Chapter 4:
Random Variables and Expectation

Problems

1. Five men and 5 women are ranked according to their scores on an examination. Assume that no two scores are alike and all 10! possible rankings are equally likely. Let X denote the highest ranking achieved by a woman (for instance, $X = 2$ if the top-ranked person was male and the next-ranked person was female). Find $P\{X = i\}, i = 1, 2, 3, \ldots, 8, 9, 10$.
2. Let X represent the difference between the number of heads and the number of tails obtained when a coin is tossed n times. What are the possible values of X?
3. In Problem 2, if the coin is assumed fair, for $n = 3$, what are the probabilities associated with the values that X can take on?
4. The distribution function of the random variable X is given

$$F(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x}{2} & 0 \le x < 1 \\ \dfrac{2}{3} & 1 \le x < 2 \\ \dfrac{11}{12} & 2 \le x < 3 \\ 1 & 3 \le x \end{cases}$$

   (a) Plot this distribution function.
   (b) What is $P\{X > \frac{1}{2}\}$?
   (c) What is $P\{2 < X \le 4\}$?
   (d) What is $P\{X < 3\}$?
   (e) What is $P\{X = 1\}$?

5. Suppose the random variable X has probability density function

$$f(x) = \begin{cases} cx^3, & \text{if } 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

   (a) Find the value of c.
   (b) Find $P\{.4 < X < .8\}$.

6. The amount of time, in hours, that a computer functions before breaking down is a continuous random variable with probability density function given by

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

What is the probability that a computer will function between 50 and 150 hours before breaking down? What is the probability that it will function less than 100 hours?

7. The lifetime in hours of a certain kind of radio tube is a random variable having a probability density function given by

$$f(x) = \begin{cases} 0 & x \leq 100 \\ \dfrac{100}{x^2} & x > 100 \end{cases}$$

What is the probability that exactly 2 of 5 such tubes in a radio set will have to be replaced within the first 150 hours of operation? Assume that the events $E_i, i = 1, 2, 3, 4, 5$, that the ith such tube will have to be replaced within this time are independent.

8. If the density function of X equals

$$f(x) = \begin{cases} c\,e^{-2x} & 0 < x < \infty \\ 0 & x < 0 \end{cases}$$

find c. What is $P\{X > 2\}$?

9. A set of five transistors are to be tested, one at a time in a random order, to see which of them are defective. Suppose that three of the five transistors are defective, and let N_1 denote the number of tests made until the first defective is spotted, and let N_2 denote the number of additional tests until the second defective is spotted. Find the joint probability mass function of N_1 and N_2.

10. The joint probability density function of X and Y is given by

$$f(x, y) = \frac{6}{7}\left(x^2 + \frac{xy}{2}\right), \quad 0 < x < 1, \quad 0 < y < 2$$

(a) Verify that this is indeed a joint density function.
(b) Compute the density function of X.
(c) Find $P\{X > Y\}$.

11. Let $X_1, X_2, \ldots, X_n$ be independent random variables, each having a uniform distribution over (0, 1). Let M = maximum $(X_1, X_2, \ldots, X_n)$. Show that the distribution function of M is given by

$$F_M(x) = x^n, \quad 0 \leq x \leq 1$$

What is the probability density function of M?

12. The joint density of X and Y is given by

$$f(x,y) = \begin{cases} x e^{-(x+y)} & x > 0,\ y > 0 \\ 0 & \text{otherwise} \end{cases}$$

(a) Compute the density of X.
(b) Compute the density of Y.
(c) Are X and Y independent?

13. The joint density of X and Y is

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

(a) Compute the density of X.
(b) Compute the density of Y.
(c) Are X and Y independent?

14. If the joint density function of X and Y factors into one part depending only on x and one depending only on y, show that X and Y are independent. That is, if

$$f(x,y) = k(x)\,h(y), \quad -\infty < x < \infty, \quad -\infty < y < \infty$$

show that X and Y are independent.

15. Is Problem 14 consistent with the results of Problems 12 and 13?

16. Suppose that X and Y are independent continuous random variables. Show that

(a) $P\{X + Y \le a\} = \int_{-\infty}^{\infty} F_X(a - y) f_Y(y)\, dy$

(b) $P\{X \le Y\} = \int_{-\infty}^{\infty} F_X(y) f_Y(y)\, dy$

where f_Y is the density function of Y, and F_X is the distribution function of X.

17. When a current I (measured in amperes) flows through a resistance R (measured in ohms), the power generated (measured in watts) is given by $W = I^2R$. Suppose that I and R are independent random variables with densities

$$f_I(x) = 6x(1 - x) \quad 0 \le x \le 1$$
$$f_R(x) = 2x \quad 0 \le x \le 1$$

Determine the density function of W.

18. Suppose that 15 percent of the families in a certain community have no children, 20 percent have 1, 35 percent have 2, and 30 percent have 3 children; suppose further that each child is equally likely (and independently) to be a boy or a girl. Determine the conditional probability mass function of the size of a randomly chosen family containing 2 girls.
19. Compute the conditional density function of X given $Y = y$ in **(a)** Problem 10 and **(b)** Problem 13.
20. Show that X and Y are independent if and only if
    (a) $p_{X|Y}(x|y) = p_X(x)$ in discrete cases; **(b)** $f_{X|Y}(x|y) = f_X(x)$ in continuous cases
21. Compute the expected value of the random variable in Problem 1.
22. Compute the expected value of the random variable in Problem 3.
23. Each night different meteorologists give us the "probability" that it will rain the next day. To judge how well these people predict, we will score each of them as follows: If a meteorologist says that it will rain with probability p, then he or she will receive a score of

$$\begin{array}{ll} 1-(1-p)^2 & \text{if it does rain} \\ 1-p^2 & \text{if it does not rain} \end{array}$$

    We will then keep track of scores over a certain time span and conclude that the meteorologist with the highest average score is the best predictor of weather. Suppose now that a given meteorologist is aware of this and so wants to maximize his or her expected score. If this individual truly believes that it will rain tomorrow with probability p^*, what value of p should he or she assert so as to maximize the expected score?
24. An insurance company writes a policy to the effect that an amount of money A must be paid if some event E occurs within a year. If the company estimates that E will occur within a year with probability p, what should it charge the customer so that its expected profit will be 10 percent of A?
25. A total of 4 buses carrying 148 students from the same school arrive at a football stadium. The buses carry, respectively, 40, 33, 25, and 50 students. One of the students is randomly selected. Let X denote the number of students that were on the bus carrying this randomly selected student. One of the 4 bus drivers is also randomly selected. Let Y denote the number of students on her bus.

    (a) Which of $E[X]$ or $E[Y]$ do you think is larger? Why?
    (b) Compute $E[X]$ and $E[Y]$.

26. Suppose that two teams play a series of games that end when one of them has won i games. Suppose that each game played is, independently, won by team A with probability p. Find the expected number of games that are played when $i = 2$. Also show that this number is maximized when $p = \frac{1}{2}$.

27. The density function of X is given by

$$f(x) = \begin{cases} a + bx^2 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

If $E[X] = \frac{3}{5}$, find a, b.

28. The lifetime in hours of electronic tubes is a random variable having a probability density function given by

$$f(x) = a^2 x e^{-ax}, \quad x \geq 0$$

Compute the expected lifetime of such a tube.

29. Let $X_1, X_2, \ldots, X_n$ be independent random variables having the common density function

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

Find **(a)** $E[\text{Max}(X_1, \ldots, X_n)]$ and **(b)** $E[\text{Min}(X_1, \ldots, X_n)]$.

30. Suppose that X has density function

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

Compute $E[X^n]$ **(a)** by computing the density of X^n and then using the definition of expectation and **(b)** by using Proposition 4.5.1.

31. The time it takes to repair a personal computer is a random variable whose density, in hours, is given by

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

The cost of the repair depends on the time it takes and is equal to $40 + 30\sqrt{x}$ when the time is x. Compute the expected cost to repair a personal computer.

32. If $E[X] = 2$ and $E[X^2] = 8$, calculate **(a)** $E[(2+4X)^2]$ and **(b)** $E[X^2+(X+1)^2]$.

33. Ten balls are randomly chosen from an urn containing 17 white and 23 black balls. Let X denote the number of white balls chosen. Compute $E[X]$

    (a) by defining appropriate indicator variables X_i, $i = 1, \ldots, 10$ so that

$$X = \sum_{i=1}^{10} X_i$$

(b) by defining appropriate indicator variables $Y_i = 1, \ldots, 17$ so that

$$X = \sum_{i=1}^{17} Y_i$$

34. If X is a continuous random variable having distribution function F, then its *median* is defined as that value of m for which

$$F(m) = 1/2$$

Find the median of the random variables with density function

(a) $f(x) = e^{-x}, \quad x \geq 0$;

(b) $f(x) = 1, \quad 0 \leq x \leq 1$.

35. The median, like the mean, is important in predicting the value of a random variable. Whereas it was shown in the text that the mean of a random variable is the best predictor from the point of view of minimizing the expected value of the square of the error, the median is the best predictor if one wants to minimize the expected value of the absolute error. That is, $E[|X - c|]$ is minimized when c is the median of the distribution function of X. Prove this result when X is continuous with distribution function F and density function f. *Hint:* Write

$$\begin{aligned}
E[|X - c|] &= \int_{-\infty}^{\infty} |x - c| f(x)\, dx \\
&= \int_{-\infty}^{c} |x - c| f(x)\, dx + \int_{c}^{\infty} |x - c| f(x)\, dx \\
&= \int_{-\infty}^{c} (c - x) f(x)\, dx + \int_{c}^{\infty} (x - c) f(x)\, dx \\
&= cF(c) - \int_{-\infty}^{c} x f(x)\, dx + \int_{c}^{\infty} x f(x)\, dx - c[1 - F(c)]
\end{aligned}$$

Now, use calculus to find the minimizing value of c.

36. We say that m_p is the *100p percentile* of the distribution function F if

$$F(m_p) = p$$

Find m_p for the distribution having density function

$$f(x) = 2e^{-2x}, \quad x \geq 0$$

37. A community consists of 100 married couples. If 50 members of the community die, what is the expected number of marriages that remain intact? Assume that the

set of people who die is equally likely to be any of the $\binom{200}{50}$ groups of size 50. *Hint:* For $i = 1, \ldots, 100$ let

$$X_i = \begin{cases} 1 & \text{if neither member of couple } i \text{ dies} \\ 0 & \text{otherwise} \end{cases}$$

38. Compute the expectation and variance of the number of successes in n independent trials, each of which results in a success with probability p. Is independence necessary?
39. Suppose that X is equally likely to take on any of the values 1, 2, 3, 4. Compute **(a)** $E[X]$ and **(b)** $\text{Var}(X)$.
40. Let $p_i = P\{X = i\}$ and suppose that $p_1 + p_2 + p_3 = 1$. If $E[X] = 2$, what values of p_1, p_2, p_3 **(a)** maximize and **(b)** minimize $\text{Var}(X)$?
41. Compute the mean and variance of the number of heads that appear in 3 flips of a fair coin.
42. Argue that for any random variable X

$$E[X^2] \geq (E[X])^2$$

When does one have equality?

43. A random variable X, which represents the weight (in ounces) of an article, has density function,

$$f(z) = \begin{cases} z - 8 & \text{for } 8 \leq z \leq 9 \\ 10 - z & \text{for } 9 < z \leq 10 \\ 0 & \text{otherwise} \end{cases}$$

(a) Calculate the mean and variance of the random variable X.

(b) The manufacturer sells the article for a fixed price of \$2.00. He guarantees to refund the purchase money to any customer who finds the weight of his article to be less than 8.25 oz. His cost of production is related to the weight of the article by the relation $x/15 + .35$. Find the expected profit per article.

44. Let X_i denote the percentage of votes cast in a given election that are for candidate i, and suppose that X_1 and X_2 have a joint density function

$$f_{X_1,X_2}(x, y) = \begin{cases} 3(x + y), & \text{if } x \geq 0, y \geq 0, 0 \leq x + y \leq 1 \\ 0, & \text{if otherwise} \end{cases}$$

(a) Find the marginal densities of X_1 and X_2.

(b) Find $E[X_i]$ and $\text{Var}(X_i)$ for $i = 1, 2$.

45. A product is classified according to the number of defects it contains and the factory that produces it. Let X_1 and X_2 be the random variables that represent the number of defects per unit (taking on possible values of 0, 1, 2, or 3) and the factory number (taking on possible values 1 or 2), respectively. The entries in the table represent the joint possibility mass function of a randomly chosen product.

X_1 \ X_2	1	2
0	$\frac{1}{8}$	$\frac{1}{16}$
1	$\frac{1}{16}$	$\frac{1}{16}$
2	$\frac{3}{16}$	$\frac{1}{8}$
3	$\frac{1}{8}$	$\frac{1}{4}$

(a) Find the marginal probability distributions of X_1 and X_2.

(b) Find $E[X_1], E[X_2]$, $\text{Var}(X_1)$, $\text{Var}(X_2)$, and $\text{Cov}(X_1, X_2)$.

The strength of the relationship between X and Y is indicated by the correlation between X andY, a dimensionless quantity obtained by dividing the covariance by the product of the standard deviations of X and Y.

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

46. Find $\text{Corr}(X_1, X_2)$ for the random variables of Problem 44.

47. Verify: $\text{Cov}(aX, Y) = a\,\text{Cov}(X, Y)$ for any constant a.

48. Prove

$$\text{Cov}\left(\sum_{i=1}^{n} X_i, Y\right) = \sum_{i=1}^{n} \text{Cov}(X_i, Y)$$

by using mathematical induction.

49. Let X have variance σ_x^2 and let Y have variance σ_y^2. Starting with

$$0 \le \text{Var}(X/\sigma_x + Y/\sigma_y)$$

show that

$$-1 \le \text{Corr}(X, Y)$$

Now using that

$$0 \le \text{Var}(X/\sigma_x - Y/\sigma_y)$$

conclude that

$$-1 \le \text{Corr}(X, Y) \le 1$$

Using the result that $\text{Var}(Z) = 0$ implies that Z is constant, argue that, if $\text{Corr}(X, Y) = 1$ or -1, then X and Y are related by

$$Y = a + bx$$

where the sign of b is positive when the correlation is 1 and negative when it is -1.

50. Consider n independent trials, each of which results in any of the outcomes i, $i = 1, 2, 3$, with respective probabilities p_1, p_2, p_3, $\sum_{i=1}^{3} p_i = 1$. Let N_i denote the number of trials that result in outcome i, and show that $\mathrm{Cov}(N_1, N_2) = -np_1p_2$. Also explain why it is intuitive that this covariance is negative. (*Hint:* For $i = 1, \ldots, n$, let

$$X_i = \begin{cases} 1 & \text{if trial } i \text{ results in outcome 1} \\ 0 & \text{if trial } i \text{ does not result in outcome 1} \end{cases}$$

Similarly, for $j = 1, \ldots, n$, let

$$Y_j = \begin{cases} 1 & \text{if trial } j \text{ results in outcome 2} \\ 0 & \text{if trial } j \text{ does not result in outcome 2} \end{cases}$$

Argue that

$$N_1 = \sum_{i=1}^{n} X_i, \quad N_2 = \sum_{j=1}^{n} Y_j$$

Then use Proposition 4.7.2 and Theorem 4.7.4.)

Proposition 4.7.2.

$$\mathrm{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{Cov}(X_i, Y_j)$$

Theorem 4.7.4. If X and Y are independent random variables, then

$$\mathrm{Cov}(X, Y) = 0$$

and so for independent $X_1, \ldots, X_n$,

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

51. A construction firm has recently sent in bids for 3 jobs worth (in profits) 10, 20, and 40 (thousand) dollars. If its probabilities of winning the jobs are respectively .2, .8, and .3, what is the firm's expected total profit?

Compute $\mathrm{Cov}(X_i, X_j)$ and use this result to show that $\mathrm{Var}(X) = 1$.

52. If X_1 and X_2 have the same probability distribution function, show that

$$\mathrm{Cov}(X_1 - X_2,\ X_1 + X_2) = 0$$

Note that independence is not being assumed.

53. Suppose that X has density function $f(x) = e^{\ x}$, $\quad x > 0$
Compute the moment generating function of X and use your result to determine its mean and variance. Check your answer for the mean by a direct calculation.

54. If the density function of X is $f(x) = 1$, $\quad 0 < x < 1$
determine $E[e^{tX}]$. Differentiate to obtain $E[X^n]$ and then check your answer.

55. Suppose that X is a random variable with mean and variance both equal to 20. What can be said about $P\{0 \le X \le 40\}$?

56. From past experience, a professor knows that the test score of a student taking her final examination is a random variable with mean 75.

(a) Give an upper bound to the probability that a student's test score will exceed 85.

Suppose in addition the professor knows that the variance of a student's test score is equal to 25.

(b) What can be said about the probability that a student will score between 65 and 85?

(c) How many students would have to take the examination so as to ensure, with probability at least .9, that the class average would be within 5 of 75?

57. Let X and Y have respective distribution functions F_X and F_Y, and suppose that for some constants a and $b > 0$,

$$F_X(x) = F_Y\left(\frac{x-a}{b}\right)$$

(a) Determine $E[X]$ in terms of $E[Y]$.

(b) Determine $\text{Var}(X)$ in terms of $\text{Var}(Y)$.

Hint: X has the same distribution as what other random variable?