CHAPTER EIGHT

Solutions for Section 8.1

1. We use the fact that the area of a rectangle is Base $\times$ Height.

(a) The fraction less than 5 meters high is the area to the left of 5, so

$$\text{Fraction} = 5 \cdot 0.05 = 0.25.$$

(b) The fraction above 6 meters high is the area to the right of 6, so

$$\text{Fraction} = (20-6)0.05 = 0.7.$$

(c) The fraction between 2 and 5 meters high is the area between 2 and 5, so

$$\text{Fraction} = (5-2)0.05 = 0.15.$$

2. We use the fact that the area of a triangle is $\frac{1}{2} \cdot \text{Base} \cdot \text{Height}$. Since $p(x)$ is a line with slope $0.1/20 = 0.005$, its equation is

$$p(x) = 0.005x.$$

(a) The fraction less than 5 meters high is the area to the left of 5. Since $p(5) = 0.005(5) = 0.025$,

$$\text{Fraction} = \frac{1}{2} \cdot 5(0.025) = 0.0625.$$

(b) The fraction more than 6 meters high is the area to the right of 6. Since $p(6) = 0.005(6) = 0.03$,

$$\begin{aligned}\text{Fraction} &= 1 - (\text{Area to left of } 6)\\ &= 1 - \frac{1}{2} \cdot 6(0.03) = 0.91.\end{aligned}$$

(c) Fraction between 2 and 5 meters high is area between 2 and 5. Since $p(2) = 0.005(2) = 0.01$,

$$\begin{aligned}\text{Fraction} &= (\text{Area to left of } 5) - (\text{Area to left of } 2)\\ &= 0.0625 - \frac{1}{2} \cdot 2 \cdot (0.01) = 0.0525.\end{aligned}$$

3. We use the fact that the area of a triangle is $\frac{1}{2} \cdot \text{Base} \cdot \text{Height}$. Since $p(x)$ is a line with slope $-0.1/20 = -0.005$ and vertical intercept 0.1, its equation is

$$p(x) = 0.1 - 0.005x.$$

(a) The fraction less than 5 meters high is the area to the left of 5. Since $p(5) = 0.1 - 0.005(5) = 0.075$,

$$\begin{aligned}\text{Fraction} &= 1 - (\text{Area to the right of } 5)\\ &= 1 - \frac{1}{2} \cdot (20-5)0.075 = 0.4375.\end{aligned}$$

(b) The fraction more than 6 meters high is the area to the right of 6. Since $p(6) = 0.1 - 0.005(6) = 0.07$,

$$\text{Fraction} = \frac{1}{2}(20-6)0.07 = 0.49.$$

(c) The fraction between 2 meters and 5 meters high is the area between 2 and 5. Since $p(2) = 0.1 - 0.005(2) = 0.09$,

$$\begin{aligned}\text{Fraction} &= (\text{Area to the right of } 2) - (\text{Area to the right of } 5)\\ &= \left(\frac{1}{2}(20-2)(0.09)\right) - \left(\frac{1}{2}(20-5)(0.075)\right)\\ &= 0.81 - 0.5625 = 0.2475\end{aligned}$$

4. We use the fact that the area of a rectangle is Base $\times$ Height.

(a) The fraction less than 5 meters high is the area to the left of 5, so

$$\text{Fraction } = 5 \cdot 0.1 = 0.5.$$

(b) The fraction above 6 meters high is the area to the right of 6, so

$$\text{Fraction } = (10 - 6) \cdot 0.1 = 0.4.$$

(c) The fraction between 2 and 5 meters high is the area between 2 and 5, so

$$\text{Fraction } = (5 - 2) \cdot 0.1 = 0.3.$$

5. For a small interval Δx around 68, the fraction of the population of American men with heights in this interval is about $(0.2)\Delta x$. For example, taking $\Delta x = 0.1$, we can say that approximately $(0.2)(0.1) = 0.02 = 2\%$ of American men have heights between 68 and 68.1 inches.

6. We can determine the fractions by estimating the area under the curve. Counting the squares for insects in the larval stage between 10 and 12 days we get 4.5 squares, with each square representing (2)·(3%) giving a total of 27% of the insects in the larval stage between 10 and 12 days.
Likewise we get 2 squares for the insects in the larval stage for less than 8 days, giving 12% of the insects in the larval stage for less than 8 days.
Likewise we get 7.5 squares for the insects in the larval stage for more than 12 days, giving 45% of the insects in the larval stage for more than 12 days.
Since the peak of the graph occurs between 12 and 13 days, the length of the larval stage is most likely to fall in this interval.

7. **(a)** The total area under the graph must be 1, so

$$\text{Area } = 5(0.01) + 5C = 1$$

So

$$\begin{aligned} 5C &= 1 - 0.05 = 0.95 \\ C &= 0.19 \end{aligned}$$

(b) The machine is more likely to break in its tenth year than first year. It is equally likely to break in its first year and second year.

(c) Since $p(t)$ is a density function,

$$\begin{aligned}
\begin{array}{c}\text{Fraction of machines}\\ \text{lasting up to 2 years}\end{array} &= \text{Area from } 0 \text{ to } 2 = 2(0.01) = 0.02. \\
\begin{array}{c}\text{Fraction of machines}\\ \text{lasting between 5 and 7 years}\end{array} &= \text{Area from } 5 \text{ to } 7 = 2(0.19) = 0.38. \\
\begin{array}{c}\text{Fraction of machines}\\ \text{lasting between 3 and 6 years}\end{array} &= \text{Area from } 3 \text{ to } 6 \\
&= \text{Area from } 3 \text{ to } 5 + \text{Area from } 5 \text{ to } 6 \\
&= 2(0.01) + 1(0.19) = 0.21.
\end{aligned}$$

8. The fact that most of the area under the graph of the density function is concentrated in two humps, centered at 8 and 12 years, indicates that most of the population belong to one of two groups, those who leave school after finishing approximately 8 years and those who finish about 12 years. There is a smaller group of people who finish approximately 16 years of school.

The percentage of adults who have completed less than ten years of school is equal to the area under the density function to the left of the vertical line at $t = 10$. (See Figure 8.1.) We know that the total area is 1, so we are estimating the percentage of the total area that lies in this shaded part shown in Figure 8.1. A rough estimate of this area is about 30%.

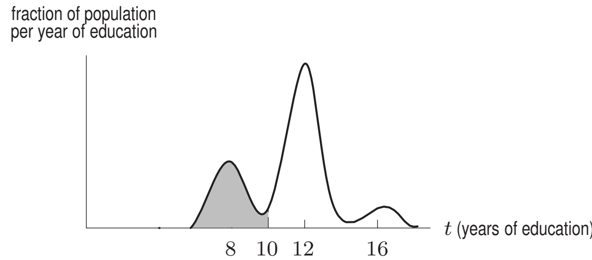


Figure 8.1: What percent has less than 10 years of education?

9. Since $p(x)$ is a density function, the area under the graph of $p(x)$ is 1, so

$$\text{Area} = \text{Base} \cdot \text{Height} = 15a = 1$$
$$a = \frac{1}{15}.$$

10. The area under the graph of $p(x)$ is the sum of two rectangles, each of base 5. Since $\int_0^{10} p(x)\,dx = 1$, we have

$$\int_0^{10} p(x)\,dx = 5 \cdot a + 5 \cdot 2a = 15a = 1$$
$$a = \frac{1}{15}.$$

11. Since $p(x)$ is a density function, the area under the graph of $p(x)$ is 1, so

$$\text{Area} = \frac{1}{2}\text{Base} \cdot \text{Height} = \frac{1}{2} \cdot 10 \cdot a = 5a = 1$$
$$a = \frac{1}{5}.$$

12. Since $p(x)$ is a density function, the area under the graph of $p(x)$ is 1, so

$$\text{Area} = \frac{1}{2} \cdot \text{Base} \cdot \text{Height} = \frac{1}{2} \cdot 100 \cdot a = 50a = 1$$
$$a = \frac{1}{50}.$$

13. If x is the yield in kg, the density function is a horizontal line at $p(x) = 1/100$ for $0 \le x \le 100$. See Figure 8.2.

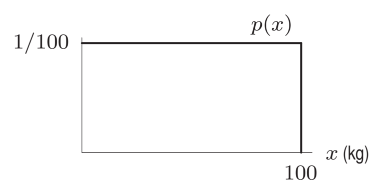


Figure 8.2

14. See Figure 8.3. Many other answers are possible.

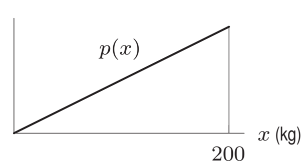


Figure 8.3

15. See Figure 8.4. Many other answers are possible.

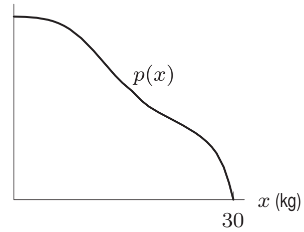


Figure 8.4

16. It is not (a) since a probability density must be a non-negative function; not (c) since the total integral of a probability density must be 1; (b) and (d) are probability density functions, but (d) is not a good model. According to (d), the probability that the next customer comes after 4 minutes is 0. In real life there should be a positive probability of not having a customer in the next 4 minutes. So (b) is the best answer.

Solutions for Section 8.2

1. (a) The cumulative distribution function $P(t)$ is defined to be the fraction of patients who get in to see the doctor within t hours. No one gets in to see the doctor in less than 0 minutes, so $P(0) = 0$. We saw in Example 2 part (c) that 60% of patients wait less than 1 hour, so $P(1) = 0.60$. We saw in part (b) of Example 2 that an additional 37.5% of patients get in to see the doctor within the second hour, so 97.5% of patients will see the doctor within 2 hours; $P(2) = 0.975$. Finally, all patients are admitted within 3 hours, so $P(3) = 1$. Notice also that $P(t) = 1$ for all values of t greater than 3. A table of values for $P(t)$ is given in Table 8.1.

Table 8.1 *Cumulative distribution function for the density function in Example 2*

t	0	1	2	3	4	$\cdots$
$P(t)$	0	0.60	0.975	1	1	$\cdots$

(b) The graph is in Figure 8.5.

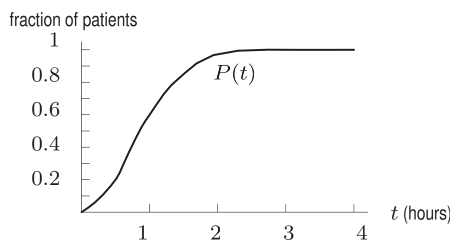


Figure 8.5

2. (a) (i) The probability density function is (III).
(ii) The cumulative distribution function is (VI).
(b) (i) The probability density function is (I).
(ii) The cumulative distribution function is (V).
(c) (i) The probability density function is (IV).
(ii) The cumulative distribution function is (II).

3. **(a)** The shaded region in Figure 8.6 represents the probability that the bus will be from 2 to 4 minutes late.

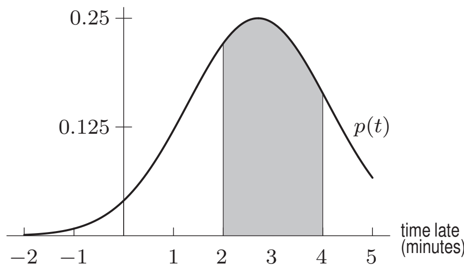


Figure 8.6

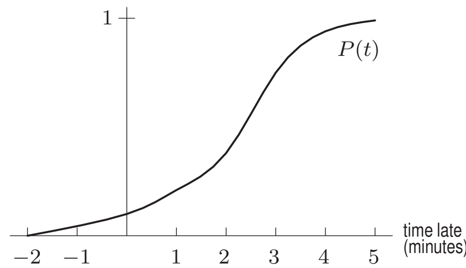


Figure 8.7

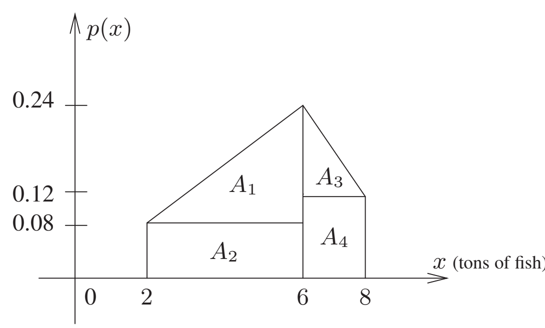
(b) The probability that the bus will be 2 to 4 minutes late (the area shaded in Figure 8.6) is $P(4) - P(2)$. The inflection point on the graph of $P(t)$ in Figure 8.7 corresponds to where $p(t)$ is a maximum. To the left of the inflection point, P is increasing at an increasing rate, while to the right of the inflection point P is increasing at a decreasing rate. Thus, the inflection point marks where the rate at which P is increasing is a maximum (i.e., where the derivative of P, which is p, is a maximum).

4. **(a)** The area under the graph of the height density function $p(x)$ is concentrated in two humps centered at 0.5 m and 1.1 m. The plants can therefore be separated into two groups, those with heights in the range 0.3 m to 0.7 m, corresponding to the first hump, and those with heights in the range 0.9 m to 1.3 m, corresponding to the second hump. This grouping of the grasses according to height is probably close to the species grouping. Since the second hump contains more area than the first, there are more plants of the tall grass species in the meadow.

(b) As do all cumulative distribution functions, the cumulative distribution function $P(x)$ of grass heights rises from 0 to 1 as x increases. Most of this rise is achieved in two spurts, the first as x goes from 0.3 m to 0.7 m, and the second as x goes from 0.9 m to 1.3 m. The plants can therefore be separated into two groups, those with heights in the range 0.3 m to 0.7 m, corresponding to the first spurt, and those with heights in the range 0.9 m to 1.3 m, corresponding to the second spurt. This grouping of the grasses according to height is the same as the grouping we made in part (a), and is probably close to the species grouping.

(c) The fraction of grasses with height less than 0.7 m equals $P(0.7) = 0.25 = 25\%$. The remaining 75% are the tall grasses.

5.

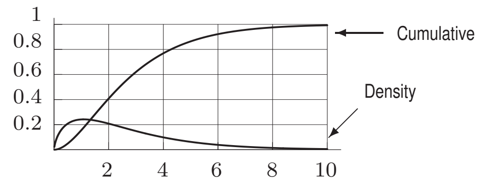


Splitting the figure into four pieces, we see that

$$\begin{aligned}\text{Area under the curve} &= A_1 + A_2 + A_3 + A_4 \\ &= \frac{1}{2}(0.16)4 + 4(0.08) + \frac{1}{2}(0.12)2 + 2(0.12) \\ &= 1.\end{aligned}$$

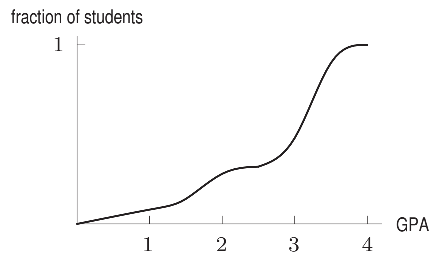
We expect the area to be 1, since $\int_{-\infty}^{\infty} p(x)\,dx = 1$ for any probability density function, and $p(x)$ is 0 except when $2 \le x \le 8$.

6. (a) The two functions are shown below. The choice is based on the fact that the cumulative distribution does not decrease.
(b) The cumulative distribution levels off to 1, so the top mark on the vertical scale must be 1.



The total area under the density function must be 1. Since the area under the density function is about 2.5 boxes, each box must have area $1/2.5 = 0.4$. Since each box has a height of 0.2, the base must be 2.

7. (a) The fraction of students passing is given by the area under the curve from 2 to 4 divided by the total area under the curve. This appears to be about $\frac{2}{3}$.
(b) The fraction with honor grades corresponds to the area under the curve from 3 to 4 divided by the total area. This is about $\frac{1}{3}$.
(c) The peak around 2 probably exists because many students work to get just a passing grade.
(d)



8. (a) $F(7) = 0.6$ tells us that 60% of the trees in the forest have height 7 meters or less.
(b) $F(7) > F(6)$. There are more trees of height less than 7 meters than trees of height less than 6 meters because every tree of height ≤ 6 meters also has height ≤ 7 meters.

9. (a) The first item is sold at the point at which the graph is first greater than zero. Thus the first item is sold at $t = 30$ or January 30. The last item is sold at the t value at which the function is first equal to 100%. Thus the last item is sold at $t = 240$ or August 28, unless its a leap year.
(b) Looking at the graph at $t = 121$ we see that roughly 65% of the inventory has been sold by May 1.
(c) The percent of the inventory sold during May and June is the difference between the percent of the inventory sold on the last day of June and the percent of the inventory sold on the first day of May. Thus, the percent of the inventory sold during May and June is roughly 25%.
(d) The percent of the inventory left after half a year is

$$100 - (\text{percent inventory sold after half year}).$$

Thus, roughly 10% of the inventory is left after half a year.
(e) The items probably went on sale on day 100 and were on sale until day 120. Roughly from April 10 until April 30.

10. (a) Let $P(x)$ be the cumulative distribution function of the heights of the unfertilized plants. As do all cumulative distribution functions, $P(x)$ rises from 0 to 1 as x increases. The greatest number of plants will have heights in the range where $P(x)$ rises the most. The steepest rise appears to occur at about $x = 1$ m. Reading from the graph we see that $P(0.9) \approx 0.2$ and $P(1.1) \approx 0.8$, so that approximately $P(1.1) - P(0.9) = 0.8 - 0.2 = 0.6 = 60\%$ of the unfertilized plants grow to heights between 0.9 m and 1.1 m. Most of the plants grow to heights in the range 0.9 m to 1.1 m.
(b) Let $P_A(x)$ be the cumulative distribution function of the plants that were fertilized with A. Since $P_A(x)$ rises the most in the range 0.7 m $\leq x \leq$ 0.9 m, many of the plants fertilized with A will have heights in the range 0.7 m to 0.9 m. Reading from the graph of P_A, we find that $P_A(0.7) \approx 0.2$ and $P_A(0.9) \approx 0.8$, so $P_A(0.9) - P_A(0.7) \approx 0.8 - 0.2 = 0.6 = 60\%$ of the plants fertilized with A have heights between 0.7 m and 0.9 m. Fertilizer A had the effect of stunting the growth of the plants.

On the other hand, the cumulative distribution function $P_B(x)$ of the heights of the plants fertilized with B rises the most in the range 1.1 m $\leq x \leq$ 1.3 m, so most of these plants have heights in the range 1.1 m to 1.3 m. Fertilizer B caused the plants to grow about 0.2 m taller than they would have with no fertilizer.

11. **(a)** The probability that a banana lasts between 1 and 2 weeks is given by

$$\int_1^2 p(t)dt = 0.25$$

Thus there is a 25% probability that the banana will last between one and two weeks.

(b) The formula given for $p(t)$ is valid for up to four weeks; for $t > 4$ we have $p(t) = 0$. So a banana lasting more than 3 weeks must last between 3 and 4 weeks. Thus the probability is

$$\int_3^4 p(t)dt = 0.325$$

32.5% of the bananas last more than 3 weeks.

(c) Since $p(t) = 0$ for $t > 4$, the probability that a banana lasts more than 4 weeks is 0.

12. **(a)** The cumulative distribution, $P(t)$, is the function whose slope is the density function $p(t)$. So $P'(t) = p(t)$. The graph of $P(t)$ starts out with a small slope at $t = 0$; its slope increases as t increases to 3. The graph of $P(t)$ levels off at 1 for $t \geq 4$. See Figure 8.8.

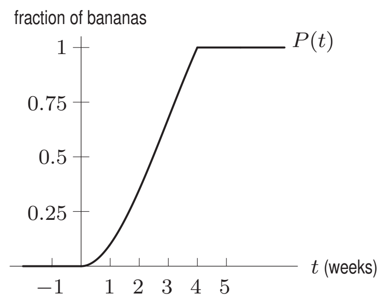


Figure 8.8

(b) The probability that a banana will last between 1 and 2 weeks is given by the difference $P(2) - P(1)$ where $P'(t) = p(t)$ and $p(t)$ is the density function. Looking at Figure 8.8 we see that the difference is roughly $0.25 = 25\%$.

13. The cumulative distribution function

$$\begin{aligned} P(t) = \int_0^t p(x)dx &= \text{Area under graph of density function } p(x) \text{ for } 0 \leq x \leq t \\ &= \text{Fraction of population who die } t \text{ years or less after treatment} \\ &= \text{Fraction of population who are dead at time } t. \end{aligned}$$

14. **(a)** Since $d(e^{-ct})/dt = ce^{-ct}$, we have

$$c\int_0^6 e^{-ct}dt = -e^{-ct}\Big|_0^6 = 1 - e^{-6c} = 0.1,$$

so

$$c = -\frac{1}{6}\ln 0.9 \approx 0.0176.$$

(b) Similarly, with $c = 0.0176$, we have

$$\begin{aligned} c\int_6^{12} e^{-ct}dt &= -e^{-ct}\Big|_6^{12} \\ &= e^{-6c} - e^{-12c} = 0.9 - 0.81 = 0.09, \end{aligned}$$

so the probability is 9%.

15. **(a)** The probability you dropped the glove within a kilometer of home is given by

$$\int_0^1 2e^{-2x}dx = -e^{-2x}\Big|_0^1 = -e^{-2}+1 \approx 0.865.$$

(b) Since the probability that the glove was dropped within y km $= \int_0^y p(x)dx = 1-e^{-2y}$, we solve

$$\begin{aligned} 1-e^{-2y} &= 0.95 \\ e^{-2y} &= 0.05 \\ y &= \frac{\ln 0.05}{-2} \approx 1.5 \text{ km.} \end{aligned}$$

Solutions for Section 8.3

1. The median daily catch is the amount of fish such that half the time a boat will bring back more fish and half the time a boat will bring back less fish. Thus the area under the curve and to the left of the median must be 0.5. There are 25 squares under the curve so the median occurs at 12.5 squares of area. Now

$$\int_2^5 p(x)dx = 10.5 \text{ squares}$$

and

$$\int_5^6 p(x)dx = 5.5 \text{ squares,}$$

so the median occurs at a little over 5 tons. We must find the value a for which

$$\int_5^a p(t)dt = 2 \text{ squares,}$$

and we note that this occurs at about $a = 0.35$. Hence

$$\begin{aligned} \int_2^{5.35} p(t)\,dt &\approx 12.5 \text{ squares} \\ &\approx 0.5. \end{aligned}$$

The median is about 5.35 tons.

2. **(a)** The median corresponds to the value of t such that $P(t) = 0.5$. Since $P(36) \approx 0.5$, the median ≈ 36.

(b) The density function is positive wherever the derivative of $P(t)$ is positive, namely from about $t = 5$ until roughly $t = 65$. The derivative function is increasing everywhere $P(t)$ is concave up. So that the density function is increasing until about $t = 35$ and is decreasing after that. The local maximum is where the function is changing from increasing to decreasing so that $t = 35$ is a local maximum.

3. **(a)** The normal distribution of car speeds with $\mu = 58$ and $\sigma = 4$ is shown in Figure 8.9.

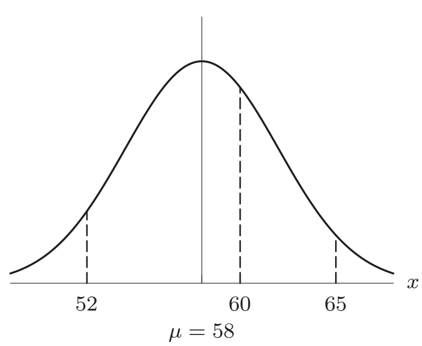


Figure 8.9

The probability that a randomly selected car is going between 60 and 65 is equal to the area under the curve from $x = 60$ to $x = 65$,

$$\text{Probability} = \frac{1}{4\sqrt{2\pi}} \int_{60}^{65} e^{-(x-58)^2/(2\cdot 4^2)}\, dx \approx 0.2685.$$

We obtain the value 0.2685 using a calculator or computer.

(b) To find the fraction of cars going under 52 km/hr, we evaluate the integral

$$\text{Fraction} = \frac{1}{4\sqrt{2\pi}} \int_{0}^{52} e^{-(x-58)^2/32} dx \approx 0.067.$$

Thus, approximately 6.7% of the cars are going less than 52 km/hr.

4. **(a)** Since $\mu = 100$ and $\sigma = 15$:

$$p(x) = \frac{1}{15\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-100}{15}\right)^2}.$$

(b) The fraction of the population with IQ scores between 115 and 120 is (integrating numerically)

$$\begin{aligned}\int_{115}^{120} p(x)\, dx &= \int_{115}^{120} \frac{1}{15\sqrt{2\pi}} e^{-\frac{(x-100)^2}{450}}\, dx \\ &= \frac{1}{15\sqrt{2\pi}} \int_{115}^{120} e^{-\frac{(x-100)^2}{450}}\, dx \\ &\approx 0.067 = 6.7\% \text{ of the population.}\end{aligned}$$

5. We know that the median is given by T such that

$$\int_{-\infty}^{T} p(t) dt = 0.5.$$

Trying different values of T, we find that $0.5 = \int_0^T p(t)dt$ for $T \approx 2.48$ weeks. Figure 8.10 supports the conclusion that $t = 2.48$ is in fact the median.

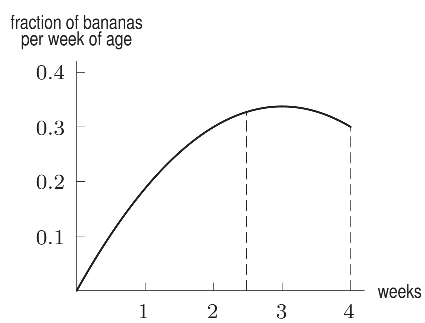


Figure 8.10

6. We know that the mean is given by

$$\int_{-\infty}^{\infty} tp(t)dt.$$

Thus we get

$$\begin{aligned}\text{Mean} &= \int_0^4 tp(t)dt \\ &= \int_0^4 (-0.0375t^3 + 0.225t^2)dt \\ &\approx 2.4\end{aligned}$$

Thus the mean is 2.4 weeks. Figure 8.11 supports the conclusion that $t = 2.4$ is in fact the mean.

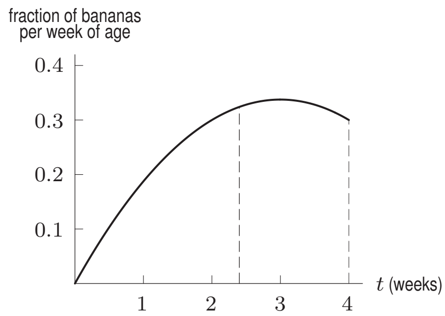


Figure 8.11

7. **(a)** We can find the proportion of students by integrating the density $p(x)$ between $x = 1.5$ and $x = 2$:

$$\begin{aligned} P(2) - P(1.5) &= \int_{1.5}^{2} \frac{x^3}{4}\,dx \\ &= \left.\frac{x^4}{16}\right|_{1.5}^{2} \\ &= \frac{(2)^4}{16} - \frac{(1.5)^4}{16} = 0.684, \end{aligned}$$

so that the proportion is $0.684 : 1$ or 68.4%.

(b) We find the mean by integrating x times the density over the relevant range:

$$\begin{aligned} \text{Mean} &= \int_0^2 x\left(\frac{x^3}{4}\right)\,dx \\ &= \int_0^2 \frac{x^4}{4}\,dx \\ &= \left.\frac{x^5}{20}\right|_0^2 \\ &= \frac{2^5}{20} = 1.6 \text{ hours.} \end{aligned}$$

(c) The median will be the time T such that exactly half of the students are finished by time T, or in other words

$$\begin{aligned} \frac{1}{2} &= \int_0^T \frac{x^3}{4}\,dx \\ \frac{1}{2} &= \left.\frac{x^4}{16}\right|_0^T \\ \frac{1}{2} &= \frac{T^4}{16} \\ T &= \sqrt[4]{8} = 1.682 \text{ hours.} \end{aligned}$$

8. (a) See Figure 8.12. The mean is larger than the median for this distribution; both are less than 15.

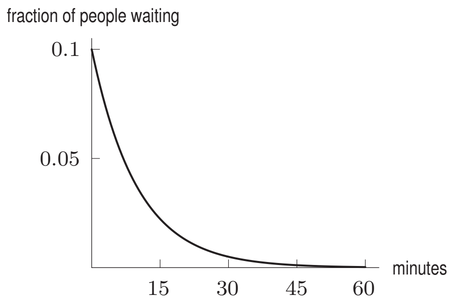


Figure 8.12

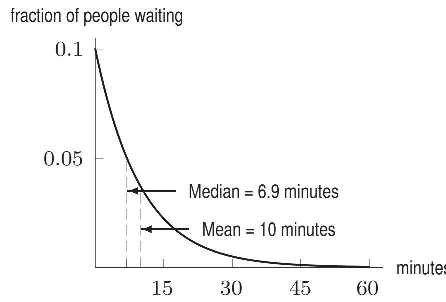


Figure 8.13

(b) We know that the median is the value T such that

$$\int_{-\infty}^{T} p(t)dt = 0.5$$

In our case this gives

$$\begin{aligned} 0.5 &= \int_0^T p(t)dt \\ &= \int_0^T 0.1e^{-0.1t}dt \end{aligned}$$

Substituting different values of T we get

$$\text{Median} = T \approx 6.9.$$

See Figure 8.13. We know that the mean is given by

$$\begin{aligned} \text{Mean} &= \int_{-\infty}^{\infty} tp(t)dt \\ &= \int_0^{60} (0.1te^{-0.1t})dt \\ &\approx 9.83. \end{aligned}$$

(c) The median tells us that exactly half of the people waiting at the stop wait less than 6.9 minutes.

The fact that the mean is 9.83 minutes can be interpreted in the following way: If all the people waiting at the stop were to wait exactly 9.83 minutes, the total time waited would be the same.

9. (a) Since P is the cumulative distribution function, the percentage of households that made between \$40,000 and \$60,000 is

$$P(60) - P(40) = 63\% - 45.4\% = 17.6\%.$$

Therefore 17.6% of the households made between \$40,000 and \$60,000.

The percentage of households making over $\$150{,}000$ is $100\% - 94.0\% = 6.0\%$.

(b) The median income is the income such that half the households make less than this amount. Looking at the table, we see that the 50% mark occurs between $\$40{,}000$ and $\$60{,}000$. Since $P(40) = 45.4\%$, we know 45.4% of the households made less than $\$40{,}000$. Assuming local linearity we calculate the slope of the line connecting $(40, 45.4)$ with $(60, 63.0)$ as $(63.0 - 45.4)/20 = 0.88$. If x is the additional income it takes to achieve the median, then $0.88x = 50.0 - 45.4$, so $x = 5.23$. Since $40 + 5.23 = 45.23$, the median income is approximately $\$45{,}200$.

(c) The percentage of households that made between $\$40{,}000$ and $\$80{,}000$ is $75.8 - 45.4 = 30.4$. Since this percentage is less than $1/3$, the statement is false.

Solutions for Chapter 8 Review

1. The two humps of probability in density (a) correspond to two intervals on which its cumulative distribution function is increasing. Thus (a) and (II) correspond.

A density function increases where its cumulative distribution funciton is concave up, and it decreases where its cumulative distribution function is concave down. Density (b) matches the distribution with both concave up and concave down sections, which is (I). Density (c) matches (III) which has a concave down section but no interval over which it is concave up.

2. Suppose x is the age of death; a possible density function is graphed in Figure 8.14.

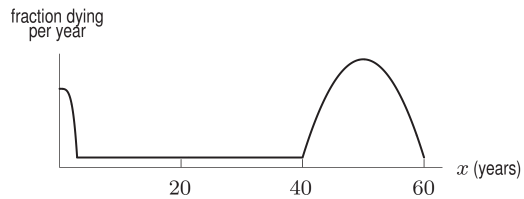


Figure 8.14

3. If x is the height in feet of each person, a possible density function is graphed in Figure 8.15. The teachers are included.

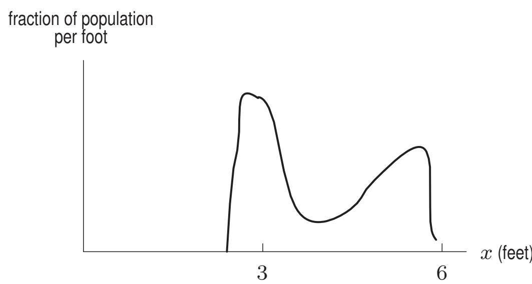


Figure 8.15

4.

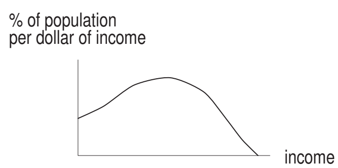


Figure 8.16: Density function

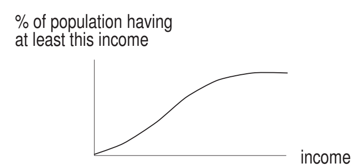


Figure 8.17: Cumulative distribution function

5.

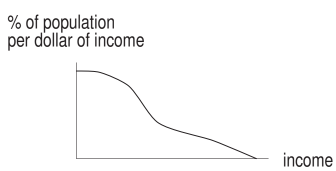


Figure 8.18: Density function

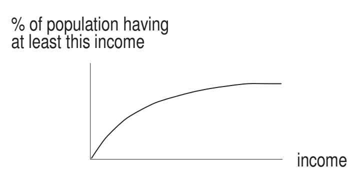


Figure 8.19: Cumulative distribution function

6.

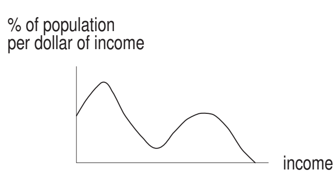


Figure 8.20: Density function

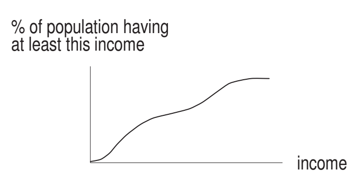


Figure 8.21: Cumulative distribution function

7. Since $p(x)$ is a density function,

$$\text{Area under graph} = \frac{1}{2}\cdot 50c = 25c = 1,$$

so $c = 1/25 = 0.04$.

8. Since $p(t)$ is a density function,

$$\text{Area under graph} = \frac{1}{2}\cdot c\cdot 0.01 = 0.005c = 1,$$

so $c = 1/0.005 = 200$.

9. Since $p(t)$ is a density function,

$$\text{Area under graph} = 50\cdot 2c + 25\cdot c = 125c = 1,$$

so $c = 1/125 = 0.008$.

10. Since $p(x) = cx$, we know $p(2) = 2c$. Since $p(x)$ is a density function,

$$\text{Area under graph} = \frac{1}{2}\cdot 2\cdot 2c = 2c = 1,$$

so $c = 1/2 = 0.5$.

11. **(a)** More insects die in the twelfth month than the first month because $P(t)$ is larger in the twelfth month. This means that the area $\int_{11}^{12} p(t)\,dt$ is larger than the area $\int_0^1 p(t)\,dt$.

(b) We want to find the fraction dying within the first 6 months. Since $p(12) = 1/6$, we have $p(6) = 1/12$, so

$$\begin{aligned}\text{Fraction living up to 6 months} &= \int_0^6 p(t)\,dt\\ &= \text{Area from } t=0 \text{ to } t=6\\ &= \frac{1}{2}\cdot 6\cdot\frac{1}{12} = \frac{1}{4}.\end{aligned}$$

So 1/4 of population lives no more than 6 months.

(c) We can first find the fraction of insects who die within the first 9 months. Using $p(9) = \frac{1}{6}\cdot\frac{9}{12} = \frac{1}{8}$, we have

$$\begin{aligned}\text{Fraction living up to 9 months} = \int_0^9 p(t)\,dt &= \text{Area from } t=0 \text{ to } t=9\\ &= \frac{1}{2}\cdot 9\cdot\frac{1}{8} = \frac{9}{16}.\end{aligned}$$

So the quantity we want is

$$\text{Fraction living more than 9 months } = 1 - \frac{9}{16} = \frac{7}{16}.$$

12. No. Though the density function has its maximum value at 50, this does not mean that a large fraction of the population receives scores near 50. The value $p(50)$ can not be interpreted as a probability. Probability corresponds to *area* under the graph of a density function. Most of the area in this case is in the broad hump covering the range $0 \le x \le 40$, very little in the peak around $x = 50$. Most people score in the range $0 \le x \le 40$.

13. **(a)** Most of the earth's surface is below sea level. Much of the earth's surface is either around 3 miles below sea level or exactly at sea level. It appears that essentially all of the surface is between 4 miles below sea level and 2 miles above sea level. Very little of the surface is around 1 mile below sea level.

(b) The fraction below sea level corresponds to the area under the curve from -4 to 0 divided by the total area under the curve. This appears to be about $\frac{3}{4}$.

14. **(a)** See Figure 8.22. This is a cumulative distribution function.

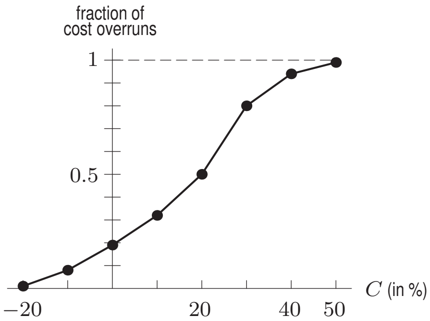


Figure 8.22

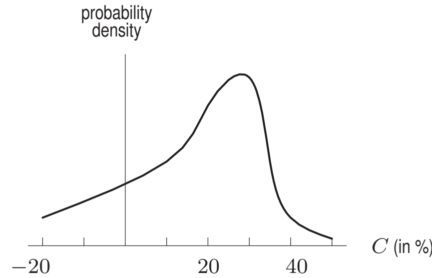


Figure 8.23

(b) The density function is the derivative of the cumulative distribution function. See Figure 8.23.

(c) Let's call the cumulative distribution function $F(C)$. The probability that there will be a cost overrun of more than 50% is $1 - F(50) = 0.01$, a 1% chance. The probability that it will be between 20% and 50% is $F(50) - F(20) = 0.99 - 0.50 = 0.49$, or 49%. The most likely amount of cost overrun occurs when the slope of the tangent line to the cumulative distribution function is a maximum. This occurs at the inflection point of the cumulative distribution graph, at about $C = 28\%$.

15. **(a)** The density function $f(r)$ will be zero outside the range $0 < r < 5$ and equal to a nonzero constant k inside this range. The area of the region under the density curve equals $5k$, which must equal 1, so $k = 0.2$. We have

$$f(r) = \begin{cases} 0 & \text{if } r \le 0 \\ 0.2 & \text{if } 0 < r < 5 \\ 0 & \text{if } 5 \le r. \end{cases}$$

The graph of $f(r)$ is given in Figure 8.24.

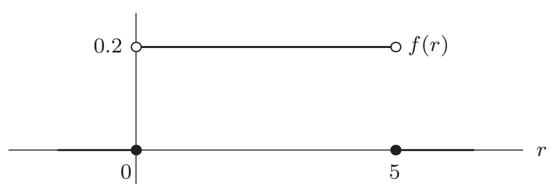


Figure 8.24

(b) The cumulative distribution function $F(r)$ equals the area of the region under the density function up to r. From the graph in Figure 8.24 we see that the area is zero if $r < 0$; for $0 \le r \le 5$ the region is rectangular of height 0.2, width r, and area $0.2r$; and for $r > 5$ the area is 1. Thus

$$F(r) = \begin{cases} 0 & \text{if } r < 0 \\ 0.2r & \text{if } 0 \le r \le 5 \\ 1 & \text{if } 5 < r. \end{cases}$$

16. Since Product D is absorbed most quickly, this is the solution. We see that 80% of the theophylline solution is absorbed within an hour, and all of it within 5 or 6 hours. As we would expect, the timed-release capsules are absorbed more slowly. Product A is absorbed slightly faster than Product B, although both are close to fully absorbed after 24 hours. However, we see very slow absorption with Product C. Even after 28 hours have passed, only about 60% of the drug has been absorbed.

17. **(a)** The percentage of calls lasting from 1 to 2 minutes is given by the integral

$$\int_1^2 p(x)\,dx \approx 0.221 = 22.1\%.$$

(b) A similar calculation (changing the limits of integration) gives the percentage of calls lasting 1 minute or less as

$$\int_0^1 p(x)\,dx \approx 0.33 = 33.0\%.$$

(c) The percentage of calls lasting 3 minutes or more is given by the improper integral

$$\int_3^\infty p(x)\,dx \approx 0.301 = 30.1\%.$$

(d) The cumulative distribution function is the integral of the probability density; thus,

$$C(h) = \int_0^h p(x)\,dx = \int_0^h 0.4e^{-0.4x}\,dx = 1 - e^{-0.4h}.$$

18. Figure 8.25 is a graph of the density function; Figure 8.26 is a graph of the cumulative distribution.

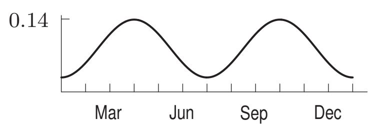


Figure 8.25: Density function

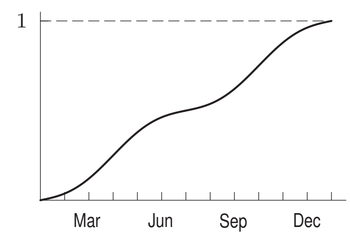


Figure 8.26: Cumulative distribution function

19. Since $p(t) = 0.04 - 0.0008t$, the cumulative distribution which satisfies $P'(t) = p(t)$ is given by

$$\begin{aligned} P(t) &= 0.04t - \frac{0.0008t^2}{2} + C \\ &= 0.04t - 0.0004t^2 + C. \end{aligned}$$

Since $P(0) = 0$, we have $C = 0$, so

$$P(t) = 0.04t - 0.0004t^2.$$

For the median T,

$$P(T) = 0.04T - 0.0004T^2 = 0.5.$$

Solving the quadratic equation

$$0.0004T^2 - 0.04T + 0.5 = 0$$

gives

$$T = \frac{0.04 \pm \sqrt{(0.04)^2 - 4(0.0004)(0.5)}}{2(0.0004)}.$$

Evaluating gives $T = 85.4$ and 14.6. Since $p(t)$ is not defined for $t > 50$ [], the median is $T = 14.6$ days.

20. The median is the value T such that

$$\int_{-\infty}^{T} p(x)dx = 0.5$$

Thus we get

$$\begin{aligned} 0.5 &= \int_{-\infty}^{T} p(x)dx \\ &= \int_0^T 0.122e^{-0.122x}\,dx \end{aligned}$$

Substituting different values for T we get $T \approx 5.68$. Thus the median occurs at 5.68 seconds
We know that the mean is

$$\int_{-\infty}^{\infty} xp(x)dx.$$

Thus we get

$$\begin{aligned}\text{Mean } &= \int_{-\infty}^{\infty} xp(x)dx \\ &= \int_{0}^{40} x(0.122e^{-0.122x})dx \approx 7.83 \text{ seconds}.\end{aligned}$$

The median tells us that fifty percent of the time gaps between cars are less than 5.68 seconds, and fifty percent of the time gaps between cars are more than 5.68 seconds. The mean tells us that over all time gaps, the average time gap between cars is 7.83 seconds.

21. False. Note that p is the density function for the population, not the cumulative density function. Thus $p(10) = 1/2$ means that the probability of x lying in a small interval of length Δx around $x = 10$ is about $(1/2)\Delta x$.

22. True. This follows directly from the definition of the cumulative density function.

23. True. The interval from $x = 9.98$ to $x = 10.04$ has length 0.06. Assuming that the value of $p(x)$ is near $1/2$ for $9.98 < x < 10.04$, the fraction of the population in that interval is $\int_{9.98}^{10.04} p(x)dx \approx (1/2)(0.06) = 0.03$.

24. False. Note that p is the density function for the population, not the cumulative density function. Thus $p(10) = p(20)$ means that x values near 10 are as likely as x values near 20.

25. True. By the definition of the cumulative distribution function, $P(20) - P(10) = 0$ is the fraction of the population having x values between 10 and 20.

CHECK YOUR UNDERSTANDING

1. False, since we also need $p(x) \geq 0$.

2. True, since both $\int_{-\infty}^{\infty} p(x)\,dx = 1$ and $p(x) \geq 0$.

3. True, since this is the interpretation of density functions given in the text.

4. True, since this is the interpretation of density functions given in the text.

5. False. The integral $\int_0^{70} p(x)\,dx$ represents the fraction of the population between the ages of 0 and 70 years old. The fraction will be significantly more than 50% of the population.

6. False, since we have to multiply $p(50)$ by an age interval to obtain a fraction of population. We can interpret $p(50)\cdot 1 = p(50)(50.5 - 49.5)$ as approximately the fraction of the population between 49.5 and 50.5 years old.

7. True, since $p(10)(10.5 - 9.5) = 0.014$, so about 1.4 percent has ages between 9.5 and 10.5 years.

8. False. For example, if $p(t)$ is an age density distribution, then $\int_0^{\infty} p(x)\,dx = 1$.

9. True, since the population with values between 0 and 1 is included in the population with values between 0 and 2.

10. False. It is possible for $p(x) = 0$ for $100 \leq x \leq 200$.

11. True, since as t increases, the fraction of the population that is less than t must stay the same or increase.

12. False. It should be $P' = p$.

13. False. $P(30)$ is the fraction of the population with values at or below 30.

14. True, since $P(20)$ is the fraction of the population with values less than 20, and $P(10)$ is the fraction of the population with values less than 10.

15. False. We have $P(30) - P(10)$ is the fraction of the population with values between 10 and 30, while $P(20)$ is the fraction of the population with values at or below 20. Clearly, the two can be different (for example, for age distribution).

16. True, since this is the definition of the cumulative distribution function.

17. False. The units of $p(x)$ are fraction of population per unit of x, while the units for $P(x)$ are fraction of population.

18. True, since the fraction having value at most t cannot be less than zero or more than one.

19. False. We know that $P(25)$ is the fraction less than 25, which we cannot determine from the information about the fraction less than 10 or less than 15.

20. True. This follows from the Fundamental Theorem of Calculus, since $P' = p$. Both expressions represent the fraction of values between 25 and 50.

21. True, this is the definition of mean value as given in the text.

22. False. The integral for the mean value uses the distribution function $p(t)$, not the cumulative distribution function.

23. False. The median T satisfies $P(T) = 0.5$.

24. True. The given integral is just $P(T)$, which is 0.5.

25. True, since the median T is the value for which one-half of the population is greater than or equal to T.

26. False. For example, if $p(x)$ is a density function with a normal distribution having mean 0, then 0 is both the median and the mean.

27. True, this is the definition of the normal distribution as specified in the text.

28. False, the given distribution has mean $\mu = 5$.

29. True, since the integrand is a normal distribution with mean $\mu = 7$ and standard deviation $\sigma = 1$.

30. False, T is the median, not the mean.

PROJECTS FOR CHAPTER EIGHT

1. (a) The area under the graph of $p(x)$ from $x = a$ to $x = b$ should be 1.
Therefore $1 = \frac{1}{2}(\text{base}) \cdot (\text{height}) = \frac{1}{2}(b-a) \cdot p(c)$, and so $p(c) = \frac{2}{b-a}$.

(b) We have $a = 6$, $b = 10$, and $c = 9$. Using $p(c) = \frac{2}{b-a}$, we have $p(9) = \frac{2}{10-6} = 0.5$.

(c) To find the equation for the first line, we note that $(6, 0)$ and $(9, 0.5)$ are points on the line, so $m_1 = \Delta y/\Delta x = (0.5-0)/(9-6) = 1/6$. And if $y = m_1 x + b_1$, substituting in $x = 6$, $y = 0$, yields $0 = (1/6)6 + b_1$, so $b_1 = -1$.

To find the equation for the second line, we note that $(10, 0)$ and $(9, 0.5)$ are points on the line, so $m_2 = \Delta y/\Delta x = (0.5-0)/(9-10) = -1/2$. And if $y = m_2 x + b_2$, substituting in $x = 10$, $y = 0$, yields $0 = (-1/2)10 + b_2$, so $b_2 = 5$.
Thus

$$p(x) = \begin{cases} \frac{x}{6} - 1 & 6 \le x \le 9 \\ -\frac{x}{2} + 5 & 9 \le x \le 10. \end{cases}$$

(d) The probability that production cost is under \$8 is given by the area under $p(x)$ from $x = 6$ to $x = 8$. The area is $\frac{1}{2}(8-6) \cdot p(8) = p(8) = \frac{8}{6} - 1 = \frac{1}{3}$.

(e) The median cost is the value $x = m$ such that the area under $p(x)$ from $x = 6$ to $x = m$ is $\frac{1}{2}$. To find m, we solve $\frac{1}{2}(m-6) \cdot p(m) = \frac{1}{2}$, or $(m-6) \cdot p(m) = 1$. We know that $6 \le m \le 9$ since the area under $p(x)$ from $x = 6$ to $x = 9$ is greater than $\frac{1}{2}$. Therefore $p(m) = \frac{m}{6} - 1$ and so

$$\begin{aligned} (m-6) \cdot \left(\frac{m}{6} - 1\right) &= 1 \\ \frac{m^2}{6} - 2m + 6 &= 1 \\ m^2 - 12m + 36 &= 6 \\ m^2 - 12m + 30 &= 0 \end{aligned}$$

By the quadratic formula,

$$m = \frac{12 \pm \sqrt{144 - 120}}{2} = 6 \pm \sqrt{6}.$$

Since $6 \le m \le 9$, we must have $m = 6 + \sqrt{6}$.

(f) The cumulative probability distribution function $P(x)$ gives the area under $p(x)$ from $x = a$ to x.

(i) For $6 \le x \le 9$, the area is given by

$$\begin{aligned} \frac{1}{2}(x-6) \cdot p(x) = \frac{1}{2}(x-6) \cdot \left(\frac{x}{6} - 1\right) &= \frac{1}{2}\left(\frac{x^2}{6} - 2x + 6\right) \\ &= \frac{x^2}{12} - x + 3 \end{aligned}$$

(ii) For $9 \leq x \leq 10$, the area is given by the area under $p(x)$ from $x = 6$ to $x = 10$, which is equal to 1, minus the area under $p(x)$ from x to $x = 10$, which is given by $\frac{1}{2} \cdot (10 - x) \cdot p(x)$. Since $9 \leq x \leq 10$, $p(x) = -\frac{x}{2} + 5$.
Therefore
$$P(x) = 1 - \frac{1}{2}(10 - x) \cdot (-\frac{x}{2} + 5) = -\frac{x^2}{4} + 5x - 24.$$
Therefore
$$P(x) = \begin{cases} \frac{x^2}{12} - x + 3 & 6 \leq x \leq 9 \\ -\frac{x^2}{4} + 5x - 24 & 9 \leq x \leq 10. \end{cases}$$
and its graph is given in Figure 8.27.

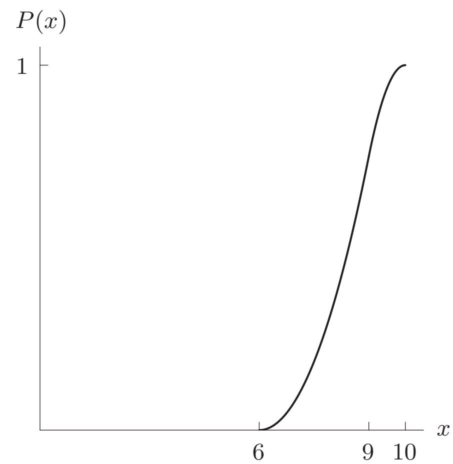


Figure 8.27