

## 3.4: Expected Value of Discrete Random Variables

In this section, and the next, we look at various *numerical* characteristics of discrete random variables. These give us a way of classifying and comparing random variables.

### Expected Value of Discrete Random Variables

We begin with the formal definition.

#### Definition 3.4.1

If  $X$  is a discrete random variable with possible values  $x_1, x_2, \dots, x_i, \dots$ , and probability mass function  $p(x)$ , then the **expected value** (or **mean**) of  $X$  is denoted  $E[X]$  and given by

$$E[X] = \sum_i x_i \cdot p(x_i). \quad (3.4.1)$$

The expected value of  $X$  may also be denoted as  $\mu_X$  or simply  $\mu$  if the context is clear.

The expected value of a random variable has many interpretations. First, looking at the formula in Definition 3.4.1 for computing expected value (Equation 3.4.1), note that it is essentially a **weighted average**. Specifically, for a discrete random variable, the expected value is computed by "weighting", or multiplying, each value of the random variable,  $x_i$ , by the probability that the random variable takes that value,  $p(x_i)$ , and then summing over all possible values. This interpretation of the expected value as a weighted average explains why it is also referred to as the **mean** of the random variable.

The expected value of a random variable is also interpreted as the **long-run value** of the random variable. In other words, if we repeat the underlying random experiment several times and take the average of the values of the random variable corresponding to the outcomes, we would get the expected value, approximately. (Note: This interpretation of expected value is similar to the [relative frequency approximation](#) for probability discussed in [Section 1.2](#).) Again, we see that the expected value is related to an average value of the random variable. Given the interpretation of the expected value as an average, either "weighted" or "long-run", the expected value is often referred to as a **measure of center** of the random variable.

Finally, the expected value of a random variable has a graphical interpretation. The expected value gives the **center of mass** of the probability mass function, which the following example demonstrates.

#### Example 3.4.1

Consider again the context of [Example 1.1.1](#), where we recorded the sequence of heads and tails in two tosses of a fair coin. In [Example 3.1.1](#) we defined the discrete random variable  $X$  to denote the number of heads obtained. In [Example 3.2.2](#) we found the pmf of  $X$ . We now apply Equation 3.4.1 from Definition 3.4.1 and compute the expected value of  $X$ :

$$\begin{aligned} E[X] &= 0 \cdot p(0) + 1 \cdot p(1) + 2 \cdot p(2) \\ &= 0 \cdot (0.25) + 1 \cdot (0.5) + 2 \cdot (0.25) \\ &= 0.5 + 0.5 = 1. \end{aligned}$$

Thus, we expect that the number of heads obtained in two tosses of a fair coin will be 1 in the long-run or on average. Figure 1 demonstrates the graphical representation of the expected value as the center of mass of the probability mass function.

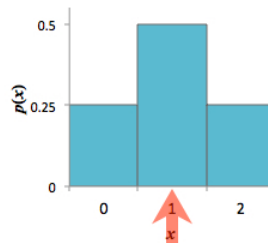


Figure 1: Histogram of  $X$ . The red arrow represents the center of mass, or the expected value of  $X$

### Example 3.4.2

Suppose we toss a fair coin three times and define the random variable  $X$  to be our winnings on a single play of a game where

- we win  $\$x$  if the first heads is on the  $x^{\text{th}}$  toss, for  $x = 1, 2, 3$ ,
- and we lose  $\$1$  if we get no heads in all three tosses.

Then  $X$  is a discrete random variable, with possible values  $x = -1, 1, 2, 3$ , and pmf given by the following table:

$x$	$p(x) = P(X = x)$
-1	$\frac{1}{8}$
1	$\frac{1}{2}$
2	$\frac{1}{4}$
3	$\frac{1}{8}$

Applying Definition 3.4.1, we find

$$\begin{aligned}
 E[X] &= \sum_i x_i \cdot p(x_i) \\
 &= (-1) \cdot \frac{1}{8} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} = \frac{5}{4} = 1.25.
 \end{aligned}$$

Thus, the expected winnings for a single play of the game is  $\$1.25$ . In other words, if we played the game multiple times, we expect the average winnings to be  $\$1.25$ .

For many of the common probability distributions, the expected value is given by a parameter of the distribution. The expected value may not be exactly equal to a parameter of the probability distribution, but rather it may be a function of the parameters. The following table gives the expected value for each of the common discrete distributions, including the Bernoulli and binomial distributions we introduced previously.

#### Expected Values for Discrete Distributions

Distribution	Expected Value
Bernoulli( $p$ )	$p$
binomial( $n, p$ )	$np$
hypergeometric( $N, n, m$ )	$\frac{nm}{N}$
geometric( $p$ )	$\frac{1}{p}$
negative binomial( $r, p$ )	$\frac{r}{p}$
Poisson( $\lambda$ )	$\lambda$

### Expected Value of Functions of Random Variables

In many applications, we may not be interested in the value of a random variable itself, but rather in a function applied to the random variable or a collection of random variables. For example, we may be interested in the value of  $X^2$ . The following theorems, which we state without proof, demonstrate how to calculate the expected value of *functions of random variables*.



### Theorem 3.4.1

Let  $X$  be a random variable and let  $g$  be a real-valued function. Define the random variable  $Y = g(X)$ .

If  $X$  is a discrete random variable with possible values  $x_1, x_2, \dots, x_i, \dots$ , and probability mass function  $p(x)$ , then the expected value of  $Y$  is given by

$$E[Y] = \sum_i g(x_i) \cdot p(x_i).$$

To put it simply, Theorem 3.4.1 states that to find the expected value of a function of a random variable, just apply the function to the possible values of the random variable in the definition of expected value. Before stating an important special case of Theorem 3.4.1, a word of caution regarding order of operations. Note that, in general,

$$E[g(X)] \neq g(E[X])! \quad (3.4.2)$$

For example,  $E[X^2] \neq (E[X])^2$ , in general. However, as the next theorem states, there are exceptions to Equation 3.4.2.

### Special Case of Theorem 3.4.1

Let  $X$  be a random variable. If  $g$  is a *linear* function, i.e.,  $g(x) = ax + b$ , then

$$E[g(X)] = E[aX + b] = aE[X] + b.$$

The above special case is referred to as the **linearity** of expected value, which implies the following properties of the expected value.

### Linearity of Expected Value

Let  $X$  be a random variable,  $c, c_1, c_2$  constants, and  $g, g_1, g_2$  real-valued functions. Then expectation  $E[\cdot]$  satisfies the following:

1. The expected value of a constant is constant:

$$E[c] = c$$

2. Constants can be factored out of expected values:

$$E[cg(X)] = cE[g(X)]$$

3. The expected value of a sum is equal to the sum of expected values:

$$E[c_1g_1(X) + c_2g_2(X)] = c_1E[g_1(X)] + c_2E[g_2(X)]$$

This page titled 3.4: Expected Value of Discrete Random Variables is shared under a [not declared](#) license and was authored, remixed, and/or curated by [Kristin Kuter](#).