Section 3.4: (Discrete) Random Variables

1 Two Examples

Definition 1 *A **random variable** is a function that assigns a numerical value to each outcome in a sample space.*

There are two types of random variables, discrete and continuous, based on the data types assigned to values. The above two examples illustrate the different ways a random variable must be defined. Accordingly, different techniques will be needed to work with the two different types of random variables.

Example 2 *Consider the experiment of rolling a pair of dice and summing the faces. The random variable X assigns to each roll its sum. The following table indicates the probabilities for each value in X.*

Sum	2	3	4	5	6	7	8	9	10	11	12

Remark 3 *Since the data is discrete, we call this function a **discrete random variable**.*

Example 4 *Let Y be the random variable that assigns to each human being, their height in inches. Given the uncountably infinite number of possible heights, it is not possible to define Y based on a table.*

Remark 5 *Since the data is continuous, we call this function a **continuous random variable**.*

Problem 6 *Consider the experiment of rolling a three sided-die and a four-sided die and summing the faces. Construct a discrete random variable X.*

Definition 7 *A random variable with only two outcomes (0-1, true-false, right-wrong, on-off, etc.) is called a* ***Bernoulli random variable****.*

Example 8 *A multiple choice test with only one correct answer per question is a Bernoulli random variable.*

Problem 9 *Formally describe a public safety officer's parking lot duty as a random variable.*

Definition 10 *A* ***probability density function (pdf) or probability mass function (pmf)*** *for a discrete random variable X is a function whose domain is all possible values of X and assigns to each $x \in X$ the probability that x occurs. Note that the sum of all probabilities in a distribution must be 1.*

Problem 11 *Is the following function a probability distribution? Explain.*

X	1	2	3
$P(X)$	0.5	0.6	-0.1

Example 12 *Consider the following probability distribution.*

X	1	2	3	4	5	6
$P(X)$	0.3	0.2	0.1	0.1		0.2

1. $P(X = 3) =$
2. $P(X = 4) =$
3. $P(X = 5) =$
4. $P(X \leq 3) =$
5. $P(X < 3) =$
6. $P(2 \leq X \leq 4) =$
7. $P(X > 5) =$

Example 13 *Consider the experiment of rolling a pair of dice and summing the faces. The random variable X assigns to each roll its sum. The following table indicates the probabilities for each value in X. Together they form the probability density function.*

Sum	2	3	4	5	6	7	8	9	10	11	12
Prob	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Problem 14 *Consider the experiment of rolling a three sided-die and a 4-sided die and summing the faces. Construct the pdf for X.*

Problem 15 *A \$2 lottery ticket offers four chances to win different amounts of money as indicated by the following probability distribution model.*

Prize	\$0	\$0.5	\$1	\$5	\$50
Prob	$\frac{71}{100}$	$\frac{15}{100}$	$\frac{10}{100}$	$\frac{3}{100}$	$\frac{1}{100}$

1. What is the probability that a ticket wins nothing?

2. What is the probability that a ticket wins money but is a losing proposition overall?

3. Two people buy one ticket each. What is the probability that they both win nothing?

4. Two people buy one ticket each. What is the probability that they both win money?

5. Ten people buy one ticket each. What is the probability that at least one person wins money?

6. Five hundred people buy one ticket each. What is the probability that at least one person wins the \$50 prize?

Problem 16 *You draw a card from a deck. If you get a club you get \$5. If you get an Ace you get \$10. For all other cards you receive nothing. Let X be the random variable of money won. Create a probability density function for this game.*

Problem 17 *You draw a card from a deck. If you get a club you get nothing. If you get red card you get \$10. If you get a spade you get \$15 and get to select another card (without replacement). If the second is another spade you receive an additional \$20. You receive nothing for any non-spade card. Create a probability density function for this game.*

A random variable with a finite number of outcomes is necessarily discrete. However, it is possible for a discrete random variable to have an infinite number of outcomes.

Example 18 *John flips a coin until he observes a tail. Construct the pdf for the number of flips of the coin for this discrete random variable.*

2 Cumulative Distribution Functions

The cumulative distribution function of a discrete pdf is the sum of all probabilities of outcomes less than or equal to some fixed value x.

Definition 19 *The cumulative distribution function F(x) of a pdf p(x) is* $\sum_{y \leq x} p(y)$.

Example 20 *Consider the experiment of rolling a pair of dice and summing the faces. The random variable X assigns to each roll its sum. The following table indicates the probabilities for each value in X. Together they form the probability density function.*

Sum	2	3	4	5	6	7	8	9	10	11	12
$p(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Now construct the cumulative distribution function of this pdf.

Sum	2	3	4	5	6	7	8	9	10	11	12
$F(x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{15}{36}$	$\frac{21}{36}$	$\frac{26}{36}$	$\frac{30}{36}$	$\frac{33}{36}$	$\frac{35}{36}$	$\frac{36}{36}$

Problem 21 *Consider the experiment of rolling a three sided-die and a 4-sided die and summing the faces. Construct $p(x)$ and $F(x)$ for X.*

Example 22 *Consider rolling a fair die until a 5 appears. Construct the pdf for the number of rolls.*
The probability of rolling a 5 (or any specific value) on a fair die is $p = \frac{1}{6}$. *Since there is never a guarantee of ever getting a 5, the set of possible values for x is the set of positive integers. Since Z^+ is infinite, one cannot write down all possible values and probabilities. Instead we need to construct a function. It can help to model a few specific values of x before writing a general form. For example, the first 5 occurs on attempt number four with probability* $\left(\frac{5}{6}\right)^3 * \frac{1}{6} = \frac{125}{1296}$. *One has to roll a number other than 5, three times and then roll a 5. Generalizing this our pdf,* $p(x) = \left(\frac{5}{6}\right)^{x-1} * \frac{1}{6}$.

Expected Value (Mean) and Standard Deviation for a Discrete Random Variable

Recall the experiment of rolling a pair of dice and summing the faces. The random variable X assigns to each roll its sum. The following table indicates the probabilities for each value in X. Together they form the probability density function.

Sum	2	3	4	5	6	7	8	9	10	11	12
$p(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

The pdf provides an easy table or function from which to acquire probabilities of individual roles. What value on average do we expect after rolling the pair of dice? What value on average do we expect after rolling the pair of dice 100 times? How much variation do we expect from that average roll? These sound like questions of a mean and standard deviation. Rather than being a mean and standard deviation of a data set, the pdf introduces the notion of uncertainty and likelihood.

Definition 23 *The* ***expected value*** *(or* ***mean****),* $E(X)$*(or* μ*), of a discrete probability distribution is given by*

$$E(X) = \sum_{x \in X} x * p(x).$$

Definition 24 *The* ***variance,*** $V(X)$, *of a discrete probability distribution is given by*

$$\sigma^2 = \sum_{x \in X} (x - E(X))^2 * p(x).$$

Example 25 *Back to the experiment of rolling a pair of dice and summing the faces. The random variable X assigns to each roll its sum. The following table is our previously constructed probability model.*

Sum	2	3	4	5	6	7	8	9	10	11	12
Prob	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

What is the expected value of the sum of a pair of dice?

$$\begin{aligned} E(X) &= \sum_{x \in X} x * p(x) \\ &= 2 * \frac{1}{36} + 3 * \frac{2}{36} + 4 * \frac{3}{36} + 5 * \frac{4}{36} + 6 * \frac{5}{36} + 7 * \frac{6}{36} \\ &\quad + 8 * \frac{5}{36} + 9 * \frac{4}{36} + 10 * \frac{3}{36} + 11 * \frac{2}{36} + 12 * \frac{1}{36} \\ &= \frac{2 + 6 + 12 + 20 + 30 + 42 + 40 + 36 + 30 + 22 + 12}{36} \\ &= 7. \end{aligned}$$

What is the standard deviation of the sum of a pair of dice?

Sum	2	3	4	5	6	7	8	9	10	11	12
$(x - E(X))^2$	25	16	9	4	1	0	1	4	9	16	25
Prob	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$$
\begin{aligned}
V(X) &= \sigma^2 = \sum_{x \in X} (x - E(X))^2 * p(x) \\
&= 25 * \frac{1}{36} + 16 * \frac{2}{36} + 9 * \frac{3}{36} + 4 * \frac{4}{36} + 1 * \frac{5}{36} + 0 * \frac{6}{36} \\
&\quad +1 * \frac{5}{36} + 4 * \frac{4}{36} + 9 * \frac{3}{36} + 16 * \frac{2}{36} + 25 * \frac{1}{36} \\
&= \frac{35}{6}.
\end{aligned}
$$

Thus, $\sigma = \sqrt{\frac{35}{6}} = 2.4152$.

Example 26 *What is the probability that the sum of two dice falls within one standard deviation of the mean?*
One standard deviation from the mean is $(7 \pm 2.4152) = (4.5848, 9.4152)$*. So we could see a sum of 5, 6, 7, 8, or 9. This occurs with probability*

$$
p = \frac{4}{36} + \frac{5}{36} + \frac{6}{36} + \frac{5}{36} + \frac{4}{36} = \frac{2}{3} = 0.66667.
$$

Exercise 27 *A \$2 lottery ticket offers four chances to win different amounts of money as indicated by the following probability distribution model. Determine the expected value of buying a single ticket. Do you wish to play this game?*

Prize	\$0	\$0.5	\$1	\$5	\$50
Prob	$\frac{71}{100}$	$\frac{15}{100}$	$\frac{10}{100}$	$\frac{3}{100}$	$\frac{1}{100}$
Value	-2	-1.5	-1	3	48

So,

$$
\begin{aligned}
E(X) &= \sum_{x \in X} x * p(x) \\
&= -2 * \frac{71}{100} - 1.5 * \frac{15}{100} - 1 * \frac{10}{100} + 3 * \frac{3}{100} + 48 * \frac{1}{100} \\
&= -1.175.
\end{aligned}
$$

I personally would not want to play this game with a negative expected value and such a small prize.

Exercise 28 *What is the expected value of buying 50 such lottery tickets?*
$-1.175 * 50 = -58.75.$

Problem 29 *You pay \$1 to play a game. The game consists of rolling a pair of dice. If you observe a sum of 7 or 11 you receive \$4. If not, you receive nothing. Compute the expected value for this game?*

Example 30 *Consider flipping an unbalanced coin that lands H 60% of the time. What is the expected value and standard deviation for this Bernoulli random variable?*

Coin	*H*	*T*
Prob	.6	.4

is the pdf. It's expected value is that the coin lands H, 60% of the time. One can view this representing a success with a 1 and a failure as a 0 for the X values.

X	1	0
$p(x)$	.6	.4

$$\begin{aligned} E(X) &= \sum_{x \in X} x * p(x) \\ &= 1 * .6 + 0 * .4 \\ &= .6 \end{aligned}$$

$$\begin{aligned} \sigma^2 &= \sum_{x \in X} (x - E(X))^2 * p(x) \\ &= (1 - .6)^2 * .6 + (0 - .6)^2 * .4 \\ &= .24 \\ \textit{Thus, } \sigma &= \sqrt{.24} \\ &= 0.489\,90 \end{aligned}$$

While not immediately obvious, it should be noted that $\sigma^2 = .6 * .4$.

At times we will want to use a given expectation as an input variable into some other function.

Example 31 *Our electronics store buys 5 game consoles at \$250 each. The pdf for the number of consoles sold in the first week is given below.*

# sold	*0*	*1*	*2*	*3*	*4*	*5*
p(x)	*.1*	*.1*	*.2*	*.3*	*.2*	*.1*

We expect to sell $0.1+1*.1+2*.2+3*.3+4*.2+5*.1 = 2.7$ consoles this week.*
This week the console is hot and can be sold for \$350 each. After this week, we must sell the remaining consoles at \$200 each. What is our expected profit? If x represents the number of consoles sold then our profit function is $f(x) = 100 * x - 50 * (5 - x) = 150x - 250$. Our expected profit is

$$
\begin{aligned}
E(f(x)) &= \sum_{i=0}^{5} f\left(i\right) * p(i) \\
&= f(0) * p(0) + f(1) * p(1) + f(2) * p(2) \\
&\quad + f(3) * p(3) + f(4) * p(4) + f(5) * p(5) \\
&= -250 * .1 - 100 * .1 + 50 * .2 + 200 * .3 + 350 * .2 + 500 * .1 \\
&= 155.
\end{aligned}
$$

Remark 32 *Note with wonder and amazement that* $150 * 2.7 - 250 = 155.0$.

3 Exercises

Navidi Section 2.4: 1-10