Random Variables

Recall that an *experiment* is a trial producing one of a possible set of outcomes and its *sample space* is the set of all possible outcomes. A *random variable* is a **function** from the sample space of an experiment to $\mathbb{R}$. Let p be a probability distribution with sample space S, and let X be a random variable on S. We measure the statistical behavior of X by:

1. The *expected value* of X on S is $E(X) = \sum_{s \in S} p(s)X(s)$, and it measures the average value of X.

2. The *variance* of X on S is $V(X) = \sum_{s \in S}(X(s) - E(X))^2 p(s)$, and it measures the spread of X. We can also express it as $V(X) = E(X^2) - E(X)^2$. The *standard deviation* is defined as $\sigma(X) = \sqrt{V(X)}$.

3. Two random variables X, Y on S are *independent* if

$$p(X(s) = a \wedge Y(s) = b) = p(X(s) = a)p(Y(s) = b)$$

Here are some important facts abouts expected value and variance:

1. $E(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E(X_i)$ for any random variables X_i.
2. $E(aX + b) = aE(X) + b$ for any $a, b \in \mathbb{R}$ and random variable X.
3. $E(XY) = E(X)E(Y)$ for independent random variables X, Y.
4. $V(X + Y) = V(X) + V(Y)$ for independent random variables X, Y.
5. $V(aX) = a^2 V(X)$ for any $a \in \mathbb{R}$ and random variable X.
6. $V(X + a) = V(X)$ for any $a \in \mathbb{R}$ and random variable X.

Exercises

1. Prove the formulas $4, 5, 6$ above.

2. Consider the experiment of flipping a fair coin ($\frac{1}{2}$ probability heads and tails). Then the sample space of a single experiment is $S = \{\text{heads, tails}\}$ and our probability distribution assigns $\frac{1}{2}$ to each.

   Consider the following random variables on S:

   (a) X defined by $X(\text{tails}) = 0$ and $X(\text{heads}) = 1$.

   (b) Y defined by $Y(\text{tails}) = -1$ and $Y(\text{heads}) = 1$.

   Compute the new random variables XY and $X + Y$ on S.

   Now compute $E[XY]$ and $E[X]E[Y]$.

   Now compute $V(X + Y)$ and $V(X) + V(Y)$.

   Are X and Y independent?

3. Consider the experiment of rolling a 6-sided die until you roll a 3. What is the sample space S? What is the probability distribution on S? How many times do you expect to roll the die on average?