6.1 Discrete Random Variables

As illustrated in previous chapters, random variables, both discrete and continuous, appear naturally in discrete-event simulation models. Because of this, it is virtually impossible to build a valid discrete-event simulation model of a *system* without a good understanding of how to construct a valid random variable model for each of the stochastic system *components*. In this chapter and the next we will develop the mathematical and computational tools for building such stochastic models. Discrete random variables are considered in this chapter; continuous random variables are considered in the next.

6.1.1 DISCRETE RANDOM VARIABLE CHARACTERISTICS

The notation and development in this section largely follows the axiomatic approach to probability. As a convention, uppercase characters X, Y, ... are used to denote random variables (discrete or continuous), the corresponding lowercase characters x, y, ... are used to denote the specific values of X, Y, ..., and calligraphic characters $\mathcal{X}$, $\mathcal{Y}$, ... are used to denote the set of all possible values (often known as the *support* of the random variable). A variety of examples are used in this section to illustrate this notation.

Definition 6.1.1 The random variable X is *discrete* if and only if its set of possible values $\mathcal{X}$ is finite or, at most, countably infinite.

In a discrete-event simulation model discrete random variables are often integers used for counting, e.g., the number of jobs in a queue or the amount of inventory demand. There is no inherent reason, however, why a discrete random variable has to be integer-valued (see Example 6.1.14).

Probability Density Function

Definition 6.1.2 A discrete random variable X is uniquely determined by its set of possible values $\mathcal{X}$ and associated *probability density function* (*pdf*), a real-valued function $f(\cdot)$ defined for each possible value $x \in \mathcal{X}$ as the probability that X has the value x

$$f(x) = \Pr(X = x).$$

By definition, $x \in \mathcal{X}$ is a possible value of X if and only if $f(x) > 0$. In addition, $f(\cdot)$ is defined so that

$$\sum_x f(x) = 1$$

where the sum is over all $x \in \mathcal{X}$.*

It is important to understand the distinction between a random variable, its set of possible values, and its pdf. The usual way to construct a *model* of a discrete random variable X is to first specify the set of possible values $\mathcal{X}$ and then, for each $x \in \mathcal{X}$, specify the corresponding probability $f(x)$. The following three examples are illustrations.

---

* The pdf of a discrete random variable is sometimes called a probability *mass* function (pmf) or probability function (pf); that terminology is not used in this book.

Example 6.1.1 If the random variable X is *Equilikely*(a, b), then $\mathcal{X}$ is the set of integers between a and b inclusive. Because $|\mathcal{X}| = b-a+1$ and each possible value is equally likely, it follows that

$$f(x) = \frac{1}{b-a+1} \qquad\qquad x = a, a+1, \ldots, b.$$

As a specific example, if we were to roll one fair die and let the random variable X be the up face, then X would be *Equilikely*$(1, 6)$.

Example 6.1.2 Roll two fair dice. If the random variable X is the sum of the two up faces, then the set of possible values is $\mathcal{X} = \{x \mid x = 2, 3, \ldots, 12\}$ and from the table in Example 2.3.1, the pdf of X is

$$f(x) = \frac{6 - |7-x|}{36} \qquad\qquad x = 2, 3, \ldots, 12.$$

Although the sum of the up faces is the usual discrete random variable for games of chance that use two dice, see Exercise 6.1.2 for an alternative.

Example 6.1.3 Suppose a coin has p as its probability of a head and suppose we agree to toss this coin until the *first* tail occurs. If X is the number of heads (i.e., the number of tosses is $X+1$), then $\mathcal{X} = \{x \mid x = 0, 1, 2, \ldots\}$ and the pdf is

$$f(x) = p^x(1-p) \qquad\qquad x = 0, 1, 2, \ldots$$

This random variable is said to be *Geometric*(p). (The coin is *fair* if $p = 0.5$.)

Because the set of possible values is *infinite*, for a *Geometric*(p) random variable some math is required to verify that $\sum_x f(x) = 1$. Fortunately, this infinite series and other similar series can be evaluated by using the following properties of *geometric series*. If $p \neq 1$ then

$$1 + p + p^2 + p^3 + \cdots + p^x = \frac{1-p^{x+1}}{1-p} \qquad\qquad x = 0, 1, 2, \ldots$$

and if $|p| < 1$ then the following three *infinite series* converge to tractable quantities:

$$1 + p + p^2 + p^3 + p^4 + \cdots = \frac{1}{1-p},$$

$$1 + 2p + 3p^2 + 4p^3 + \cdots = \frac{1}{(1-p)^2},$$

$$1 + 2^2p + 3^2p^2 + 4^2p^3 + \cdots = \frac{1+p}{(1-p)^3}.$$

Although the three infinite series converge for any $|p| < 1$, negative values of p have no meaning in Example 6.1.3. From the first of these infinite series we have that

$$\sum_x f(x) = \sum_{x=0}^{\infty} p^x(1-p) = (1-p)(1 + p + p^2 + p^3 + p^4 + \cdots) = 1$$

as required.

As the previous examples illustrate, discrete random variables have possible values that are determined by the outcomes of a random experiment. Therefore, a Monte Carlo simulation program can be used to generate these possible values consistent with their probability of occurrence — see, for example, program `galileo` in Section 2.3. Provided the number of replications is large, a histogram of the values generated by replication should agree well with the random variable's pdf. Indeed, in the limit as the number of replications becomes infinite, the discrete-data histogram and the discrete random variable's pdf should agree *exactly*, as illustrated in Section 4.2.

Cumulative Distribution Function

Definition 6.1.3 The *cumulative distribution function* (*cdf*) of the discrete random variable X is the real-valued function $F(\cdot)$ defined for each $x \in \mathcal{X}$ as

$$F(x) = \Pr(X \leq x) = \sum_{t \leq x} f(t)$$

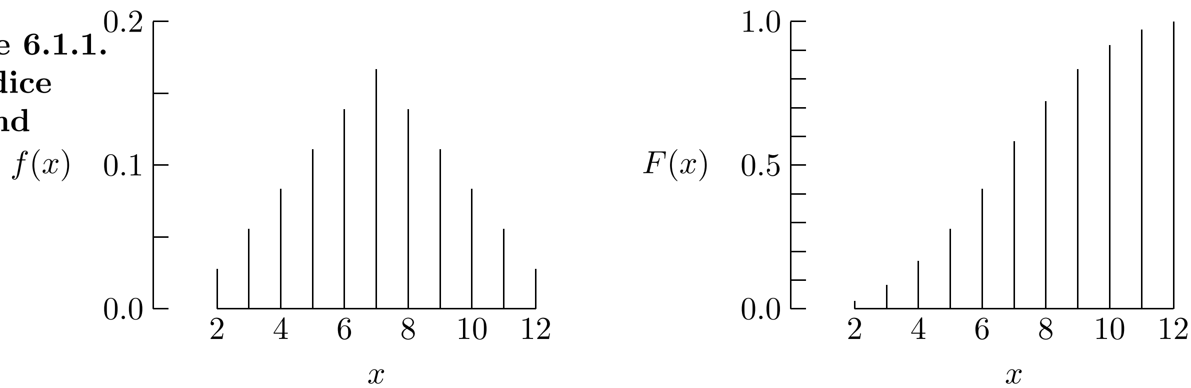
where the sum is over all $t \in \mathcal{X}$ for which $t \leq x$.

Example 6.1.4 If X is an *Equilikely*(a, b) random variable (Example 6.1.1), then the cdf is

$$F(x) = \sum_{t=a}^{x} 1/(b - a + 1) = (x - a + 1)/(b - a + 1) \qquad x = a, a + 1, \ldots, b.$$

Example 6.1.5 For the sum-of-two-dice random variable in Example 6.1.2 there is no simple equation for $F(x)$. That is of no real significance, however, because $|\mathcal{X}|$ is small enough that the cumulative pdf values are easily tabulated to yield the cdf. Figure 6.1.1 shows the pdf on the left and the cdf on the right (the vertical scale is different for the two figures). Any of the four styles for plotting a cdf given in Figure 4.2.8 is acceptable.

Figure 6.1.1. Two dice pdf and cdf.



Example 6.1.6 If X is a *Geometric*(p) random variable (Example 6.1.3), then the cdf is

$$F(x) = \sum_{t=0}^{x} p^t(1 - p) = (1 - p)(1 + p + p^2 + \cdots + p^x) = 1 - p^{x+1} \qquad x = 0, 1, 2, \ldots$$

The cdf of a discrete random variable can always be generated from its corresponding pdf by recursion. If the possible values of X are the consecutive integers $x = a, a+1, \ldots, b$, for example, then

$$\begin{aligned} F(a) &= f(a) \\ F(x) &= F(x-1) + f(x) \qquad\qquad x = a+1, a+2, \ldots, b. \end{aligned}$$

Similarly, a pdf can always be generated from its corresponding cdf by subtracting consecutive terms

$$\begin{aligned} f(a) &= F(a) \\ f(x) &= F(x) - F(x-1) \qquad\qquad x = a+1, a+2, \ldots, b. \end{aligned}$$

Therefore, a model of a discrete random variable can be defined by specifying *either* the pdf *or* the cdf and then computing the other — there is no need to specify both.

As illustrated in Example 6.1.5, the cdf of a discrete random variable is strictly monotone increasing — if x_1 and x_2 are possible values of X with $x_1 < x_2$, then $F(x_1) < F(x_2)$. Moreover, since $F(x)$ is defined as a probability, $0 \le F(x) \le 1$. The monotonicity of $F(\cdot)$ is fundamentally important; in the next section we will use it as the basis for a method to generate discrete random variates.

Mean and Standard Deviation

Recall from Definition 4.2.2 that for a discrete-data sample the sample mean $\bar{x}$ and standard deviation s can be computed from the discrete-data histogram as

$$\bar{x} = \sum_x x\hat{f}(x) \qquad \text{and} \qquad s = \sqrt{\sum_x (x-\bar{x})^2 \hat{f}(x)},$$

respectively. Moreover, the histogram relative frequencies $\hat{f}(x)$ converge to the corresponding probabilities $f(x)$ as the sample size becomes infinite. These two observations motivate the following definition.

Definition 6.1.4 The (population) *mean* μ and the corresponding (population) *standard deviation* σ are

$$\mu = \sum_x x f(x) \qquad \text{and} \qquad \sigma = \sqrt{\sum_x (x-\mu)^2 f(x)},$$

where the summations are over all $x \in \mathcal{X}$. The population *variance* is σ^2. An alternative, algebraically equivalent expression for σ is

$$\sigma = \sqrt{\left(\sum_x x^2 f(x)\right) - \mu^2}.$$

The population mean μ is a fixed constant, whereas the sample mean $\bar{x}$ is a random variable that varies from sample to sample. They are fundamentally different, although analogous, quantities. Since $\bar{x}$ is a random variable, it has a pdf and cdf that describe its distribution.

Example 6.1.7 If X is an *Equilikely*(a, b) random variable (Example 6.1.1), then

$$\mu = \frac{a+b}{2} \qquad \text{and} \qquad \sigma = \sqrt{\frac{(b-a+1)^2 - 1}{12}}$$

The derivation of these two equations is left as an exercise. In particular, if X is the result of rolling one fair die, then X is *Equilikely*$(1, 6)$ so that

$$\mu = 3.5 \qquad \text{and} \qquad \sigma = \sqrt{\frac{35}{12}} \cong 1.708,$$

illustrating the fact that μ need not necessarily be a member of $\mathcal{X}$. The geometric interpretation of the mean is the horizontal center of mass of the distribution.

Example 6.1.8 If X is the sum-of-two-dice random variable (Example 6.1.2), then

$$\mu = \sum_{x=2}^{12} x f(x) = \cdots = 7 \qquad \text{and} \qquad \sigma^2 = \sum_{x=2}^{12} (x-\mu)^2 f(x) = \cdots = 35/6.$$

Therefore the population standard deviation is $\sigma = \sqrt{35/6} \cong 2.415$.

Example 6.1.9 If X is a *Geometric*(p) random variable (Example 6.1.3), then

$$\mu = \sum_{x=0}^{\infty} x f(x) = \sum_{x=1}^{\infty} x p^x (1-p) = \cdots = \frac{p}{1-p}$$

using the infinite series following Example 6.1.3, and

$$\sigma^2 = \left(\sum_{x=0}^{\infty} x^2 f(x) \right) - \mu^2 = \left(\sum_{x=1}^{\infty} x^2 p^x (1-p) \right) - \frac{p^2}{(1-p)^2} = \cdots = \frac{p}{(1-p)^2}$$

so that $\sigma = \sqrt{p}/(1-p)$. The derivation of these equations is left as an exercise. In particular, tossing a fair ($p = 0.5$) coin until the first tail occurs generates a *Geometric*(0.5) random variable (X is the number of heads) with

$$\mu = 1 \qquad \text{and} \qquad \sigma = \sqrt{2} \cong 1.414.$$

The population mean and the population variance are two special cases of a more general notion known as "expected value".

Expected Value

Definition 6.1.5 The mean of a random variable (discrete or continuous) is also known as the *expected value*. It is conventional to denote the expected value as $E[\,\cdot\,]$. That is, the expected value of the discrete random variable X is

$$E[X] = \sum_{x} x f(x) = \mu$$

where the summation is over all $x \in \mathcal{X}$.*

* The expected value may not exist if there are infinitely many possible values.

If we were to use a Monte Carlo simulation to generate a large random variate sample $x_1, x_2, \ldots, x_n$ corresponding to the random variable X and then calculate the sample mean $\bar{x}$, we would expect to find that $\bar{x} \to E[X] = \mu$ as $n \to \infty$. Thus, the expected value of X is really the "expected average" and in that sense the term "expected value" is potentially misleading. The expected value (the mean) is not necessarily the *most likely* possible value [which is the *mode*, the element in $\mathcal{X}$ corresponding to the largest value of $f(x)$].

Example 6.1.10 If a fair coin is tossed until the first tail appears, then the *most likely* number of heads is 0 and the *expected* number of heads is 1 (see Example 6.1.9). In this case, 0 occurs with probability 1/2 and 1 occurs with probability 1/4. Thus the most likely value (the mode) is twice as likely as the expected value (the mean). On the other hand, for some random variables the mean and mode may be the same. For example, if X is the sum-of-two-dice random variable then the expected value and the most likely value are both 7 (see Examples 6.1.5 and 6.1.8).

Definition 6.1.6 If $h(\cdot)$ is a function defined for all possible values of X, then as x takes on all possible values in $\mathcal{X}$ the equation $y = h(x)$ defines the set $\mathcal{Y}$ of possible values for a *new* random variable $Y = h(X)$. The expected value of Y is

$$E[Y] = E[h(X)] = \sum_x h(x)f(x)$$

where the sum is over all $x \in \mathcal{X}$.*

Example 6.1.11 If $y = (x - \mu)^2$ with $\mu = E[X]$, then from Definitions 6.1.4 and 6.1.6

$$E[Y] = E[(X - \mu)^2] = \sum_x (x - \mu)^2 f(x) = \sigma^2.$$

That is, the variance σ^2 is the expected value of the squared difference about the mean. Similarly, if $y = x^2 - \mu^2$ then

$$E[Y] = E[X^2 - \mu^2] = \sum_x (x^2 - \mu^2)f(x) = \left(\sum_x x^2 f(x)\right) - \mu^2 = \sigma^2$$

so that

$$\sigma^2 = E[X^2] - E[X]^2.$$

This last equation demonstrates that the two operations $E[\cdot]$ and $(\cdot)^2$ do *not* commute; the expected value of X^2 is not equal to the square of $E[X]$. Indeed, $E[X^2] \geq E[X]^2$ with equality if and only if X is not really random at all, i.e., $\sigma^2 = 0$, often known as a *degenerate* distribution.

---

* Definition 6.1.6 can be established as a theorem. When presented as a definition it is sometimes called the "law of the unconscious statistician". If the set of possible values $\mathcal{X}$ is infinite then $E[Y]$ may not exist.

Example 6.1.12 If $Y = aX + b$ for constants a and b, then

$$E[Y] = E[aX + b] = \sum_x (ax + b)f(x) = a\left(\sum_x xf(x)\right) + b = aE[X] + b.$$

In particular, suppose that X is the number of heads before the first tail and that you are playing a game with a fair coin where you win \$2 for every head. Let the random variable Y be the amount you win. The possible values of Y are defined by

$$y = h(x) = 2x \qquad\qquad x = 0, 1, 2, \ldots$$

and your *expected winnings* (for each play of the game) are

$$E[Y] = E[2X] = 2E[X] = 2.$$

If you play this game repeatedly and pay more than \$2 per game to do so, then in the long run expect to lose money.

6.1.2 DISCRETE RANDOM VARIABLE MODELS

Let X be any discrete random *variable*. In the next section we will consider a unified algorithmic approach to generating possible values of X. The values so generated are random *variates*. The distinction between a random variable (discrete or continuous) and a corresponding random variate is subtle, but important. The former is an abstract, but well-defined, mathematical function that maps the outcome of an experiment to a real number (see Definition 6.1.2); the latter is an algorithmically generated possible value (realization) of the former (see Definition 2.3.2). For example, the functions `Equilikely` (Definition 2.3.4) and `Geometric` (Definition 3.1.2) generate random *variates* corresponding to *Equilikely*(a, b) and *Geometric*(p) random *variables*, respectively.

Bernoulli Random Variable

Example 6.1.13 The discrete random *variable* X with possible values $\mathcal{X} = \{0, 1\}$ is said to be *Bernoulli*(p) if $X = 1$ with probability p and $X = 0$ otherwise (i.e., with probability $1 - p$). In effect, X is a Boolean random variable with 1 as `true` and 0 as `false`. The pdf for a *Bernoulli*(p) random variable is $f(x) = p^x(1-p)^{1-x}$ for $x \in \mathcal{X}$. The corresponding cdf is $F(x) = (1-p)^{1-x}$ for $x \in \mathcal{X}$. The mean is $\mu = 0 \cdot (1-p) + 1 \cdot p = p$ and the variance is $\sigma^2 = (0-p)^2(1-p) + (1-p)^2 p = p(1-p)$. Therefore, the standard deviation is $\sigma = \sqrt{p(1-p)}$. We can generate a corresponding *Bernoulli*(p) random *variate* as follows.

```
if (Random() < 1.0 - p)
   return 0;
else
   return 1;
```

As illustrated in Section 2.3, no matter how sophisticated or computationally complex, a Monte Carlo simulation that uses n replications to estimate an (unknown) probability p is equivalent to generating an *iid* sequence of n *Bernoulli*(p) random variates.

Example 6.1.14 A popular state lottery game *Pick-3* requires players to pick a 3-digit number from the 1000 numbers between 000 and 999. It costs \$1 to play the game. If the 3-digit number picked by a player matches the 3-digit number chosen, at random, by the state then the player wins \$500, minus the original \$1 investment, for a net yield of +\$499. Otherwise, the player's yield is −\$1. (See Exercise 6.1.7 for another way to play this game.) Let the discrete random variable X represent the result of playing the game with the convention that $X = 1$ denotes a win and $X = 0$ denotes a loss. Then X is a *Bernoulli*(p) random variable with $p = 1/1000$. In addition, let the discrete random variable $Y = h(X)$ be the player's yield where

$$h(x) = \begin{cases} -1 & x = 0 \\ 499 & x = 1. \end{cases}$$

From Definition 6.1.6, the player's *expected* yield is

$$E[Y] = \sum_{x=0}^{1} h(x)f(x) = h(0)(1-p) + h(1)p = -1 \cdot \frac{999}{1000} + 499 \cdot \frac{1}{1000} = -0.5.$$

In this case, Y has just two possible values — one is 1000 times more likely than the other and neither is the expected value. Even though the support values for Y are far apart (−1 and 499), the value of $E[Y]$ shows that playing *Pick-3* is the equivalent of a voluntary 50 cent tax to the state for every dollar bet.

Because it has only two possible values, a *Bernoulli*(p) random variable may seem to have limited applicability. That is not the case, however, because this simple random variable can be used to construct more sophisticated stochastic models, as illustrated by the following examples.

Binomial Random Variable

Example 6.1.15 In the spirit of Example 6.1.3, suppose a coin has p as its probability of tossing a head and suppose we toss this coin n times. Let X be the number of heads; in this case X is said to be a *Binomial*(n, p) random variable. The set of possible values is $\mathcal{X} = \{0, 1, 2, \ldots, n\}$ and the associated pdf is

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} \qquad x = 0, 1, 2, \ldots, n.$$

That is, $p^x(1-p)^{n-x}$ is the probability of x heads and $n - x$ tails and the binomial coefficient accounts for the number of different sequences in which these heads and tails can occur. Equivalently, n tosses of the coin generate an *iid* sequence X_1, X_2, ..., X_n of *Bernoulli*(p) random variables ($X_i = 1$ corresponds to a head on the i^{th} toss) and

$$X = X_1 + X_2 + \cdots + X_n.$$

Although it may be intuitive that the pdf in Example 6.1.15 is correct, when building a discrete random variable model it is necessary to confirm that, in fact, $f(x) > 0$ for all $x \in \mathcal{X}$ (which is obvious in this case) and that $\sum_x f(x) = 1$ (which is not obvious). To verify that the pdf sum is 1, we can use the *binomial equation*

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}.$$

In the particular case where $a = p$ and $b = 1 - p$

$$1 = (1)^n = \big(p + (1-p)\big)^n = \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x}$$

which is equivalent to $f(0) + f(1) + \cdots + f(n) = 1$ as desired. To determine the mean of a binomial random variable:

$$\begin{aligned}
\mu = E[X] &= \sum_{x=0}^{n} x f(x) \\
&= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} x \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} x \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!\,(n-x)!} p^x (1-p)^{n-x} \\
&= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\,(n-x)!} p^{x-1} (1-p)^{n-x}.
\end{aligned}$$

To evaluate the last sum, let $m = n - 1$ and $t = x - 1$ so that $m - t = n - x$. Then from the binomial equation

$$\mu = np \sum_{t=0}^{m} \frac{m!}{t!\,(m-t)!} p^t (1-p)^{m-t} = np\big(p + (1-p)\big)^m = np(1)^m = np.$$

That is, the mean of a *Binomial*(n, p) random variable is $\mu = np$. In a similar way, it can be shown that the variance is

$$\sigma^2 = E[X^2] - \mu^2 = \left(\sum_{x=0}^{n} x^2 f(x)\right) - \mu^2 = \cdots = np(1-p)$$

so that the standard deviation of a *Binomial*(n, p) random variable is $\sigma = \sqrt{np(1-p)}$.

Pascal Random Variable

Example 6.1.16 As a second example of using *Bernoulli*(p) random variables to build a more sophisticated stochastic model, suppose a coin has p as its probability of a head and suppose we toss this coin until the n^{th} tail occurs. If X is the number of heads (i.e., the number of tosses is $X + n$), then X is said to be a *Pascal*(n, p) random variable. The set of possible values is $\mathcal{X} = \{0, 1, 2, \ldots\}$ and the associated pdf is

$$f(x) = \binom{n+x-1}{x} p^x (1-p)^n \qquad\qquad x = 0, 1, 2, \ldots$$

That is, $p^x(1-p)^n$ is the probability of x heads and n tails and the binomial coefficient accounts for the number of different sequences in which these $n + x$ heads and tails can occur, given that the last coin toss must be a tail.

Although it may be intuitive that the *Pascal*(n, p) pdf equation is correct, it is necessary to prove that the infinite pdf sum does, in fact, converge to 1. The proof of this property is based on another (negative exponent) version of the *binomial equation* — for any positive integer n

$$(1-p)^{-n} = 1 + \binom{n}{1} p + \binom{n+1}{2} p^2 + \cdots + \binom{n+x-1}{x} p^x + \cdots$$

provided $|p| < 1$. (See Example 6.1.3 for $n = 1$ and $n = 2$ versions of this equation.) By using the negative-exponent binomial equation we see that*

$$\sum_{x=0}^{\infty} \binom{n+x-1}{x} p^x (1-p)^n = (1-p)^n \sum_{x=0}^{\infty} \binom{n+x-1}{x} p^x = (1-p)^n (1-p)^{-n} = 1$$

which confirms that $f(0) + f(1) + f(2) + \cdots = 1$. Moreover, in a similar way it can be shown that the mean is

$$\mu = E[X] = \sum_{x=0}^{\infty} x f(x) = \cdots = \frac{np}{1-p}$$

and the variance is

$$\sigma^2 = E[X^2] - \mu^2 = \left(\sum_{x=0}^{\infty} x^2 f(x) \right) - \mu^2 = \cdots = \frac{np}{(1-p)^2}$$

so that the standard deviation is $\sigma = \sqrt{np}/(1-p)$. The details of this derivation are left as an exercise.

---

* Because the negative-exponent binomial equation is applicable to the mathematical analysis of its characteristics, a *Pascal*(n, p) random variable is also known as a *negative binomial* random variable.

Example 6.1.17 As a third example of using *Bernoulli*(p) random variables to build a more sophisticated stochastic model, a *Geometric*(p) random variable is a special case of a *Pascal*(n, p) random variable when $n = 1$. If instead $n > 1$ and if X_1, X_2, ..., X_n is an *iid* sequence of n *Geometric*(p) random variables, then the sum

$$X = X_1 + X_2 + \cdots + X_n$$

is a *Pascal*(n, p) random variable. For example, if $n = 4$ and if p is large then a typical head/tail sequence might look like

$$\underbrace{hhhhhht}_{X_1=6}\;\underbrace{hhhhhhhhht}_{X_2=9}\;\underbrace{hhhht}_{X_3=4}\;\underbrace{hhhhhhht}_{X_4=7}$$

where X_1, X_2, X_3, X_4 count the number of heads in each $h \ldots ht$ sequence and, in this case,

$$X = X_1 + X_2 + X_3 + X_4 = 26.$$

The number of heads in each $h \cdots ht$ *Bernoulli*(p) sequence is an independent realization of a *Geometric*(p) random variable. In this way we see that a *Pascal*(n, p) random variable is the sum of n *iid* *Geometric*(p) random variables. From Example 6.1.9 we know that if X is *Geometric*(p) then $\mu = p/(1 - p)$ and $\sigma = \sqrt{p}/(1 - p)$.

Poisson Random Variable

A *Poisson*(μ) random variable is a limiting case of a *Binomial*$(n, \mu/n)$ random variable. That is, let X be a *Binomial*(n, p) random variable with $p = \mu/n$. Fix the values of μ and x and consider what happens in the limit as $n \to \infty$. The pdf of X is

$$f(x) = \frac{n!}{x!(n-x)!}\left(\frac{\mu}{n}\right)^x\left(1-\frac{\mu}{n}\right)^{n-x} = \frac{\mu^x}{x!}\left(\frac{n!n^x}{(n-x)!(n-\mu)^x}\right)\left(1-\frac{\mu}{n}\right)^n$$

for $x = 0, 1, \ldots, n$. It can be shown that

$$\lim_{n\to\infty}\left(\frac{n!n^x}{(n-x)!(n-\mu)^x}\right) = 1 \qquad \text{and} \qquad \lim_{n\to\infty}\left(1-\frac{\mu}{n}\right)^n = \exp(-\mu)$$

so that

$$\lim_{n\to\infty} f(x) = \frac{\mu^x \exp(-\mu)}{x!}.$$

This limiting case is the motivation for defining a *Poisson*(μ) random variable. For large values of n, *Binomial*$(n, \mu/n)$ and *Poisson*(μ) random variables are virtually identical, particularly if μ is small. As an exercise you are asked to prove that the parameter μ in the definition of the *Poisson*(μ) pdf

$$f(x) = \frac{\mu^x \exp(-\mu)}{x!} \qquad x = 0, 1, 2, \ldots$$

is in fact the mean and that the standard deviation is $\sigma = \sqrt{\mu}$.

6.1.3 EXERCISES

Exercise 6.1.1 (a) Simulate rolling a pair of dice 360 times with five different seeds and generate five histograms of the resulting sum of the two up faces. Compare the histogram mean, standard deviation and relative frequencies with the corresponding population mean, standard deviation and pdf. (b) Repeat for 3600, 36 000, and 360 000 replications. (c) Comment.

Exercise 6.1.2 Repeat the previous exercise *except* that the random variable of interest is the absolute value of the *difference* between the two up faces.

Exercise 6.1.3 Derive the equations for μ and σ in Example 6.1.7. (See Exercise 6.1.5.)

Exercise 6.1.4 Prove that $\sum_x (x-\mu)^2 f(x) = \big(\sum_x x^2 f(x)\big) - \mu^2$.

Exercise 6.1.5 X is a discrete random variable with possible values $x = 1, 2, \ldots, n$. (a) If the pdf of X is $f(x) = \alpha x$ then what is α (as a function of n)? (b) Determine the cdf, mean, and standard deviation of X. *Hint*:

$$\sum_{x=1}^{n} x = \frac{n(n+1)}{2} \qquad \sum_{x=1}^{n} x^2 = \frac{n(n+1)(2n+1)}{6} \qquad \sum_{x=1}^{n} x^3 = \frac{n^2(n+1)^2}{4}.$$

Exercise 6.1.6 Fill in the $= \cdots =$'s in Example 6.1.9.

Exercise 6.1.7 As an alternative to Example 6.1.14, another way to play Pick-3 is for the player to opt for a win *in any order*. That is, for example, if the player's number is 123 then the player will win (the same amount) if the state draws any of 123, 132, 231, 213, 321, or 312. Because this is an easier win, the pay-off is suitably smaller, namely \$80 for a net yield of +\$79. (a) What is the player's expected yield (per game) if this variation of the game is played? (Assume the player is bright enough to pick a 3-digit number with three different digits.) (b) Construct a Monte Carlo simulation to supply "convincing numerical evidence" of the correctness of your solution.

Exercise 6.1.8[a] An urn is initially filled with 1 amber ball and 1 black ball. Each time a ball is drawn, at random, if it is a black ball then it *and* another black ball are put back in the urn. Let X be the number of random draws required to find the amber ball. (a) What is $E[X]$? (b) Construct a Monte Carlo simulation to estimate $E[X]$ based on 1000, 10 000, 100 000, and 1 000 000 replications. (c) Comment.

Exercise 6.1.9[a] The location of an interval of fixed length $r > 0$ is selected at random on the real number line. Let X be the number of integers within the interval. Find the mean and standard deviation of X as a function of (r, p) where $p = r - \lfloor r \rfloor$.

Exercise 6.1.10 If X is *Poisson*(μ) random variable, prove that the mean of X is the parameter μ and that the standard deviation is $\sqrt{\mu}$.