

Lecture 8. Random Variables (continued), Expected Value, Variance

Example 4.

A biased coin will land on heads with probability p . Let us flip such a coin k times. An outcome ω is a sequence k of H s and T s. Now, the outcomes are not equally probable. We get all H s with probability p^k , and all T s with probability $(1-p)^k$. A moment of thought shows that the probability of an outcome depends on k , p and the number x of heads, but not on order in which the x heads and $k-x$ tails occur. Indeed, the probability of any such outcome is $p^x(1-p)^{k-x}$. Now, there are $\binom{k}{x}$ ways to place x heads in a sequence of length k , so the probability of flipping x heads in a sequence of k flips is:

$$b(x, k, p) = \binom{k}{x} p^x (1-p)^{k-x}.$$

If we define X to be the random variable that counts the number of heads in a sequence of k flips of a biased coin—i.e., $X(\omega) =$ the number of heads in ω —then $P(X = x) = b(x, k, p)$.

Problems 4. (See Mathematica Notebook 8.)

- In a Mathematica notebook, define a Mathematica function that calculates $b(x, k, p)$, the probability of getting x heads, in k flips of a coin that has probability p of landing on heads. (You should view $b(x, k, p)$ as a descriptive of a family of different probability spaces, which vary with the **parameters** k and p .)
- Using the function you have defined, find $P(30 < X \leq 40)$ when the parameters are $k = 100$ and $p = .25$.
- Compare to the results of Problem 3.

A philosophical interlude

Note that in Example 4, as well as in Example 3 previously, x denoted a fixed but unspecified number, while X was a function. This is a difference that makes a difference—a whopping big one, and it deserves to be thought through carefully and understood. Because x is unspecified, we may envision it having any value that we choose. The variation in x , however, is not part of the model itself, but is a choice we make in posing a question that we put to the model. On the other hand, the function X expresses the variation that occurs *within* the model. Within the model itself, X truly has many values, each being determined by evaluating X at a point of Ω .

Random variables are used to model the variation that we see in the data we have obtained through many many acts of measurement. The *occasions* on which we act to record data form a collection of some sort that lies “out there” in the real world. A body of data of sufficient size enables us to estimate the probability of obtaining specific data, when we act on random occasions to record observations. In probability modeling, we substitute a probability space for this collection of occasions, we substitute a formal pmf for the relative frequencies that we experience and we use the function X itself to model the process that goes from an occasion to the data obtained by performing a measurement on that occasion.

When we have a model, it is seldom that we can solve a problem with it simply by evaluating a random variable at a single point of sample space. Much more commonly, we wish to examine the global behavior of the random variable. (When we refer to “global” behavior of a function, we mean properties of the function that cannot be observed except by surveying its entire domain.) For example, to compute $P(X = x)$, we work with the set $\{\omega \in \Omega \mid X(\omega) = x\}$. We can never find out anything about this set by evaluating X at one element of Ω . To create a list of the elements of $\{\omega \in \Omega \mid X(\omega) = x\}$, we must sift through *all* the elements of Ω , check the value of X at each one, and based on whether or not $X(\omega) = x$ or not, put ω on our list, or leave it off.

The idea that X speaks about global aspects of the probability space comes to the fore in the concept of expected value, which we will look at next.

Expected value

Definition. Let X be a random variable on a discrete probability space Ω with pmf $f : \Omega \rightarrow [0, 1]$. The *expected value* of X , denoted $E(X)$, is defined as follows:

$$E(X) := \sum_{\omega \in \Omega} X(\omega)f(\omega).$$

In other words, the expected value of X is a weighted average of the values of X at all the outcomes, where the weights are the probability masses of the outcomes.

Example 1. Two dice are rolled. Let X be the maximum of the numbers obtained. Find $E(X)$.

Solution. The number of outcomes in the event $X = x$ is given in the following table:

$x =$	1	2	3	4	5	6
$\#(X = x) =$	1	3	5	7	9	11

Thus,

$$\begin{aligned} E(X) &= 1 \cdot \frac{1}{36} + 2 \cdot \frac{3}{36} + 3 \cdot \frac{5}{36} + 4 \cdot \frac{7}{36} + 5 \cdot \frac{9}{36} + 6 \cdot \frac{11}{36} \\ &= (1 + 6 + 15 + 28 + 45 + 66) \frac{1}{36} \\ &= \frac{161}{36}. \end{aligned}$$

Problem. Suppose the example is altered by rolling three dice instead of two. Find $E(X)$.

Example 2. The coupon problem. Six coupons are available in boxes of cereal. Each box contains 1 of the six, and each box is equally likely to hold any one of the six, regardless of what boxes have previously been sold. On average, how many boxes must you buy to get all six coupons?

Solution. We are going to approach this problem by programming a simulation. After we examine the output of the simulation, we will tackle it theoretically. See Mathematica Notebook 8.