

Section 4.1, Probability Distributions

A **random variable** x represents a numerical value associated with each outcome of a probability experiment. (Note: Many other books make a distinction between an unknown random variable X and the value that it takes x .)

There are two kind of random variables: discrete and continuous. The kind of random variable determines how we will analyze it.

A random variable is **discrete** if it has a finite number of possible outcomes, or a countable number (i.e. the integers are infinite, but are able to be counted). For example, the number of heads you get when flip a coin 100 times is discrete, since it can only be a whole number between 0 and 100.

A random variable is **continuous** if it has an uncountable number of possible outcomes, such as an interval on a number line (the interval does not need to be bounded). For example, the amount of time you spend studying for this class is a continuous random variable, since it can be any number that is at least zero hours, and you can spend any percentage of an hour or minute studying.

Examples

Determine whether each of the following variables x is discrete or continuous:

1. Let x represent the number of times you do laundry this month. **Discrete**
2. Let x represent your annual salary given to the nearest cent. **Discrete**
3. Let x represent a car's speed as it drives past the 3300 South exit on northbound I-15. **Continuous**
4. Let x represent your height at age 10. **Continuous**
5. Let x represent number of math classes that you have taken in your life. **Discrete**

We'll be spending the rest of this chapter analyzing discrete random variables. In Chapter 5 we will be working with a very important class of continuous random variables called Normal random variables.

1 Discrete Probability Distributions

A **discrete probability distribution** lists each possible value that a random variable can take, along with its probability. It has the following properties:

- The probability of each value of the discrete random variable is between 0 and 1, so $0 \leq P(x) \leq 1$.
- The sum of all the probabilities is 1, so $\sum P(x) = 1$.

Examples

Determine if each of the following tables represents a probability distribution:

1.

x	-5	6	9
$P(x)$	0.5	0.25	0.25

Yes, this is a probability distribution, since all of the probabilities are between 0 and 1, and they add to 1.

2.

x	1	2	3	4
$P(x)$	0.4	0.4	0.4	0.2

This is not a probability distribution, since the probabilities add to 1.4, not 1.

3.

x	1	2	3	4
$P(x)$	0.4	0.4	0.4	-0.2

This is not a probability distribution, since $P(4) = -0.2$, and probabilities cannot be negative.

Using a frequency distribution, you can make a probability distribution by using the relative frequencies for the probabilities.

Example

Make a probability distribution from the following frequency distribution represent the number of fish caught in a 6-hour period:

Number of fish caught	0	1	2	3	4
Frequency	88	72	30	8	2

To find the probabilities, we need to divide the frequencies by the total number of observations, 200, so we get:

Number of fish caught	0	1	2	3	4
Probability	0.44	0.36	0.15	0.04	0.01

The **mean**, or **expected value** of a discrete random variable, μ , is given by

$$\mu = E(x) = \sum xP(x).$$

The **variance** of a discrete random variable, σ^2 is given by

$$\sigma^2 = \sum (x - \mu)^2 P(x),$$

and its **standard deviation**, σ is given

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum (x - \mu)^2 P(x)}.$$

Examples

Calculate the expected value, variance, and standard deviation for each of the following probability distributions:

1.

x	-5	6	9
$P(x)$	0.5	0.25	0.25

First, we need to calculate the expected value (as it is used in the rest of the calculations), so

$$\mu = -5 \cdot 0.5 + 6 \cdot 0.25 + 9 \cdot 0.25 = 1.25.$$

Then the variance is

$$\sigma^2 = (-5 - 1.25)^2 \cdot 0.5 + (6 - 1.25)^2 \cdot 0.25 + (9 - 1.25)^2 \cdot 0.25 = 40.1875,$$

so the standard deviation is $\sigma = \sqrt{40.1875} = 6.33936$.

2.

Number of fish caught	0	1	2	3	4
Probability	0.44	0.36	0.15	0.04	0.01

The expected value is

$$\mu = 0 \cdot 0.44 + 1 \cdot 0.36 + 2 \cdot 0.15 + 3 \cdot 0.04 + 4 \cdot 0.01 = 0.82.$$

The variance is

$$\begin{aligned} \sigma^2 &= (0 - 0.82)^2 \cdot 0.44 + (1 - 0.82)^2 \cdot 0.36 + (2 - 0.82)^2 \cdot 0.15 + (3 - 0.82)^2 \cdot 0.04 \\ &\quad + (4 - 0.82)^2 \cdot 0.01 \\ &= 0.8076 \end{aligned}$$

and the standard deviation is $\sigma = \sqrt{0.8076} = 0.8987$.