Definitions

Discrete Random Variables

- In this chapter and for most of the remainder of this book, we examine probability models that assign numbers to the outcomes in the sample space.

- When we observe one of these numbers, we refer to the observation as a *random variable*.

- In our notation, the name of a random variable is always a capital letter, for example, X.

- The set of possible values of X is the *range* of X.

- Since we often consider more than one random variable at a time, we denote the range of a random variable by the letter S with a subscript that is the name of the random variable.

- Thus S_X is the range of random variable X, S_Y is the range of random variable Y, and so forth.

Example 3.1

The experiment is to attach a photo detector to an optical fiber and count the number of photons arriving in a one-microsecond time interval. Each observation is a random variable X. The range of X is $S_X = \{0, 1, 2, \ldots\}$. In this case, S_X, the range of X, and the sample space S are identical.

Example 3.2

The experiment is to test six integrated circuits and after each test observe whether the circuit is accepted (a) or rejected (r). Each observation is a sequence of six letters where each letter is either a or r. For example, $s_8 =$ aaraaa. The sample space S consists of the 64 possible sequences. A random variable related to this experiment is N, the number of accepted circuits. For outcome s_8, $N = 5$ circuits are accepted. The range of N is $S_N = \{0, 1, \ldots, 6\}$.

Example 3.3

In Example 3.2, the net revenue R obtained for a batch of six integrated circuits is \$5 for each circuit accepted minus \$7 for each circuit rejected. (This is because for each bad circuit that goes out of the factory, it will cost the company \$7 to deal with the customer's complaint and supply a good replacement circuit.) When N circuits are accepted, $6 - N$ circuits are rejected so that the net revenue R is related to N by the function

$$R = g(N) = 5N - 7(6 - N) = 12N - 42 \text{ dollars.} \tag{3.1}$$

Since $S_N = \{0, \ldots, 6\}$, the range of R is

$$S_R = \{-42, -30, -18, -6, 6, 18, 30\}. \tag{3.2}$$

The revenue associated with $s_8 =$ aaraaa and all other outcomes for which $N = 5$ is

$$g(5) = 12 \times 5 - 42 = 18 \text{ dollars} \tag{3.3}$$

Definition 3.1 Random Variable

A random variable *consists of an experiment with a probability measure* P$[\cdot]$ *defined on a sample space* S *and a function that assigns a real number to each outcome in the sample space of the experiment.*

Example 3.4 Problem

The procedure of an experiment is to fire a rocket in a vertical direction from Earth's surface with initial velocity V km/h. The observation is T seconds, the time elapsed until the rocket returns to Earth. Under what conditions is the experiment improper?

Example 3.4 Solution

At low velocities, V, the rocket will return to Earth at a random time T seconds that depends on atmospheric conditions and small details of the rocket's shape and weight. However, when $V > v^* \approx 40{,}000$ km/hr, the rocket will not return to Earth. Thus, the experiment is improper when $V > v^*$ because it is impossible to perform the specified observation.

Notation

- On occasion, it is important to identify the random variable X by the function $X(s)$ that maps the sample outcome s to the corresponding value of the random variable X.

- As needed, we will write $\{X = x\}$ to emphasize that there is a set of sample points $s \in S$ for which $X(s) = x$.

- That is, we have adopted the shorthand notation

$$\{X = x\} = \{s \in S | X(s) = x\}. \tag{3.4}$$

Definition 3.2 Discrete Random Variable

X is a discrete *random variable if the range of X is a countable set*

$$S_X = \{x_1, x_2, \ldots\}.$$

Quiz 3.1

A student takes two courses. In each course, the student will earn either a B or a C. To calculate a grade point average (GPA), a B is worth 3 points and a C is worth 2 points. The student's GPA G_2 is the sum of the points earned for each course divided by 2. Make a table of the sample space of the experiment and the corresponding values of the GPA, G_2.

Quiz 3.1 Solution

The sample space, probabilities and corresponding grades for the experiment are

Outcomes	BB	BC	CB	CC
G_2	3.0	2.5	2.5	2.0

Probability Mass Function

Probability Mass Function

Definition 3.3 (PMF)

The probability mass function *(PMF) of the discrete random variable X is*

$$P_X(x) = \mathsf{P}[X = x]$$

Example 3.5 Problem

When the basketball player Wilt Chamberlain shot two free throws, each shot was equally likely either to be good (g) or bad (b). Each shot that was good was worth 1 point. What is the PMF of X, the number of points that he scored?

Example 3.5 Solution

There are four outcomes of this experiment: gg, gb, bg, and bb. A simple tree diagram indicates that each outcome has probability $1/4$. The sample space and probabilities of the experiment and the corresponding values of X are given in the table:

Outcomes	bb	bg	gb	gg
$P[\cdot]$	$1/4$	$1/4$	$1/4$	$1/4$
X	0	1	1	2

The random variable X has three possible values corresponding to three events:

$$\{X = 0\} = \{bb\}, \qquad \{X = 1\} = \{gb, bg\}, \qquad \{X = 2\} = \{gg\}. \tag{3.6}$$

Since each outcome has probability $1/4$, these three events have probabilities

$$P[X = 0] = 1/4, \qquad P[X = 1] = 1/2, \qquad P[X = 2] = 1/4. \tag{3.7}$$

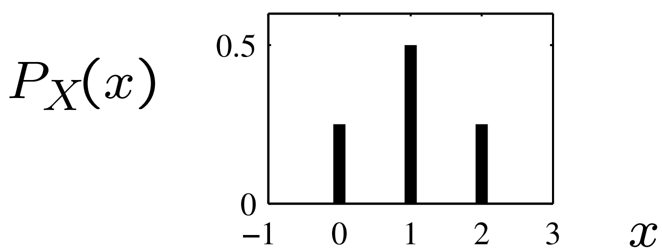
[Continued]

Example 3.5 Solution (Continued 2)

We can express the probabilities of these events in terms of the probability mass function

$$P_X(x) = \begin{cases} 1/4 & x = 0, \\ 1/2 & x = 1, \\ 1/4 & x = 2, \\ 0 & \text{otherwise.} \end{cases} \tag{3.8}$$

It is often useful or convenient to depict $P_X(x)$ in two other display formats: as a bar plot or as a table.



x	0	1	2
$P_X(x)$	1/4	1/2	1/4

Each PMF display format has its uses. The function definition (3.8) is best when $P_X(x)$ is given in terms of algebraic functions of x for various subsets of S_X. The bar plot is best for visualizing the probability masses. The table can be a convenient compact representation when the PMF is a long list of sample values and corresponding probabilities.

Theorem 3.1

For a discrete random variable X with PMF $P_X(x)$ and range S_X:

(a) For any x, $P_X(x) \geq 0$.

(b) $\sum_{x \in S_X} P_X(x) = 1$.

(c) For any event $B \subset S_X$, the probability that X is in the set B is

$$\mathrm{P}[B] = \sum_{x \in B} P_X(x).$$

Proof: Theorem 3.1

All three properties are consequences of the axioms of probability (Section 1.2). First, $P_X(x) \geq 0$ since $P_X(x) = \mathrm{P}[X = x]$. Next, we observe that every outcome $s \in S$ is associated with a number $x \in S_X$. Therefore, $\mathrm{P}[x \in S_X] = \sum_{x \in S_X} P_X(x) = \mathrm{P}[s \in S] = \mathrm{P}[S] = 1$. Since the events $\{X = x\}$ and $\{X = y\}$ are mutually exclusive when $x \neq y$, B can be written as the union of mutually exclusive events $B = \bigcup_{x \in B}\{X = x\}$. Thus we can use Axiom 3 (if B is countably infinite) or Theorem 1.2 (if B is finite) to write

$$\mathrm{P}[B] = \sum_{x \in B} \mathrm{P}[X = x] = \sum_{x \in B} P_X(x). \tag{3.9}$$

Quiz 3.2

The random variable N has PMF

$$P_N(n) = \begin{cases} c/n & n = 1, 2, 3, \\ 0 & \text{otherwise.} \end{cases} \tag{3.10}$$

Find

(a) The value of the constant c

(b) $\mathrm{P}[N = 1]$

(c) $\mathrm{P}[N \geq 2]$

(d) $\mathrm{P}[N > 3]$

Quiz 3.2 Solution

(a) To find c, we recall that the PMF must sum to 1. That is,

$$\sum_{n=1}^{3} P_N(n) = c\left(1 + \frac{1}{2} + \frac{1}{3}\right) = 1. \tag{1}$$

This implies $c = 6/11$. Now that we have found c, the remaining parts are straightforward.

(b) $\mathrm{P}[N = 1] = P_N(1) = c = 6/11$.

(c) $\mathrm{P}[N \geq 2] = P_N(2) + P_N(3)$
$= c/2 + c/3 = 5/11$.

(d) $\mathrm{P}[N > 3] = \sum_{n=4}^{\infty} P_N(n) = 0$.

Section 3.3

Families of Discrete Random Variables

Families of Random Variables

- In practical applications, certain families of random variables appear over and over again in many experiments.
- In each family, the probability mass functions of all the random variables have the same mathematical form.
- They differ only in the values of one or two parameters.
- Depending on the family, the PMF formula contains one or two parameters.
- By assigning numerical values to the parameters, we obtain a specific random variable.
- Our nomenclature for a family consists of the family name followed by one or two parameters in parentheses.
- For example, *binomial* (n, p) refers in general to the family of binomial random variables.
- *Binomial* $(7, 0.1)$ refers to the binomial random variable with parameters $n = 7$ and $p = 0.1$.

Example 3.6

Consider the following experiments:

- Flip a coin and let it land on a table. Observe whether the side facing up is heads or tails. Let X be the number of heads observed.

- Select a student at random and find out her telephone number. Let $X = 0$ if the last digit is even. Otherwise, let $X = 1$.

- Observe one bit transmitted by a modem that is downloading a file from the Internet. Let X be the value of the bit (0 or 1).

All three experiments lead to the probability mass function

$$P_X(x) = \begin{cases} 1/2 & x = 0, \\ 1/2 & x = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.11}$$

Definition 3.4 Bernoulli (p) Random Variable

X *is a* Bernoulli (p) *random variable if the PMF of* X *has the form*

$$P_X(x) = \begin{cases} 1-p & x=0, \\ p & x=1, \\ 0 & \textit{otherwise}, \end{cases}$$

where the parameter p *is in the range* $0 < p < 1$*.*

Example 3.7 Problem

Test one circuit and observe X, the number of rejects. What is $P_X(x)$ the PMF of random variable X?

Example 3.7 Solution

Because there are only two outcomes in the sample space, $X = 1$ with probability p and $X = 0$ with probability $1 - p$,

$$P_X(x) = \begin{cases} 1-p & x = 0, \\ p & x = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.12}$$

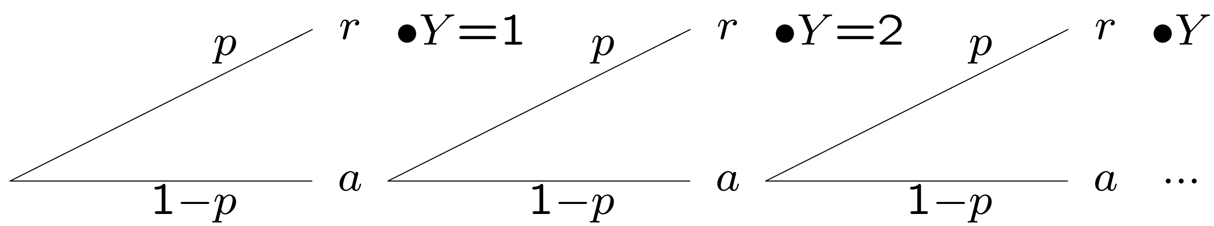
Therefore, the number of circuits rejected in one test is a Bernoulli (p) random variable.

Example 3.8 Problem

In a sequence of independent tests of integrated circuits, each circuit is rejected with probability p. Let Y equal the number of tests up to and including the first test that results in a reject. What is the PMF of Y?

Example 3.8 Solution

The procedure is to keep testing circuits until a reject appears. Using a to denote an accepted circuit and r to denote a reject, the tree is



From the tree, we see that $\text{P}[Y = 1] = p$, $\text{P}[Y = 2] = p(1-p)$, $\text{P}[Y = 3] = p(1-p)^2$, and, in general, $\text{P}[Y = y] = p(1-p)^{y-1}$. Therefore,

$$P_Y(y) = \begin{cases} p(1-p)^{y-1} & y = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases} \tag{3.13}$$

Y is referred to as a *geometric random variable* because the probabilities in the PMF constitute a geometric series.

Definition 3.5 Geometric (p) Random Variable

X *is a* geometric (p) *random variable if the PMF of* X *has the form*

$$P_X(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, \ldots \\ 0 & \textit{otherwise.} \end{cases}$$

where the parameter p *is in the range* $0 < p < 1$.

Example 3.9 Problem

In a sequence of n independent tests of integrated circuits, each circuit is rejected with probability p. Let K equal the number of rejects in the n tests. Find the PMF $P_K(k)$.

Example 3.9 Solution

Adopting the vocabulary of Section 2.3, we call each discovery of a defective circuit a *success*, and each test is an independent trial with success probability p. The event $K = k$ corresponds to k successes in n trials. We refer to Theorem 2.8 to determine that the PMF of K is

$$P_K(k) = \binom{n}{k} p^k (1-p)^{n-k}. \tag{3.14}$$

K is an example of a *binomial random variable*.

Definition 3.6 Binomial (n, p) Random Variable

X *is a* binomial (n, p) *random variable if the PMF of* X *has the form*

$$P_X(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

where $0 < p < 1$ *and* n *is an integer such that* $n \geq 1$.

Example 3.10 Problem

Perform independent tests of integrated circuits in which each circuit is rejected with probability p. Observe L, the number of tests performed until there are k rejects. What is the PMF of L?

Example 3.10 Solution

For large values of k, it is not practical to draw the tree. In this case, $L = l$ if and only if there are $k-1$ successes in the first $l-1$ trials *and* there is a success on trial l so that

$$\mathrm{P}[L = l] = \mathrm{P}\left[\underbrace{k-1 \text{ rejects in } l-1 \text{ attempts}}_{A}, \underbrace{\text{reject on attempt } l}_{B}\right] \tag{3.15}$$

The events A and B are independent since the outcome of attempt l is not affected by the previous $l-1$ attempts. Note that $\mathrm{P}[A]$ is the binomial probability of $k-1$ successes (i.e., rejects) in $l-1$ trials so that

$$\mathrm{P}[A] = \binom{l-1}{k-1} p^{k-1}(1-p)^{l-1-(k-1)}. \tag{3.16}$$

Finally, since $\mathrm{P}[B] = p$,

$$P_L(l) = \mathrm{P}[A]\,\mathrm{P}[B] = \binom{l-1}{k-1} p^{k}(1-p)^{l-k} \tag{3.17}$$

L is an example of a *Pascal* random variable.

Definition 3.7 Pascal (k, p) Random Variable

X is a Pascal (k, p) *random variable if the PMF of* X *has the form*

$$P_X(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$$

where $0 < p < 1$ *and* k *is an integer such that* $k \geq 1$.

Example 3.11

In an experiment with equiprobable outcomes, the random variable N has the range $S_N = \{k, k+1, k+2, \cdots, l\}$, where k and l are integers with $k < l$. The range contains $l - k + 1$ numbers, each with probability $1/(l - k + 1)$. Therefore, the PMF of N is

$$P_N(n) = \begin{cases} 1/(l-k+1) & n = k, k+1, k+2, \ldots, l \\ 0 & \text{otherwise} \end{cases} \tag{3.18}$$

N is an example of a *discrete uniform* random variable.

Definition 3.8 Discrete Uniform (k, l) Random Variable

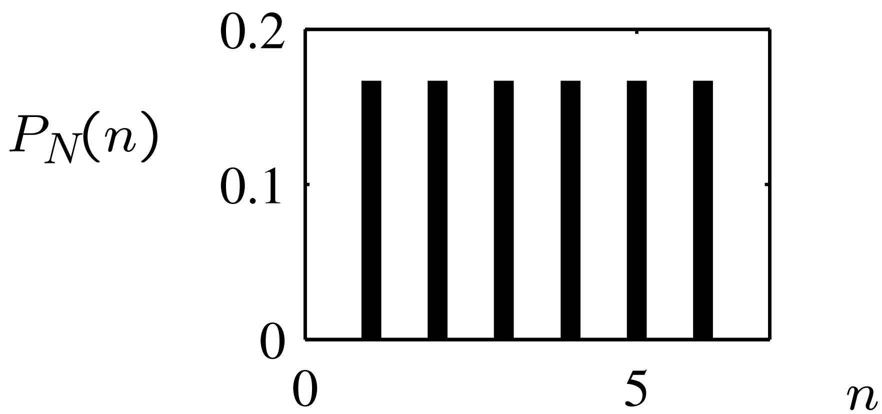
X *is a* discrete uniform (k, l) *random variable if the PMF of* X *has the form*

$$P_X(x) = \begin{cases} 1/(l-k+1) & x = k, k+1, k+2, \ldots, l \\ 0 & \text{otherwise} \end{cases}$$

where the parameters k *and* l *are integers such that* $k < l$.

Example 3.12

Roll a fair die. The random variable N is the number of spots on the side facing up. Therefore, N is a discrete uniform $(1, 6)$ random variable with PMF



$$P_N(n) = \begin{cases} 1/6 & n = 1, 2, 3, 4, 5, 6, \\ 0 & \text{otherwise.} \end{cases} \tag{3.19}$$

Definition 3.9 Poisson (α) Random Variable

X *is a* Poisson (α) *random variable if the PMF of* X *has the form*

$$P_X(x) = \begin{cases} \alpha^x e^{-\alpha}/x! & x = 0, 1, 2, \ldots, \\ 0 & \textit{otherwise}, \end{cases}$$

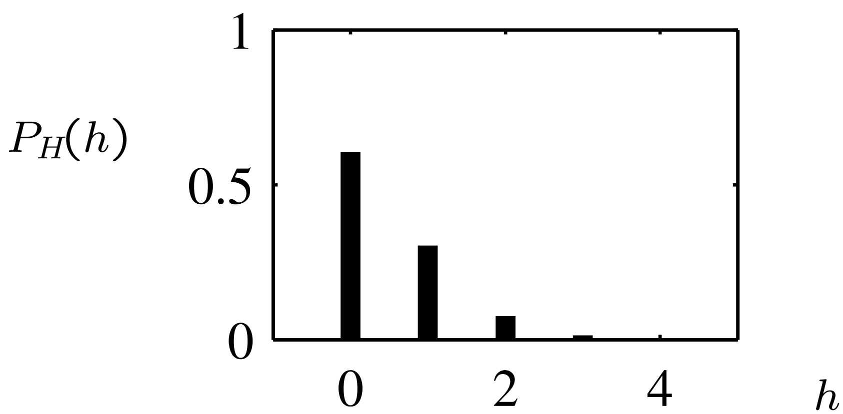
where the parameter α *is in the range* $\alpha > 0$.

Example 3.13 Problem

The number of hits at a website in any time interval is a Poisson random variable. A particular site has on average $\lambda = 2$ hits per second. What is the probability that there are no hits in an interval of 0.25 seconds? What is the probability that there are no more than two hits in an interval of one second?

Example 3.13 Solution

In an interval of 0.25 seconds, the number of hits H is a Poisson random variable with $\alpha = \lambda T = (2 \text{ hits/s}) \times (0.25 \text{ s}) = 0.5$ hits. The PMF of H is

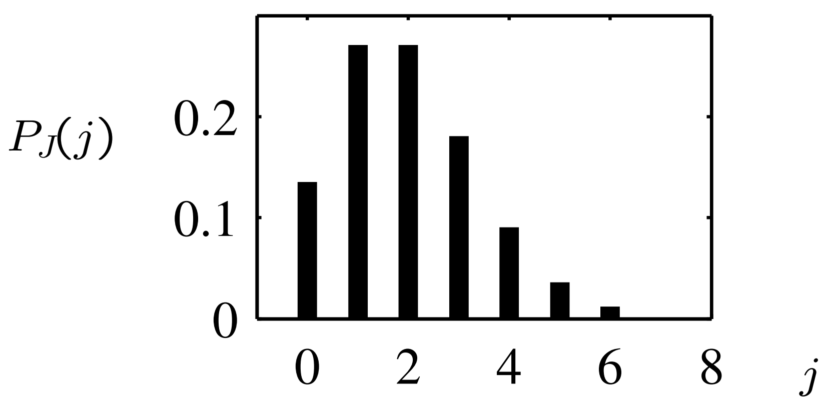


$$P_H(h) = \begin{cases} 0.5^h e^{-0.5}/h! & h = 0, 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

The probability of no hits is

$$P[H = 0] = P_H(0) = (0.5)^0 e^{-0.5}/0! = 0.607. \qquad (3.20)$$

In an interval of 1 second, $\alpha = \lambda T = (2 \text{ hits/s}) \times (1 \text{ s}) = 2$ hits. Letting J denote the number of hits in one second, the PMF of J is



$$P_J(j) = \begin{cases} 2^j e^{-2}/j! & j = 0, 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

[Continued]

Example 3.13 Solution (Continued 2)

To find the probability of no more than two hits, we note that

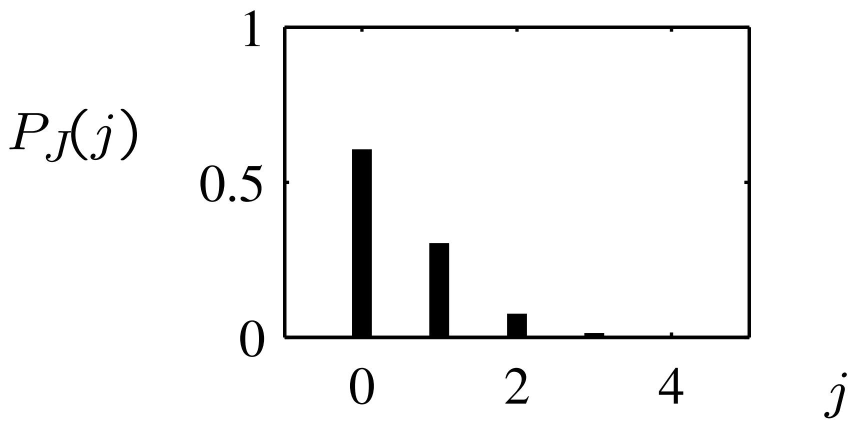
$$\{J \leq 2\} = \{J = 0\} \cup \{J = 1\} \cup \{J = 2\} \tag{3.21}$$

is the union of three mutually exclusive events. Therefore,

$$\begin{aligned} \mathrm{P}[J \leq 2] &= \mathrm{P}[J = 0] + \mathrm{P}[J = 1] + \mathrm{P}[J = 2] \\ &= P_J(0) + P_J(1) + P_J(2) \\ &= e^{-2} + 2^1 e^{-2}/1! + 2^2 e^{-2}/2! = 0.677. \end{aligned} \tag{3.22}$$

Example 3.14

Calls arrive at random times at a telephone switching office with an average of $\lambda = 0.25$ calls/second. The PMF of the number of calls that arrive in a $T = 2$-second interval is the Poisson (0.5) random variable with PMF



$$P_J(j) = \begin{cases} (0.5)^j e^{-0.5}/j! & j = 0, 1, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

Note that we obtain the same PMF if we define the arrival rate as $\lambda = 60 \cdot 0.25 = 15$ calls per minute and derive the PMF of the number of calls that arrive in $2/60 = 1/30$ minutes.

Quiz 3.3

Each time a modem transmits one bit, the receiving modem analyzes the signal that arrives and decides whether the transmitted bit is 0 or 1. It makes an error with probability p, independent of whether any other bit is received correctly.

(a) If the transmission continues until the receiving modem makes its first error, what is the PMF of X, the number of bits transmitted?

(b) If $p = 0.1$, what is the probability that $X = 10$? What is the probability that $X \geq 10$?

(c) If the modem transmits 100 bits, what is the PMF of Y, the number of errors?

(d) If $p = 0.01$ and the modem transmits 100 bits, what is the probability of $Y = 2$ errors at the receiver? What is the probability that $Y \leq 2$?

(e) If the transmission continues until the receiving modem makes three errors, what is the PMF of Z, the number of bits transmitted?

(f) If $p = 0.25$, what is the probability of $Z = 12$ bits transmitted until the modem makes three errors?

Quiz 3.3 Solution

Decoding each transmitted bit is an independent trial where we call a bit error a "success." Each bit is in error, that is, the trial is a success, with probability p. Now we can interpret each experiment in the generic context of independent trials.

(a) The random variable X is the number of trials up to and including the first success. Similar to Example 3.8, X has the geometric PMF

$$P_X(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

(b) If $p = 0.1$, then the probability exactly 10 bits are sent is

$$P_X(10) = (0.1)(0.9)^9 = 0.0387. \tag{1}$$

The probability that at least 10 bits are sent is

$$\mathrm{P}[X \geq 10] = \sum_{x=10}^{\infty} P_X(x). \tag{2}$$

This sum is not too hard to calculate. However, its even easier to observe that $X \geq 10$ if the first 10 bits are transmitted correctly. That is,

$$\mathrm{P}[X \geq 10] = \mathrm{P}[\text{first 10 bits correct}] = (1-p)^{10}. \tag{3}$$

For $p = 0.1$,

$$\mathrm{P}[X \geq 10] = 0.9^{10} = 0.3487. \tag{4}$$

(c) The random variable Y is the number of successes in 100 independent trials. Just as in Example 3.9, Y has the binomial PMF

$$P_Y(y) = \binom{100}{y} p^y (1-p)^{100-y}. \tag{5}$$

Quiz 3.3 Solution (Continued 2)

If $p = 0.01$, the probability of exactly 2 errors is

$$P_Y(2) = \binom{100}{2}(0.01)^2(0.99)^{98} = 0.1849. \tag{6}$$

(d) The probability of no more than 2 errors is

$$\begin{aligned} \mathrm{P}[Y \leq 2] &= P_Y(0) + P_Y(1) + P_Y(2) \\ &= (0.99)^{100} + 100(0.01)(0.99)^{99} + \binom{100}{2}(0.01)^2(0.99)^{98} \\ &= 0.9207. \end{aligned} \tag{7}$$

(e) Random variable Z is the number of trials up to and including the third success. Thus Z has the Pascal PMF (see Example 3.10)

$$P_Z(z) = \binom{z-1}{2} p^3(1-p)^{z-3}. \tag{8}$$

Note that $P_Z(z) > 0$ for $z = 3, 4, 5, \ldots$.

(f) If $p = 0.25$, the probability that the third error occurs on bit 12 is

$$P_Z(12) = \binom{11}{2}(0.25)^3(0.75)^9 = 0.0645. \tag{9}$$

Cumulative Distribution Function (CDF)

Definition 3.10 Cumulative Distribution Function (CDF)

The cumulative distribution function *(CDF) of random variable X is*

$$F_X(x) = \mathsf{P}\left[X \leq x\right].$$

Theorem 3.2

For any discrete random variable X with range $S_X = \{x_1, x_2, \ldots\}$ satisfying $x_1 \leq x_2 \leq \ldots$,

(a) $F_X(-\infty) = 0$ and $F_X(\infty) = 1$.

(b) For all $x' \geq x$, $F_X(x') \geq F_X(x)$.

(c) For $x_i \in S_X$ and ϵ, an arbitrarily small positive number,

$$F_X(x_i) - F_X(x_i - \epsilon) = P_X(x_i).$$

(d) $F_X(x) = F_X(x_i)$ for all x such that $x_i \leq x < x_{i+1}$.

3.4 Comment: Theorem 3.2

Each property of Theorem 3.2 has an equivalent statement in words:

(a) Going from left to right on the x-axis, $F_X(x)$ starts at zero and ends at one.

(b) The CDF never decreases as it goes from left to right.

(c) For a discrete random variable X, there is a jump (discontinuity) at each value of $x_i \in S_X$. The height of the jump at x_i is $P_X(x_i)$.

(d) Between jumps, the graph of the CDF of the discrete random variable X is a horizontal line.

Theorem 3.3

For all $b \geq a$,

$$F_X(b) - F_X(a) = \mathsf{P}\left[a < X \leq b\right].$$

Proof: Theorem 3.3

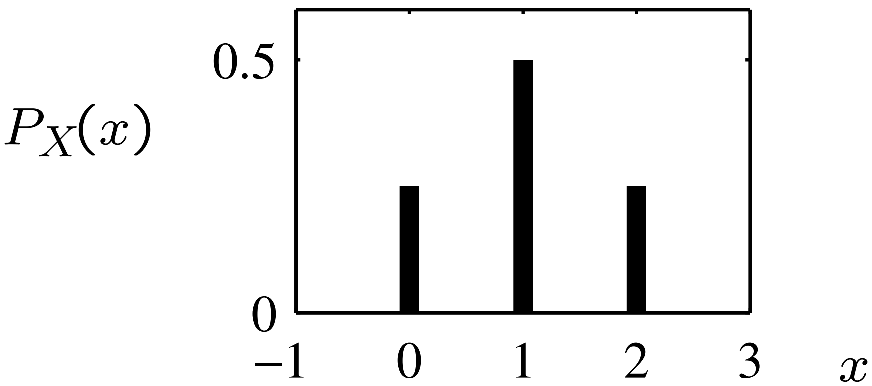
To prove this theorem, express the event $E_{ab} = \{a < X \leq b\}$ as a part of a union of mutually exclusive events. Start with the event $E_b = \{X \leq b\}$. Note that E_b can be written as the union

$$E_b = \{X \leq b\} = \{X \leq a\} \cup \{a < X \leq b\} = E_a \cup E_{ab} \tag{3.23}$$

Note also that E_a and E_{ab} are mutually exclusive so that $\mathrm{P}[E_b] = \mathrm{P}[E_a] + \mathrm{P}[E_{ab}]$. Since $\mathrm{P}[E_b] = F_X(b)$ and $\mathrm{P}[E_a] = F_X(a)$, we can write $F_X(b) = F_X(a) + \mathrm{P}[a < X \leq b]$. Therefore, $\mathrm{P}[a < X \leq b] = F_X(b) - F_X(a)$.

Example 3.15 Problem

In Example 3.5, random variable X has PMF

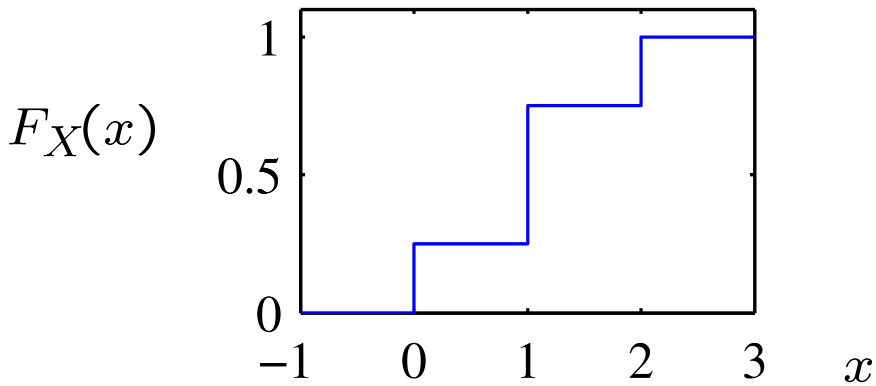


$$P_X(x) = \begin{cases} 1/4 & x = 0, \\ 1/2 & x = 1, \\ 1/4 & x = 2, \\ 0 & \text{otherwise.} \end{cases} \tag{3.24}$$

Find and sketch the CDF of random variable X.

Example 3.15 Solution

Referring to the PMF $P_X(x)$, we derive the CDF of random variable X:



$$F_X(x) = \mathrm{P}[X \leq x] = \begin{cases} 0 & x < 0, \\ 1/4 & 0 \leq x < 1, \\ 3/4 & 1 \leq x < 2 \\ 1 & x \geq 2. \end{cases}$$

Keep in mind that at the discontinuities $x = 0$, $x = 1$ and $x = 2$, the values of $F_X(x)$ are the upper values: $F_X(0) = 1/4$, $F_X(1) = 3/4$ and $F_X(2) = 1$. Math texts call this the *right hand limit* of $F_X(x)$.

Example 3.16 Problem

In Example 3.8, let the probability that a circuit is rejected equal $p = 1/4$. The PMF of Y, the number of tests up to and including the first reject, is the geometric (1/4) random variable with PMF

$$P_Y(y) = \begin{cases} (1/4)(3/4)^{y-1} & y = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases} \tag{3.25}$$

What is the CDF of Y?

Example 3.16 Solution

Random variable Y has nonzero probabilities for all positive integers. For any integer $n \geq 1$, the CDF is

$$F_Y(n) = \sum_{j=1}^{n} P_Y(j) = \sum_{j=1}^{n} \frac{1}{4}\left(\frac{3}{4}\right)^{j-1}. \tag{3.26}$$

Equation (3.26) is a geometric series. Familiarity with the geometric series is essential for calculating probabilities involving geometric random variables. Appendix B summarizes the most important facts. In particular, Math Fact B.4 implies $(1-x)\sum_{j=1}^{n} x^{j-1} = 1 - x^n$. Substituting $x = 3/4$, we obtain

$$F_Y(n) = 1 - \left(\frac{3}{4}\right)^{n}. \tag{3.27}$$

The complete expression for the CDF of Y must show $F_Y(y)$ for all integer *and noninteger* values of y.

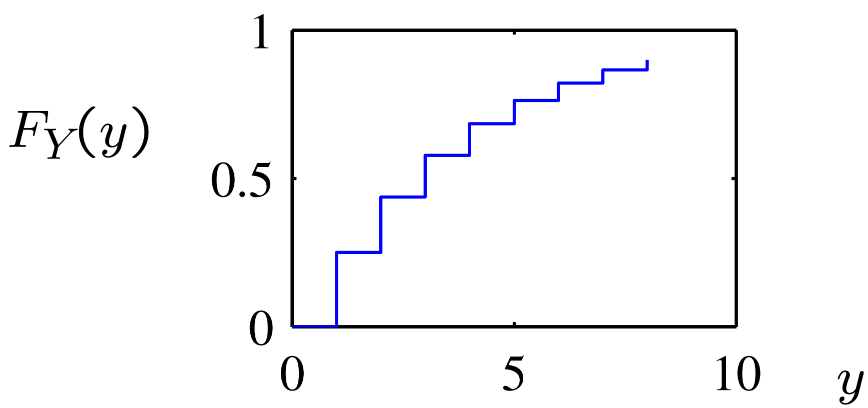
[Continued]

Example 3.16 Solution (Continued 2)

For an integer-valued random variable Y, we can do this in a simple way using the *floor function* $\lfloor y \rfloor$, which is the largest integer less than or equal to y. In particular, if $n \le y < n-1$ for some integer n, then $\lfloor y \rfloor = n$ and

$$F_Y(y) = \mathrm{P}[Y \le y] = \mathrm{P}[Y \le n] = F_Y(n) = F_Y(\lfloor y \rfloor). \tag{3.28}$$

In terms of the floor function, we can express the CDF of Y as



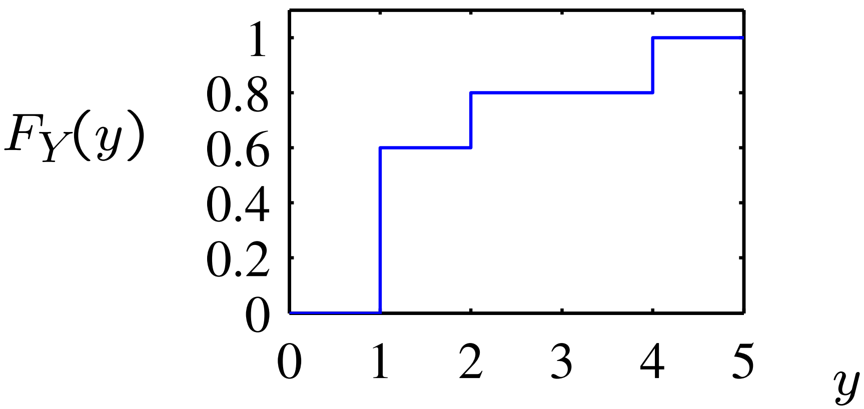
$$F_Y(y) = \begin{cases} 0 & y < 1, \\ 1 - (3/4)^{\lfloor y \rfloor} & y \ge 1. \end{cases} \tag{3.29}$$

To find the probability that Y takes a value in the set $\{4, 5, 6, 7, 8\}$, we refer to Theorem 3.3 and compute

$$\mathrm{P}[3 < Y \le 8] = F_Y(8) - F_Y(3) = (3/4)^3 - (3/4)^8 = 0.322. \tag{3.30}$$

Quiz 3.4

Use the CDF $F_Y(y)$ to find the following probabilities:



(a) $\mathrm{P}[Y < 1]$

(b) $\mathrm{P}[Y \leq 1]$

(c) $\mathrm{P}[Y > 2]$

(d) $\mathrm{P}[Y \geq 2]$

(e) $\mathrm{P}[Y = 1]$

(f) $\mathrm{P}[Y = 3]$

Quiz 3.4 Solution

Each of these probabilities can be read from the graph of the CDF $F_Y(y)$. However, we must keep in mind that when $F_Y(y)$ has a discontinuity at y_0, $F_Y(y)$ takes the upper value $F_Y(y_0^+)$.

(a) $\mathrm{P}[Y < 1] = F_Y(1^-) = 0$.

(b) $\mathrm{P}[Y \leq 1] = F_Y(1) = 0.6$.

(c) $\mathrm{P}[Y > 2] = 1 - \mathrm{P}[Y \leq 2] = 1 - F_Y(2) = 1 - 0.8 = 0.2$.

(d) $\mathrm{P}[Y \geq 2] = 1 - \mathrm{P}[Y < 2] = 1 - F_Y(2^-) = 1 - 0.6 = 0.4$.

(e) $\mathrm{P}[Y = 1] = \mathrm{P}[Y \leq 1] - \mathrm{P}[Y < 1] = F_Y(1^+) - F_Y(1^-) = 0.6$.

(f) $\mathrm{P}[Y = 3] = \mathrm{P}[Y \leq 3] - \mathrm{P}[Y < 3] = F_Y(3^+) - F_Y(3^-) = 0.8 - 0.8 = 0$.

Section 3.5

Averages and Expected Value

Example 3.17 Problem

For one quiz, 10 students have the following grades (on a scale of 0 to 10):

$$9, 5, 10, 8, 4, 7, 5, 5, 8, 7 \tag{3.31}$$

Find the mean, the median, and the mode.

Example 3.17 Solution

The sum of the ten grades is 68. The mean value is $68/10 = 6.8$. The median is 7, because there are four grades below 7 and four grades above 7. The mode is 5, because three students have a grade of 5, more than the number of students who received any other grade.

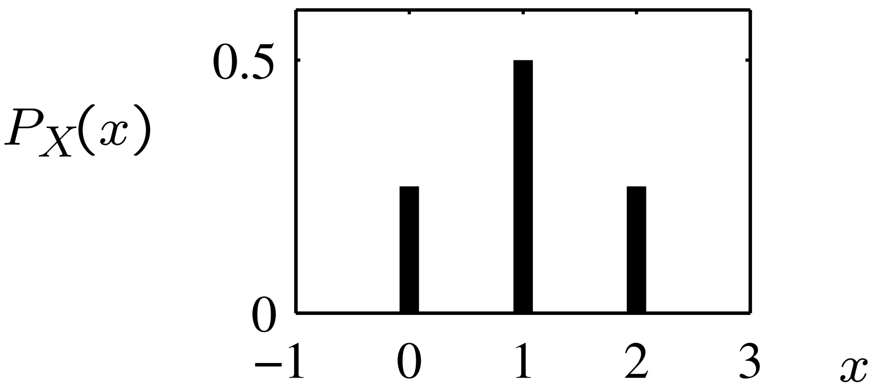
Definition 3.11 Expected Value

The expected value *of* X *is*

$$\mathsf{E}[X] = \mu_X = \sum_{x \in S_X} x P_X(x).$$

Example 3.18 Problem

Random variable X in Example 3.5 has PMF



$$P_X(x) = \begin{cases} 1/4 & x = 0, \\ 1/2 & x = 1, \\ 1/4 & x = 2, \\ 0 & \text{otherwise.} \end{cases} \tag{3.32}$$

What is E[X]?

Example 3.18 Solution

$$\begin{aligned}
\mathsf{E}[X] = \mu_X &= 0 \cdot P_X(0) + 1 \cdot P_X(1) + 2 \cdot P_X(2) \\
&= 0(1/4) + 1(1/2) + 2(1/4) = 1.
\end{aligned} \tag{3.33}$$

Theorem 3.4

The Bernoulli (p) random variable X has expected value $\mathsf{E}[X] = p$.

Proof: Theorem 3.4

$\mathsf{E}[X] = 0 \cdot P_X(0) + 1 P_X(1) = 0(1-p) + 1(p) = p.$

Theorem 3.5

The geometric (p) random variable X has expected value $\mathsf{E}[X] = 1/p$.

Proof: Theorem 3.5

Let $q = 1 - p$. The PMF of X becomes

$$P_X(x) = \begin{cases} pq^{x-1} & x = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases} \tag{3.39}$$

The expected value $\mathrm{E}[X]$ is the infinite sum

$$\mathrm{E}[X] = \sum_{x=1}^{\infty} x P_X(x) = \sum_{x=1}^{\infty} xpq^{x-1}. \tag{3.40}$$

Applying the identity of Math Fact B.7, we have

$$\mathrm{E}[X] = p \sum_{x=1}^{\infty} xq^{x-1} = \frac{p}{q} \sum_{x=1}^{\infty} xq^{x} = \frac{p}{q} \frac{q}{1-q^2} = \frac{p}{p^2} = \frac{1}{p}. \tag{3.41}$$

Theorem 3.6

The Poisson (α) random variable in Definition 3.9 has expected value $\mathrm{E}[X] = \alpha$.

Proof: Theorem 3.6

$$\mathrm{E}[X] = \sum_{x=0}^{\infty} x P_X(x) = \sum_{x=0}^{\infty} x \frac{\alpha^x}{x!} e^{-\alpha}. \tag{3.42}$$

We observe that $x/x! = 1/(x-1)!$ and also that the $x = 0$ term in the sum is zero. In addition, we substitute $\alpha^x = \alpha \cdot \alpha^{x-1}$ to factor α from the sum to obtain

$$\mathrm{E}[X] = \alpha \sum_{x=1}^{\infty} \frac{\alpha^{x-1}}{(x-1)!} e^{-\alpha}. \tag{3.43}$$

Next we substitute $l = x - 1$, with the result

$$\mathrm{E}[X] = \alpha \underbrace{\sum_{l=0}^{\infty} \frac{\alpha^l}{l!} e^{-\alpha}}_{1} = \alpha. \tag{3.44}$$

We can conclude that the sum in this formula equals 1 either by referring to the identity $e^\alpha = \sum_{l=0}^{\infty} \alpha^l / l!$ or by applying Theorem 3.1(b) to the fact that the sum is the sum of the PMF of a Poisson random variable L over all values in S_L and $\mathrm{P}[S_L] = 1$.

Theorem 3.7

(a) For the binomial (n, p) random variable X of Definition 3.6,

$$E[X] = np.$$

(b) For the Pascal (k, p) random variable X of Definition 3.7,

$$E[X] = k/p.$$

(c) For the discrete uniform (k, l) random variable X of Definition 3.8,

$$E[X] = (k + l)/2.$$

Theorem 3.8

Perform n Bernoulli trials. In each trial, let the probability of success be α/n, where $\alpha > 0$ is a constant and $n > \alpha$. Let the random variable K_n be the number of successes in the n trials. As $n \to \infty$, $P_{K_n}(k)$ converges to the PMF of a Poisson (α) random variable.

Proof: Theorem 3.8

We first note that K_n is the binomial $(n, \alpha n)$ random variable with PMF

$$P_{K_n}(k) = \binom{n}{k} (\alpha/n)^k \left(1 - \frac{\alpha}{n}\right)^{n-k}. \tag{3.45}$$

For $k = 0, \ldots, n$, we can write

$$P_K(k) = \frac{n(n-1)\cdots(n-k+1)}{n^k} \frac{\alpha^k}{k!} \left(1 - \frac{\alpha}{n}\right)^{n-k}. \tag{3.46}$$

Notice that in the first fraction, there are k terms in the numerator. The denominator is n^k, also a product of k terms, all equal to n. Therefore, we can express this fraction as the product of k fractions, each of the form $(n-j)/n$. As $n \to \infty$, each of these fractions approaches 1. Hence,

$$\lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{n^k} = 1. \tag{3.47}$$

Furthermore, we have

$$\left(1 - \frac{\alpha}{n}\right)^{n-k} = \frac{\left(1 - \frac{\alpha}{n}\right)^n}{\left(1 - \frac{\alpha}{n}\right)^k}. \tag{3.48}$$

[Continued]

Proof: Theorem 3.8 (Continued 2)

As n grows without bound, the denominator approaches 1 and, in the numerator, we recognize the identity $\lim_{n\to\infty}(1-\alpha/n)^n = e^{-\alpha}$. Putting these three limits together leads us to the result that for any integer $k \geq 0$,

$$\lim_{n\to\infty} P_{K_n}(k) = \begin{cases} \alpha^k e^{-\alpha}/k! & k = 0, 1, \ldots \\ 0 & \text{otherwise,} \end{cases} \tag{3.49}$$

which is the Poisson PMF.

Quiz 3.5

In a pay-as-you go cellphone plan, the cost of sending an SMS text message is 10 cents and the cost of receiving a text is 5 cents. For a certain subscriber, the probability of sending a text is 1/3 and the probability of receiving a text is 2/3. Let C equal the cost (in cents) of one text message and find

(a) The PMF $P_C(c)$

(b) The expected value $\text{E}[C]$

(c) The probability that four texts are received before a text is sent.

(d) The expected number of texts received before a text is sent.

Quiz 3.5 Solution

(a) With probability $1/3$, the subscriber sends a text and the cost is $C = 10$ cents. Otherwise, with probability $2/3$, the subscriber receives a text and the cost is $C = 5$ cents. This corresponds to the PMF

$$P_C(c) = \begin{cases} 2/3 & c = 5, \\ 1/3 & c = 10, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

(b) The expected value of C is

$$\mathrm{E}[C] = (2/3)(5) + (1/3)(10) = 6.67 \text{ cents.} \tag{2}$$

(c) For the next two parts we think of each text as a Bernoulli trial such that the trial is a "success" if the subscriber sends a text. The success probability is $p = 1/3$. Let R denote the number of texts received before sending a text. In terms of Bernoulli trials, R is the number of failures before the first success. R is similar to a geometric random variable except $R = 0$ is possible if the first text is sent rather than received. In general $R = r$ if the first r trials are failures (i.e. the first r texts are received) and trial $r + 1$ is a success. Thus R has PMF

$$P_R(r) = \begin{cases} (1-p)^r p & r = 0, 1, 2 \ldots \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

[Continued]

Quiz 3.5 Solution (Continued 2)

The probability of receiving four texts before sending a text is

$$P_R(4) = (1-p)^4 p. \tag{4}$$

(d) The expected number of texts received before sending a text is

$$\mathrm{E}[R] = \sum_{r=0}^{\infty} r P_R(r) = \sum_{r=0}^{\infty} r(1-p)^r p. \tag{5}$$

Letting $q = 1-p$ and observing that the $r = 0$ term in the sum is zero,

$$\mathrm{E}[R] = p \sum_{r=1}^{\infty} r q^r. \tag{6}$$

Using Math Fact B.7, we have

$$\mathrm{E}[R] = p \frac{q}{(1-q)^2} = \frac{1-p}{p} = 2. \tag{7}$$

Section 3.6

Functions of a Random Variable

Definition 3.12 Derived Random Variable

Each sample value y *of a* derived random variable Y *is a mathematical function* $g(x)$ *of a sample value* x *of another random variable* X*. We adopt the notation* $Y = g(X)$ *to describe the relationship of the two random variables.*

Example 3.19 Problem

A parcel shipping company offers a charging plan: \$1.00 for the first pound, \$0.90 for the second pound, etc., down to \$0.60 for the fifth pound, with rounding up for a fraction of a pound. For all packages between 6 and 10 pounds, the shipper will charge \$5.00 per package. (It will not accept shipments over 10 pounds.) Find a function $Y = g(X)$ for the charge in cents for sending one package.

Example 3.19 Solution

When the package weight is an integer $X \in \{1, 2, \ldots, 10\}$ that specifies the number of pounds with rounding up for a fraction of a pound, the function

$$Y = g(X) = \begin{cases} 105X - 5X^2 & X = 1, 2, 3, 4, 5 \\ 500 & X = 6, 7, 8, 9, 10. \end{cases} \tag{3.50}$$

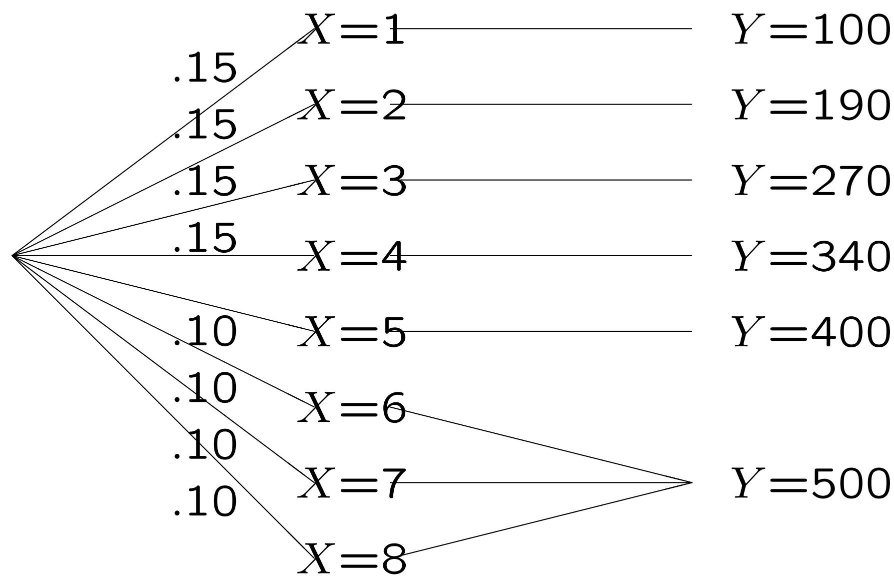
corresponds to the charging plan.

Theorem 3.9

For a discrete random variable X, the PMF of $Y = g(X)$ is

$$P_Y(y) = \sum_{x:g(x)=y} P_X(x) .$$

Figure 3.1



The derived random variable $Y = g(X)$ for Example 3.21.

Example 3.20 Problem

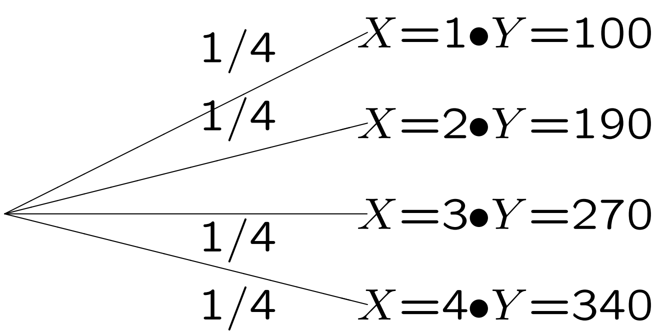
In Example 3.19, suppose all packages weigh 1, 2, 3, or 4 pounds with equal probability. Find the PMF and expected value of Y, the shipping charge for a package.

Example 3.20 Solution

From the problem statement, the weight X has PMF

$$P_X(x) = \begin{cases} 1/4 & x = 1, 2, 3, 4, \\ 0 & \text{otherwise.} \end{cases} \tag{3.52}$$

The charge for a shipment, Y, has range $S_Y = \{100, 190, 270, 340\}$ corresponding to $S_X = \{1, \ldots, 4\}$. The experiment can be described by the following tree. Here each value of Y derives from a unique value of X. Hence, we can use Equation (3.51) to find $P_Y(y)$.



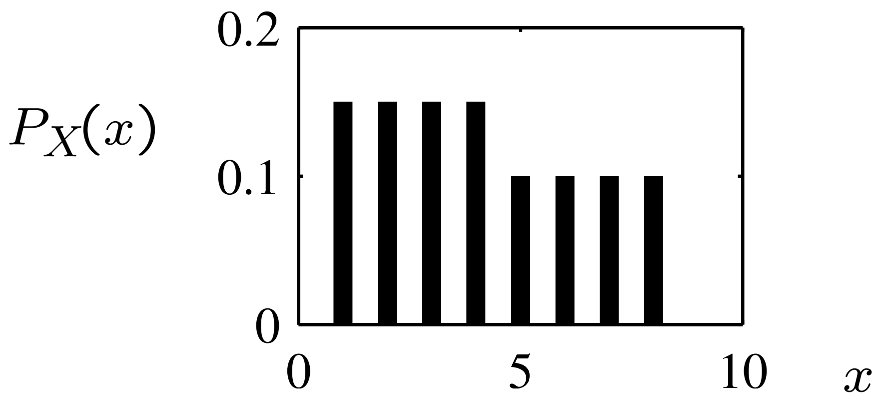
$$P_Y(y) = \begin{cases} 1/4 & y = 100, 190, 270, 340, \\ 0 & \text{otherwise.} \end{cases}$$

The expected shipping bill is

$$\begin{aligned} \mathrm{E}[Y] &= \frac{1}{4}(100 + 190 + 270 + 340) \\ &= 225 \text{ cents.} \end{aligned}$$

Example 3.21 Problem

Suppose the probability model for the weight in pounds X of a package in Example 3.19 is



$$P_X(x) = \begin{cases} 0.15 & x = 1,2,3,4, \\ 0.1 & x = 5,6,7,8, \\ 0 & \text{otherwise.} \end{cases}$$

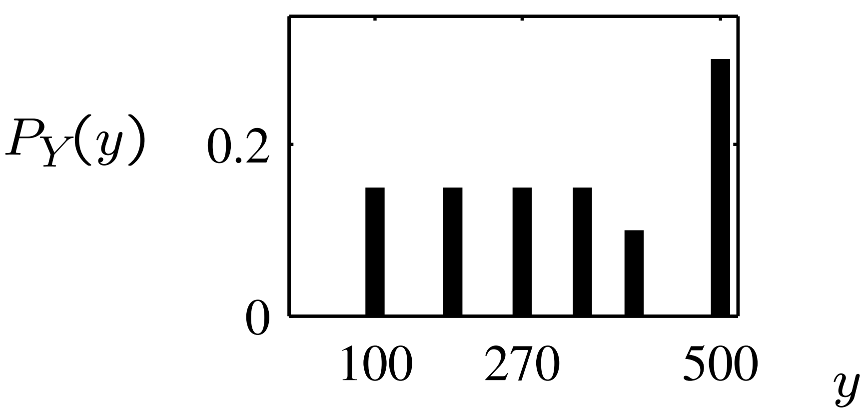
For the pricing plan given in Example 3.19, what is the PMF and expected value of Y, the cost of shipping a package?

Example 3.21 Solution

Now we have three values of X, specifically $(6, 7, 8)$, transformed by $g(\cdot)$ into $Y = 500$. For this situation we need the more general view of the PMF of Y, given by Theorem 3.9. In particular, $y_6 = 500$, and we have to add the probabilities of the outcomes $X = 6$, $X = 7$, and $X = 8$ to find $P_Y(500)$. That is,

$$P_Y(500) = P_X(6) + P_X(7) + P_X(8) = 0.30. \tag{3.53}$$

The steps in the procedure are illustrated in the diagram of Figure 3.1. Applying Theorem 3.9, we have



$$P_Y(y) = \begin{cases} 0.15 & y = 100, 190, 270, 340, \\ 0.10 & y = 400, \\ 0.30 & y = 500, \\ 0 & \text{otherwise.} \end{cases}$$

For this probability model, the expected cost of shipping a package is

$$\mathrm{E}[Y] = 0.15(100 + 190 + 270 + 340) + 0.10(400) + 0.30(500) = 325 \text{ cents}$$

Quiz 3.6

Monitor three customers purchasing smartphones at the Phonesmart store and observe whether each buys an Apricot phone for \$450 or a Banana phone for \$300. The random variable N is the number of customers purchasing an Apricot phone. Assume N has PMF

$$P_N(n) = \begin{cases} 0.4 & n = 0, \\ 0.2 & n = 1, 2, 3, \\ 0 & \text{otherwise.} \end{cases} \tag{3.54}$$

M dollars is the amount of money paid by three customers.

(a) Express M as a function of N.

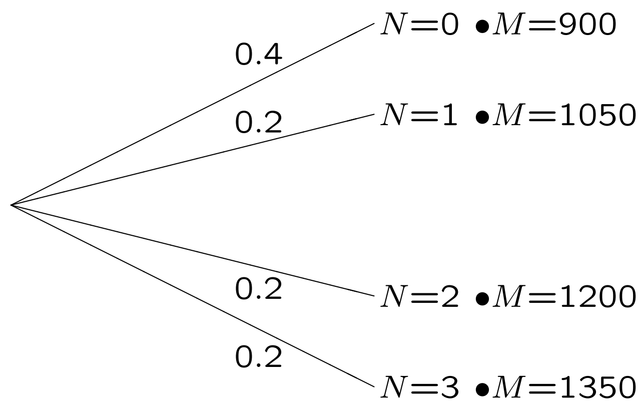
(b) Find $P_M(m)$ and $\mathsf{E}[M]$.

Quiz 3.6 Solution

(a) As a function of N, the money spent by the tree customers is

$$M = 450N + 300(3 - N) = 900 + 150N.$$

(b) To find the PMF of M, we can draw the following tree and map the outcomes to values of M:



From this tree,

$$P_M(m) = \begin{cases} 0.4 & m = 900, \\ 0.2 & m = 1050, 1200, 1350 \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

From the PMF $P_M(m)$, the expected value of M is

$$\begin{aligned} \mathrm{E}[M] &= 900P_M(900) + 1050P_M(1050) \\ &\quad + 1200P_M(1200) + 1350P_M(1350) \\ &= (900)(0.4) + (1050 + 1200 + 1350)(0.2) = 1080. \end{aligned} \tag{2}$$

Expected Value of a Derived Random Variable

Theorem 3.10

Given a random variable X with PMF $P_X(x)$ and the derived random variable $Y = g(X)$, the expected value of Y is

$$E[Y] = \mu_Y = \sum_{x \in S_X} g(x) P_X(x).$$

Proof: Theorem 3.10

From the definition of $\mathsf{E}[Y]$ and Theorem 3.9, we can write

$$\mathsf{E}\left[Y\right] = \sum_{y \in S_Y} y P_Y(y) = \sum_{y \in S_Y} y \sum_{x : g(x) = y} P_X(x) = \sum_{y \in S_Y} \sum_{x : g(x) = y} g(x) P_X(x)\,, \tag{3.55}$$

where the last double summation follows because $g(x) = y$ for each x in the inner sum. Since $g(x)$ transforms each possible outcome $x \in S_X$ to a value $y \in S_Y$, the preceding double summation can be written as a single sum over all possible values $x \in S_X$. That is,

$$\mathsf{E}\left[Y\right] = \sum_{x \in S_X} g(x) P_X(x)\,. \tag{3.56}$$

Example 3.22 Problem

In Example 3.20,

$$P_X(x) = \begin{cases} 1/4 & x = 1, 2, 3, 4, \\ 0 & \text{otherwise}, \end{cases}$$

$$g(X) = \begin{cases} 105X - 5X^2 & 1 \le X \le 5, \\ 500 & 6 \le X \le 10. \end{cases}$$

What is $\mathrm{E}[Y] = \mathrm{E}[g(X)]$?

Example 3.22 Solution

Applying Theorem 3.10 we have

$$\begin{aligned}
\mathrm{E}[Y] &= \sum_{x=1}^{4} P_X(x)\, g(x) \\
&= (1/4)[(105)(1) - (5)(1)^2] + (1/4)[(105)(2) - (5)(2)^2] \\
&\qquad + (1/4)[(105)(3) - (5)(3)^2] + (1/4)[(105)(4) - (5)(4)^2] \\
&= (1/4)[100 + 190 + 270 + 340] = 225 \text{ cents.} \qquad (3.58)
\end{aligned}$$

Theorem 3.11

For any random variable X,

$$\mathsf{E}\left[X - \mu_X\right] = 0.$$

Proof: Theorem 3.11

Defining $g(X) = X - \mu_X$ and applying Theorem 3.10 yields

$$\mathrm{E}\left[g(X)\right] = \sum_{x \in S_X} (x - \mu_X) P_X(x) = \sum_{x \in S_X} x P_X(x) - \mu_X \sum_{x \in S_X} P_X(x) . \quad (3.60)$$

The first term on the right side is μ_X by definition. In the second term, $\sum_{x \in S_X} P_X(x) = 1$, so both terms on the right side are μ_X and the difference is zero.

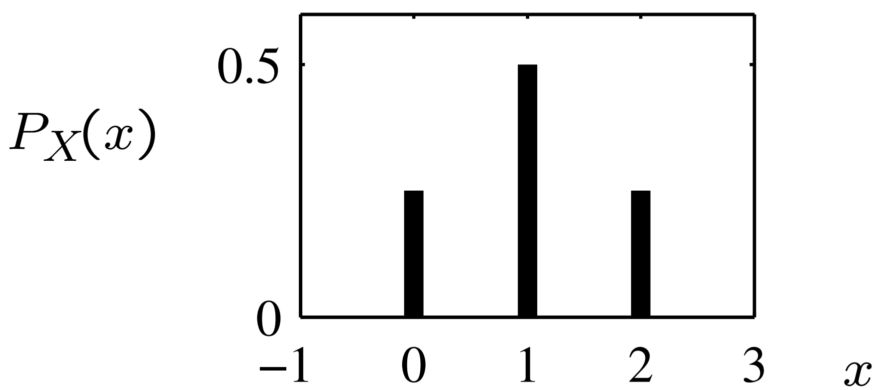
Theorem 3.12

For any random variable X,

$$\mathsf{E}[aX + b] = a\,\mathsf{E}[X] + b.$$

Example 3.23 Problem

Recall from Examples 3.5 and 3.18 that X has PMF



$$P_X(x) = \begin{cases} 1/4 & x = 0, \\ 1/2 & x = 1, \\ 1/4 & x = 2, \\ 0 & \text{otherwise.} \end{cases} \tag{3.61}$$

What is the expected value of $V = g(X) = 4X + 7$?

Example 3.23 Solution

From Theorem 3.12,

$$\mathrm{E}[V] = \mathrm{E}[g(X)] = \mathrm{E}[4X + 7] = 4\,\mathrm{E}[X] + 7 = 4(1) + 7 = 11. \qquad (3.62)$$

We can verify this result by applying Theorem 3.10:

$$\begin{aligned} \mathrm{E}[V] &= g(0)P_X(0) + g(1)P_X(1) + g(2)P_X(2) \\ &= 7(1/4) + 11(1/2) + 15(1/4) = 11. \end{aligned} \qquad (3.63)$$

Example 3.24 Problem

Continuing Example 3.23, let $W = h(X) = X^2$. What is $\mathrm{E}[W]$?

Example 3.24 Solution

Theorem 3.10 gives

$$\mathsf{E}[W] = \sum h(x) P_X(x) = (1/4)0^2 + (1/2)1^2 + (1/4)2^2 = 1.5. \qquad (3.64)$$

Note that this is not the same as $h(\mathsf{E}[W]) = (1)^2 = 1$.

Quiz 3.7

The number of memory chips M needed in a personal computer depends on how many application programs, A, the owner wants to run simultaneously. The number of chips M and the number of application programs A are described by

$$M = \begin{cases} 4 & \text{chips for 1 program,} \\ 4 & \text{chips for 2 programs,} \\ 6 & \text{chips for 3 programs,} \\ 8 & \text{chips for 4 programs,} \end{cases} \qquad P_A(a) = \begin{cases} 0.1(5-a) & a = 1,2,3,4, \\ 0 & \text{otherwise.} \end{cases} \tag{3.65}$$

(a) What is the expected number of programs $\mu_A = \mathrm{E}[A]$?

(b) Express M, the number of memory chips, as a function $M = g(A)$ of the number of application programs A.

(c) Find $\mathrm{E}[M] = \mathrm{E}[g(A)]$. Does $\mathrm{E}[M] = g(\mathrm{E}[A])$?

Quiz 3.7 Solution

(a) Using Definition 3.11, the expected number of applications is

$$E[A] = \sum_{a=1}^{4} a P_A(a)$$
$$= 1(0.4) + 2(0.3) + 3(0.2) + 4(0.1) = 2. \quad (1)$$

(b) The number of memory chips is

$$M = g(A) = \begin{cases} 4 & A = 1, 2, \\ 6 & A = 3, \\ 8 & A = 4. \end{cases} \quad (2)$$

(c) By Theorem 3.10, the expected number of memory chips is

$$E[M] = \sum_{a=1}^{4} g(A) P_A(a)$$
$$= 4(0.4) + 4(0.3) + 6(0.2) + 8(0.1) = 4.8. \quad (3)$$

Since $E[A] = 2$,

$$g(E[A]) = g(2) = 4.$$

However, $E[M] = 4.8 \neq g(E[A])$. The two quantities are different because $g(A)$ is not of the form $\alpha A + \beta$.

Section 3.8

Variance and Standard Deviation

Definition 3.13 Variance

The variance *of random variable* X *is*

$$\text{Var}[X] = \text{E}\left[(X - \mu_X)^2\right].$$

Definition 3.14 Standard Deviation

The standard deviation *of random variable X is*

$$\sigma_X = \sqrt{\mathrm{Var}[X]}.$$

Theorem 3.13

In the absence of observations, the minimum mean square error estimate of random variable X is

$$\hat{x} = \mathsf{E}[X].$$

Proof: Theorem 3.13

After substituting $\widehat{X} = \widehat{x}$, we expand the square in Equation (3.66) to write

$$e = \mathrm{E}\left[X^2\right] - 2\widehat{x}\,\mathrm{E}\left[X\right] + \widehat{x}^2. \tag{3.67}$$

To minimize e, we solve

$$\frac{de}{d\widehat{x}} = -2\,\mathrm{E}\left[X\right] + 2\widehat{x} = 0, \tag{3.68}$$

yielding $\widehat{x} = \mathrm{E}[X]$.

Theorem 3.14

$$\mathsf{Var}\left[X\right] = \mathsf{E}\left[X^2\right] - \mu_X^2 = \mathsf{E}\left[X^2\right] - \left(\mathsf{E}\left[X\right]\right)^2.$$

Proof: Theorem 3.14

Expanding the square in (3.70), we have

$$\begin{aligned}
\text{Var}[X] &= \sum_{x\in S_X} x^2 P_X(x) - \sum_{x\in S_X} 2\mu_X x P_X(x) + \sum_{x\in S_X} \mu_X^2 P_X(x) \\
&= \text{E}[X^2] - 2\mu_X \sum_{x\in S_X} x P_X(x) + \mu_X^2 \sum_{x\in S_X} P_X(x) \\
&= \text{E}[X^2] - 2\mu_X^2 + \mu_X^2. \qquad (3.71)
\end{aligned}$$

Definition 3.15 Moments

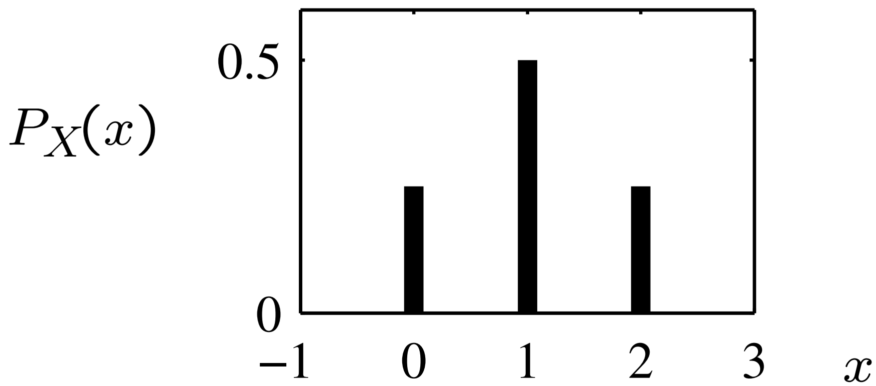
For random variable X:

(a) The nth moment *is* $\mathsf{E}[X^n]$.

(b) The nth central moment *is* $\mathsf{E}[(X - \mu_X)^n]$.

Example 3.25 Problem

Continuing Examples 3.5, 3.18, and 3.23, we recall that X has PMF



$$P_X(x) = \begin{cases} 1/4 & x = 0, \\ 1/2 & x = 1, \\ 1/4 & x = 2, \\ 0 & \text{otherwise,} \end{cases} \tag{3.72}$$

and expected value $E[X] = 1$. What is the variance of X?

Example 3.25 Solution

In order of increasing simplicity, we present three ways to compute Var$[X]$.

- From Definition 3.13, define

$$W = (X - \mu_X)^2 = (X-1)^2. \tag{3.73}$$

  We observe that $W = 0$ if and only if $X = 1$; otherwise, if $X = 0$ or $X = 2$, then $W = 1$. Thus $\text{P}[W = 0] = P_X(1) = 1/2$ and $\text{P}[W = 1] = P_X(0) + P_X(2) = 1/2$. The PMF of W is

$$P_W(w) = \begin{cases} 1/2 & w = 0, 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.74}$$

  Then

$$\text{Var}[X] = \text{E}[W] = (1/2)(0) + (1/2)(1) = 1/2. \tag{3.75}$$

- Recall that Theorem 3.10 produces the same result without requiring the derivation of $P_W(w)$.

$$\begin{aligned} \text{Var}[X] &= \text{E}\left[(X - \mu_X)^2\right] \\ &= (0-1)^2 P_X(0) + (1-1)^2 P_X(1) + (2-1)^2 P_X(2) \\ &= 1/2. \end{aligned} \tag{3.76}$$

- To apply Theorem 3.14, we find that

$$\text{E}\left[X^2\right] = 0^2 P_X(0) + 1^2 P_X(1) + 2^2 P_X(2) = 1.5. \tag{3.77}$$

  Thus Theorem 3.14 yields

$$\text{Var}[X] = \text{E}\left[X^2\right] - \mu_X^2 = 1.5 - 1^2 = 1/2. \tag{3.78}$$

Theorem 3.15

$$\mathsf{Var}\left[aX + b\right] = a^2 \, \mathsf{Var}\left[X\right].$$

Proof: Theorem 3.15

We let $Y = aX + b$ and apply Theorem 3.14. We first expand the second moment to obtain

$$\mathsf{E}\left[Y^2\right] = \mathsf{E}\left[a^2X^2 + 2abX + b^2\right] = a^2\,\mathsf{E}\left[X^2\right] + 2ab\mu_X + b^2. \tag{3.80}$$

Expanding the right side of Theorem 3.12 yields

$$\mu_Y^2 = a^2\mu_X^2 + 2ab\mu_x + b^2. \tag{3.81}$$

Because $\mathrm{Var}[Y] = \mathsf{E}[Y^2] - \mu_Y^2$, Equations (3.80) and (3.81) imply that

$$\mathrm{Var}\left[Y\right] = a^2\,\mathsf{E}\left[X^2\right] - a^2\mu_X^2 = a^2(\mathsf{E}\left[X^2\right] - \mu_X^2) = a^2\,\mathrm{Var}\left[X\right]. \tag{3.82}$$

Example 3.26 Problem

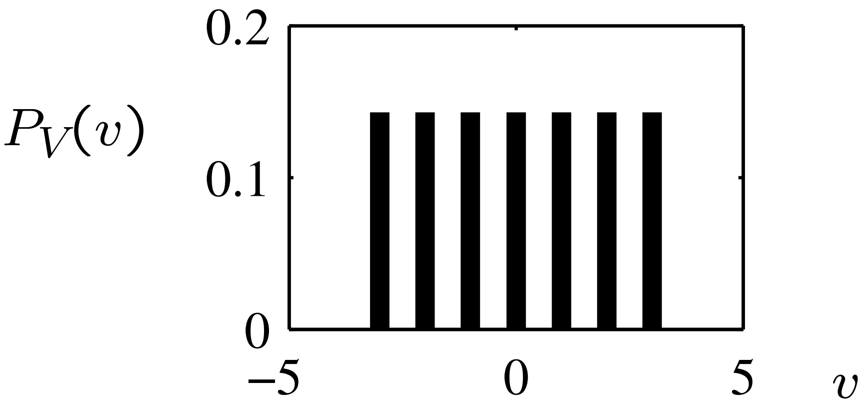
A printer automatically prints an initial cover page that precedes the regular printing of an X page document. Using this printer, the number of printed pages is $Y = X + 1$. Express the expected value and variance of Y as functions of $\mathrm{E}[X]$ and $\mathrm{Var}[X]$.

Example 3.26 Solution

The expected number of transmitted pages is $\mathsf{E}[Y] = \mathsf{E}[X] + 1$. The variance of the number of pages sent is $\mathsf{Var}[Y] = \mathsf{Var}[X]$.

Example 3.27 Problem

The amplitude V (volts) of a sinusoidal signal is a random variable with PMF



$$P_V(v) = \begin{cases} 1/7 & v = -3, -2, \ldots, 3, \\ 0 & \text{otherwise.} \end{cases}$$

A new voltmeter records the amplitude U in millivolts. Find the variance and standard deviation of U.

Example 3.27 Solution

Note that $U = 1000V$. To use Theorem 3.15, we first find the variance of V. The expected value of the amplitude is

$$\mu_V = 1/7[-3 + (-2) + (-1) + 0 + 1 + 2 + 3] = 0 \text{ volts.} \tag{3.83}$$

The second moment is

$$\begin{aligned} \mathrm{E}\left[V^2\right] &= 1/7[(-3)^2 + (-2)^2 + (-1)^2 + 0^2 + 1^2 + 2^2 + 3^2] \\ &= 4 \text{ volts}^2. \end{aligned} \tag{3.84}$$

Therefore the variance is $\mathrm{Var}[V] = \mathrm{E}[V^2] - \mu_V^2 = 4$ volts^2. By Theorem 3.15,

$$\mathrm{Var}\,[U] = 1000^2\,\mathrm{Var}[V] = 4{,}000{,}000 \text{ millivolts}^2, \tag{3.85}$$

and thus $\sigma_U = 2000$ millivolts.

Theorem 3.16

(a) If X is Bernoulli (p), then

$$\text{Var}[X] = p(1-p).$$

(b) If X is geometric (p), then

$$\text{Var}[X] = (1-p)/p^2.$$

(c) If X is binomial (n, p), then

$$\text{Var}[X] = np(1-p).$$

(d) If X is Pascal (k, p), then

$$\text{Var}[X] = k(1-p)/p^2.$$

(e) If X is Poisson (α), then

$$\text{Var}[X] = \alpha.$$

(f) If X is discrete uniform (k, l),

$$\text{Var}[X] = (l-k)(l-k+2)/12.$$

Quiz 3.8

In an experiment with three customers entering the Phonesmart store, the observation is N, the number of phones purchased. The PMF of N is

$$P_N(n) = \begin{cases} (4-n)/10 & n = 0,1,2,3 \\ 0 & \text{otherwise.} \end{cases} \tag{3.86}$$

Find

(a) The expected value $\mathrm{E}[N]$

(b) The second moment $\mathrm{E}[N^2]$

(c) The variance $\mathrm{Var}[N]$

(d) The standard deviation σ_N

Quiz 3.8 Solution

For this problem, it is helpful to wrote out the PMF of N in the table

n	0	1	2	3
$P_N(n)$	0.4	0.3	0.2	0.1

The PMF $P_N(n)$ allows us to calculate each of the desired quantities.

(a) The expected value is

$$\begin{aligned} \mathsf{E}[N] &= \sum_{n=0}^{3} n P_N(n) \\ &= 0(0.4) + 1(0.3) + 2(0.2) + 3(0.1) = 1. \end{aligned} \tag{1}$$

(b) The second moment of N is

$$\begin{aligned} \mathsf{E}\left[N^2\right] &= \sum_{n=0}^{3} n^2 P_N(n) \\ &= 0^2(0.4) + 1^2(0.3) + 2^2(0.2) + 3^2(0.1) = 2. \end{aligned} \tag{2}$$

(c) The variance of N is

$$\mathsf{Var}[N] = \mathsf{E}\left[N^2\right] - (\mathsf{E}[N])^2 = 2 - 1^2 = 1. \tag{3}$$

(d) The standard deviation is $\sigma_N = \sqrt{\mathsf{Var}[N]} = 1$.

Matlab

Matlab: Finite Random Variables

- In Matlab, we represent the range S_X by the column vector $\mathbf{s} = \begin{bmatrix} s_1 & \cdots & s_n \end{bmatrix}'$ and the corresponding probabilities by the vector $\mathbf{p} = \begin{bmatrix} p_1 & \cdots & p_n \end{bmatrix}'$.*

- The function `y=finitepmf(sx,px,x)` generates the probabilities of the elements of the m-dimensional vector $\mathbf{x} = \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}'$.

- The output is $\mathbf{y} = \begin{bmatrix} y_1 & \cdots & y_m \end{bmatrix}'$ where $y_i = P_X(x_i)$.

- That is, for each requested x_i, `finitepmf` returns the value $P_X(x_i)$.

- If x_i is not in the sample space of X, $y_i = 0$.

*Although column vectors are supposed to appear as columns, we generally write a column vector $\mathbf{x}$ in the form of a transposed row vector $\begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}'$ to save space.

Example 3.28 Problem

In Example 3.21, the random variable X, the weight of a package, has PMF

$$P_X(x) = \begin{cases} 0.15 & x = 1,2,3,4, \\ 0.1 & x = 5,6,7,8, \\ 0 & \text{otherwise.} \end{cases} \tag{3.87}$$

Write a Matlab function that calculates $P_X(x)$. Calculate the probability of an x_i pound package for $x_1 = 2$, $x_2 = 2.5$, and $x_3 = 6$.

Example 3.28 Solution

The Matlab function `shipweightpmf(x)` implements $P_X(x)$. We can then use `shipweightpmf` to calculate the desired probabilities:

```
function y=shipweightpmf(x)
s=(1:8)';
p=[0.15*ones(4,1); 0.1*ones(4,1)];
y=finitepmf(s,p,x);
```

```
>> shipweightpmf([2 2.5 6])'
ans =
   0.1500   0   0.1000
```

Example 3.29 Problem

Write a Matlab function `geometricpmf(p,x)` to calculate, for the sample values in vector `x`, $P_X(x)$ for a geometric (p) random variable.

Example 3.29 Solution

```
function pmf=geometricpmf(p,x)
%geometric(p) rv X
%out: pmf(i)=Prob[X=x(i)]
x=x(:);
pmf= p*((1-p).^(x-1));
pmf= (x>0).*(x==floor(x)).*pmf;
```

In `geometricpmf.m`, the last line ensures that values $x_i \notin S_X$ are assigned zero probability. Because `x=x(:)` reshapes `x` to be a column vector, the output `pmf` is always a column vector.

Example 3.30 Problem

Write a Matlab function that calculates the Poisson (α) PMF.

Example 3.30 Solution

For an integer x, we could calculate $P_X(x)$ by the direct calculation

```
px= ((alpha^x)*exp(-alpha*x))/factorial(x)
```

This will yield the right answer as long as the argument x for the factorial function is not too large. In Matlab version 6, `factorial(171)` causes an overflow. In addition, for $a > 1$, calculating the ratio $a^x/x!$ for large x can cause numerical problems because both a^x and $x!$ will be very large numbers, possibly with a small quotient. Another shortcoming of the direct calculation is apparent if you want to calculate $P_X(x)$ for the set of possible values $\mathbf{x} = [0, 1, \ldots, n]$. Calculating factorials is a lot of work for a computer and the direct approach fails to exploit the fact that if we have already calculated $(x-1)!$, we can easily compute $x! = x \cdot (x-1)!$. [Continued]

Example 3.30 Solution (Continued 2)

A more efficient calculation makes use of the observation

$$P_X(x) = \frac{a^x e^{-a}}{x!} = \frac{a}{x} P_X(x-1)\,. \tag{3.88}$$

The `poissonpmf.m` function uses Equation (3.88) to calculate $P_X(x)$. Even this code is not perfect because Matlab has limited range.

```
function pmf=poissonpmf(alpha,x)
%output: pmf(i)=P[X=x(i)]
x=x(:); k=(1:max(x))';
ip=[1;((alpha*ones(size(k)))./k)];
pb=exp(-alpha)*cumprod(ip);
  %pb= [P(X=0)...P(X=n)]
pmf=pb(x+1); %pb(1)=P[X=0]
pmf=(x>=0).*(x==floor(x)).*pmf;
  %pmf(i)=0 for zero-prob x(i)
```

In Matlab, `exp(-alpha)` returns zero for `alpha` $>$ 745.13. For these large values of `alpha`,

`poissonpmf(alpha,x)`

returns zero for all `x`. Problem 3.9.9 outlines a solution that is used in the version of `poissonpmf.m` on the companion website.

Example 3.31 Problem

Write a Matlab function that calculates the CDF of a Poisson random variable.

Example 3.31 Solution

```
function cdf=poissoncdf(alpha,x)
%output cdf(i)=Prob[X<=x(i)]
x=floor(x(:));
sx=0:max(x);
cdf=cumsum(poissonpmf(alpha,sx));
  %cdf from 0 to max(x)
okx=(x>=0);%x(i)<0 -> cdf=0
x=(okx.*x);%set negative x(i)=0
cdf= okx.*cdf(x+1);
  %cdf=0 for x(i)<0
```

Here we present the Matlab code for the Poisson CDF. Since the sample values of a Poisson random variable X are integers, we observe that $F_X(x) = F_X(\lfloor x \rfloor)$ where $\lfloor x \rfloor$, equivalent to the Matlab `floor(x)` function, denotes the largest integer less than or equal to x.

Example 3.32 Problem

In Example 3.13 a website has on average $\lambda = 2$ hits per second. What is the probability of no more than 130 hits in one minute? What is the probability of more than 110 hits in one minute?

Example 3.32 Solution

Let M equal the number of hits in one minute (60 seconds). Note that M is a Poisson (α) random variable with $\alpha = 2 \times 60 = 120$ hits. The PMF of M is

$$P_M(m) = \begin{cases} (120)^m e^{-120}/m! & m = 0, 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases} \tag{3.89}$$

```
>> poissoncdf(120,130)
ans =
    0.8315
>> 1-poissoncdf(120,110)
ans =
    0.8061
```

The Matlab solution shown on the left executes the following math calculations:

$$\mathrm{P}[M \leq 130] = \sum_{m=0}^{130} P_M(m), \tag{3.90}$$

$$\begin{aligned} \mathrm{P}[M > 110] &= 1 - \mathrm{P}[M \leq 110] \\ &= 1 - \sum_{m=0}^{110} P_M(m). \end{aligned} \tag{3.91}$$

Example 3.33 Problem

Write a function that generates m samples of a binomial (n, p) random variable.

Example 3.33 Solution

```
function x=binomialrv(n,p,m)
% m binomial(n,p) samples

r=rand(m,1);
cdf=binomialcdf(n,p,0:n);
x=count(cdf,r);
```

For vectors `x` and `y`, `c=count(x,y)` returns a vector `c` such that `c(i)` is the number of elements of `x` that are less than or equal to $y(i)$. In terms of our earlier pseudocode, $k^* =$ `count(cdf,r)`.

If `count(cdf,r)` $= 0$, then $r \leq P_X(0)$ and $k^* = 0$.

Example 3.34 Problem

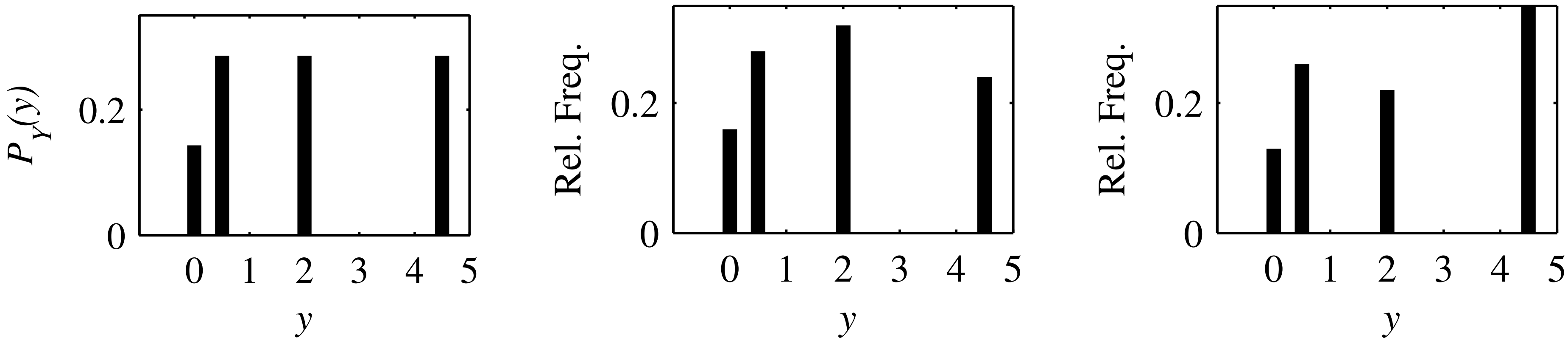
The amplitude V (volts) of a sinusoidal signal is a discrete uniform $(-3, 3)$ random variable. Let $Y = V^2/2$ watts denote the power of the transmitted signal. Simulate $n = 100$ trials of the experiment producing the power measurement Y. Plot the relative frequency of each $y \in S_Y$.

Example 3.34 Solution

```
function voltpower(n)
v=duniformrv(-3,3,n);
y=(v.^2)/2;
yrange=0:max(y);
yfreq=(hist(y,yrange)/n)';
pmfplot(yrange,yfreq);
```

In `voltpower.m`, we calculate $Y = V^2/2$ for each of n samples of the voltage V. The function `unique(y)` assembles a vector of the unique values of `y`. As in Example 2.20, the function `hist(y,yrange)` produces a vector with jth element equal to the number of occurrences of `yrange(j)` in the vector `y`. The function `pmfplot.m` is a utility for producing PMF bar plots in the style of this text. Figure 3.2 shows the exact PMF $P_Y(y)$, which Problem 3.6.3 asks you to derive, along with the results of two runs of `voltpower(100)`.

Figure 3.2



PMF $P_Y(y)$ Sample Run 1 Sample Run 2

The PMF of Y and the relative frequencies found in two sample runs of `voltpower(100)`. Note that in each run, the relative frequencies are close to (but not exactly equal to) the corresponding PMF.

Matlab: **Derived Random Variables**

- For derived random variables, we exploit a feature of `finitepmf(sx,px,x)` that allows the elements of `sx` to be repeated.

- Essentially, we use (`sx`, `px`), or equivalently $(\mathbf{s}_X, \mathbf{p}_X)$, to represent a random variable X described by the following experimental procedure:

> <u>`Finite sample space`</u>
>
> `Roll an` n`-sided die such that side` i `has probability` p_i`.`
>
> `If side` j `appears, set` $X = x_j$`.`

A consequence of this approach is that if $x_2 = 3$ and $x_5 = 3$, then the probability of observing $X = 3$ is $P_X(3) = p_2 + p_5$.

Example 3.35

```
>> sx=[1 3 5 7 3];
>> px=[0.1 0.2 0.2 0.3 0.2];
>> pmfx=finitepmf(sx,px,1:7);
>> pmfx'
ans =
   0.10 0 0.40 0 0.20 0 0.30
```

`finitepmf()` accounts for multiple occurrences of a sample value. In the example on the left,

`pmfx(3)=px(2)+px(5)=0.4`.

Example 3.36 Problem

Recall that in Example 3.21 the weight in pounds X of a package and the cost $Y = g(X)$ of shipping a package were described by

$$P_X(x) = \begin{cases} 0.15 & x = 1, 2, 3, 4, \\ 0.1 & x = 5, 6, 7, 8, \\ 0 & \text{otherwise}, \end{cases} \qquad Y = \begin{cases} 105X - 5X^2 & 1 \leq X \leq 5, \\ 500 & 6 \leq X \leq 10. \end{cases}$$

Write a function `y=shipcostrv(m)` that outputs m samples of the shipping cost Y.

Example 3.36 Solution

```
function y=shipcostrv(m)
sx=(1:8)';
px=[0.15*ones(4,1); ...
        0.1*ones(4,1)];
gx=(sx<=5).* ...
  (105*sx-5*(sx.^2))...
   + ((sx>5).*500);
y=finiterv(gx,px,m);
```

The vector `gx` is the mapping $g(x)$ for each $x \in S_X$. In `gx`, the element 500 appears three times, corresponding to $x = 6$, $x = 7$, and $x = 8$. The function `y=finiterv(gx,px,m))` produces m samples of the shipping cost Y.

```
>> shipcostrv(9)'
ans =
  270 190 500 270 500 190 190 100 500
```

Quiz 3.9

In Section 3.5, it was argued that the average

$$m_n = \frac{1}{n} \sum_{i=1}^{n} x(i) \qquad (3.94)$$

of samples $x(1), x(2), \ldots, x(n)$ of a random variable X will converge to $\mathsf{E}[X]$ as n becomes large. For a discrete uniform $(0, 10)$ random variable X, use Matlab to examine this convergence.

(a) For 100 sample values of X, plot the sequence $m_1, m_2, \ldots, m_{100}$. Repeat this experiment five times, plotting all five m_n curves on common axes.

(b) Repeat part (a) for 1000 sample values of X.

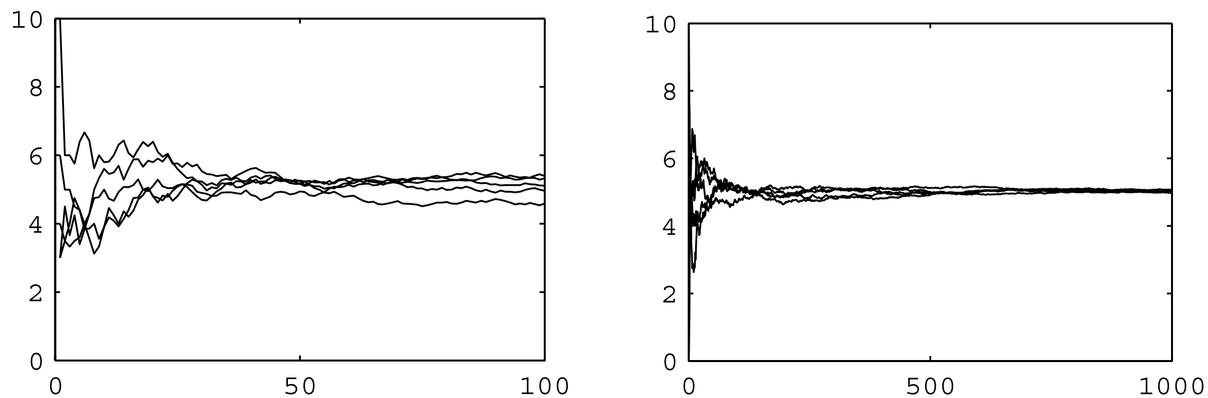
Quiz 3.9 Solution

The function `samplemean(k)` generates and plots five m_n sequences for $n = 1, 2, \ldots, k$. The ith column `M(:,i)` of `M` holds a sequence $m_1, m_2, \ldots, m_k$.

```
function M=samplemean(k);
K=(1:k)';
M=zeros(k,5);
for i=1:5,
    X=duniformrv(0,10,k);
    M(:,i)=cumsum(X)./K;
end;
plot(K,M);
```

Here are two examples of `samplemean`: [Continued]

Quiz 3.9 Solution (Continued 2)



(a) `samplemean(100)` **(b)** `samplemean(1000)`

Each time `samplemean(k)` is called produces a random output. What is observed in these figures is that for small n, m_n is fairly random but as n gets large, m_n gets close to $\mathrm{E}[X] = 5$. Although each sequence $m_1, m_2, \ldots$ that we generate is random, the sequences always converges to $\mathrm{E}[X]$. This random convergence is analyzed in Chapter 10.