12 \end{cases}$$

Definition 4 *Given a probability space* $(\Omega, \mathbb{P})$ *and a random variable* X *on it, the* ***expected value*** *of* X *is the number* $\mathbb{E}[X]$ *given by*

$$\mathbb{E}[x] := \sum_{\textit{possible values of } X} x \cdot p_X(x) = \sum_{\textit{possible values of } X} x \cdot \mathbb{P}(X = x)$$

The expected value of a random variable is a "weighted average" of all of its possible values. So while it's impossible to expect any sort of result from a random experiment, we can determine the result we should expect on average.

Problem 8 *Find the expected value of each random variable from Problems 5, 6, and 7.*

Solution: $1.5, 3.5$, and 7.

Problem 9 *Let's see what it means for a result to be what we expect on average. Ask your instructor for a six-sided die. Roll it ten times and write down your results below. Average all of them. Then roll it twenty times (or more, if you wish), and average all of those results.*

Problem 10 *Based on all the dice you've rolled, hypothesize the value of* $\mathbb{E}[X_n]$*, where* X_n *is the random variable that is the sum of the faces of the dice when you roll it* n *times.*

Let's prove (or disprove) your conjecture from the last problem.

Problem 11 *Show that*

$$\mathbb{E}[X] = \sum_{o \in \Omega} X(o)\mathbb{P}(\{o\})$$

(That is, instead of adding up the values of X, we can add up the values on the outcomes separately.)

Solution: Intuitively, we can gather the outcomes into the events $X = x$, and the sum doesn't change if we add in different groups. Mathematically, we can write

$$\begin{aligned}\sum_{o \in \Omega} X(o)\mathbb{P}(\{o\}) &= \sum_{\text{possible values of } X} \sum_{o \in \{X=x\}} X(o)\mathbb{P}(\{o\}) \\ &= \sum_{\text{possible values of } X} x \sum_{o \in \{X=x\}} \mathbb{P}(\{o\}) = \sum_{\text{possible values of } X} x \cdot p_X(x) = \mathbb{E}[x]\end{aligned}$$

Problem 12 *(Linearity of Expectation) Show that for any two random variables X_1 and X_2 on the same probability space*

$$\mathbb{E}[X_1 + X_2] = \mathbb{E}[X_1] + \mathbb{E}[X_2]$$

Use this fact to solve Problem 10*.*

Solution: By Problem 11

$$\mathbb{E}[X_1 + X_2] = \sum_{o \in \Omega} (X_1(o) + X_2(o))\mathbb{P}(\{o\}) = \mathbb{E}[X_1] + \mathbb{E}[X_2]$$

When rolling n dice, we can let X_i be the value on the face of the i^{th} dice, so that the sum of the n dice is $\sum X_i$. Since each X_i is one roll, each $\mathbb{E}[X_i] = 3.5$, so $\mathbb{E}[\sum X_i] = \sum \mathbb{E}[X_i] = 3.5n$, as (probably) expected.

3 Variance and More Experiments

Different random variables can have the same expected value, which can be achieved in different ways. To illustrate this, we'll do an experiment involving rolling dice.

Problem 13 *Let X be the random variable which is the sum of three rolls of a fair six-sided die. Let Y be the random variable which is the value of one roll of a fair twenty-sided die. Verify that both $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ are* 10.5

Solution: Three six-sided dice is given by linearity as in Problem 10. For the twenty-sided die, each integer from 1 to 20 has an equal 1/20 chance of being rolled, so that

$$\mathbb{E}[Y] = \sum_{i=1}^{20} i \cdot \frac{1}{20} = \frac{20 \times 21}{2 \times 20} = 10.5$$

Problem 14 *Ask your instructor for a six-sided die and a twenty-sided die. Roll the six-sided die 120 times (in 40 groups of 3, each which you will add together as a group) and the twenty-sided die 40 times, and record every result on this page and the next. What can you say about the distribution of the results?*

(This page left blank for your experimentation.)

As we see in the previous example, even though two random variables may have the same expected value, one of them may take values closer to that expected value more often than the other. We can measure this with the random variable $X - \mathbb{E}[X]$, but since this variable can be positive or negative (and we just want to measure how far away it is), we should take its absolute value. For calculus reasons (not relevant to us right now), it is better to use a continuous function, so we'll take $(X - \mathbb{E}[X])^2$, which accomplishes the same thing.

Problem 15 *Find the variance of each random variable from Problem 5, 6, and 7.*

Solution: $3/4, 35/12$, and $35/6$.

Problem 16 *Show that also*

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

(This is the most commonly written form of the formula.)

Solution: By linearity

$$Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2 - 2X\mathbb{E}[X] + \mathbb{E}[X]^2] = \mathbb{E}[X^2] - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Problem 17 *Find the variance of the sum of three six-sided dice rolls. Then find the variance of a single twenty-sided die roll. (Hint: This will require finding or at least describing the PMF of each random variable.)*

Solution: 8.75 and 33.25

Problem 18 *Check that the random variable Z that measures the number of heads in 21 flips of a fair coin also has expected value 10.5, and calculate its variance. If you repeated the experiment from Problem 12 with this random variable, how would you expect the distribution of the results to look? (Note: Due to time constraints, it will be impossible to run this experiment at the Math Circle. But feel free to try it at home if you're really bored.)*

Solution: The expected value follows by linearity. For the variance, we should find the PMF. For any integer $0 \leq i \leq 21$, the probability of i flips being heads is $\binom{21}{i}/2^{21}$, since there are $\binom{21}{i}$ ways to have a string of length 21 with i heads, and each string (hence outcome) has probability $1/2^{21}$. Calculating the variance from this information gives 5.25, which is smaller than both 8.75 and 33.25, so we should expect that the coins will tend to give results closer to 10.5 than any of the dice.

Problem 19 *(Bonus) We take a look at a different application of linearity of expectation, to a problem where there's no random variables to begin with.* ***Buffon's Needle Problem*** *is as follows. Suppose we have a needle of length 1, and an infinite sequence of parallel lines covering the plane that are each distance 1 from their closest neighbors. If the needle is randomly dropped onto the plane, what is the probability that it intersects one of the lines?*

- *Let X be the number of times the needle intersects a line. Describe the PMF of X.*
- *Let X_1 be the number of times the front half of the needle intersects a line, and X_2 the number of times the back half of the needle intersects a line. Does the expected number of total intersections change if the needle is broken at the midpoint?*
- *The difficulty in solving the original problem is the dependence on a random angle where the needle lands. What shape doesn't care which angle it lands at? Can you break the needle to make (or approximate) that shape?*
- *Find the expected number of times that a circle of radius $r < 1/2$ (let's say) intersects one of these lines.*
- *Solve the original problem.*
- *From your solution, can you devise a way to experimentally calculate the value of π? If you have extra time, you can even try performing this experiment, but if not, we suggest that you try this at home as well (though maybe not with a needle).*