

## 2.4 Concluding Remarks

This chapter introduced the concept of a random variable, one of the fundamental ideas of probability theory. A fully rigorous discussion of random variables requires a background in measure theory. The development here is sufficient for the needs of this course.

Discrete and continuous random variables have been defined, and it should be mentioned that more general random variables can also be defined and are useful on occasion. In particular, it makes sense to consider random variables that have both a discrete and a continuous component. For example, the lifetime of a transistor might be 0 with some probability  $p > 0$  if it does not function at all; if it does function, the lifetime could be modeled as a continuous random variable.

## 2.5 Problems

1. Suppose that  $X$  is a discrete random variable with  $P(X = 0) = .25$ ,  $P(X = 1) = .125$ ,  $P(X = 2) = .125$ , and  $P(X = 3) = .5$ . Graph the frequency function and the cumulative distribution function of  $X$ .
2. An experiment consists of throwing a fair coin four times. Find the frequency function and the cumulative distribution function of the following random variables: (a) the number of heads before the first tail, (b) the number of heads following the first tail, (c) the number of heads minus the number of tails, and (d) the number of tails times the number of heads.
3. The following table shows the cumulative distribution function of a discrete random variable. Find the frequency function.

$k$	$F(k)$
0	0
1	.1
2	.3
3	.7
4	.8
5	1.0

4. If  $X$  is an integer-valued random variable, show that the frequency function is related to the cdf by  $p(k) = F(k) - F(k - 1)$ .
5. Show that  $P(u < X \leq v) = F(v) - F(u)$  for any  $u$  and  $v$  in the cases that (a)  $X$  is a discrete random variable and (b)  $X$  is a continuous random variable.
6. Let  $A$  and  $B$  be events, and let  $I_A$  and  $I_B$  be the associated indicator random variables. Show that

$$I_{A \cap B} = I_A I_B = \min(I_A, I_B)$$

and

$$I_{A \cup B} = \max(I_A, I_B)$$

7. Find the cdf of a Bernoulli random variable.
8. Show that the binomial probabilities sum to 1.
9. For what values of  $p$  is a two-out-of-three majority decoder better than transmission of the message once?
10. Appending three extra bits to a 4-bit word in a particular way (a Hamming code) allows detection and correction of up to one error in any of the bits. If each bit has probability .05 of being changed during communication, and the bits are changed independently of each other, what is the probability that the word is correctly received (that is, 0 or 1 bit is in error)? How does this probability compare to the probability that the word will be transmitted correctly with no check bits, in which case all four bits would have to be transmitted correctly for the word to be correct?
11. Consider the binomial distribution with  $n$  trials and probability  $p$  of success on each trial. For what value of  $k$  is  $P(X = k)$  maximized? This value is called the **mode** of the distribution. (*Hint*: Consider the ratio of successive terms.)
12. Which is more likely: 9 heads in 10 tosses of a fair coin or 18 heads in 20 tosses?
13. A multiple-choice test consists of 20 items, each with four choices. A student is able to eliminate one of the choices on each question as incorrect and chooses randomly from the remaining three choices. A passing grade is 12 items or more correct.
  - a. What is the probability that the student passes?
  - b. Answer the question in part (a) again, assuming that the student can eliminate two of the choices on each question.
14. Two boys play basketball in the following way. They take turns shooting and stop when a basket is made. Player A goes first and has probability  $p_1$  of making a basket on any throw. Player B, who shoots second, has probability  $p_2$  of making a basket. The outcomes of the successive trials are assumed to be independent.
  - a. Find the frequency function for the total number of attempts.
  - b. What is the probability that player A wins?
15. Two teams, A and B, play a series of games. If team A has probability .4 of winning each game, is it to its advantage to play the best three out of five games or the best four out of seven? Assume the outcomes of successive games are independent.
16. Show that if  $n$  approaches  $\infty$  and  $r/n$  approaches  $p$  and  $m$  is fixed, the hypergeometric frequency function tends to the binomial frequency function. Give a heuristic argument for why this is true.
17. Suppose that in a sequence of independent Bernoulli trials, each with probability of success  $p$ , the number of failures up to the first success is counted. What is the frequency function for this random variable?
18. Continuing with Problem 17, find the frequency function for the number of failures up to the  $r$ th success.

19. Find an expression for the cumulative distribution function of a geometric random variable.
20. If  $X$  is a geometric random variable with  $p = .5$ , for what value of  $k$  is  $P(X \leq k) \approx .99$ ?
21. If  $X$  is a geometric random variable, show that

$$P(X > n + k - 1 | X > n - 1) = P(X > k)$$

In light of the construction of a geometric distribution from a sequence of independent Bernoulli trials, how can this be interpreted so that it is “obvious”?

22. Three identical fair coins are thrown simultaneously until all three show the same face. What is the probability that they are thrown more than three times?
23. In a sequence of independent trials with probability  $p$  of success, what is the probability that there are  $r$  successes before the  $k$ th failure?
24. (Banach Match Problem) A pipe smoker carries one box of matches in his left pocket and one box in his right. Initially, each box contains  $n$  matches. If he needs a match, the smoker is equally likely to choose either pocket. What is the frequency function for the number of matches in the other box when he first discovers that one box is empty?
25. The probability of being dealt a royal straight flush (ace, king, queen, jack, and ten of the same suit) in poker is about  $1.3 \times 10^{-8}$ . Suppose that an avid poker player sees 100 hands a week, 52 weeks a year, for 20 years.
  - a. What is the probability that she is never dealt a royal straight flush dealt?
  - b. What is the probability that she is dealt exactly two royal straight flushes?
26. The university administration assures a mathematician that he has only 1 chance in 10,000 of being trapped in a much-maligned elevator in the mathematics building. If he goes to work 5 days a week, 52 weeks a year, for 10 years, and always rides the elevator up to his office when he first arrives, what is the probability that he will never be trapped? That he will be trapped once? Twice? Assume that the outcomes on all the days are mutually independent (a dubious assumption in practice).
27. Suppose that a rare disease has an incidence of 1 in 1000. Assuming that members of the population are affected independently, find the probability of  $k$  cases in a population of 100,000 for  $k = 0, 1, 2$ .
28. Let  $p_0, p_1, \dots, p_n$  denote the probability mass function of the binomial distribution with parameters  $n$  and  $p$ . Let  $q = 1 - p$ . Show that the binomial probabilities can be computed recursively by  $p_0 = q^n$  and

$$p_k = \frac{(n - k + 1)p}{kq} p_{k-1}, \quad k = 1, 2, \dots, n$$

Use this relation to find  $P(X \leq 4)$  for  $n = 9000$  and  $p = .0005$ .

29. Show that the Poisson probabilities  $p_0, p_1, \dots$  can be computed recursively by  $p_0 = \exp(-\lambda)$  and

$$p_k = \frac{\lambda}{k} p_{k-1}, \quad k = 1, 2, \dots$$

Use this scheme to find  $P(X \leq 4)$  for  $\lambda = 4.5$  and compare to the results of Problem 28.

30. Suppose that in a city, the number of suicides can be approximated by a Poisson process with  $\lambda = .33$  per month.
- Find the probability of  $k$  suicides in a year for  $k = 0, 1, 2, \dots$ . What is the most probable number of suicides?
  - What is the probability of two suicides in one week?
31. Phone calls are received at a certain residence as a Poisson process with parameter  $\lambda = 2$  per hour.
- If Diane takes a 10-min. shower, what is the probability that the phone rings during that time?
  - How long can her shower be if she wishes the probability of receiving no phone calls to be at most .5?
32. For what value of  $k$  is the Poisson frequency function with parameter  $\lambda$  maximized? (*Hint:* Consider the ratio of consecutive terms.)
33. Let  $F(x) = 1 - \exp(-\alpha x^\beta)$  for  $x \geq 0, \alpha > 0, \beta > 0$ , and  $F(x) = 0$  for  $x < 0$ . Show that  $F$  is a cdf, and find the corresponding density.
34. Let  $f(x) = (1 + \alpha x)/2$  for  $-1 \leq x \leq 1$  and  $f(x) = 0$  otherwise, where  $-1 \leq \alpha \leq 1$ . Show that  $f$  is a density, and find the corresponding cdf. Find the quartiles and the median of the distribution in terms of  $\alpha$ .
35. Sketch the pdf and cdf of a random variable that is uniform on  $[-1, 1]$ .
36. If  $U$  is a uniform random variable on  $[0, 1]$ , what is the distribution of the random variable  $X = [nU]$ , where  $[t]$  denotes the greatest integer less than or equal to  $t$ ?
37. A line segment of length 1 is cut once at random. What is the probability that the longer piece is more than twice the length of the shorter piece?
38. If  $f$  and  $g$  are densities, show that  $\alpha f + (1 - \alpha)g$  is a density, where  $0 \leq \alpha \leq 1$ .
39. The **Cauchy** cumulative distribution function is

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(x), \quad -\infty < x < \infty$$

- Show that this is a cdf.
  - Find the density function.
  - Find  $x$  such that  $P(X > x) = .1$ .
40. Suppose that  $X$  has the density function  $f(x) = cx^2$  for  $0 \leq x \leq 1$  and  $f(x) = 0$  otherwise.
- Find  $c$ .
  - Find the cdf.
  - What is  $P(.1 \leq X < .5)$ ?

41. Find the upper and lower quartiles of the exponential distribution.
42. Find the probability density for the distance from an event to its nearest neighbor for a Poisson process in the plane.
43. Find the probability density for the distance from an event to its nearest neighbor for a Poisson process in three-dimensional space.
44. Let  $T$  be an exponential random variable with parameter  $\lambda$ . Let  $X$  be a discrete random variable defined as  $X = k$  if  $k \leq T < k + 1$ ,  $k = 0, 1, \dots$ . Find the frequency function of  $X$ .
45. Suppose that the lifetime of an electronic component follows an exponential distribution with  $\lambda = .1$ .
  - a. Find the probability that the lifetime is less than 10.
  - b. Find the probability that the lifetime is between 5 and 15.
  - c. Find  $t$  such that the probability that the lifetime is greater than  $t$  is .01.
46.  $T$  is an exponential random variable, and  $P(T < 1) = .05$ . What is  $\lambda$ ?
47. If  $\alpha > 1$ , show that the gamma density has a maximum at  $(\alpha - 1)/\lambda$ .
48. Show that the gamma density integrates to 1.
49. The gamma function is a generalized factorial function.
  - a. Show that  $\Gamma(1) = 1$ .
  - b. Show that  $\Gamma(x + 1) = x\Gamma(x)$ . (*Hint*: Use integration by parts.)
  - c. Conclude that  $\Gamma(n) = (n - 1)!$ , for  $n = 1, 2, 3, \dots$ .
  - d. Use the fact that  $\Gamma(\frac{1}{2}) = \sqrt{\pi}$  to show that, if  $n$  is an odd integer,

$$\Gamma\left(\frac{n}{2}\right) = \frac{\sqrt{\pi}(n-1)!}{2^{n-1}\left(\frac{n-1}{2}\right)!}$$

50. Show by a change of variables that

$$\begin{aligned}\Gamma(x) &= 2 \int_0^\infty t^{2x-1} e^{-t^2} dt \\ &= \int_{-\infty}^\infty e^{xt} e^{-e^t} dt\end{aligned}$$

51. Show that the normal density integrates to 1. (*Hint*: First make a change of variables to reduce the integral to that for the standard normal. The problem is then to show that  $\int_{-\infty}^\infty \exp(-x^2/2) dx = \sqrt{2\pi}$ . Square both sides and reexpress the problem as that of showing

$$\left(\int_{-\infty}^\infty \exp(-x^2/2) dx\right) \left(\int_{-\infty}^\infty \exp(-y^2/2) dy\right) = 2\pi$$

Finally, write the product of integrals as a double integral and change to polar coordinates.)

52. Suppose that in a certain population, individuals' heights are approximately normally distributed with parameters  $\mu = 70$  and  $\sigma = 3$  in.
- What proportion of the population is over 6 ft. tall?
  - What is the distribution of heights if they are expressed in centimeters? In meters?
53. Let  $X$  be a normal random variable with  $\mu = 5$  and  $\sigma = 10$ . Find (a)  $P(X > 10)$ , (b)  $P(-20 < X < 15)$ , and (c) the value of  $x$  such that  $P(X > x) = .05$ .
54. If  $X \sim N(\mu, \sigma^2)$ , show that  $P(|X - \mu| \leq .675\sigma) = .5$ .
55.  $X \sim N(\mu, \sigma^2)$ , find the value of  $c$  in terms of  $\sigma$  such that  $P(\mu - c \leq X \leq \mu + c) = .95$ .
56. If  $X \sim N(0, \sigma^2)$ , find the density of  $Y = |X|$ .
57.  $X \sim N(\mu, \sigma^2)$  and  $Y = aX + b$ , where  $a < 0$ , show that  $Y \sim N(a\mu + b, a^2\sigma^2)$ .
58. If  $U$  is uniform on  $[0, 1]$ , find the density function of  $\sqrt{U}$ .
59. If  $U$  is uniform on  $[-1, 1]$ , find the density function of  $U^2$ .
60. Find the density function of  $Y = e^Z$ , where  $Z \sim N(\mu, \sigma^2)$ . This is called the **lognormal density**, since  $\log Y$  is normally distributed.
61. Find the density of  $cX$  when  $X$  follows a gamma distribution. Show that only  $\lambda$  is affected by such a transformation, which justifies calling  $\lambda$  a scale parameter.
62. Show that if  $X$  has a density function  $f_X$  and  $Y = aX + b$ , then

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

63. Suppose that  $\Theta$  follows a uniform distribution on the interval  $[-\pi/2, \pi/2]$ . Find the cdf and density of  $\tan \Theta$ .
64. A particle of mass  $m$  has a random velocity,  $V$ , which is normally distributed with parameters  $\mu = 0$  and  $\sigma$ . Find the density function of the kinetic energy,  $E = \frac{1}{2}mV^2$ .
65. How could random variables with the following density function be generated from a uniform random number generator?

$$f(x) = \frac{1 + \alpha x}{2}, \quad -1 \leq x \leq 1, \quad -1 \leq \alpha \leq 1$$

66. Let  $f(x) = \alpha x^{-\alpha-1}$  for  $x \geq 1$  and  $f(x) = 0$  otherwise, where  $\alpha$  is a positive parameter. Show how to generate random variables from this density from a uniform random number generator.
67. The **Weibull** cumulative distribution function is

$$F(x) = 1 - e^{-(x/\alpha)^\beta}, \quad x \geq 0, \quad \alpha > 0, \quad \beta > 0$$

- Find the density function.

- b. Show that if  $W$  follows a Weibull distribution, then  $X = (W/\alpha)^\beta$  follows an exponential distribution.
  - c. How could Weibull random variables be generated from a uniform random number generator?
68. If the radius of a circle is an exponential random variable, find the density function of the area.
69. If the radius of a sphere is an exponential random variable, find the density function of the volume.
70. Let  $U$  be a uniform random variable. Find the density function of  $V = U^{-\alpha}$ ,  $\alpha > 0$ . Compare the rates of decrease of the tails of the densities as a function of  $\alpha$ . Does the comparison make sense intuitively?
71. This problem shows one way to generate discrete random variables from a uniform random number generator. Suppose that  $F$  is the cdf of an integer-valued random variable; let  $U$  be uniform on  $[0, 1]$ . Define a random variable  $Y = k$  if  $F(k-1) < U \leq F(k)$ . Show that  $Y$  has cdf  $F$ . Apply this result to show how to generate geometric random variables from uniform random variables.
72. One of the most commonly used (but not one of the best) methods of generating pseudorandom numbers is the linear congruential method, which works as follows. Let  $x_0$  be an initial number (the “seed”). The sequence is generated recursively as

$$x_n = (ax_{n-1} + c) \bmod m$$

- a. Choose values of  $a$ ,  $c$ , and  $m$ , and try this out. Do the sequences “look” random?
- b. Making good choices of  $a$ ,  $c$ , and  $m$  involves both art and theory. The following are some values that have been proposed: (1)  $a = 69069$ ,  $c = 0$ ,  $m = 2^{31}$ ; (2)  $a = 65539$ ,  $c = 0$ ,  $m = 2^{31}$ . The latter is an infamous generator called RANDU. Try out these schemes, and examine the results.