3.15: Problems

Problems

1. At each toss of a die, the die lands with one face on top. This face is distinguished from the other five faces by the number of dots that appear on it. Tossing a die produces data. What is the distribution? What is the random variable of this distribution? What outcomes are possible for this distribution? How would we collect a sample of ten values of the random variable of this distribution?

2. Suppose that we toss a die three times and average the results observed. How would you describe the distribution from which this average is derived? What is the random variable of this distribution? What outcomes are possible for this distribution? What would we do to collect a sample of ten values of the random variable of this distribution?

3. Suppose that we toss three dice simultaneously and average the results observed. How would you describe the distribution from which this average is derived? What is the random variable of this distribution? What outcomes are possible for this distribution? What would we do to collect a sample of ten values of the random variable of this distribution? Suppose that some third party collects a set, call it A, of ten values from this distribution and a second set, call it B, of values from the distribution in problem 2. If we are given the data in each set but are not told which label goes with which set of data, can we analyze the data to determine which set is A and which is B?

4. The manufacturing process for an electronic component produces 3 bad components in every 1000 components produced. The bad components appear randomly. What is the probability that

(a) a randomly selected component is bad?

(b) a randomly selected component is good?

(c) 2 bad components are produced in succession?

(d) 100 good components are produced in succession?

5. A product incorporates two of the components in the previous problem. What is the probability that

(a) both components are good?

(b) both components are bad?

(c) one component is good and one component is bad?

(d) at least one component is good?

6. A card is selected at random from a well-shuffled deck. A second card is then selected at random from among the remaining 51 cards. What is the probability that

(a) the first card is a heart?

(b) the second card is a heart?

(c) neither card is a heart?

(d) both cards are hearts?

(e) at least one card is a heart?

7. A graduating class has 70 men and 77 women. How many combinations of homecoming king and queen are possible?

8. After the queen is selected from the graduating class of problem 7, one woman is selected to “first attendant” to the homecoming queen. Thereafter, another woman is selected to be “second attendant.” After the queen is selected, how many ways can two attendants be selected?

9. A red die and a green die are rolled. What is the probability that

(a) both come up 3?

(b) both come up the same?

(c) they come up different?

(d) the red die comes up less than the green die?

(e) the red die comes up exactly two less than the green die?

(f) together they show 5?

10. A television game show offers a contestant a new car as the prize for correctly guessing which of three doors the car is behind. After the contestant selects a door, the game-show host opens an incorrect door. The host then gives the contestant the option of switching from the door he originally chose to the other door that remains unopened. Should the contestant change his selection?

[Hint: Consider the final set of outcomes to result from a sequence of three choices. First, the game-show producer selects a door and places the car behind this door. Diagram the possibilities. What is the probability of each? Second, the contestant selects a door. There are now nine possible outcomes. Diagram them. What is the probability of each? Third, the host opens a door. There are now twelve possible outcomes. Diagram them. What is the probability of each? Note that these twelve possibilities are not all equally probable.]

11. For a particular distribution, possible values of the random variable, x, range from zero to one. The probability density function for this distribution is $df/dx=1$.

(a) Show that the probability of finding x in the range $0 \le x \le 1$ is one.

(b) What is the mean of this distribution?

(c) What is the variance of this distribution? The standard deviation?

(d) A quantity, g, is a function of x: $g\left(x\right)=x^2$. What is the expected value of g?

12. For a particular distribution, possible values of the random variable, x, range from one to three. The probability density function for this distribution is $df/dx=cx$, where c is a constant.

(a) What is the value of the constant, c?

(b) What is the mean of this distribution?

(c) What is the variance of this distribution? The standard deviation?

(d) If $g\left(x\right)=x^2$, what is the expected value of g?

13. For a particular distribution, possible values of the random variable, x, range from two to four. The probability density function for this distribution is $df/dx=cx^3$, where c is a constant.

(a) What is the value of the constant, c?

(b) What is the mean of this distribution?

(c) What is the variance of this distribution? The standard deviation?

(d) If $g\left(x\right)=x^2$, what is the expected value of g?

14. For a particular distribution, possible values of the random variable, x, range from zero to four. For $0 \le x \le \le 1$, the probability density function is $df/dx=x/2$. For $1, theprobabilitydensityfunctionis\((df/dx=(4-x)/6$.

(a) Show that the area under this probability distribution function is one.

(b) What is the mean of this distribution?

(c) What is the variance of this distribution? The standard deviation?

(d) If $g\left(x\right)=x^2$, what is the expected value of g?

15. The following values, x_i, of the random variable, x, are drawn from a distribution: 9.63, 9.00, 11.87, 10.13, 10.83, 9.50, 10.40, 9.83, and 10.09.

(a) Arrange these values in increasing order and calculate the “rank probability,” $i/(N+1)$, associated with each of the x_i values.

(b) Plot the rank probability (on the ordinate) versus the random-variable value (on the abscissa). Sketch a smooth curve through the points on this plot.

(c) What function is approximated by the curve sketched in part b?

(d) Plot the data points along a horizontal axis. Then create a bar graph (histogram) by erecting bars of equal area between each pair of data points.

(e) What function is approximated by the tops of the bars erected in part d?

16. For a particular distribution, possible values of the random variable range from zero to four. The following values of the random variable are drawn from this distribution: 0.1, 1.0, 1.1, 1.5, 2.1. Sketch an approximate probability density function for this distribution.

17. The possible values for the random variable of a particular distribution lie in the range $0 \le x \le 10$. In six trials, the following values are obtained: 1.0, 1.9, 2.3, 2.7, 3.0, 3.8.

(a) Sketch an approximate probability density function for this distribution.

(b) What is the best estimate we can make of the mean of this distribution?

(c) What is the best estimate we can make of the variance of this distribution?

(d) What is the best estimate we can make of the variance of averages-of-six drawn from this distribution?

(e) What is the best estimate we can make of the variance of averages-of-sixteen drawn from this distribution?

18. A computer program generates numbers from a normal distribution with a mean of zero and a standard deviation of 10. Also, for any integer N, the program will generate and average N values from this distribution. It will repeat this operation until it has produced 100 such averages. It will then compute the estimated standard deviation of these 100 average values. The table below gives various values of N and the estimated standard deviation, s, that was found for 100 averages of that N. Plot these data in a way that tests the validity of the central limit theorem.

N	s
4	5.182
9	2.794
16	2.206
25	2.152
36	1.689
49	1.092
64	1.001
81	1.004
100	1.074
144	0.601
196	0.546
256	0.690
324	0.545

19. If $f(u)$ is the cumulative probability distribution function for a distribution, what is the expected value of $f(u)$? What interpretation can you place on this result?

20. Five replications of a volumetric analysis yield concentration estimates of 0.3000, 0.3008, 0.3012, 0.3014, and 0.3020 mol L^{-1}. Calculate the rank probability of each of these results. Sketch, over the concentration range \(0.3000<0.3020\)> mol L^{-1}, an approximation of the cumulative probability distribution function for the distribution that yielded these data.

21. The Louisville Mudhens play on a square baseball field that measures 100 meters on a side. Casey's hits always fall on the field. (He never hits a foul ball or hits one out of the park.) The probability density function for the distance that a Casey hit goes parallel to the first-base line is $df_x(x)/dx = \left(2 \times 10^{-4}\right) x$. (That is, we take the first-base line as our x-axis; the third-base line as

our y-axis; and home plate is at the origin. $df_x(x)/dx$ is independent of the distance that the hit goes parallel to the third-base line, our y-axis.) The probability density function for the distance that a Casey hit goes parallel to the third-base line is $df_y(y)/dy = \left(3\times 10^{-6}\right) y^2$. ($df_y(y)/dy$ is independent of the distance that the hit goes parallel to the first-base line, our x-axis.)

(a) What is the probability that a Casey hit lands at a point (x,y) such that \(x^*<x^*+dx\)> and \(y^*<y^*+dy\)> ?

(b) What is the two-dimensionally probability density function that describes Casey's hits, expressed in this Cartesian coordinate system?

(c) Recall that polar coordinates transform to Cartesian coordinates according to $x = r\cos\theta$ and $y = r\sin\theta$. What is the probability density function for Casey's hits expressed using polar coordinates?

(d) Recall that the differential element of area in polar coordinates is $rdrd\theta$. Find the probability that a Casey hit lands within the pie-shaped area bounded by \(0<50\)> m and $0<\theta<\pi/4$.

22. In Chapter 2, we derived the Barometric Formula, $\eta(h) = \eta(0)\exp(-mgh/kT)$ for molecules of mass m in an isothermal atmosphere at a height h above the surface of the earth. $\eta(h)$ is the number of molecules per unit volume at height h ; $\eta(0)$ is the number of molecules per unit volume at the earth's surface, where $h=0$. Consider a vertical cylinder of unit cross-sectional area, extending from the earth's surface to an infinite height. Let $f(h)$ be the fraction of the molecules in this cylinder that is at a height less than h. Prove that the probability density function is $df/dh = (mg/kT)\exp(-mgh/kT)$.

23. A particular distribution has six outcomes. These outcomes and their probabilities are a [GrindEQ__0_1_]; b [GrindEQ__0_2_]; c [GrindEQ__0_3_]; d [GrindEQ__0_2_]; e [GrindEQ__0_1_]; and f [GrindEQ__0_1_].

(a) Partitioning I assigns these outcomes to a set of three events: Event $A = a$ or b or c; Event $B = d$; and Event $C = e$ or f. What are the probabilities of Events A, B, and C?

(b) Partitioning II assigns the outcomes to two events: Event $D = a$ or b or c; and Event $E = d$ or e or f. What are the probabilities of Events D and E? Express the probabilities of Events D and E in terms of the probabilities of Events A, B, and C.

(c) Partitioning III assigns the outcomes to three events: Event $F = a$ or b; Event $G = c$ or d; and Event $H = e$ or f. What are the probabilities of Events F, G, and H? Can the probabilities of Events F, G, and H be expressed in terms of the probabilities of Events A, B, and C?

24. Consider a partitioning of outcomes into events that is not exhaustive; that is, not every outcome is assigned to an event. What problem arises when we want to describe the probabilities of these events?

25. Consider a partitioning of outcomes into events that is not mutually exclusive; that is, one (or more) outcome is assigned to two (or more) events. What problem arises when we want to describe the probabilities of these events?

26. For integer values of p $(p\neq 1)$, we find

$$\int x^p \ln(x)\, dx = \left(\frac{x^{p+1}}{p+1}\right)\ln(x) - \frac{x^{p+1}}{(p+1)^2}$$

(a) Sketch the function, $h(x) = df(x)/dx = -4x\ln(x)$, over the interval $0\le x\le 1$.

(b) Show that we can consider $h(x) = df(x)/dx = -4x\ln(x)$ to be a probability density function over this interval; that is, show $f(1) - f(0) = 1$. Let us name the corresponding distribution "Sam."

(c) What is the mean, μ, of Sam?

(d) What is the variance, σ^2, of Sam?

(e) What is the standard deviation, σ, of Sam?

(f) What is the variance of averages-of-four samples taken from Sam?

(g) The following four values are obtained in random sampling of an unknown distribution: 0.050; 0.010; 0.020; and 0.040. Estimate the mean, μ, variance (σ^2or s^2), and the standard deviation (σ or s) for this unknown distribution.

(h) What is the probability that a single sample drawn from Sam will lie in the interval $0\le x\le 0.10$? Note: The upper limit of this interval is 0.10, not 1.0 as in part (a).

(i) Is it likely that the unknown distribution sampled in part g is in fact the distribution we named Sam? Why, or why not?

27. We define the mean, μ, as the expected value of the random variable: $\mu=\int_{-\infty}^{\infty} u\left(df/du\right)du$. Define $\overline{u}=\sum_{i=1}^{N}\left(u_i/N\right)$, where the u_i are N independent values of the random variable. Show that the expected value of $\overline{u}$ is μ.

28. A box contains a large number of plastic balls. An integer, W, in the range $1\le W\le 20$ is printed on each ball. There are many balls printed with each integer. The integer specifies the mass of the ball in grams. Six random samples of three balls each are drawn from the box. The balls are replaced and the box is shaken between drawings. The numbers on the balls in drawings I through VI are:

I: 3, 4, 9

II: 1, 6, 17

III: 2, 5, 8

IV: 2, 6, 7

V: 3, 5, 6

VI: 2, 3, 10

(a) What are the population sets represented by the samples I through VI?

(b) Sketch the probability density function as estimated from sample I.

(c) Sketch the probability density function as estimated from sample II

(d) Using the data from samples I through VI, estimate the probability of drawing a ball of each mass in a single trial.

(e) Sketch the probability density function as estimated from the probability values in part (d).

(f) From the data in sample I, estimate the average mass of a ball in the box.

(g) From the data in sample II, estimate the average mass of a ball in the box.

(h) From the probability values calculated in part (d), estimate the average mass of a ball in the box.