Probability Part I - Experiments with Random Variables

Yan Tao, Edited by Terry Wang

Advanced 1 - Feb 2025

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.*

[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. Significantly more technical ordinance is needed to describe this theory rigorously as well. 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* $\mathbb{P}(\{o \in \Omega \mid o \textit{ has at least two heads}\})$.

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

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

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?*

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

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.*

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.

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.*

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.*

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*

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

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.*

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 (or, alternatively, use a computer program from the internet to roll dice). 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.)

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.

Problem 13 *(Other Properties of Expectation) Show that for any random variable X on the same probability space and a constant real number c*

$$\mathbb{E}[X] = c\mathbb{E}[X_2]$$
$$\mathbb{E}[c] = c$$

3 Variance and Additional Problems

Definition 5 *The **variance*** $Var(X)$ *for a random variable* X *is given by*

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

Problem 14 *Find the variance of each random variable from Problem 5, 6, and 7.*

Problem 15 *Show that also*

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

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

Problem 16 *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.)*

Problem 17 *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.)*

Problem 18 *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*, as well as its expected value.*
- *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).*

Problem 19 *(Monty Hall Problem) Suppose you are on a game show, and there are three doors, behind two are goats and behind the other one is a car. Your goal is to choose the one with the car behind it. Suppose you pick a door, and the game show host opens one of the other doors to reveal a goat. He then offers you the opportunity to switch your choice to the remaining other door (that you did not choose and that he did not reveal). Do you or do you not change your choice? What are your initial chances for winning versus if you switch?*

Problem 20 *Suppose there are two boxes in front of you, each with a random number between 1 and 1000 of dollar coins* [2] *in them. Your goal is to end up with the envelope with more coins.*

Suppose you choose an initial box, call the number of coins in this box A. Suppose you also have a 6 sided die and you throw it 10 times and add up the sum of the throws, call this number X. You now have the opportunity to look inside your chosen box and you will be forced to switch boxes if $A \leq X$, otherwise you stick with your initial choice. Do you increase your chances of winning by taking this opportunity and why or why not?

Problem 21 *(Drunk Boarding Problem) There are 100 passengers waiting to board an aircraft, numbered* 1 *through* 100*. Let's say passenger n has assigned seating n and passengers board in order according to their number (*1 *goes first, then* 2*, and so on). The issue is that passenger* 1 *is sick and cannot find his seat and thus takes a random seat. Each subsequent passenger either takes their assigned seating if it is unoccupied, or a random unoccupied seat if someone is already sitting in their assigned seat. You are passenger* 100*, what are the chances you get your seat?*

[2] My friend just told me that these are apparently these are rare in the US? We have them in Canadian currency.

Problem 22 *Suppose you have events $A, B, C \subseteq \Omega$, prove the following identities from the basic definitions:*

$$\mathbb{P}(A \cap B) \geq \mathbb{P}(A) + \mathbb{P}(B) - 1$$

$$\mathbb{P}(A \cup B \cup C) \leq \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C)$$

Problem 23 *(Markov's inequality) Let X be a random variable, prove that if X is always non-negative, then for any positive number a:*

$$\mathbb{P}(X \geq a) \leq \frac{1}{a}\mathbb{E}[X]$$

Problem 24 *(Lyapunov inequality and Jensen's inequality) Let $\lambda_1, \ldots, \lambda_n$ be real numbers in the unit interval $(0, 1)$ such that they add up to 1. Let $x_1, \ldots, x_n$ be real numbers as well, prove that:*

$$\sum_{k=1}^{n} \lambda_k x_k^2 \geq \left(\sum_{k=1}^{n} \lambda_k x_k \right)^2$$

Now, using the above result, let X be a random variable, prove that if X is always non-negative, then:

$$\mathbb{E}[X^2]^{\frac{1}{2}} \geq \mathbb{E}[X]$$