ACH2053 – Introdução à Estatística

Exercícios recomendados
Capítulo 3: Variáveis Aleatórias e Distribuições

Fonte: Morris DeGroot, Mark Schervish. Probability and Statistics. 4th Ed.

Seção 3.1:

1. Suppose that a random variable X has the uniform distribution on the integers $10, \ldots, 20$. Find the probability that X is even.

2. Suppose that a random variable X has a discrete distribution with the following p.f.:

$$f(x) = \begin{cases} cx & \text{for } x = 1, \ldots, 5, \\ 0 & \text{otherwise.} \end{cases}$$

Determine the value of the constant c.

3. Suppose that two balanced dice are rolled, and let X denote the absolute value of the difference between the two numbers that appear. Determine and sketch the p.f. of X.

4. Suppose that a fair coin is tossed 10 times independently. Determine the p.f. of the number of heads that will be obtained.

5. Suppose that a box contains seven red balls and three blue balls. If five balls are selected at random, without replacement, determine the p.f. of the number of red balls that will be obtained.

6. Suppose that a random variable X has the binomial distribution with parameters $n = 15$ and $p = 0.5$. Find $\Pr(X < 6)$.

Dica: para cálculo da distribuição binomial, utilize a função `dbinom` do R. Para saber como chamar a função, digite

```
?dbinom
```

no shell do R ou do RStudio.

8. If 10 percent of the balls in a certain box are red, and if 20 balls are selected from the box at random, with replacement, what is the probability that more than three red balls will be obtained?

9. Suppose that a random variable X has a discrete distribution with the following p.f.:

$$f(x) = \begin{cases} \frac{c}{2^x} & \text{for } x = 0, 1, 2, \ldots, \\ 0 & \text{otherwise.} \end{cases}$$

Find the value of the constant c.

Dica: Apêndice A do Livro "Introduction to Algorithms" (Cormen, Leiserson, Rivest, Stein) contém limites para algumas séries tradicionais (incluindo o somatório com o termo $c/2^x$ acima).

Seção 3.2:

3. Suppose that the p.d.f. of a random variable X is as follows:

$$f(x) = \begin{cases} \frac{1}{36}(9 - x^2) & \text{for } -3 \leq x \leq 3, \\ 0 & \text{otherwise.} \end{cases}$$

Sketch this p.d.f. and determine the values of the following probabilities: **a.** $\Pr(X < 0)$ **b.** $\Pr(-1 \leq X \leq 1)$
c. $\Pr(X > 2)$.

4. Suppose that the p.d.f. of a random variable X is as follows:

$$f(x) = \begin{cases} cx^2 & \text{for } 1 \leq x \leq 2, \\ 0 & \text{otherwise.} \end{cases}$$

a. Find the value of the constant c and sketch the p.d.f.

b. Find the value of $\Pr(X > 3/2)$.

5. Suppose that the p.d.f. of a random variable X is as follows:

$$f(x) = \begin{cases} \frac{1}{8}x & \text{for } 0 \leq x \leq 4, \\ 0 & \text{otherwise.} \end{cases}$$

a. Find the value of t such that $\Pr(X \leq t) = 1/4$.

b. Find the value of t such that $\Pr(X \geq t) = 1/2$.

7. Suppose that a random variable X has the uniform distribution on the interval $[-2, 8]$. Find the p.d.f. of X and the value of $\Pr(0 < X < 7)$.

8. Suppose that the p.d.f. of a random variable X is as follows:

$$f(x) = \begin{cases} ce^{-2x} & \text{for } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

a. Find the value of the constant c and sketch the p.d.f.

b. Find the value of $\Pr(1 < X < 2)$.

12. In Example 3.1.3 on page 94, determine the distribution of the random variable Y, the electricity demand. Also, find $\Pr(Y < 50)$.

13. An ice cream seller takes 20 gallons of ice cream in her truck each day. Let X stand for the number of gallons that she sells. The probability is 0.1 that $X = 20$. If she doesn't sell all 20 gallons, the distribution of X follows a continuous distribution with a p.d.f. of the form

$$f(x) = \begin{cases} cx & \text{for } 0 < x < 20, \\ 0 & \text{otherwise,} \end{cases}$$

where c is a constant that makes $\Pr(X < 20) = 0.9$. Find the constant c so that $\Pr(X < 20) = 0.9$ as described above.

Seção 3.3:

1. Suppose that a random variable X has the Bernoulli distribution with parameter $p = 0.7$. (See Definition 3.1.5.) Sketch the c.d.f. of X.

2. Suppose that a random variable X can take only the values -2, 0, 1, and 4, and that the probabilities of these values are as follows: $\Pr(X = -2) = 0.4$, $\Pr(X = 0) = 0.1$, $\Pr(X = 1) = 0.3$, and $\Pr(X = 4) = 0.2$. Sketch the c.d.f. of X.

3. Suppose that a coin is tossed repeatedly until a head is obtained for the first time, and let X denote the number of tosses that are required. Sketch the c.d.f. of X.

4. Suppose that the c.d.f. F of a random variable X is as sketched in Fig. 3.9. Find each of the following probabilities:

a. $\Pr(X = -1)$ b. $\Pr(X < 0)$
c. $\Pr(X \le 0)$ d. $\Pr(X = 1)$
e. $\Pr(0 < X \le 3)$ f. $\Pr(0 < X < 3)$
g. $\Pr(0 \le X \le 3)$ h. $\Pr(1 < X \le 2)$
i. $\Pr(1 \le X \le 2)$ j. $\Pr(X > 5)$
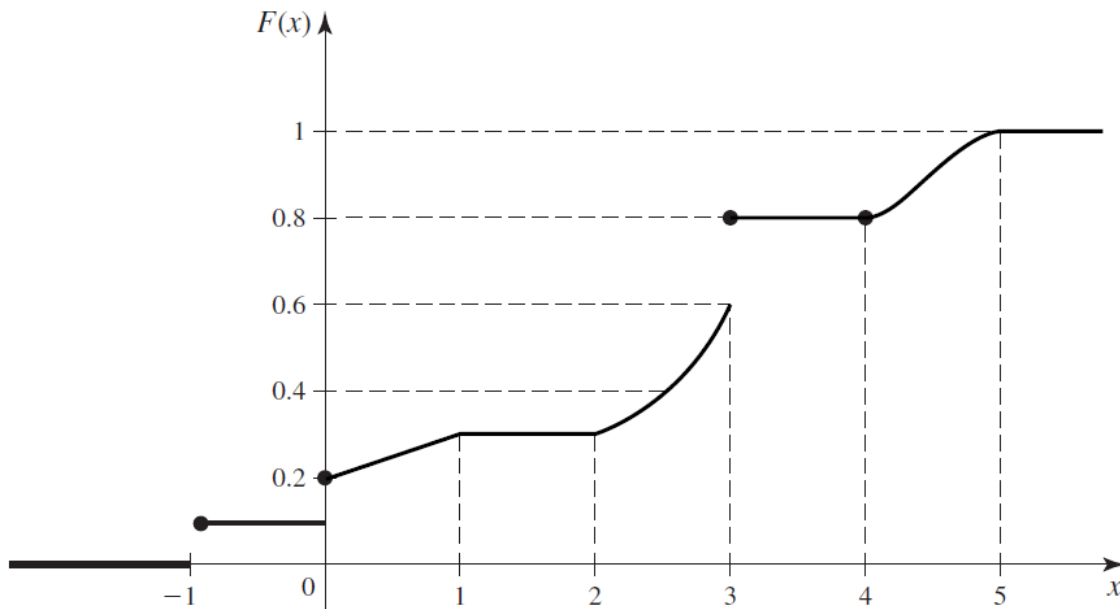k. $\Pr(X \ge 5)$ l. $\Pr(3 \le X \le 4)$

Figure 3.9 The c.d.f. for Exercise 4.

5. Suppose that the c.d.f. of a random variable X is as follows:

$$F(x) = \begin{cases} 0 & \text{for } x \le 0, \\ \frac{1}{9}x^2 & \text{for } 0 < x \le 3, \\ 1 & \text{for } x > 3. \end{cases}$$

Find and sketch the p.d.f. of X.

6. Suppose that the c.d.f. of a random variable X is as follows:

$$F(x) = \begin{cases} e^{x-3} & \text{for } x \le 3, \\ 1 & \text{for } x > 3. \end{cases}$$

Find and sketch the p.d.f. of X.

7. Suppose, as in Exercise 7 of Sec. 3.2, that a random variable X has the uniform distribution on the interval $[-2, 8]$. Find and sketch the c.d.f. of X.

9. Suppose that X has the uniform distribution on the interval $[0, 5]$ and that the random variable Y is defined by $Y = 0$ if $X \leq 1$, $Y = 5$ if $X \geq 3$, and $Y = X$ otherwise. Sketch the c.d.f. of Y.

10. For the c.d.f. in Example 3.3.4, find the quantile function.

11. For the c.d.f. in Exercise 5, find the quantile function.

12. For the c.d.f. in Exercise 6, find the quantile function.

13. Suppose that a broker believes that the change in value X of a particular investment over the next two months has the uniform distribution on the interval $[-12, 24]$. Find the value at risk VaR for two months at probability level 0.95.

14. Find the quartiles and the median of the binomial distribution with parameters $n = 10$ and $p = 0.2$.

15. Suppose that X has the p.d.f.

$$f(x) = \begin{cases} 2x & \text{if } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Find and sketch the c.d.f. or X.

16. Find the quantile function for the distribution in Example 3.3.1.

20. In Exercise 13 of Sec. 3.2, draw a sketch of the c.d.f. F of X and find $F(10)$.

Seção 3.4:

1. Suppose that the joint p.d.f. of a pair of random variables (X, Y) is constant on the rectangle where $0 \leq x \leq 2$ and $0 \leq y \leq 1$, and suppose that the p.d.f. is 0 off of this rectangle.

   a. Find the constant value of the p.d.f. on the rectangle.

   b. Find $\Pr(X \geq Y)$.

2. Suppose that in an electric display sign there are three light bulbs in the first row and four light bulbs in the second row. Let X denote the number of bulbs in the first row that will be burned out at a specified time t, and let Y denote the number of bulbs in the second row that will be burned out at the same time t. Suppose that the joint p.f. of X and Y is as specified in the following table:

			Y		
X	0	1	2	3	4
0	0.08	0.07	0.06	0.01	0.01
1	0.06	0.10	0.12	0.05	0.02
2	0.05	0.06	0.09	0.04	0.03
3	0.02	0.03	0.03	0.03	0.04

Determine each of the following probabilities:

a. $\Pr(X = 2)$ **b.** $\Pr(Y \geq 2)$

c. $\Pr(X \leq 2 \text{ and } Y \leq 2)$ **d.** $\Pr(X = Y)$

e. $\Pr(X > Y)$

3. Suppose that X and Y have a discrete joint distribution for which the joint p.f. is defined as follows:

$$f(x, y) = \begin{cases} c|x + y| & \text{for } x = -2, -1, 0, 1, 2 \text{ and} \\ & \quad y = -2, -1, 0, 1, 2, \\ 0 & \text{otherwise.} \end{cases}$$

Determine **(a)** the value of the constant c; **(b)** $\Pr(X = 0 \text{ and } Y = -2)$; **(c)** $\Pr(X = 1)$; **(d)** $\Pr(|X - Y| \leq 1)$.

4. Suppose that X and Y have a continuous joint distribution for which the joint p.d.f. is defined as follows:

$$f(x, y) = \begin{cases} cy^2 & \text{for } 0 \leq x \leq 2 \text{ and } 0 \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Determine **(a)** the value of the constant c; **(b)** $\Pr(X + Y > 2)$; **(c)** $\Pr(Y < 1/2)$; **(d)** $\Pr(X \leq 1)$; **(e)** $\Pr(X = 3Y)$.

7. Suppose that a point (X, Y) is to be chosen from the square S in the xy-plane containing all points (x, y) such that $0 \le x \le 1$ and $0 \le y \le 1$. Suppose that the probability that the chosen point will be the corner $(0, 0)$ is 0.1, the probability that it will be the corner $(1, 0)$ is 0.2, the probability that it will be the corner $(0, 1)$ is 0.4, and the probability that it will be the corner $(1, 1)$ is 0.1. Suppose also that if the chosen point is not one of the four corners of the square, then it will be an interior point of the square and will be chosen according to a constant p.d.f. over the interior of the square. Determine **(a)** $\Pr(X \le 1/4)$ and **(b)** $\Pr(X + Y \le 1)$.

8. Suppose that X and Y are random variables such that (X, Y) must belong to the rectangle in the xy-plane containing all points (x, y) for which $0 \le x \le 3$ and $0 \le y \le 4$. Suppose also that the joint c.d.f. of X and Y at every point (x, y) in this rectangle is specified as follows:

$$F(x, y) = \frac{1}{156}xy(x^2 + y).$$

Determine **(a)** $\Pr(1 \le X \le 2$ and $1 \le Y \le 2)$; **(b)** $\Pr(2 \le X \le 4$ and $2 \le Y \le 4)$; **(c)** the c.d.f. of Y; **(d)** the joint p.d.f. of X and Y; **(e)** $\Pr(Y \le X)$.

9. In Example 3.4.5, compute the probability that water demand X is greater than electric demand Y.

11. Consider the clinical trial of depression drugs in Example 2.1.4. Suppose that a patient is selected at random from the 150 patients in that study and we record Y, an indicator of the treatment group for that patient, and X, an indicator of whether or not the patient relapsed. Table 3.3 contains the joint p.f. of X and Y.

   a. Calculate the probability that a patient selected at random from this study used Lithium (either alone or in combination with Imipramine) and did not relapse.

   b. Calculate the probability that the patient had a relapse (without regard to the treatment group).

Table 3.3 Proportions in clinical depression study for Exercise 11

Response (X)	Treatment group (Y)			
	Imipramine (1)	Lithium (2)	Combination (3)	Placebo (4)
Relapse (0)	0.120	0.087	0.146	0.160
No relapse (1)	0.147	0.166	0.107	0.067

Seção 3.5:

1. Suppose that X and Y have a continuous joint distribution for which the joint p.d.f. is

$$f(x, y) = \begin{cases} k & \text{for } a \leq x \leq b \text{ and } c \leq y \leq d, \\ 0 & \text{otherwise,} \end{cases}$$

where $a < b$, $c < d$, and $k > 0$. Find the marginal distributions of X and Y.

2. Suppose that X and Y have a discrete joint distribution for which the joint p.f. is defined as follows:

$$f(x, y) = \begin{cases} \frac{1}{30}(x + y) & \text{for } x = 0, 1, 2 \text{ and } y = 0, 1, 2, 3, \\ 0 & \text{otherwise.} \end{cases}$$

    a. Determine the marginal p.f.'s of X and Y.

    b. Are X and Y independent?

3. Suppose that X and Y have a continuous joint distribution for which the joint p.d.f. is defined as follows:

$$f(x, y) = \begin{cases} \frac{3}{2}y^2 & \text{for } 0 \leq x \leq 2 \text{ and } 0 \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

    a. Determine the marginal p.d.f.'s of X and Y.

    b. Are X and Y independent?

    c. Are the event $\{X < 1\}$ and the event $\{Y \geq 1/2\}$ independent?

4. Suppose that the joint p.d.f. of X and Y is as follows:

$$f(x, y) = \begin{cases} \frac{15}{4}x^2 & \text{for } 0 \leq y \leq 1 - x^2, \\ 0 & \text{otherwise.} \end{cases}$$

   a. Determine the marginal p.d.f.'s of X and Y.

   b. Are X and Y independent?

5. A certain drugstore has three public telephone booths. For $i = 0, 1, 2, 3$, let p_i denote the probability that exactly i telephone booths will be occupied on any Monday evening at 8:00 P.M.; and suppose that $p_0 = 0.1$, $p_1 = 0.2$, $p_2 = 0.4$, and $p_3 = 0.3$. Let X and Y denote the number of booths that will be occupied at 8:00 P.M. on two independent Monday evenings. Determine: **(a)** the joint p.f. of X and Y; **(b)** $\Pr(X = Y)$; **(c)** $\Pr(X > Y)$.

11. Suppose that two persons make an appointment to meet between 5 P.M. and 6 P.M. at a certain location, and they agree that neither person will wait more than 10 minutes for the other person. If they arrive independently at random times between 5 P.M. and 6 P.M., what is the probability that they will meet?
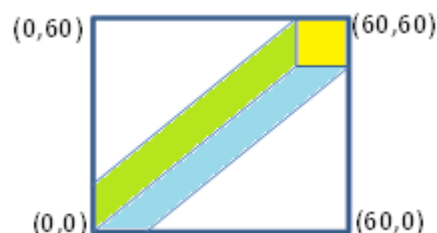
Dica: Denote por X e Y, respectivamente, os horários em que as pessoas 1 e 2 chegam após as 5h ($0 \leq X \leq 60$, $0 \leq Y \leq 60$). Por exemplo, se a pessoa 1 chega às 5:15 e pessoa 2 chega às 5:20, então X=15, Y=20.
Sob essa notação, (X,Y) tem distribuição uniforme no quadrado com limites $\{(0,0), (60,60)\}$ representado pela figura abaixo (com borda externa azul), ou seja:

$$f(X,Y) = \begin{cases} c \text{ para } 0 \leq X \leq 60, \ 0 \leq Y \leq 60 \\ 0 \text{ caso contrário} \end{cases},$$

onde você deve calcular a constante de normalização c.

Dizer que nenhum dos dois esperará mais do que 10 minutos pode ser representado matematicamente pela condição $|X - Y| \leq 10$. Essa condição pode ser satisfeita em três situações, representadas pelas áreas amarela, azul e verde na figura:
   a) Ambos chegam entre 5:50 e 6:00 (área amarela, cuja probabilidade pode ser calculada diretamente pela área do quadrado correspondente);
   b) Pessoa 1 chega entre 5:00 e 5:50, e pessoa 2 chega até 10 minutos depois (área verde): essa probabilidade pode ser calculada diretamente sobre o paralelepípedo correspondente, ou então por uma integral dupla - com a integral de x nos limites (0,50) e a integral de y nos limites (x, x+10).
   c) Pessoa 2 chega entre 5:00 e 5:50, e pessoa 1 chega até 10 minutos depois (área azul): essa probabilidade pode ser calculada de forma análoga ao item (b).

6. Suppose that in a certain drug the concentration of a particular chemical is a random variable with a continuous distribution for which the p.d.f. g is as follows:

$$g(x) = \begin{cases} \frac{3}{8}x^2 & \text{for } 0 \le x \le 2, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that the concentrations X and Y of the chemical in two separate batches of the drug are independent random variables for each of which the p.d.f. is g. Determine **(a)** the joint p.d.f. of X and Y; **(b)** $\Pr(X = Y)$; **(c)** $\Pr(X > Y)$; **(d)** $\Pr(X + Y \le 1)$.

10. Suppose that a point (X, Y) is chosen at random from the circle S defined as follows:

$$S = \{(x, y) : x^2 + y^2 \le 1\}.$$

a. Determine the joint p.d.f. of X and Y, the marginal p.d.f. of X, and the marginal p.d.f. of Y.

b. Are X and Y independent?

11. Suppose that two persons make an appointment to meet between 5 P.M. and 6 P.M. at a certain location, and they agree that neither person will wait more than 10 minutes for the other person. If they arrive independently at random times between 5 P.M. and 6 P.M., what is the probability that they will meet?

15. A painting process consists of two stages. In the first stage, the paint is applied, and in the second stage, a protective coat is added. Let X be the time spent on the first stage, and let Y be the time spent on the second stage. The first stage involves an inspection. If the paint fails the inspection, one must wait three minutes and apply the paint again. After a second application, there is no further inspection. The joint p.d.f. of X and Y is

$$f(x, y) = \begin{cases} \frac{1}{3} & \text{if } 1 < x < 3 \text{ and } 0 < y < 1, \\ \frac{1}{6} & \text{if } 6 < x < 8 \text{ and } 0 < y < 1, \\ 0 & \text{otherwise.} \end{cases}$$

  a. Sketch the region where $f(x, y) > 0$. Note that it is not exactly a rectangle.
  b. Find the marginal p.d.f.'s of X and Y.
  c. Show that X and Y are independent.

This problem does not contradict Theorem 3.5.6. In that theorem the conditions, including that the set where $f(x, y) > 0$ be rectangular, are sufficient but not necessary.

Seção 3.6:

2. Each student in a certain high school was classified according to her year in school (freshman, sophomore, junior, or senior) and according to the number of times that she had visited a certain museum (never, once, or more than once). The proportions of students in the various classifications are given in the following table:

	Never	Once	More than once
Freshmen	0.08	0.10	0.04
Sophomores	0.04	0.10	0.04
Juniors	0.04	0.20	0.09
Seniors	0.02	0.15	0.10

  a. If a student selected at random from the high school is a junior, what is the probability that she has never visited the museum?
  b. If a student selected at random from the high school has visited the museum three times, what is the probability that she is a senior?

3. Suppose that a point (X, Y) is chosen at random from the disk S defined as follows:

$$S = \{(x, y) : (x - 1)^2 + (y + 2)^2 \leq 9\}.$$

Determine **(a)** the conditional p.d.f. of Y for every given value of X, and **(b)** $\Pr(Y > 0 | X = 2)$.

4. Suppose that the joint p.d.f. of two random variables X and Y is as follows:

$$f(x, y) = \begin{cases} c(x + y^2) & \text{for } 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Determine **(a)** the conditional p.d.f. of X for every given value of Y, and **(b)** $\Pr(X < \frac{1}{2} | Y = \frac{1}{2})$.

8. Suppose that a person's score X on a mathematics aptitude test is a number between 0 and 1, and that his score Y on a music aptitude test is also a number between 0 and 1. Suppose further that in the population of all college students in the United States, the scores X and Y are distributed according to the following joint p.d.f.:

$$f(x, y) = \begin{cases} \frac{2}{5}(2x + 3y) & \text{for } 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

   a. What proportion of college students obtain a score greater than 0.8 on the mathematics test?

   b. If a student's score on the music test is 0.3, what is the probability that his score on the mathematics test will be greater than 0.8?

   c. If a student's score on the mathematics test is 0.3, what is the probability that his score on the music test will be greater than 0.8?

Dicas:
Para o item (a), obtenha a f.d.p. marginal $f_1(x)$; a partir desta, calcule a f.d.a. marginal $F_1(x)$; em seguida, calcule $\Pr(X \geq 0.8) = 1 - F_1(x)$.
Para o item (b), calcule a f.d.p. condicional de x dado y=0.3, $g_1(x|y=0.3) = f(x,y) / f_2(y)$, substituindo y por 0.3; em seguida, obtenha a f.d.a. condicional $G_1(x|y=0.3)$; finalmente, calcule $\Pr(X \geq 0.8 | Y=0.3) = 1 - G_1(x|y=0.3)$.
Para o item (c), o raciocínio é análogo ao item (b), exceto que, nesse caso, interessa calcular a f.d.p. condicional de y dado x=0.3, $g_2(y|x=0.3) = f(x,y) / f_1(x)$; em seguida, deve-se obter a f.d.a. condicional $G_2(y|x=0.3)$, a partir da qual se poderá calcular $\Pr(Y \geq 0.8 | X=0.3)$.

9. Suppose that either of two instruments might be used for making a certain measurement. Instrument 1 yields a measurement whose p.d.f. h_1 is

$$h_1(x) = \begin{cases} 2x & \text{for } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Instrument 2 yields a measurement whose p.d.f. h_2 is

$$h_2(x) = \begin{cases} 3x^2 & \text{for } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that one of the two instruments is chosen at random and a measurement X is made with it.

   a. Determine the marginal p.d.f. of X.

   b. If the value of the measurement is $X = 1/4$, what is the probability that instrument 1 was used?

Seção 3.8:

1. Suppose that the p.d.f. of a random variable X is as follows:

$$f(x) = \begin{cases} 3x^2 & \text{for } 0 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Also, suppose that $Y = 1 - X^2$. Determine the p.d.f. of Y.

2. Suppose that a random variable X can have each of the seven values $-3, -2, -1, 0, 1, 2, 3$ with equal probability. Determine the p.f. of $Y = X^2 - X$.

3. Suppose that the p.d.f. of a random variable X is as follows:

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

Also, suppose that $Y = X(2 - X)$. Determine the c.d.f. and the p.d.f. of Y.

6. Suppose that the p.d.f. of X is as given in Exercise 3. Determine the p.d.f. of $Y = 3X + 2$.

7. Suppose that a random variable X has the uniform distribution on the interval $[0, 1]$. Determine the p.d.f. of **(a)** X^2, **(b)** $-X^3$, and **(c)** $X^{1/2}$.

11. Explain how to use a uniform pseudo-random number generator to generate four independent values from a distribution for which the p.d.f. is

$$g(y) = \begin{cases} \frac{1}{2}(2y + 1) & \text{for } 0 < y < 1, \\ 0 & \text{otherwise.} \end{cases}$$

13. Let Z be the rate at which customers are served in a queue. Assume that Z has the p.d.f.

$$f(z) = \begin{cases} 2e^{-2z} & \text{for } z > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Find the p.d.f. of the average waiting time $T = 1/Z$.