TOPIC 4
PROBABILITY AND STATISTICS

SUBTOPIC 4.7
DISCRETE RANDOM VARIABLES

Overview
(/study/app
122-
cid-
754029/

Teacher view

Index

4. Probability and statistics / 4.7 Discrete random variables

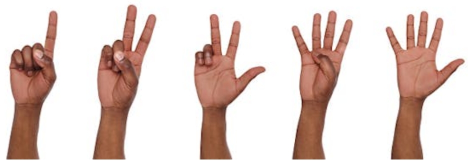
The big picture

We have already learned about two key concepts in probability. First, the set of all the possible outcomes of an experiment is called the sample space. Second, that if you add up the probabilities of each of the outcomes in the sample space, the total will always equal 1.

These two concepts are the basis of what we call a probability distribution, the set of all possible outcomes of an experiment paired with the probabilities of those outcomes. Probability distributions can be either discrete or continuous, depending on the independent variable involved. A discrete random variable is one that has distinct values, which are usually the results of counting experiments. An example that we will explore thoroughly in subtopic 4.8 (/study/app/m/sid-122-cid-754029/book/the-big-picture-id-26260/) is the binomial distribution, seen where repeated independent events have only two possible outcomes, and the discrete random variable is the number of times one of those outcomes occurs out of the total number of trials. Two other discrete probability distributions are the geometric distribution and the Poisson distribution. On the other hand, a continuous random variable is the outcome of experiments that involve measurements. A rather common example that we will study in depth in subtopic 4.9 (/study/app/m/sid-122-cid-754029/book/the-big-picture-id-26265/) is the normal distribution, but other continuous distributions include the exponential distribution, the gamma distribution and the chi-square distribution. Since the independent variable is continuous, the probability of getting any exact value is 0, but we can use integration (with technology) to find the probability of falling within a range of values.

Discrete and continuous random variables are associated with the outcomes of different kinds of experiments.

Student view



A discrete random variable is often the outcome of a counting experiment.

Credit: Milkos GettyImages

A continuous random variable is often the outcome of a measurement.

Credit: Tetiana Rostopira GettyImages

Before exploring specific distributions, we will examine the basic concept of what a probability distribution is and how it can be used to calculate the expected value of a discrete random variable.

Theory of Knowledge

Challenge the givens: what happens if one side of a coin has more weight than the other due to the minted reliefs? What happens if the dice rolled has more than six sides?

City design (i.e. engineering) and conceptual design (i.e. architecture and interior design) are some real-life situations that highlight TOK-type thinking to challenge the status quo. For example, see the following TED Talk by Kent Larson, which shows how design variables may not be discrete when city planners change the buildings themselves.

Overview
(/study/ap
122-
cid-
754029/

Kent Larson: Brilliant designs to fit more people in every city

Concept

The study of probability distributions is really a study of a **system** within mathematics. As we have already stated, there are numerous probability distributions, but in this subtopic you will learn about the common characteristics that these distributions share. While the distributions may have different ways of finding certain measures or may apply to different types of data, their similarities make it possible to learn the general system first so that you have a framework of knowledge to build on as you explore specific distributions in the future.

4. Probability and statistics / 4.7 Discrete random variables

Probability distributions

Probability distributions and discrete random variables

Consider the simple experiment of tossing a coin three times. The number of heads can be designated as the random variable, indicated by X. Random variables are usually capitalised, whereas the values they take are indicated by the corresponding lower case,

Student
view

e.g. X and x . The values of random variables (here, the number of heads) are determined by chance.

If a random variable can take exactly N values, each of which corresponds to a unique outcome in the sample space of an experiment, then this variable is a discrete random variable. Thus, the following are all examples of discrete random variables:

- number of heads obtained in coin tosses
- number of seeds germinating
- number of calls received per minute
- number of boxes containing fewer than 100 matches.

For our coin-tossing experiment, if H and T are the events of obtaining a head or a tail, respectively, then the sample space is $\{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$.

The possible values of X are $0, 1, 2, 3$, which we write as $X = x$, $x \in \{0, 1, 2, 3\}$.

The probability that X takes on any of these values, written as $\mathrm{P(X = x)}$, follows from the probability rules underlying the experiment of tossing a coin, for example, $\mathrm{P(X = 2) = P(HHT) + P(HTH) + P(THH)} = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$.

A probability distribution is a combination of the sample space of a random experiment with the probabilities of each of the events in the sample space. For X, the (discrete) random variable of the number of heads obtained in three tosses, we obtain the probability distribution function (PDF) shown in the table below.

The probability distribution for the (discrete) random variable X, which is the number of heads obtained when tossing a fair coin three times.

x	0	1	2	3
$\mathrm{P}(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Thus, the probability that 1 head is obtained is written as $\mathrm{P}(X = 1) = \frac{3}{8}$.

There are two key properties of every probability distribution:

✓ **Important**

For each value, x_i , of the random variable X , we have that $0 \leqslant \mathrm{P}(\mathrm{X}=\mathrm{x_i}) \leq 1$.

As X may take on any of the N values of the sample space, the sum of the probabilities of each of them as an outcome of the experiment must equal 1:

$$\sum_{i=1}^{N} \mathrm{P}(\mathrm{X}=\mathrm{x_i}) = \mathrm{P}(\mathrm{X}=\mathrm{x_1}) + \mathrm{P}(\mathrm{X}=\mathrm{x_2}) + \ldots + \mathrm{P}(\mathrm{X}=\mathrm{x_N}) = 1$$

In other words, all of the individual probabilities must be some value between 0 and 1 inclusive, and the sum of all the probabilities must equal 1. When a probability distribution possesses both of these properties, we call it a **well-defined** probability distribution.

Example 1

★☆☆

Data are collected from a biathlon competition to find how many of the 5 targets at each station the competitors missed. The organisers know that none of the competitors missed all of the targets, but have lost the data for how many times the competitors missed 3. The data collected enables us to find the experimental probabilities of each number of misses and give the probability distribution in the table below. Find the value of k .

The probability distribution for X with unknown k.

x	0	1	2	3	4
$\mathrm{P}(X=x)$	0.25	0.15	0.15	k	0.25

Overview
(/study/ap
122-
cid-
754029/

A biathlete approaching the shooting range.

Credit: technotr Getty Images

Since we know
$\sum_{i=1}^{N} \mathrm{P}(\mathrm{X} = \mathrm{x_i}) = \mathrm{P}(\mathrm{X} = \mathrm{x_1}) + \mathrm{P}(\mathrm{X} = \mathrm{x_2}) + \ldots + \mathrm{P}(\mathrm{X} = \mathrm{x_N}) = 1$, we can write the following equation and solve it for k:

$$\begin{aligned} 0.25 + 0.15 + 0.15 + k + 0.25 &= 1 \\ 0.8 + k &= 1 \\ k &= 0.2 \end{aligned}$$

Example 2

★★☆

The probability distribution function for a random variable X is given by $\mathrm{P}(\mathrm{X} = \mathrm{i}) = \frac{\mathrm{k}}{2\mathrm{i}}$, for $i = 1, 2, 3, 4, 5, 6.$ Find k.

Since we know
$\sum_{i=1}^{N} \mathrm{P}(\mathrm{X} = \mathrm{x_i}) = \mathrm{P}(\mathrm{X} = \mathrm{x_1}) + \mathrm{P}(\mathrm{X} = \mathrm{x_2}) + \ldots + \mathrm{P}(\mathrm{X} = \mathrm{x_N}) = 1$, we can write the following equation and solve it for k:

Student
view

Overview
(/study/app/m/sid-122-cid-754029/)

$$1 = \frac{k}{2 \times 1} + \frac{k}{2 \times 2} + \frac{k}{2 \times 3} + \frac{k}{2 \times 4} + \frac{k}{2 \times 5} + \frac{k}{2 \times 6}$$
$$1 = \frac{147k}{120}$$
$$k = \frac{120}{147}$$
$$k = \frac{40}{49}$$

The following example involves a discrete distribution on an infinite sample space. If you read this section as part of the Applications and interpretation SL book, you can skip this example. Infinite discrete random variables are not part of the Applications and interpretation SL syllabus.

Example 3

★★★

The random variable X has a probability distribution function given by $\mathrm{P(X = i) = k}\left(\frac{1}{3}\right)^{\mathrm{i}}$, where $i = 0, 1, 2, 3, \ldots$

Find k and $\mathrm{P(X \geqslant 2)}$.

If we start writing down the probabilities to add them up to 1, we see a unique challenge because of the fact that i goes to infinity.

$$1 = \sum_{i=0}^{\infty} k\left(\frac{1}{3}\right)^{i}$$
$$1 = k\left[1 + \frac{1}{3} + \left(\frac{1}{3}\right)^{2} + \left(\frac{1}{3}\right)^{3} + \left(\frac{1}{3}\right)^{4} + \ldots\right]$$

A closer look inside the square brackets reveals that we have an infinite geometric series with an initial value $a_1 = 1$ and a common ratio $r = \frac{1}{3}$.

Since $|r| < 1$, the series converges and we can use the formula for the sum of an infinite geometric series:

Student view

$$S_\infty = \frac{a_1}{1-r} = \frac{1}{1-\frac{1}{3}} = \frac{3}{2}.$$

Overview (/study/ap 122- cid- 754029/

Hence, we can substitute this into our original sum to get the following:

$$\begin{aligned} 1 &= k\left[1+\frac{1}{3}+\left(\frac{1}{3}\right)^2+\left(\frac{1}{3}\right)^3+\left(\frac{1}{3}\right)^4+\ldots\right] \\ 1 &= k\times\frac{3}{2} \\ k &= \frac{2}{3} \end{aligned}$$

Now that you know the value of k, you can find $\mathrm{P}(\mathrm{X} \geqslant 2)$ using the complement principle.

$$\begin{aligned} \mathrm{P}(\mathrm{X} \geqslant 2) &= 1-\mathrm{P}(\mathrm{X}=0)-\mathrm{P}(\mathrm{X}=1) \\ &= 1-\frac{2}{3}\times 1-\frac{2}{3}\times\frac{1}{3} \\ &= 1-\frac{2}{3}-\frac{2}{9} \\ &= \frac{1}{9} \end{aligned}$$

Be aware

It is easy for simple arithmetic or algebra errors to give you impossible results, such as the probability of an event being less than zero or greater than one. Watch for these unreasonable answers and search out your mistake.

Exam tip

When completing problems with multiple parts on the exam, you can still earn full marks on one part if you make a mistake on the part before, as long as you use your incorrect answer properly. One exception to this is if your incorrect answer is something impossible, such as the negative probability described above.

Student view

3 section questions

Overview (/study/app/m/sid-122-cid-754029/

4. Probability and statistics / 4.7 Discrete random variables

Expected value of a discrete random variable

Calculating expected value

In section 4.5.2 (/study/app/m/sid-122-cid-754029/book/expected-outcomes-id-26247/) we used probabilities to determine the expected number of a particular outcome. The expected value, or expected mean, is the value you would expect to obtain on average if you performed an experiment many times.

Important

The expected value of a random variable X with a probability distribution function $\mathrm{P(X = x)}$ is written as $\mathrm{E(X)}$, or μ, and is given by the formula

$$\begin{aligned}\mathrm{E(X)} = \mu &= \sum_{i=1}^{N} x_i \mathrm{P(X = x_i)} \\ &= x_1\mathrm{P(X = x_1)} + \mathrm{x_2 P(X = x_2)} + \mathrm{x_3 P(X = x_3)} + \ldots + \mathrm{x_N P(X = x_N)}\end{aligned}$$

Making connections

The formula for $\mathrm{E(X)}$ may seem familiar, because it is very similar to the mean of a frequency distribution that you learned how to find in section 4.2.1 (/study/app/m/sid-122-cid-754029/book/grouped-data-and-quartiles-id-26228/). In this case, we have a proportion of the whole (the probability) instead of a frequency, but the mechanics of the two formulae are quite similar.

Example 1

★★☆

Student view

Overview (/study/app/m/sid-122-cid-754029/

Consider the probability distribution from the biathlon in section 4.7.1 (/study/app/m/sid-122-cid-754029/book/probability-distributions-id-26256/) Example 1, where X represented the number of targets biathletes miss at a station. Now that we have solved for k, the probability distribution looks like this:

x	0	1	2	3	4
$\mathrm{P}(X = x)$	0.25	0.15	0.15	0.2	0.25

Find the expected mean number of targets a biathlete will miss at a station.

To calculate the expected mean, the formula tells us to multiply each value of with its corresponding probability.

$$\begin{aligned}
\mathrm{E(X)} = \mu &= \sum_{i=1}^{N} x_i \mathrm{P(X = x_i)} \\
&= 0 \times \mathrm{P}(0) + 1 \times \mathrm{P}(1) + 2 \times \mathrm{P}(2) + 3 \times \mathrm{P}(3) + 4 \times \mathrm{P}(4) \\
&= 0 \times 0.25 + 1 \times 0.15 + 2 \times 0.15 + 3 \times 0.2 + 4 \times 0.25 \\
&= 0 + 0.15 + 0.3 + 0.6 + 1 \\
&= 2.05
\end{aligned}$$

Therefore, we can expect that based on our distribution a shooter in a biathlon will miss just over two targets at each station.

Be aware

Note that in discrete distributions it is common for the expected value **not** to be one of the actual outcomes of the experiment. How do you think you could interpret the mean if it is impossible for the variable to take on that value?

Activity

Student view

Overview (/study/app/m/sid-122-cid-754029/

As we have stated, expected value is another method for finding the mean of a situation. If we transform the outcomes of an experiment in the same way we changed data in section 4.3.3 (/study/app/m/sid-122-cid-754029/book/data-transformation-id-26236/), how do you think it will affect $\mathrm{E}(\mathrm{X})$? Use the data from Example 4 and transform it several ways.

Expected value and fairness

Game designers have to be very careful to create games that are balanced between being challenging and being beatable. In games between two parties where the probability of winning is not equal, a game designer can adjust the pay-outs in order to provide balance that makes it more even. A fair game is perfectly balanced so that the expected mean pay-out is 0 for both players.

Example 2

★★☆

A game for two players uses a 20-sided dice. If Player A rolls a prime number, they win 5 points. Otherwise, Player B wins k points. If the random variable X represents how many points Player A is winning by, find the value of k that makes this a fair game.

First, construct a table to examine the probability of each outcome, noting that there are 8 prime numbers on the 20-sided dice: $\{2, 3, 5, 7, 11, 13, 17, 19\}$.

	Player A wins (a prime number is rolled)	**Player B wins (a different number is rolled)**
Pay-out for Player A	$+5$	$-k$
Probability	$\frac{8}{20} = \frac{2}{5}$	$\frac{3}{5}$

The expected mean outcome for the game is then

$$2 - \frac{3}{5}k$$

Student view

For the game to be fair, $\mathrm{E}(\mathrm{X}) = 0$.

Overview
(/study/ap
122-
cid-
754029/

Thus, $2-\frac{3}{5}k=0$ and $k=\frac{10}{3}$.

The applet below lets you explore the concept of expected value using a dice. If you think of the value on the dice as the pay-out, each value would have an equal probability, so the expected value would be 3.5. This is theoretical, but if you were to change the probability of rolling one value or another, it would change the expectation. The applet allows you to simulate rolling a dice with different probabilities. What happens to the experimental mean as the number of rolls increases?

Interactive 1. Adjust the Probabilities for a Dice to See Changes in the Expected Value.

Credit: GeoGebra (https://www.geogebra.org/m/JHg7VJUk) Juan Carlos Ponce Campuzano

More information for interactive 1

This interactive allows users to explore the concept of expected value (EV) using a simulated dice roll. The simulation includes controls such as "Drag," "Roll die," "Stop," "Speed-up," and "New Sample," enabling users to manipulate the experiment dynamically. The x-axis represents the number of rolls, while the y-axis displays the average value of the outcomes. Users can adjust the theoretical probability of each die outcome to observe its impact on the average and expected value. For example, if the theoretical probability of rolling a "6" is increased from $\frac{1}{6}$ to $\frac{1}{3}$, the expected value will rise, and the experimental mean will converge to this new EV as the number of rolls increases.

The simulation visually demonstrates how the experimental mean approaches the expected value over time, reinforcing the law of large numbers. Users can drag black points on the graph to examine specific values and

Section Student... (0/0) Feedback Print (/study/app/m/sid-122-cid-754029/book/expected-value-of-a-discrete-random-variable-id-26257/print/) Assign

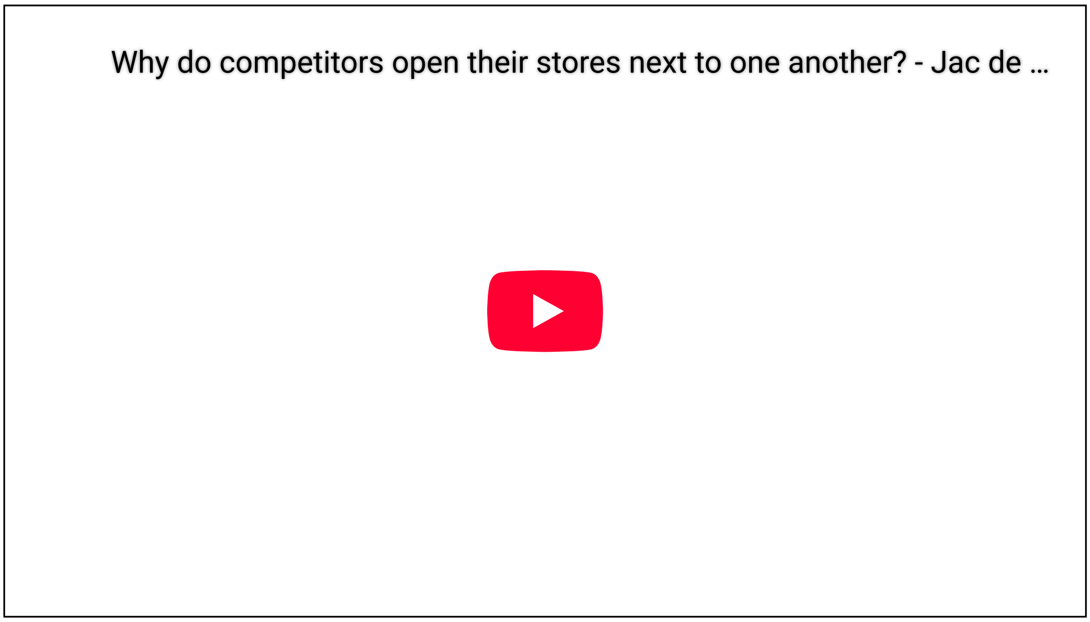
Student view

Overview
(/study/ap
122-
cid-
754029/

outcomes, providing a hands-on understanding of probability distributions. Additionally, changing the theoretical probability for one outcome (e.g., setting $P(1) = 0.5$ while keeping others equal) shifts the EV, illustrating how biased probabilities affect results. The expected value is calculated as the sum of each outcome multiplied by its probability (for instance, for a fair die: $EV = 1 \cdot \left(\frac{1}{6}\right) + 2 \cdot \left(\frac{1}{6}\right) + \ldots + 6 \cdot \left(\frac{1}{6}\right) = 3.5$). This interactive tool effectively bridges theoretical probability with empirical observations, making it ideal for learning statistical concepts.

International Mindedness

The concept of fairness is one of the concepts explored in the field of game theory. Game theory is used in actual games, of course, but it is also the mathematics behind all sorts of decision-making and negotiations. For example, government officials use game theory when negotiating trade agreements between countries. How do you think international trade agreements get made to ensure they are mutually beneficial?



Theory of Knowledge

Mathematics can be used to make predictions based on probability and expected value. As a result, probability calculations are often used to dictate public policy.

Psychologists Amos Tversky and Daniel Kahnemann investigated the role of language in human interpretation of probability. As outlined in the video below, they found that **language** had a large impact on how we process probabilities. They coined this 'framing effects' and found that our reason can be hijacked by **language.** Watch the video and consider the knowledge question 'To what extent does language affect mathematical knowledge?'

Section Student... (0/0) Feedback Print (/study/app/m/sid-122-cid-754029/book/probability-distributions-id-26256/print/) Assign

Student view

Overview
(/study/ap
122-
cid-
754029/

The Framing Effect (Intro Psych Tutorial #94)

2 section questions

4. Probability and statistics / 4.7 Discrete random variables

Checklist

Section Student... (0/0) Feedback Print (/study/app/m/sid-122-cid-754029/book/checklist-id-26258/print/) Assign

What you should know

By the end of this subtopic you should be able to:

- recall and explain the two key characteristics of a probability distribution
- use a probability distribution function to calculate probabilities of various outcomes
- calculate the expected value of a probability distribution of a discrete random variable.

Student view

4. Probability and statistics / 4.7 Discrete random variables

Overview
(/study/ap
122-
cid-
754029/

Investigation

Section Student... (0/0) Feedback Print (/study/app/m/sid-122-cid-754029/book/investigation-id-26259/print/) Assign

Create a game of your own using either two dice or a deck of cards. Make a set of rules that determine whether you win or lose the game and determine what the probabilities of each outcome would be. Assign a pay-out to each outcome and calculate the expected value of the game. Is it fair? If not, adjust the pay-out to make it fair.

Once you have created a basic game with a win or lose outcome, try creating a new game with three or more outcomes. Assign **different** pay-outs to each outcome. For example, in a game with four outcomes, one could result in winning a lot, one in winning a little, one would have a pay-out of 0 and the last one would make you lose a lot.

Play the game several times and see if the mean pay-out you receive is close to the 0 as you planned it. Even though you designed it to be a fair game, is there any strategy you can think of to make it end in your favour?

Student
view