Mathematical Statistics

Gregg Waterman
Oregon Institute of Technology

Contents

2 Probability Distributions

Performance Criteria:

2. Understand random variables and basic principles of discrete and continuous probability distributions.

    (a) Give the range of a random variable; determine the value of a random variable for a given outcome.

    (b) Describe an event in terms of a random variable.

    (c) Find the probability of an event that is described in terms of a random variable.

    (d) Give values of a discrete probability distribution function and a discrete cumulative discrete probability function associated with a random variable.

    (e) Find probabilities from either a discrete probability function or a discrete cumulative probability distribution.

    (f) Determine the value of a constant that makes a function defined on a discrete set a probability distribution function.

    (g) Given a discrete probability distribution function, find the associated cumulative distribution, or vice-versa.

    (h) Determine probabilities from a histogram.

    (i) Find probabilities from either a continuous probability density function or a continuous cumulative probability distribution.

    (j) Determine the value of a constant that makes a function defined on a continuous set a probability distribution function.

    (k) Given a continuous probability density function, find the associated cumulative distribution, or vice-versa.

    (l) Determine the mean, variance or standard deviation of a random variable.

We will now begin our study of random variables and probability functions, which are the basis for "modern" probability theory.

2.1 Random Variables

Performance Criteria:

2. (a) Give the range of a random variable; determine the value of a random variable for a given outcome.
   (b) Describe an event in terms of a random variable.
   (c) Find the probability of an event that is described in terms of a random variable.

For many experiments there is some logical way to assign a number to each outcome.

◇ **Example 2.1(a):** For the experiment of flipping a coin three times in a row we could assign to each outcome the number of heads obtained.

◇ **Example 2.1(b):** For the experiment of rolling a pair of dice, one red and one green, we could assign to each outcome the number showing on the red die.

◇ **Example 2.1(c):** For the experiment of flipping a coin repeatedly until the first head is obtained, we could assign to each outcome the number of the flip on which the head is obtained.

◇ **Example 2.1(d):** For the experiment of randomly selecting a resistor from a batch of resistors from the manufacturer, we assign to each outcome (resistor) its resistance, in ohms.

A function that assigns a real number to each outcome in a sample space is called a **random variable**. Note that the domain of such a function is the sample space of the experiment. The range of the random variable is a set of real numbers; in the case of the first example above, the range is the set $\{0, 1, 2, 3\}$.

1. (a) Give the range of the random variable in Example 2.1(b).
   (b) Give the range of the random variable in Example 2.1(c).
   (c) Give the range of the random variable in Example 2.1(d). (For those who are not electronics folks, the resistance must be greater than zero.)

If the range of a random variable is a finite or countably infinite set of numbers, we call it a **discrete random variable**. If it is some interval of the real line, we call it a **continuous random variable**. The random variables from Example 2.1(a), (b) and (c) are discrete, the random variable from Example 2.1(d) is continuous.

NOTE: Due to the fact that we can only measure with so much accuracy, *practically speaking* all random variables are discrete. We will act as if we could measure to any degree of accuracy when determining whether a random variable is discrete or continuous.

In an algebra or calculus class the most common letter for denoting a function is f. It is common practice to denote random variables by the *upper case* letter X. When we need more letters, we use Y and Z; for the first example above we would write $X(HTH) = 2$.

2. For Example 2.1(b), use the notation (a, b) for the outcomes, where a is the value rolled on the red die and b is the value rolled on the green. Let X be the random variable described in the example, and find $X((2, 5))$.

3. For the same experiment we can define different random variables. For the experiment from Example 2.1(b), define the random variable Y to be the sum of the numbers on the two die.

   (a) What is $Y((2, 5))$?

   (b) Are there other outcomes for which the random variable Y has the same value as you obtained in (a)? If so, list them all.

   (c) List all outcomes for which the random variable Y has value 12.

   (d) List all outcomes for which $Y = 3$.

4. Consider the random variable X for Example 2.1(c), where a coin is flipped repeatedly until the first head is obtained.

   (a) Find $X(TTTTTH)$.

   (b) Are there any two outcomes for which the random variable takes the same value?

We can see from Exercise 3(b) that a random variable, like functions that you are used to (think $f(x) = x^2$) can assign the same value to more than one outcome of an experiment. In some cases they don't, as shown in Exercise 4(b).

At this point you are probably wondering what purpose random variables have! Their purpose will become more apparent in the next section, but for now we make the following observation: *Many events for a given experiment can be described very precisely in terms of a well defined random variable.* Consider the experiment and random variable of Example 2.1(a). The event of getting exactly two heads is described by the statement $X = 2$, and the event of getting at least two heads can be described by $X \geq 2$. For the experiment and random variable of Example 2.1(c), the event of getting the first head somewhere between the fifth and tenth flip (inclusive) is $5 \leq X \leq 10$.

5. For the experiment and random variable of Example 2.1(a), write each of the following events in terms of the random variable.

   (a) Getting one or two heads.

   (b) Getting all tails.

6. For the experiment and random variable of Example 2.1(c), write each of the following events in terms of the random variable.

   (a) Getting the first heads in less than six flips.

   (b) Getting the first heads in exactly three flips.

With the convenience of describing events using random variables, we can also concisely state certain probabilities. Again using Example 2.1(a), the probability of obtaining *exactly* two heads in the three flips of the coin is denoted by $P(X = 2)$. Since there are eight outcomes to the experiment and three of them have exactly two heads, $P(X = 2) = \frac{3}{8}$.

7. Consider again the experiment of rolling a pair of dice, one red and one green, and again let Y be the random variable that assigns to any outcome the sum of the numbers on the two die. Make a six by six table with the numbers 1, 2, .., 6 across the top and down the left side. In each cell, record the value of the random variable Y for the outcome of getting the row number on the red die and the column number on the green die.

8. Use your table from the previous exercise to find each of the following probabilities.

(a) $P(Y = 4)$ (b) $P(Y = 7)$ (c) $P(Y \leq 4)$

(d) $P(Y > 4)$ (e) $P(Y = 15)$ (f) $P(Y \neq 10)$

9. An urn contains three red marbles and seven blue marbles. An experiment consists of randomly selecting four marbles from the urn, with replacement. Let X be the random variable that assigns to each outcome the number of red marbles selected.

   (a) What is $\text{Ran}(X)$? (Recall that this means "range of X.")

   (b) Find $P(X = 3)$, $P(X \geq 1)$.

10. Consider the same urn as used in the previous exercise, but with the experiment being that four marbles are selected randomly without replacement. Let X again be the random variable that assigns to each outcome the number of red marbles selected.

    (a) What is $\text{Ran}(X)$?

    (b) Find $P(X = 3)$, $P(X \geq 1)$.

2.2 Discrete Probability Distributions

Performance Criteria:

2. (d) Give values of a discrete probability distribution function and a discrete cumulative discrete probability function associated with a random variable.

   (e) Find probabilities from either a discrete probability function or a discrete cumulative probability distribution.

   (f) Determine the value of a constant that makes a function defined on a discrete set a probability distribution function.

   (g) Given a discrete probability distribution function, find the associated cumulative distribution, or vice-versa.

In this section we introduce two functions associated with a random variable; these functions give us any probability of interest for that random variable. The reason for developing such functions is that they then allow us to use tools from algebra and calculus in dealing with probabilities.

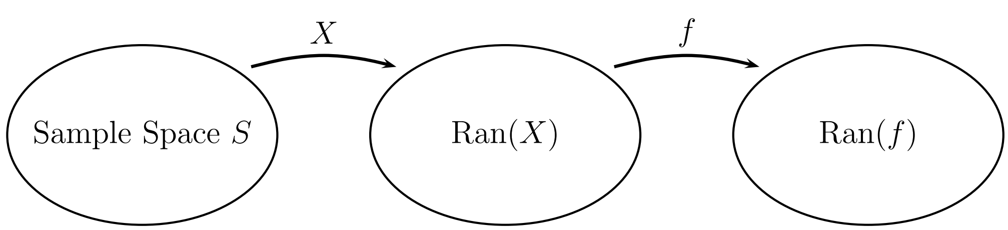
Let X be a discrete random variable defined on the sample space S of an experiment. We define the **probability distribution function** f of X by

$$f(x) = P(X = x)$$

for each x in the range of X.

We will often just say “probability distribution” when we mean “probability distribution function”.

The notation here is a bit confusing when you are first learning the concepts of a random variable and probability distribution function. I’ll try to make some sense of it, but you should expect to have to give it some deep thought as well! Let’s try starting with a picture:



So X is a function that assigns a number to each element of the sample space, and the numbers assigned by X are referred to with the variable x. f is then another function that assigns to each possible value of x another number, which is the probability of getting that value x for the random variable.

◇ **Example 2.2(a):** Suppose that an experiment consists of flipping a coin twice in a row, and let X be the random variable that assigns to each outcome the number of heads. Give the sample space, the range of the random variable X, and the value of f for each value in the range of X.

The sample space is $S = \{TT, TH, HT, HH\}$ and the range of X is $\mathrm{Ran}(X) = \{0, 1, 2\}$. Since the probability of no heads is $\frac{1}{4}$, the above definition gives us that $f(0) = \frac{1}{4}$. Similarly, $f(2) = \frac{1}{4}$ also and $f(1) = \frac{1}{2}$. We will usually summarize a discrete probability function with a table of values like this:

x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

(We could write such a table vertically, like you did in an algebra class, but I'll write them horizontally to save space.) Once we have such a table, we can use f to find various probabilities. For example. $P(X = 1) = f(1) = \frac{1}{2}$ and $P(X \geq 1) = f(1) + f(2) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$.

There is a very useful way of thinking of probability distribution functions. With the previous example in mind, consider the number line as an "infinite ruler" that has no mass of its own. (To make sure that we are all thinking of the same picture, the positive end of our number line is to the right.) We then place a mass of $\frac{1}{4}$ unit at zero on the number line, a mass of $\frac{1}{2}$ at one, and a mass of $\frac{1}{4}$ at two. Note that the total mass is one unit. (People will sometimes talk about a probability mass function, which just means a probability distribution.)

Using this idea, let's define another function that we will name F. (*Note that our f's are now case sensitive!*) F will have a value for any real number x. To find $F(x)$ for some particular x, we go to the point x on our infinite number line and total up the mass at, and to the left of, x. So, for the above example, $F(-1) = 0$ because there is no mass at $x = -1$ or to its left. $F(\frac{3}{2}) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$ because there is no mass at $x = \frac{3}{2}$, but there are masses of $\frac{1}{4}$ and $\frac{1}{2}$ at $x = 0$ and $x = 1$, both of which are to the left of $\frac{3}{2}$.

NOTE: f is only defined for values in the range of the random variable X, which might just be a few numbers, like in the example above. F, on the other hand, *is defined for all real numbers!*

1. For the above example, find each of the following.

   (a) $F(1)$ (b) $F(.83)$ (c) $F(1.99999)$ (d) $F(2)$ (e) $F(73)$

The usefulness of the function F is not necessarily apparent with finite random variables like the one in the above example, but it can be handy for discrete random variables with infinite ranges. The real value of it will be most apparent in the next section, when we develop analogous ideas for continuous random variables.

We will now formalize the previous discussion. Note that the probability distribution function f is defined only on the range of the random variable X. That is, the domain of f is the range of X. We could extend the domain of f to all real numbers by simply defining $f(x) = 0$ for all values of x not in the range of X. Assume that we have done that before we make the following definition.

Let X be a discrete random variable with probability distribution f. We define the **cumulative distribution function** of X by

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

for each $x \in (-\infty, \infty)$.

Again, F *is defined for ALL real numbers.* In the discrete case F can only be described in a piecewise manner. The method for determining a description of F is given in the next example.

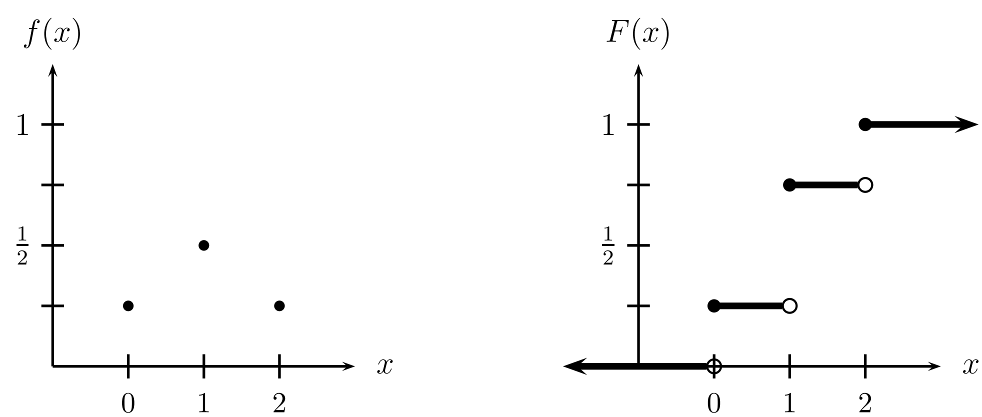
◇ **Example 2.2(b):** Determine the cumulative distribution function F for the random variable and probability distribution from Example 2.2(a).

F must be defined for all real numbers, so let's consider a negative value of x first. To find $F(x)$ we add up all values of $f(t)$ for $t \leq x$. But since zero is the smallest value of t for which $f(t)$, $f(t) = 0$ for all $t \leq x$ and we have $F(x) = 0$ for any negative value of x.

Now if x is greater than or equal to zero, but less than one, $f(t) = 0$ for all $t \leq x$ except $t = 0$, where $f(0) = \frac{1}{4}$. Thus the sum of the values of $f(t)$ for $t \leq x$ is also $\frac{1}{4}$, so $F(x) = \frac{1}{4}$ for $0 \leq x < 1$. Notice that this does not include $x = 1$, where F "picks up" another $\frac{1}{2}$ since $f(1) = \frac{1}{2}$. To summarize, F can be described as follows:

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{4} & \text{for } 0 \leq x < 1 \\ \frac{3}{4} & \text{for } 1 \leq x < 2 \\ 1 & \text{for } x \geq 2 \end{cases}$$

Spend enough time thinking about this example to develop a complete understanding of it. if it is not pretty clear as is, try to work through the details from the the values of f *and the definition of* F*.* It is possible to graph f and F, and doing so might illuminate what is going on here. The graphs are shown below, with the graph of f to the left and the graph of F to the right.



2. Consider the experiment of flipping a coin 4 times, with the sample space denoted by

$$\{TTTT, HTTT, THTT, ..., HHHT, HHHH\}.$$

Note that the sample space has $2 \cdot 2 \cdot 2 \cdot 2 = 2^4 = 16$ outcomes. Define the random variable X to be the number of heads obtained in four tosses.

(a) $X(THTH) =$ __________ (b) $X(HHHH) =$ __________

(c) List all possible values of X, from smallest to largest: ________________________

(d) $f(3) = P(X$ ______ $) =$ ______ (e) $F(3) = P(X$ ______ $) =$ ______

(f) $P(X \le 2) = f(\_\_\_\_) + f(\_\_\_\_) + f(\_\_\_\_) = F(\_\_\_\_\_\_\_\_) = \_\_\_\_\_\_\_\_\_\_\_\_$

(g) $P(X \ge 2) = f(\_\_\_\_) + f(\_\_\_\_) + f(\_\_\_\_) = 1 - F(\_\_\_\_\_\_\_\_) = \_\_\_\_\_\_\_\_\_\_\_\_$

(h) Write the probability that X is odd in terms of f, and find its value.

(i) Write $P(1 \le X \le 3)$ in terms of f, and find its value.

(j) Write $P(1 \le X \le 3)$ in terms of F, and find its value.

3. Use the experiment from the previous exercise for the following.

(a) Give a table of values for the probability density function f.

(b) Sketch the graph of f.

(c) Give a piecewise definition of the cumulative distribution function F.

(d) Sketch the graph of F.

The result of Exercise 2(j) illustrates a useful idea.

Theorem 2.1: Let X be a discrete random variable with values $x_0, x_1, x_2, ..., x_n$ and having probability distributions f and F. Then for $1 \le i < j \le n$,

$$P(x_i \le X \le x_j) = F(x_j) - F(x_{i-1})\,.$$

In particular,

$$P(X = x_j) = f(x_j) = F(x_j) - F(x_{j-1})\,.$$

4. A random variable X has the values $x = 0, 1, 2, 3$. The cumulative distribution function $F(x)$ has the following values:

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{3}{15} & \text{for } 0 \le x < 1 \\ \frac{10}{15} & \text{for } 1 \le x < 2 \\ \frac{14}{15} & \text{for } 2 \le x < 3 \\ 1 & \text{for } x \ge 3 \end{cases}$$

Give the probability distribution function f for the random variable X in table form, like done in the first example. The above theorem might be found useful here!

5. Find each of the following for the random variable from Exercise 4.

(a) $P(X \leq 2)$ (b) $P(X < 2)$ (c) $P(X = 2)$

(d) $P(X \geq 1)$ (e) $P(X > 1)$ (f) $P(X \leq -2)$

(g) $P(X \geq 5)$ (h) $P(X \leq 5)$ (i) $P(1 \leq X \leq 3)$

6. An experiment consists of flipping a coin repeatedly until a head is obtained; The possible outcomes are

$$H, TH, TTH, TTTH, TTTTH, TTTTTH, ...$$

The random variable X assigns to each outcome the number of the flip on which the first head occurred so, for example, $X(TTH) = 3$.

(a) Give a table of values for the probability function f for values of x up through 5. Then draw a neat and reasonably sized graph of f for those values.

(b) Give a piecewise definition of the cumulative probability function F for values of x up through 3, then draw its graph.

(c) Give a formula for $f(x)$, $x = 1, 2, 3, 4, ...$

(d) Let n be any natural number greater than or equal to one. Give a formula for $F(x)$ when $n \leq x < n + 1$.

7. Prove the first part of Theorem 2.1.

It should not be hard to believe the following.

Theorem 2.2: Let X be a discrete random variable having probability distribution f. Then

1) $f(x) \geq 0$ for every $x \in \text{Ran}(X)$

2) $\sum_{\text{Ran}(X)} f(x) = 1$

The above two conditions must be satisfied by the probability distribution function for the discrete random variable X. Conversely, any function satisfying the above two conditions is in fact a probability distribution for the random variable X for which $P(X = x) = f(x)$ for every x in the range of X.

8. Determine the value of c so that $f(x) = cx$, $x = 1, 2, 3, 4$ is a probability distribution function.

9. Determine the value of c so that $f(x) = \dfrac{c}{x}$, $x = 1, 2, 3, 4$ is a probability distribution function.

10. Determine the value of c so that $f(x) = c\begin{pmatrix} 3 \\ x \end{pmatrix}\begin{pmatrix} 5 \\ 2-x \end{pmatrix}$, $x = 0,\ 1,\ 2$ is a probability distribution function for a discrete random variable X.

The symbols $-\infty$ and ∞ do not represent numbers, so it makes no sense to write something like $f(\infty)$. In the interest of brevity we will define $f(\infty) = \lim_{x\to\infty} f(x)$, and similarly for $-\infty$ and F. Using this notation,

Theorem 2.3: Let X be a discrete random variable with probability distributions f and F. Then

1) $f(-\infty) = f(\infty) = 0$ and $F(-\infty) = 0$, $F(\infty) = 1$

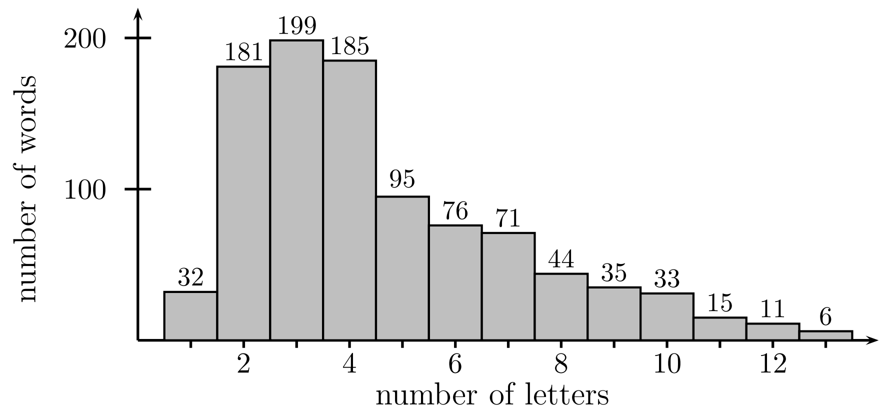
2) for any $a < b$, $F(a) \leq F(b)$.

2.3 Data and Histograms

Performance Criteria:

2. (h) Determine probabilities from a histogram.

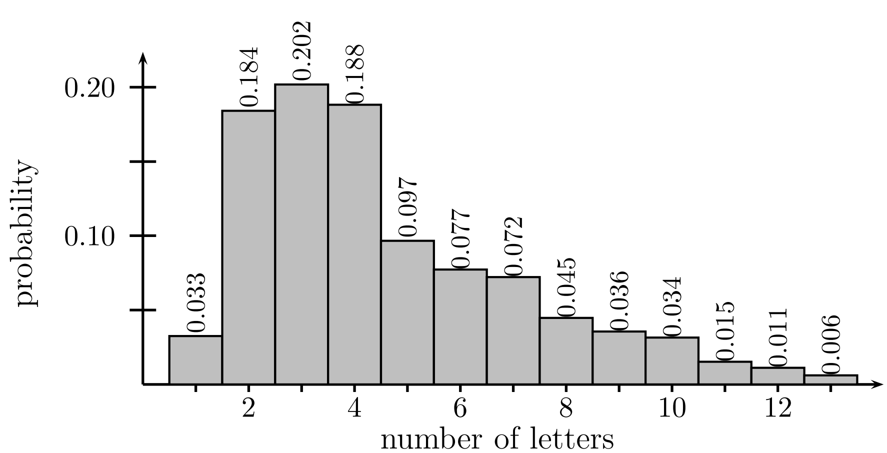
In a display of incredible fortitude, Mr. Waterman counted the number of letters in each of 983 randomly selected words from a statistics book. The graph below shows how many words of each word length there were - such a graph is called a histogram. The number at the top of each column indicates the number of words with that many letters.



1. If we were to randomly select one of the 983 words, what is the probability that it would have

   (a) six letters? (b) more than eight letters? (c) three or more letters?

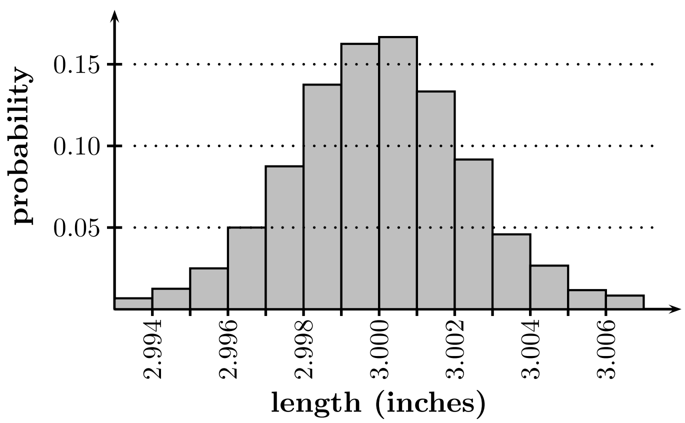
Selecting one of the 983 words randomly is an experiment, with the outcomes being the entire set of words. There is an obvious random variable here; it assigns to each outcome (word) the number of letters in the word. It is a discrete random variable, with values 1, 2, 3, ..., 13. If we were to divide the number of words of each length by the total number of words we would see that $P(X=1)=\frac{32}{983}\approx 0.0326$, $P(X=2)=\frac{181}{983}\approx 0.1841$, and so on. (From now on, when stating decimal approximations to probabilities we will simply use =, with the understanding that the value is likely a rounded approximation.) We can create a new histogram, called a probability histogram, where the height of each column indicates the probability of selecting a word with that many letters, rather than the number of words with that many letters:



This, of course, is a graph of the probability distribution function f for the random variable that assigns to a randomly chosen word the number of letters in that word. *Now here is a key idea - rather than thinking of the height of each column indicating the probability of selecting a word with the corresponding number of letters, we want to think of the* <u>*area*</u> *of each column as being the probability!*

2. With this interpretation, what is the total shaded area for the graph above?

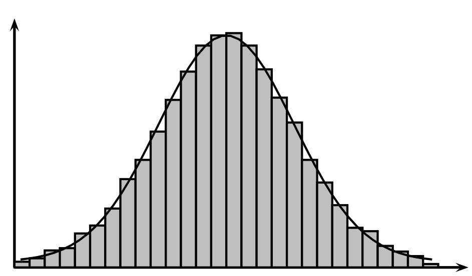
The random variable for the previous situation was a discrete random variable. Now imagine this scenario: We go to a plant where three inch long bolts are made, and we consider the experiment of randomly selecting many bolts, one at a time. We then let X be the random variable that assigns to each bolt its length in inches. Then this is, at least theoretically, a continuous random variable, since its domain is the set of all real numbers greater than zero. Now we certainly can't create a graph with a bar above every real number, like above. Instead we make a probability histogram by designating "classes", or ranges of lengths, and graphing the probabilities of randomly selecting a bolt in each of the classes. Such a histogram is shown above and to the right. *We again think of the* <u>*area*</u> *of each column in the histogram as representing the probability of selecting a bolt in that class.*



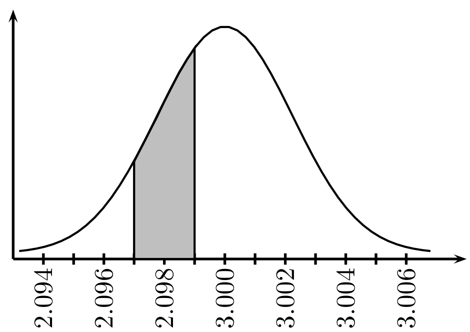
3. *If possible*, determine the approximate probability that the randomly selected bolt has length in the given range. *Assume that all bolts selected are represented in the graph.*

   (a) at least 3.003 inches.
   (b) less than 3.003 inches.
   (c) less than 2.0985 inches.
   (d) between 2.097 and 2.099 inches.
   (e) between 2.093 and 3.007 inches.

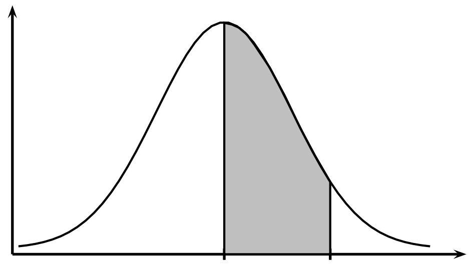
Note that we can draw a continuous curve that approximates the shape of the probability distribution from the above exercise, as shown below and to the left. If we were to gather more data and increase the number of classes so that every bar representing a class was very narrow, we would find that the resulting probability distribution would look even more like the smooth curve, as shown below and to the right.



Recall that the probability of randomly selecting a bolt with length between 2.097 inches and 2.099 inches can be thought of as the sum of the areas of the two bars representing lengths in those ranges. If we were to replace the histogram with a continuous curve, as shown to the right, the shaded area below the continuous curve and between those two values should be a good approximation of the probability of randomly selecting a bolt with a length between 2.097 and 2.099 inches. The continuous curve is an example of what is called a **continuous probability distribution**. The total area under a continuous probability distribution is one, and the probability of randomly obtaining a data value between any two given values is the area under the curve and between those two values.



4. A number of adults have assembled an "easy to assemble" children's toy. The probability distribution for random variable that assigns to a randomly selected adult the length of time it takes them to assemble the toy is shown to the right. *The distribution is symmetric about 43 minutes.* The shaded area is 0.4. Using the distribution, what is the probability that a randomly selected adult will be able to assemble the toy in



(a) between 43 and 51 minutes?

(b) less than 43 minutes?

(c) less than 51 minutes?

(d) more than 51 minutes?

Considering the situation from the above exercise, what would happen to the probabilities you would find for assembling the toy in 43 to 51 minutes, then 43 to 47 minutes, 43 to 44 minutes, 43 to 43.1 minutes, and so on? Clearly they will get smaller and smaller, approaching zero. The probability of randomly selecting an adult who takes *exactly* 43 minutes (or 35 minutes, or 49 minutes, etc.) to assemble the toy is zero.

2.4 Continuous Probability Distributions

Performance Criteria:

2. (i) Find probabilities from either a continuous probability density function or a continuous cumulative probability distribution.
   (j) Determine the value of a constant that makes a function defined on a continuous set a probability distribution function.
   (k) Given a continuous probability density function, find the associated cumulative distribution, or vice-versa.

Based on the discussion at the end of the previous section we see that when working with continuous probability distributions, *we do not want have a distribution function f whose value at any x gives the probability of the random variable taking that value.* Such a function would be zero at all values of x! Instead we define a continuous probability distribution function as follows.

Let X be a continuous random variable defined on the sample space S of an experiment. The **probability density function** of X is a function f for which

$$P(a \le X \le b) = \int_a^b f(x)\,dx$$

for all real numbers a and b with $a \le b$.

Because the probability of a continuous random variable taking an exact value is zero, we have

$$P(a \le X \le b) = P(a \le X < b) = P(a < X \le b) = P(a < X < b)\,.$$

Recall the physical model of masses on a number line for a discrete probability distribution. The analog for this situation is a number line with a piece of wire, of perhaps varying thickness, laying on it. No single point on the number line has any mass on it; mass only exists in sections of the wire of some length. If the density were constant, then the mass of any piece of the wire would simply be the (linear) density (measured in some units like grams per centimeter) times the length of the piece. Since the density is perhaps variable, we have to compute the mass of any section of the wire using an integral, as defined above. This is the reason for the term probability *density* function.

1. A continuous random variable X has the probability density function

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{2}x & \text{for } 0 \le x \le 2 \\ 0 & \text{for } x > 2 \end{cases}$$

Find each of the following.

(a) $P(0 \le X \le 1)$ (b) $P(\frac{1}{2} \le X < 2)$ (c) $P(X = 1)$

(d) $P(-1 \le X \le 1)$ Note that since the definition of f changes at $x = 0$, one must compute $\int_{-1}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx$ to find this probability.

(e) $P(X \le 3)$ (f) $P(X \le -2)$

2. A continuous random variable X has the probability density function

$$f(x) = \begin{cases} 0 & \text{for } x < 1 \\ \dfrac{1}{x^2} & \text{for } x \ge 1 \end{cases}$$

Find each of the following.

(a) $P(2 \le X \le 4)$ (b) $P(0 \le X < 2)$ (c) $P(X \ge 1)$

(d) $P(X \le 5)$ (e) $P(X > 3)$ (f) $P(X = 7)$

The following theorem is the continuous analog to Theorem 2.2.

Theorem 2.4: Let X be a continuous random variable having probability density f. Then

1) $f(x) \ge 0$ for every x for all real numbers x,

2) $\int_{-\infty}^{\infty} f(x)\,dx = 1$.

Like Theorem 2.2, this indicates that any function that satisfies the above two conditions can serve as a probability density function for a continuous random variable.

3. Determine a value of C so that the function is a probability density function.

$$f(x) = \begin{cases} C & \text{for } 0 \le x \le 5 \\ 0 & \text{otherwise} \end{cases}$$

4. Determine a value of C so that the function is a probability density function.

$$f(x) = \begin{cases} Cx & \text{for } 1 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$$

Consider the probability distribution function from Exercise 2, and let a and b be constants with $1 \le a < b$. Then

$$P(a \le X \le b) = \int_a^b f(x)\,dx = \int_a^b \frac{1}{x^2}\,dx = -\frac{1}{x}\Big]_a^b = \left(-\frac{1}{b}\right) - \left(-\frac{1}{a}\right)$$

This shows that finding the probability $P(a \le X \le b)$ amounts to simply finding the difference between the antiderivative of f at a and at b. That is, if F is such that $F'(x) = f(x)$, then $P(a \le X \le b) = F(b) - F(a)$. This function F is (almost) the cumulative distribution function of the continuous random variable X.

Let X be a continuous random variable with probability density function f. We define the **cumulative distribution function** of X by

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$$

for each $x \in (-\infty, \infty)$.

Note that $P(X \geq x) = 1 - P(X < x) = 1 - P(X \leq x)$.

◇ **Example 2.4(a):** Find the cumulative probability function F for the density function of Exercise 2.

For $x < 1$ we have $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{x} 0\,dt = 0\,.$

For $x \geq 1$, $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{1} 0\,dt + \int_{1}^{x} \frac{1}{t^2}\,dt = 0 + \left[-\frac{1}{t}\right]_1^x = 1 - \frac{1}{x}\,.$

In conclusion then, $F(x) = \begin{cases} 0 & \text{for } x < 1 \\ 1 - \dfrac{1}{x} & \text{for } x \geq 1 \end{cases}$

---

5. Find the cumulative probability distribution for the probability density function of

(a) $f(x) = \begin{cases} \frac{1}{5} & \text{for } 0 \leq x \leq 5 \\ 0 & \text{otherwise} \end{cases}$ (b) $f(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{2}x & \text{for } 0 \leq x \leq 2 \\ 0 & \text{for } x > 2 \end{cases}$

The following result follows easily from the definition of the cumulative distribution.

Theorem 2.5: Let X be a continuous random variable having probability density f and cumulative distribution F. Then

$$P(a \leq X \leq b) = F(b) - F(a)\,.$$

6. A continuous random variable X has the cumulative distribution function

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \dfrac{\sqrt{x}}{2} & \text{for } 0 \leq x \leq 4 \\ 1 & \text{for } x > 4 \end{cases}$$

Give each of the following probabilities **in exact form**. (No decimals! Some of your answers will contain square roots.)

(a) $P(X < 3)$ (b) $P(X \geq 1)$ (c) $P(1 < X \leq 3)$
(d) $P(X = 2)$ (e) $P(X \geq -1)$ (f) $P(X \leq -1)$

7. Consider again the probability density function $f(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{2}x & \text{for } 0 \leq x \leq 2 \\ 0 & \text{for } x > 2 \end{cases}$

   (a) Graph the density function f, from $x = -2$ to $x = 4$.

   (b) Graph the cumulative distribution function F, from $x = -2$ to $x = 4$.

8. Graph the density function $f(x) = \begin{cases} \frac{1}{5} & \text{for } 0 \leq x \leq 5 \\ 0 & \text{otherwise} \end{cases}$ and its cumulative distribution $F(x)$.

The graphs from the previous exercises can perhaps illuminate what is going on here. The function F evaluated at any point x simply gives the area under f to the left of x. So when x is less than zero, there is no area under the graph of f. As x crosses zero (moving right) it starts picking up area . It continues to pick up area until $x = 2$, after which no new area is accumulated as x continues to move to the right. This illustrates the following theorem.

Theorem 2.6: Let X be a continuous random variable with probability distributions f and F. Then

1) $f(-\infty) = f(\infty) = 0$ and $F(-\infty) = 0,\ F(\infty) = 1$

2) for any $a < b$, $F(a) \leq F(b)$.

9. A continuous random variable X has the cumulative distribution function

$$F(x) = \begin{cases} 0 & \text{for } x \leq 2 \\ c(x-2)^2 & \text{for } 2 < x < 5 \\ 1 & \text{for } x \geq 5 \end{cases}$$

   Find the value of c. *This does not require integration!*

10. Find each of the following for the random variable and cumulative distribution function F from the previous exercise.

   (a) $P(X \leq 4)$ (b) $P(X < 4)$ (c) $P(X = 4)$

   (d) $P(X \geq 3)$ (e) $P(3 \leq X < 4)$ (f) $P(1 < X < 4)$

   (g) $P(X \geq 7)$ (h) $P(X \leq 7)$

The definition of the cumulative distribution function F tells us how to obtain it from the probability density f. Since F is obtained by integrating f, one might suspect that f is obtained from F by differentiating. This is in fact the case.

Theorem 2.7: Let X be a continuous random variable with probability distributions f and F. For all values of x for which F is differentiable we have $f(x) = F'(x)$.

11. Find the probability density function f for the cumulative probability function F from the previous two exercises. Check it by using it to compute $P(3 \le X < 4)$ and seeing if it agrees with your answer to (e) above.

12. Find the probability density function for the cumulative distribution from Exercise 6.

2.5 Mean and Variance of a Distribution

Performance Criteria:

2. (l) Determine the mean, variance or standard deviation of a random variable.

1. Consider the discrete random variable X with probability distribution

$$\begin{array}{rcccc} x: & 0 & 1 & 2 & 3 \\ f(x): & \frac{1}{16} & \frac{4}{16} & \frac{8}{16} & \frac{3}{16} \end{array}$$

Give the answers to the following in exact (fraction) form.

(a) Find the average of the four numbers 0, 1, 2, 3.

(b) We define the **mean** or **expected value** μ of the discrete random variable X with probability distribution f by

$$\mu = E(X) = \sum_x x\, f(x).$$

Find the expected value of the random variable given.

(c) Sketch the graph of f. Put a vertical arrow $\uparrow$ pointing at the position of μ on the x-axis.

(d) Once we have the expected value of a random variable, we can compute something called the **variance** σ^2 of the random variable, defined by

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

Find the variance of the random variable given.

(e) The positive square root of the variance, denoted by σ, is called the **standard deviation** of the random variable X. Find the standard deviation of the given random variable.

So what does all this mean? Suppose that we had an urn with red, blue, green and yellow marbles in it, *equal numbers of each*. You are going to draw one marble at random from the urn; if you draw a red marble you will be given \$0, if you draw a blue you are given \$1, and you are given \$2 for a green and \$3 by a yellow. If you were to draw a whole bunch of times, your average "winnings" would be what you computed in (a) of the above exercise.

Now suppose instead that the urn you were drawing from had one red marble, four blue marbles, eight green marbles and three yellow marbles. (How many marbles are there in all? Note how this situation relates to the probability distribution function from the above exercise.) Then your winnings, on average, would be higher than the previous arrangement, because the probability of winning \$2 or \$3 is higher than the probability of winning \$0 or \$1. The expected value is then the average winnings that you could expect, hence the name "expected value".

Going back to our physical model of a discrete probability distribution as a set of masses located at single points on a number line, the expected value of the distribution is the center of mass, or "balance point" of the number line.

The next exercise should illuminate what the variance is about.

2. Consider the following two different probability distributions g and h for the random variable X that takes the values $x = 0,\ 1,\ 2,\ 3$.

$x:$	0	1	2	3
$g(x):$	$\frac{1}{16}$	$\frac{7}{16}$	$\frac{7}{16}$	$\frac{1}{16}$

$x:$	0	1	2	3
$h(x):$	$\frac{3}{16}$	$\frac{5}{16}$	$\frac{5}{16}$	$\frac{3}{16}$

   (a) Find the expected value for the first distribution. Does the result make sense? What do you think the expected value for the second distribution will be?

   (b) Find the variances of the two distributions. Which probability distribution has the larger variance? Can you see what it indicates about the distribution?

Hopefully you recognized that the variance indicates how "spread out" the probabilities are over the values of the random variable.

Let's get the definitions of the expected value and variance formalized a bit. Note that we also call the expected value the **mean**.

> Let X be a discrete random variable with probability distribution f. Then the **mean** or **expected value** of X is
>
> $$\mu = E(X) = \sum_x x\, f(x)$$
>
> and the **variance** of X is
>
> $$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2\, f(x)\,.$$
>
> The positive square root of σ^2 is called the **standard deviation**, denoted by σ.

In situations where it might not be clear what the random variable is, we sometimes use μ_X and σ_X^2 instead of μ and σ^2.

◇ **Example 2.5(a):** Suppose that for some course an instructor assigns the following *weights* to the various components of students' grades: assignments, 15%; quizzes, 15%; exams, 40%; final exam, 30%. If an individual student has grades in each area of 93%, 81%, 84% and 78%, respectively, what is their overall course grade?

Their grade is $93(0.15) + 81(0.15) + 84(0.40) + 78(0.30) = 83.1\%$.

What has been done here is the students grade has been computed as what we call a **weighted average**. Rather than just average the percentage grades in each area "straight out", we average them in a way that makes some "count" more than others. Note that we could compute an "ordinary average" by replacing the **weights** 0.15, 0.15, 0.40, 0.30 each with 0.25. Notice that either way, *the weights are all positive and their sum is one!* this means that the weights are a discrete probability distribution.

This is exactly what we are doing when we compute an expected value; we are averaging the values of a random variable, weighting each according to its probability. To understand what we are doing when we compute the variance, let's introduce some terminology. We will call the quantity $x-\mu$ the **deviation** of the random variable value x (from the mean). When we compute the variance we are also computing a weighted average, but we are averaging the squares of the deviations.

3. The point of the variance is to give some idea of how the values of a random variable vary. It would make more sense to average the deviations of the values of the random variable rather than their squares. That is, it would seem that we should compute

$$\sum_x (x-u)\,f(x)\,.$$

Do this for the random variable from Exercise 1.

This exercise illustrates why averaging the deviations is not a good idea. One solution to the problem would be to average the absolute values of the deviations. It turns out, however, that averaging the squares of the deviations is more mathematically advantageous. We will not go into why that is here, just take my word for it!

4. Compute the mean and the variance for the random variable with probability distribution given by

x	0	1	2	3
$f(x)$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$

5. Find the mean and variance for the experiment consisting of flipping a coin four times in a row, with the random variable X that assigns to each outcome the number of heads obtained. How many heads should one expect to get on average when performing this experiment repeatedly?

We now make a very convenient observation:

$$\begin{aligned}
\sum_x (x-\mu)^2 f(x) &= \sum_x (x^2-2x\mu+\mu^2)\cdot f(x)\\
&= \sum_x x^2 f(x) - 2\mu\sum_x x f(x) + \mu^2\sum_x f(x)\\
&= E(X^2) - 2\mu E(X) + \mu^2\\
&= E(X^2) - 2\mu^2 + \mu^2\\
&= E(X^2) - \mu^2
\end{aligned}$$

Theorem 2.8: If X is a discrete random variable, then the variance of X can be computed by

$$\sigma^2 = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2\,.$$

6. Use this formula to find the variance for the probability distribution from Exercise 4.

7. We will define a **Bernoulli random variable** to be a random variable X with range 0, 1, with $P(X=1)=p$.

   (a) What is $P(X=0)$?

   (b) Find the mean and variance of this random variable.

For continuous random variables we define the expected value and variance in essentially the same way as for discrete random variables, replacing summation with integration.

Let X be a continuous random variable with probability distribution f. Then the **expected value** of X is

$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

and the **variance** of X is

$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (x-\mu)^2\, f(x)\, dx\,.$$

The argument preceding Theorem 2.8 can be repeated for continuous probability functions, replacing summations with integrals.

Theorem 2.9: Let X be a continuous random variable. Then the variance of X can be computed by

$$\sigma^2 = E(X^2) - \mu^2 = E(X^2) - [E(X)]^2\,,$$

where $E(X^2) = \displaystyle\int_{-\infty}^{\infty} x^2\, f(x)\, dx\,.$

8. Find the mean and variance for the probability density function $f(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{2}x & \text{for } 0 \le x \le 2 \\ 0 & \text{for } x > 2 \end{cases}$

9. Find the mean and variance for the probability density function $f(x) = \begin{cases} \frac{1}{5} & \text{for } 0 \le x \le 5 \\ 0 & \text{otherwise} \end{cases}$

10. Find the mean and variance for the probability density function $f(x) = \begin{cases} \frac{2}{15}x & \text{for } 1 \le x \le 4 \\ 0 & \text{otherwise} \end{cases}$

11. Try finding the mean for the probability density function $f(x) = \begin{cases} 0 & \text{for } x < 1 \\ \dfrac{1}{x^2} & \text{for } x \ge 1 \end{cases}$

    What goes wrong?

Theorem 2.10: If X is a discrete random variable with expected value $E(X)$ and variance σ, then

$$E(aX+b) = aE(X)+b \qquad \text{and} \qquad \operatorname{var}(aX+b) = a^2\sigma^2$$

2.6 Chapter 2 Exercises

Give all answers in exact form - NO DECIMALS!

1. An experiment consists of first flipping a coin and then, if the coin comes up heads, flipping the coin again. If the coin comes up tails on the first flip a four-sided die is rolled. A random variable X is assigned to the outcomes as follows: A sum is obtained for each outcome, with tails counting as zero and heads as one, and with the numbers on the die counting their value. So, for example, tails followed by a three is assigned $0+3=3$ and heads followed by heads is assigned $1+1=2$. (Of course there are multiple ways to obtain some range values of the random variable.)

   (a) Abbreviating the outcomes as HH, HT, T3, etc., give the value of X for each outcome *using function notation.*

   (b) Give $\text{Ran}(X)$.

   (c) Define $f : \text{Ran}(X) \to \mathbb{R}$ and sketch its graph.

   (d) Define $F : \mathbb{R} \to \mathbb{R}$ and sketch its graph.

2. *If possible*, find values of c so that each of the following is a probability density function for a *continuous* random variable X.

   (a) $f(x) = \begin{cases} 0 & \text{for } x < 1 \\ \dfrac{c}{\sqrt{x}} & \text{for } x \geq 1 \end{cases}$ (b) $f(x) = \begin{cases} 0 & \text{for } x < 1 \\ \dfrac{c}{x^3} & \text{for } x \geq 1 \end{cases}$

3. Find the mean for the probability function from part (b) of the previous exercise.

4. Find the cumulative probability function F for the function f from Exercise 2(b).

5. A *discrete* random variable X has values $x = 1, 2, 3, 4, 5$. It is known that $F(3) = \frac{8}{17}$ and $F(4) = \frac{14}{17}$. Find any of the following that you can; *a number of them cannot be determined from the information given.*

   (a) $P(X < 3)$ (b) $P(X \leq 3)$ (c) $P(X = 3)$

   (d) $P(X \geq 4)$ (e) $P(3 < X \leq 4)$ (f) $P(X = 4)$

   (g) $P(X > 3)$ (h) $P(X > 4)$ (i) $P(3 \leq X \leq 4)$

6. For a *continuous* random variable X, $F(3) = \frac{8}{17}$ and $F(4) = \frac{14}{17}$. Find any of the values from Exercise 3 that you can.

7. A *continuous* random variable X has the cumulative distribution function

$$F(x) = \begin{cases} 0 & \text{for } x < 2 \\ 1 - \dfrac{2}{x} & \text{for } x \geq 2 \end{cases}$$

   Find any of the values from Exercise 4 that you can for this cumulative distribution function.

8. Find the probability density function f for the cumulative probability function from the previous exercise. *Be sure to define it for all values of* x.

9. A continuous random variable X has the cumulative density function

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 - e^{-x} & \text{for } x \geq 0 \end{cases}$$

Find each of the following. Again, your answers should be in exact form, so they will contain terms like e^{-3}. *Simplify when possible.*

(a) $P(X \leq -2) =$ ______ (b) $P(X \geq -2) =$ ______

(c) $P(1 < X \leq 5) =$ ______ (d) $P(X < 7) =$ ______

(e) $P(X = 7) =$ ______ (f) $P(X \geq 7) =$ ______

10. Consider the probability density function f for the *continuous* random variable X given by

$$f(x) = \begin{cases} \frac{1}{3} & \text{for } 0 \leq x < 2 \\ \frac{1}{9} & \text{for } 2 \leq x \leq 5 \\ 0 & \text{elsewhere} \end{cases}$$

(a) Find $P(X < 1)$. (b) Find $P(1 \leq X \leq 4)$.

(c) Fill in the blanks:

$$F(x) = \begin{cases} \_\_\_\_\_\_ & \text{for } \_\_\_\_\_\_\_\_\_\_\_\_ \\ \_\_\_\_\_\_ & \text{for } \_\_\_\_\_\_\_\_\_\_\_\_ \\ \frac{1}{9}x + \frac{4}{9} & \text{for } \_\_\_\_\_\_\_\_\_\_\_\_ \\ \_\_\_\_\_\_ & \text{for } \_\_\_\_\_\_\_\_\_\_\_\_ \end{cases}$$

11. Find the mean and variance for the distribution from the previous exercise.

12. A probability distribution is given by

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{5}x & \text{for } 0 < x \leq 5 \\ 1 & \text{for } x > 5 \end{cases}$$

(a) Is this distribution for a discrete random variable, or for a continuous random variable? Is the distribution cumulative?

(b) Give the probability function f.

13. Consider a random variable X.

(a) Under what conditions is $P(a \leq X \leq b) = F(b) - F(a)$?

(b) Under what conditions is $P(a \leq X \leq b) = F(b) - F(a-1)$?

(c) In the context of parts (a) and (b), discuss $P(a \leq X < b)$ and $P(a < X \leq b)$.

14. Find the expected value and variance for each of the following.

(a) The discrete probability distribution function $\begin{array}{cccc} x: & 0 & 1 & 2 \\ f(x): & \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \end{array}$

(b) The continuous probability density function $f(x) = \begin{cases} -\frac{1}{2}x + 1 & \text{for } 0 \le x \le 2 \\ 0 & \text{elsewhere} \end{cases}$

15. Consider the continuous probability density function $f(x) = \begin{cases} \frac{1}{15} & \text{for } 0 \le x < 3 \\ \frac{1}{9} & \text{for } 3 \le x \le 7 \\ 0 & \text{elsewhere} \end{cases}$

(a) Find $P(2 \le X \le 6)$. *You should not need to integrate, since the function is constant on intervals.*

(b) Sketch the graph of $F(x)$.

16. Which of the graphs below represents

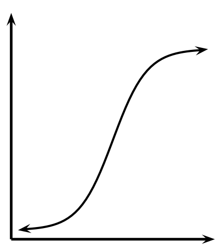
(a) f for a discrete random variable?

(b) F for a discrete random variable?

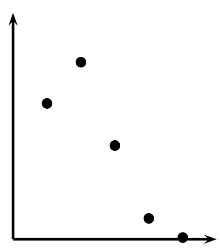
(c) f for a continuous random variable?

(d) F for a continuous random variable?

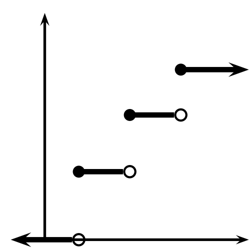
(i)



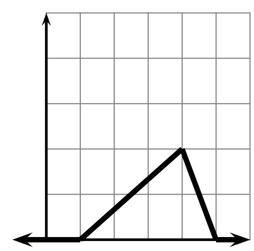
(ii)



(iii)



(iv)



17. Consider the continuous probability distribution function

$$f(x) = \begin{cases} -\frac{1}{2}x + 1 & \text{for } 0 \le x \le 2 \\ 0 & \text{elsewhere} \end{cases}$$

(a) Why can't $\mu = 2\frac{1}{2}$?

(b) Is μ less than 1, greater than 1, or is it 1?

18. (a) Is it possible for a distribution to have an expected value of zero? If so, make such a distribution.

(b) is it possible for a distribution to have a variance of zero? If so, make one.

H Solutions to Exercises

H.2 Chapter 2 Solutions

Section 2.1

1. (a) $\{0, 1, 2, 3, 4, 5, 6\}$ (b) $\{1, 2, 3, ...\}$ (c) $(0, \infty)$

2. $X((2, 5)) = 2$

3. (a) $Y(2, 5) = 7$ (b) $(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)$
   (c) $(6, 6)$ (d) $(1, 2), (2, 1)$

4. (a) $X(\text{TTTTTH}) = 6$
   (b) No, each value of the random variable results from only one outcome.

5. (a) $1 \le X \le 2$ (b) $X = 0$

6. (a) $X \le 5$ or $X < 6$ (b) $X = 3$

7. See table to the right. Every value of the random variable does not occur the same number of times.

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

8. (a) $P(Y = 4) = \frac{3}{36} = \frac{1}{12}$ (b) $P(Y = 7) = \frac{6}{36} = \frac{1}{6}$ (c) $P(Y \le 4) = \frac{6}{36} = \frac{1}{6}$
   (d) $P(Y > 4) = 1 - P(Y \le 4) = 1 - \frac{1}{6} = \frac{5}{6}$ (e) $P(Y = 15) = 0$
   (f) $P(Y \ne 10) = 1 - P(Y = 10) = 1 - \frac{3}{36} = \frac{33}{36} = \frac{11}{12}$

9. (a) $\text{Ran}(X) = \{0, 1, 2, 3\}$ (b) $P(X = 3) = (\frac{3}{10})^3(\frac{7}{10}) \cdot 4 = \frac{756}{10000} = \frac{189}{2500}$
   (c) $P(X \ge 1) = 1 - P(X = 0) = 1 - (\frac{7}{10})^4 = 1 - \frac{2401}{10000} = \frac{7599}{10000}$

10. (a) $\text{Ran}(X) = \{0, 1, 2, 3\}$ (b) $P(X = 3) = \frac{3 \cdot 2 \cdot 1 \cdot 7}{10 \cdot 9 \cdot 8 \cdot 7} \cdot 4 = \frac{168}{5040} = \frac{7}{210}$
    $P(X \ge 1) = 1 - P(X = 0) = 1 - \frac{7 \cdot 6 \cdot 5 \cdot 4}{10 \cdot 9 \cdot 8 \cdot 7} = 1 - \frac{1}{6} = \frac{5}{6}$

Section 2.2

1. (a) (a) $F(1) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$ (b) $F(.83) = \frac{1}{4}$ (c) $F(1.99999) = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$
   (d) $F(2) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$ (e) $F(73) = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$

2. (a) $X(\text{THTH}) = 2$ (b) $X(\text{HHHH}) = 4$
   (c) $0, 1, 2, 3, 4$ (d) $f(3) = P(X = 3) = \frac{4}{16} = \frac{1}{4}$
   (e) $F(3) = P(X \le 3) = 1 - P(X = 4) = 1 - \frac{1}{16} = \frac{15}{16}$
   (f) $P(X \le 2) = f(0) + f(1) + f(2) = F(2) = \frac{1}{16} + \frac{4}{16} + \frac{6}{16} = \frac{11}{16}$
   (g) $P(X \ge 2) = f(2) + f(3) + f(4) = 1 - F(1) = 1 - (\frac{1}{16} + \frac{4}{16}) = \frac{11}{16}$
   (h) $f(1) + f(3) = \frac{4}{16} + \frac{4}{16} = \frac{1}{2}$

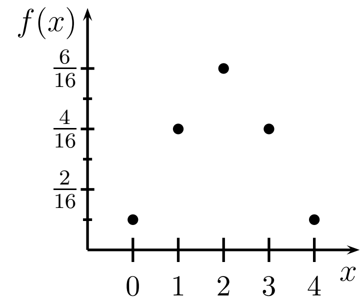
(i) $P(1 \le X \le 3) = f(1) + f(2) + f(3) = \frac{4}{16} + \frac{6}{16} + \frac{4}{16} = \frac{14}{16} = \frac{7}{8}$

(j) $P(1 \le X \le 3) = F(3) - F(0) = \frac{15}{16} - \frac{1}{16} = \frac{14}{16} = \frac{7}{8}$

3. (a)

x	0	1	2	3	4
$f(x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

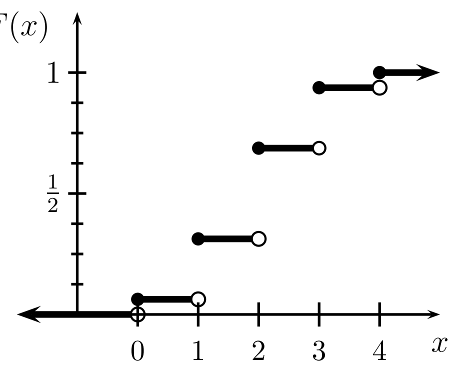
(b)



(c)

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{16} & \text{for } 0 \le x < 1 \\ \frac{5}{16} & \text{for } 1 \le x < 2 \\ \frac{11}{16} & \text{for } 2 \le x < 3 \\ \frac{15}{16} & \text{for } 3 \le x < 4 \\ 1 & \text{for } x \ge 4 \end{cases}$$

(d)



4.

x	0	1	2	3
$f(x)$	$\frac{3}{15}$	$\frac{7}{10}$	$\frac{4}{15}$	$\frac{1}{15}$

5. (a) $P(X \le 2) = F(2) = \frac{14}{15}$ (b) $P(X < 2) = F(1) = \frac{10}{15}$

(c) $P(X = 2) = F(2) - F(1) = f(2) = \frac{4}{15}$ (d) $P(X \ge 1) = 1 - F(0) = 1 - \frac{3}{15} = \frac{12}{15}$

(e) $P(X > 1) = 1 - F(1) = 1 - \frac{10}{15} = \frac{5}{15}$ (f) $P(X \le -2) = F(-2) = 0$

(g) $P(X \ge 5) = 0$ (h) $P(X \le 5) = F(5) = 1$

(i) $P(1 \le X \le 3) = F(3) - F(0) = 1 - \frac{3}{15} = \frac{12}{15}$

6. (a)

x	1	2	3	4	$\cdots$
$f(x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\cdots$

(b)

$$F(x) = \begin{cases} 0 & \text{for } x < 1 \\ \frac{1}{2} & \text{for } 1 \le x < 2 \\ \frac{3}{4} & \text{for } 2 \le x < 3 \\ \frac{7}{8} & \text{for } 3 \le x < 4 \\ \frac{15}{16} & \text{for } 4 \le x < 5 \\ \vdots & \qquad \vdots \end{cases}$$

(c) $f(x) = \dfrac{1}{2^x}$ (d) $F(x) = \dfrac{2^n - 1}{2^n}$ for $n \le x < n + 1$

8. $1 = f(1) + f(2) + f(3) + f(4) = 1c + 2c + 3c + 4c = 10c$, so $c = \frac{1}{10}$

9. $1 = f(1) + f(2) + f(3) + f(4) = \frac{c}{1} + \frac{c}{2} + \frac{c}{3} + \frac{c}{4} = \frac{25c}{12}$, so $c = \frac{12}{25}$

10. $1 = f(0) + f(1) + f(2) = c\binom{3}{0}\binom{5}{2} + c\binom{3}{1}\binom{5}{1} + c\binom{3}{2}\binom{5}{0}$

$$= 10c + 15c + 3c = 28c, \text{ so } c = \tfrac{1}{28}$$

Section 2.3

1. (a) $\frac{76}{983}$ (b) $\frac{35+33+15+11+6}{983}$ (c) $1 - \frac{181+32}{983} = \frac{770}{983}$

2. 1, because the total probability is one

3. (a) $0.045 + 0.03 + 0.015 + 0.01 = 0.10$ (b) $1 - 0.10 = 0.90$
   (c) can't tell (d) $0.085 + 0.135 = 0.22$ (e) 1

4. (a) 0.4 (b) 0.5 (c) 0.9 (d) 0.1

Section 2.4

1. (a) $P(0 \le X \le 1) = \displaystyle\int_0^1 \tfrac{1}{2}x\, dx = \tfrac{1}{4}x^2\Big]_0^1 = \tfrac{1}{4}(1)^2 - \tfrac{1}{4}(0)^2 = \tfrac{1}{4}$
   (b) $P(\frac{1}{2} \le X \le 2) = \displaystyle\int_{\frac{1}{2}}^2 \tfrac{1}{2}x\, dx = \tfrac{1}{4}x^2\Big]_{\frac{1}{2}}^2 = \tfrac{1}{4}(2)^2 - \tfrac{1}{4}(\tfrac{1}{2})^2 = 1 - \tfrac{1}{16} = \tfrac{15}{16}$
   (c) $P(X = 1) = 0$ (d) $P(-1 \le X \le 1) = \displaystyle\int_{-1}^0 0\, dx + \int_0^1 \tfrac{1}{2}x\, dx = 0 + \tfrac{1}{4} = \tfrac{1}{4}$
   (e) $P(X \le 3) = 1$ since the only non-zero probability is from zero to two.
   (f) $P(X \le -2) = 0$ for the same reason.

2. (a) $P(2 \le X \le 4) = \displaystyle\int_2^4 \frac{1}{x^2}\, dx = -\frac{1}{x}\Big]_2^4 = -\frac{1}{4} - \left(-\frac{1}{2}\right) = -\frac{1}{4} + \frac{1}{2} = \frac{1}{4}$
   (b) $P(0 \le X \le 2) = P(1 \le X \le 2) = \displaystyle\int_1^2 \frac{1}{x^2}\, dx = -\frac{1}{x}\Big]_1^2 = -\frac{1}{2} - \left(-\frac{1}{1}\right) = \frac{1}{2}$
   (c) $P(X \ge 1) = 1$ since all of the probability is from one on.
   (d) $P(X \le 5) = P(1 \le X \le 5) = \displaystyle\int_1^5 \frac{1}{x^2}\, dx = -\frac{1}{x}\Big]_1^5 = -\frac{1}{5} - \left(-\frac{1}{1}\right) = \frac{4}{5}$
   (e) $P(X > 3) = \displaystyle\int_3^\infty \frac{1}{x^2}\, dx = -\frac{1}{x}\Big]_3^\infty = -\frac{1}{\infty} - \left(-\frac{1}{3}\right) = 0 + \frac{1}{3} = \frac{1}{3}$
   (f) $P(X = 7) = 0$

3. $1 = \displaystyle\int_0^5 C\, dx = Cx\Big]_0^5 = 5C - 0C = 5C$, so $C = \frac{1}{5}$.

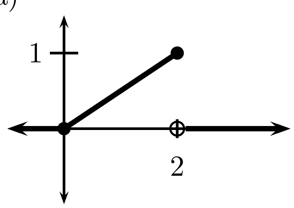
4. $1 = \displaystyle\int_1^4 Cx\, dx = \tfrac{1}{2}Cx^2\Big]_1^4 = 8C - \tfrac{1}{2}C = \tfrac{15}{2}$, so $C = \frac{2}{15}$.

5. (a) $F(x) = 0$ for $x \le 0$. For $0 < x \le 5$, $F(x) = \displaystyle\int_{-\infty}^x f(t)\, dt = \int_0^x \tfrac{1}{5}\, dt = \tfrac{1}{5}t\Big]_0^x = \tfrac{1}{5}x$. For $x > 5$, $F(x) = 1$. Thus $F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{5}x & \text{for } 0 \le x \le 5 \\ 1 & \text{for } x > 5 \end{cases}$

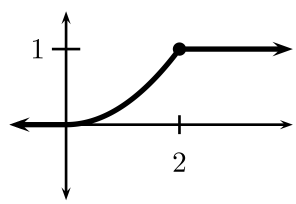
(b) $F(x) = 0$ for $x \le 0$. For $0 < x \le 2$, $F(x) = \int_{-\infty}^{x} f(t)\ dt = \int_0^x \frac{1}{2}t\ dt = \frac{1}{4}t^2\Big]_0^x = \frac{1}{4}x^2$. For $x > 2$, $F(x) = 1$. Therefore $F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{4}x^2 & \text{for } 0 \le x \le 2 \\ 1 & \text{for } x > 2 \end{cases}$

6. (a) $P(X < 3) = P(X \le 3) = F(3) = \frac{\sqrt{3}}{2}$
   (b) $P(X \ge 1) = 1 - P(X < 1) = 1 - P(X \le 1) = 1 - F(1) = 1 - \frac{\sqrt{1}}{2} = \frac{1}{2}$
   (c) $P(1 < X \le 3) = F(3) - F(1) = \frac{\sqrt{3}}{2} - \frac{1}{2} = \frac{\sqrt{3}-1}{2}$ (d) $P(X = 2) = 0$
   (e) $P(X \ge -1) = 1 - P(X \le -1) = 1 - F(-1) = 1 - 0 = 1$
   (f) $P(X = -1) = F(-1) = 0$

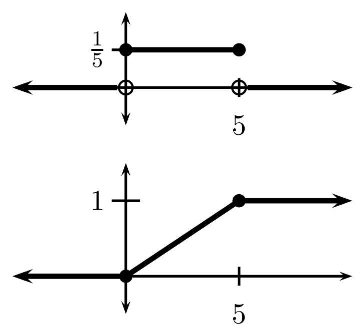
7. (a)



(b)



8.



9. Since F "picks up" probability as x increases, and it has reached value one by the time it reaches $x = 5$, we must have $c(5-2)^2 = 1$, so $c = \frac{1}{9}$.

10. (a) $P(X \le 4) = F(4) = \frac{4}{9}$ (b) $P(X < 4) = F(4) = \frac{4}{9}$
    (c) $P(X = 4) = 0$ (d) $P(X \ge 3) = 1 - P(X \le 3) = 1 - F(3) = 1 - \frac{1}{9} = \frac{8}{9}$
    (e) $P(3 \le X < 4) = F(4) - F(3) = \frac{4}{9} - \frac{1}{9} = \frac{3}{9}$
    (f) $P(1 < X < 4) = F(4) - F(1) = \frac{4}{9} - 0 = \frac{4}{9}$
    (g) $P(X \ge 7) = 1 - P(X \le 7) = 1 - F(7) = 1 - 1 = 0$ (h) $P(X \le 7) = F(7) = 1$

11. $f(x) = F'(x) = \frac{2}{9}(x-2)$ for $2 < x < 5$ and $f(x) = 0$ otherwise.

$$P(3 \le X < 4) = \int_3^4 \frac{2}{9}(x-2)\,dx = \int_3^4 \left(\frac{2}{9}x - \frac{4}{9}\right) dx = \frac{1}{9}x^2 - \frac{4}{9}x\Big]_3^4 = \left(\frac{16}{9} - \frac{16}{9}\right) - \left(\frac{9}{9} - \frac{12}{9}\right) = \frac{3}{9}$$

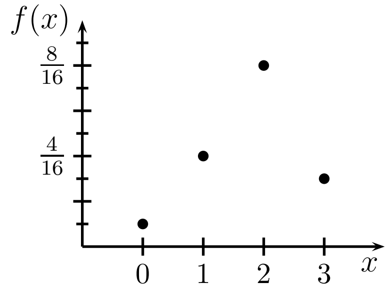
12. Note that $F(x) = \frac{1}{2}x^{\frac{1}{2}}$, so $f(x) = F'(x) = \frac{1}{4}x^{-\frac{1}{2}} = \frac{1}{4\sqrt{x}}$ for $0 \le x \le 4$ and $f(x) = 0$ otherwise.

Section 2.5

1. (a) $\frac{0+1+2+3}{4} = \frac{3}{2} = 1.5$

   (b) $\mu = E(X) = 0 \cdot \frac{1}{16} + 1 \cdot \frac{4}{16} + 2 \cdot \frac{8}{16} + 3 \cdot \frac{3}{16} = \frac{4}{16} + \frac{16}{16} + \frac{9}{16} = \frac{29}{16}$

   (c)



   (d) $\sigma^2 = (0 - \frac{29}{16})^2 \cdot \frac{1}{16} + (1 - \frac{29}{16})^2 \cdot \frac{4}{16} + (2 - \frac{29}{16})^2 \cdot \frac{8}{16} + (3 - \frac{29}{16})^2 \cdot \frac{3}{16} = \frac{83}{128} \approx 0.65$

   (e) $\sigma = \sqrt{\frac{83}{128}} \approx 0.81$

3. $E(X - \mu) = 0$

4. $\frac{11}{8},\ \frac{47}{64}$

5. $\mu = 2,\ \sigma^2 = 1,\ 2$

7. $\mu = p,\ p - p^2 = p(1-p)$

8. $\mu = \frac{4}{3},\ \sigma^2 = \frac{2}{9}$

9. $\mu = \frac{5}{2},\ \sigma^2 = \frac{25}{12}$

10. $\mu = \frac{14}{5},\ \sigma^2 = \frac{33}{50}$

11. The distribution has no expected value, since the integral diverges.