4.2 Probability Distributions for Discrete Random Variables

LEARNING OBJECTIVES

1. To learn the concept of the probability distribution of a discrete random variable.
2. To learn the concepts of the mean, variance, and standard deviation of a discrete random variable, and how to compute them.

Probability Distributions

Associated to each possible value x of a discrete random variable X is the probability $P(x)$ that X will take the value x in one trial of the experiment.

Definition

*The **probability distribution** of a discrete random variable X is a list of each possible value of X together with the probability that X takes that value in one trial of the experiment.*

The probabilities in the probability distribution of a random variable *X* must satisfy the following two conditions:

1. Each probability $P(x)$ must be between 0 and 1: $0 \leq P(x) \leq 1$.
2. The sum of all the probabilities is 1: $\Sigma P(x) = 1$.

EXAMPLE 1

A fair coin is tossed twice. Let *X* be the number of heads that are observed.

a. Construct the probability distribution of *X*.
b. Find the probability that at least one head is observed.

Solution:

a. The possible values that *X* can take are 0, 1, and 2. Each of these numbers corresponds to an event in the sample space $S = \{hh, ht, th, tt\}$ of equally likely outcomes for this experiment: *X* = 0 to $\{tt\}$, *X* = 1 to $\{ht, th\}$, and *X* = 2 to $\{hh\}$. The probability of each of these events, hence of the corresponding value of *X*, can be found simply by counting, to give

x	0	1	2
$P(x)$	0.25	0.50	0.25

This table is the probability distribution of *X*.

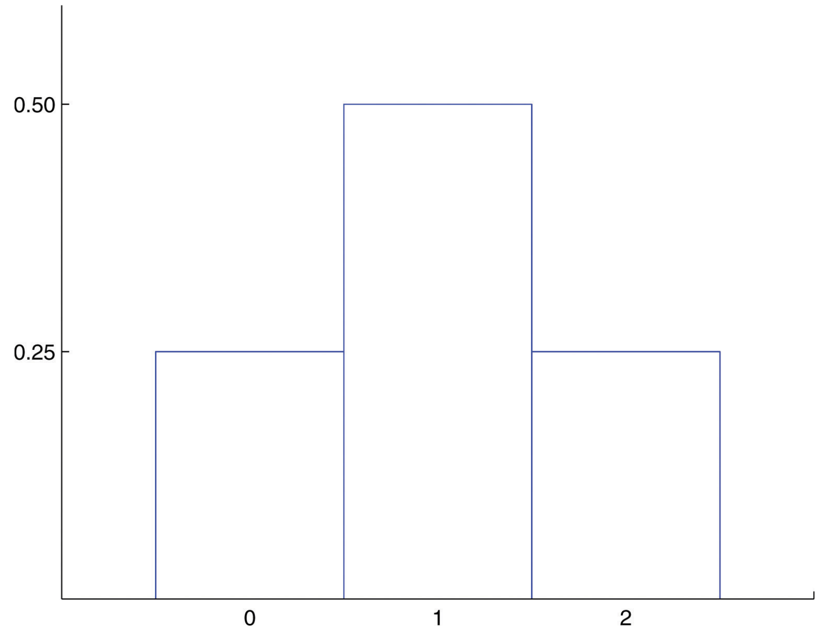
b. "At least one head" is the event *X* ≥ 1, which is the union of the mutually exclusive events $X = 1$ and $X = 2$. Thus

$$P(X \geq 1) = P(1) + P(2) = 0.50 + 0.25 = 0.75$$

A histogram that graphically illustrates the probability distribution is given in Figure 4.1 "Probability Distribution for Tossing a Fair Coin Twice".

Figure 4.1

Probability Distribution for Tossing a Fair Coin Twice



EXAMPLE 2

A pair of fair dice is rolled. Let X denote the sum of the number of dots on the top faces.

a. Construct the probability distribution of X.
b. Find $P(X \geq 9)$.
c. Find the probability that X takes an even value.

Solution:

The sample space of equally likely outcomes is

11	12	13	14	15	16
21	22	23	24	25	26
31	32	33	34	35	36
41	42	43	44	45	46
51	52	53	54	55	56
61	62	63	64	65	66

a. The possible values for X are the numbers 2 through 12. $X = 2$ is the event {11}, so $P(2) = 1/36$. $X = 3$ is the event {12,21}, so $P(3) = 2/36$. Continuing this way we obtain the table

x	2	3	4	5	6	7	8	9	10	11	12
$P(x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

This table is the probability distribution of X.

b. The event $X \geq 9$ is the union of the mutually exclusive events $X = 9$, $X = 10$, $X = 11$, and $X = 12$. Thus

$$P(X \geq 9) = P(9) + P(10) + P(11) + P(12) = \frac{4}{36} + \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{10}{36} = 0.2\overline{7}$$

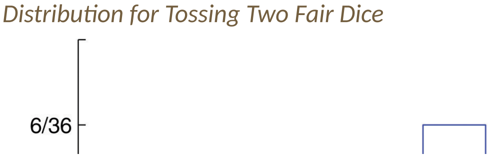
c. Before we immediately jump to the conclusion that the probability that X takes an even value must be 0.5, note that X takes six different even values but only five different odd values. We compute

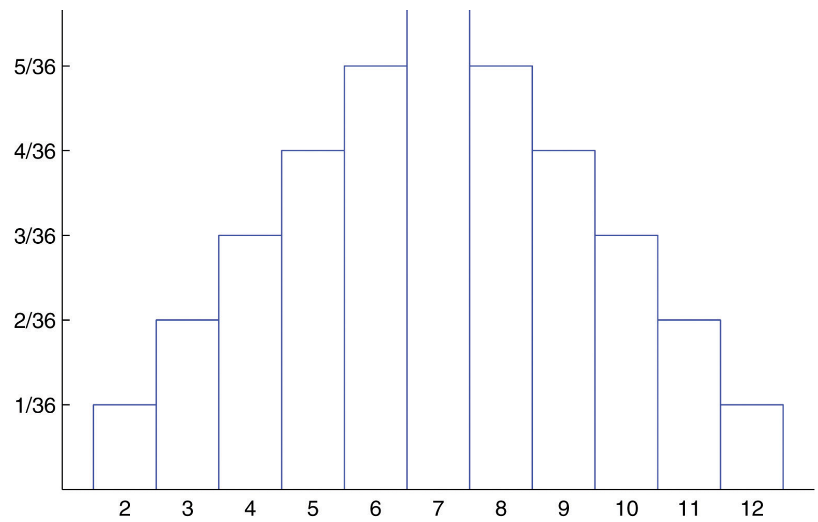
$$\begin{aligned} P(X \text{ is even}) &= P(2) + P(4) + P(6) + P(8) + P(10) + P(12) \\ &= \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{5}{36} + \frac{3}{36} + \frac{1}{36} = \frac{18}{36} = 0.5 \end{aligned}$$

A histogram that graphically illustrates the probability distribution is given in Figure 4.2 "Probability Distribution for Tossing Two Fair Dice".

Figure 4.2

Probability Distribution for Tossing Two Fair Dice





The Mean and Standard Deviation of a Discrete Random Variable

Definition

The **mean** *(also called the* **expected value***) of a discrete random variable X is the number*

$$\mu = E(X) = \Sigma\, x\, P(x)$$

The mean of a random variable may be interpreted as the average of the values assumed by the random variable in repeated trials of the experiment.

EXAMPLE 3

Find the mean of the discrete random variable *X* whose probability distribution is

x	−2	1	2	3.5
$P(x)$	0.21	0.34	0.24	0.21

Solution:

The formula in the definition gives

$$\begin{aligned}\mu &= \Sigma x\, P(x) \\ &= (-2)\cdot 0.21 + (1)\cdot 0.34 + (2)\cdot 0.24 + (3.5)\cdot 0.21 = 1.135\end{aligned}$$

EXAMPLE 4

A service organization in a large town organizes a raffle each month. One thousand raffle tickets are sold for $1 each. Each has an equal chance of winning. First prize is $300, second prize is $200, and third prize is $100. Let *X* denote the net gain from the purchase of one ticket.

a. Construct the probability distribution of *X*.
b. Find the probability of winning any money in the purchase of one ticket.
c. Find the expected value of *X*, and interpret its meaning.

Solution:

a. If a ticket is selected as the first prize winner, the net gain to the purchaser is the $300 prize less the $1 that was paid for the ticket, hence *X* = 300 − 1 = 299. There is one such ticket, so *P*(299) = 0.001. Applying the same "income minus outgo" principle to the second and third prize winners and to the 997 losing tickets yields the probability distribution:

x	299	199	99	−1
$P(x)$	0.001	0.001	0.001	0.997

b. Let *W* denote the event that a ticket is selected to win one of the prizes. Using the table

$$P(W) = P(299) + P(199) + P(99) = 0.001 + 0.001 + 0.001 = 0.003$$

c. Using the formula in the definition of expected value,

$$E(X) = 299 \cdot 0.001 + 199 \cdot 0.001 + 99 \cdot 0.001 + (-1) \cdot 0.997 = -0.4$$

The negative value means that one loses money on the average. In particular, if someone were to buy tickets repeatedly, then although he would win now and then, on average he would lose 40 cents per ticket purchased.

The concept of expected value is also basic to the insurance industry, as the following simplified example illustrates.

EXAMPLE 5

A life insurance company will sell a $200,000 one-year term life insurance policy to an individual in a particular risk group for a premium of $195. Find the expected value to the company of a single policy if a person in this risk group has a 99.97% chance of surviving one year.

Solution:

Let *X* denote the net gain to the company from the sale of one such policy. There are two possibilities: the insured person lives the whole year or the insured person dies before the year is up. Applying the “income minus outgo” principle, in the former case the value of *X* is 195 − 0; in the latter case it is $195 - 200{,}000 = -199{,}805$. Since the probability in the first case is 0.9997 and in the second case is $1 - 0.9997 = 0.0003$, the probability distribution for *X* is:

x	195	$-199{,}805$
$P(x)$	0.9997	0.0003

Therefore

$$E(X) = \Sigma\, x\, P(x) = 195 \cdot 0.9997 + (-199{,}805) \cdot 0.0003 = 135$$

Occasionally (in fact, 3 times in 10,000) the company loses a large amount of money on a policy, but typically it gains $195, which by our computation of $E(X)$ works out to a net gain of $135 per policy sold, on average.

Definition

The **variance**, σ^2, *of a discrete random variable* X *is the number*

$$\sigma^2 = \Sigma(x - \mu)^2 \, P(x)$$

which by algebra is equivalent to the formula

$$\sigma^2 = \left[\Sigma \, x^2 \, P(x)\right] - \mu^2$$

Definition

The **standard deviation**, σ, *of a discrete random variable* X *is the square root of its variance, hence is given by the formulas*

$$\sigma = \sqrt{\Sigma(x - \mu)^2 \, P(x)} = \sqrt{\left[\Sigma \, x^2 \, P(x)\right] - \mu^2}$$

The variance and standard deviation of a discrete random variable X may be interpreted as measures of the variability of the values assumed by the random variable in repeated trials of the experiment. The units on the standard deviation match those of X.

EXAMPLE 6

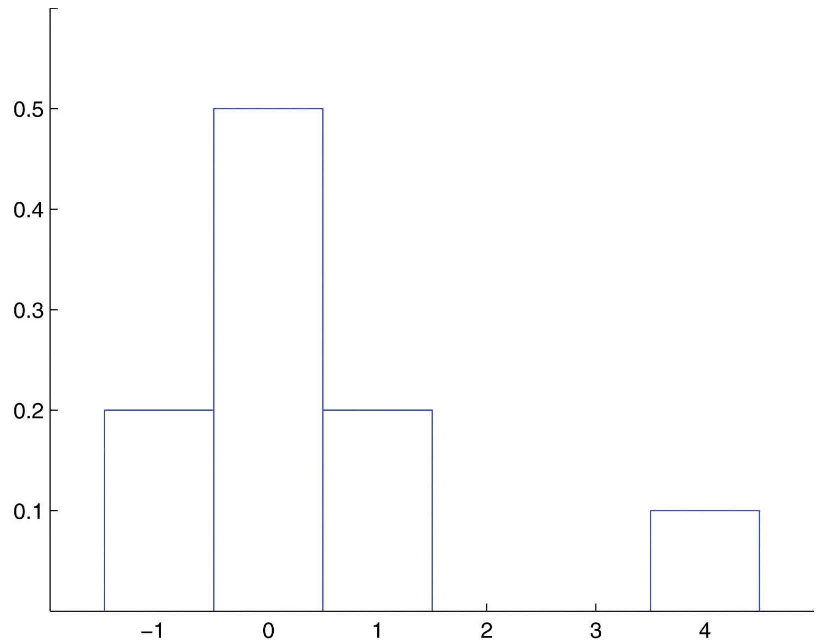
A discrete random variable X has the following probability distribution:

x	-1	0	1	4
$P(x)$	0.2	0.5	a	0.1

A histogram that graphically illustrates the probability distribution is given in Figure 4.3 "Probability Distribution of a Discrete Random Variable".

Figure 4.3

Probability Distribution of a Discrete Random Variable



Compute each of the following quantities.

a. a.

b. $P(0)$.

c. $P(X > 0)$.

d. $P(X \geq 0)$.

e. $P(X \leq -2)$.

f. The mean μ of X.

g. The variance σ^2 of X.

h. The standard deviation σ of X.

Solution:

a. Since all probabilities must add up to 1, $a = 1 - (0.2 + 0.5 + 0.1) = 0.2$.

b. Directly from the table, $P(0) = 0.5$.

c. From the table, $P(X > 0) = P(1) + P(4) = 0.2 + 0.1 = 0.3$.

d. From the table, $P(X \geq 0) = P(0) + P(1) + P(4) = 0.5 + 0.2 + 0.1 = 0.8$.

e. Since none of the numbers listed as possible values for *X* is less than or equal to −2, the event $X \leq -2$ is impossible, so $P(X \leq -2) = 0$.

f. Using the formula in the definition of μ,

$$\mu = \Sigma x\, P(x) = (-1) \cdot 0.2 + 0 \cdot 0.5 + 1 \cdot 0.2 + 4 \cdot 0.1 = 0.4$$

g. Using the formula in the definition of σ^2 and the value of μ that was just computed,

$$\begin{aligned}\sigma^2 &= \Sigma(x - \mu)^2 P(x) \\ &= (-1 - 0.4)^2 \cdot 0.2 + (0 - 0.4)^2 \cdot 0.5 + (1 - 0.4)^2 \cdot 0.2 + (4 - 0.4)^2 \cdot 0.1 \\ &= 1.84\end{aligned}$$

h. Using the result of part (g), $\sigma = \sqrt{1.84} = 1.3565$.

KEY TAKEAWAYS

- The probability distribution of a discrete random variable *X* is a listing of each possible value *x* taken by *X* along with the probability $P(x)$ that *X* takes that value in one trial of the experiment.
- The mean μ of a discrete random variable *X* is a number that indicates the average value of *X* over numerous trials of the experiment. It is computed using the formula $\mu = \Sigma x\, P(x)$.
- The variance σ^2 and standard deviation σ of a discrete random variable *X* are numbers that indicate the variability of *X* over numerous trials of the experiment. They may be computed using the formula $\sigma^2 = \left[\Sigma x^2\, P(x)\right] - \mu^2$, taking the square root to obtain σ.

EXERCISES

BASIC

1. Determine whether or not the table is a valid probability distribution of a discrete random variable. Explain fully.

   a.

x	−2	0	2	4
$P(x)$	0.3	0.5	0.2	0.1

   b.

x	0.5	0.25	0.25
$P(x)$	−0.4	0.6	0.8

   c.

x	1.1	2.5	4.1	4.6	5.3
$P(x)$	0.16	0.14	0.11	0.27	0.22

2. Determine whether or not the table is a valid probability distribution of a discrete random variable. Explain fully.

   a.

x	0	1	2	3	4
$P(x)$	−0.25	0.50	0.35	0.10	0.30

   b.

x	1	2	3
$P(x)$	0.325	0.406	0.164

   c.

x	25	26	27	28	29
$P(x)$	0.13	0.27	0.28	0.18	0.14

3. A discrete random variable *X* has the following probability distribution:

x	77	78	79	80	81
$P(x)$	0.15	0.15	0.20	0.40	0.10

Compute each of the following quantities.

a. $P(80)$.

b. $P(X > 80)$.

c. $P(X \le 80)$.

d. The mean μ of X.

e. The variance σ^2 of X.

f. The standard deviation σ of X.

4. A discrete random variable X has the following probability distribution:

x	13	18	20	24	27
$P(x)$	0.22	0.25	0.20	0.17	0.16

Compute each of the following quantities.

a. $P(18)$.

b. $P(X > 18)$.

c. $P(X \le 18)$.

d. The mean μ of X.

e. The variance σ^2 of X.

f. The standard deviation σ of X.

5. If each die in a pair is “loaded” so that one comes up half as often as it should, six comes up half again as often as it should, and the probabilities of the other faces are unaltered, then the probability distribution for the sum X of the number of dots on the top faces when the two are rolled is

x	2	3	4	5	6	7
$P(x)$	$\frac{1}{144}$	$\frac{4}{144}$	$\frac{8}{144}$	$\frac{12}{144}$	$\frac{16}{144}$	$\frac{22}{144}$

x	8	9	10	11	12
$P(x)$	$\frac{24}{144}$	$\frac{20}{144}$	$\frac{16}{144}$	$\frac{12}{144}$	$\frac{9}{144}$

Compute each of the following.

a. $P(5 \le X \le 9)$.

b. $P(X \ge 7)$.

c. The mean μ of X. (For fair dice this number is 7.)

d. The standard deviation σ of X. (For fair dice this number is about 2.415.)

APPLICATIONS

6. Borachio works in an automotive tire factory. The number *X* of sound but blemished tires that he produces on a random day has the probability distribution

x	2	3	4	5
$P(x)$	0.48	0.36	0.12	0.04

   a. Find the probability that Borachio will produce more than three blemished tires tomorrow.
   b. Find the probability that Borachio will produce at most two blemished tires tomorrow.
   c. Compute the mean and standard deviation of *X*. Interpret the mean in the context of the problem.

7. In a hamster breeder's experience the number *X* of live pups in a litter of a female not over twelve months in age who has not borne a litter in the past six weeks has the probability distribution

x	3	4	5	6	7	8	9
$P(x)$	0.04	0.10	0.26	0.31	0.22	0.05	0.02

   a. Find the probability that the next litter will produce five to seven live pups.
   b. Find the probability that the next litter will produce at least six live pups.
   c. Compute the mean and standard deviation of *X*. Interpret the mean in the context of the problem.

8. The number *X* of days in the summer months that a construction crew cannot work because of the weather has the probability distribution

x	6	7	8	9	10
$P(x)$	0.03	0.08	0.15	0.20	0.19

x	11	12	13	14
$P(x)$	0.16	0.10	0.07	0.02

   a. Find the probability that no more than ten days will be lost next summer.
   b. Find the probability that from 8 to 12 days will be lost next summer.
   c. Find the probability that no days at all will be lost next summer.
   d. Compute the mean and standard deviation of *X*. Interpret the mean in the context of the problem.

9. Let *X* denote the number of boys in a randomly selected three-child family. Assuming that boys

and girls are equally likely, construct the probability distribution of *X*.

10. Let *X* denote the number of times a fair coin lands heads in three tosses. Construct the probability distribution of *X*.

11. Five thousand lottery tickets are sold for $1 each. One ticket will win $1,000, two tickets will win $500 each, and ten tickets will win $100 each. Let *X* denote the net gain from the purchase of a randomly selected ticket.

    a. Construct the probability distribution of *X*.
    b. Compute the expected value $E(X)$ of *X*. Interpret its meaning.
    c. Compute the standard deviation σ of *X*.

12. Seven thousand lottery tickets are sold for $5 each. One ticket will win $2,000, two tickets will win $750 each, and five tickets will win $100 each. Let *X* denote the net gain from the purchase of a randomly selected ticket.

    a. Construct the probability distribution of *X*.
    b. Compute the expected value $E(X)$ of *X*. Interpret its meaning.
    c. Compute the standard deviation σ of *X*.

13. An insurance company will sell a $90,000 one-year term life insurance policy to an individual in a particular risk group for a premium of $478. Find the expected value to the company of a single policy if a person in this risk group has a 99.62% chance of surviving one year.

14. An insurance company will sell a $10,000 one-year term life insurance policy to an individual in a particular risk group for a premium of $368. Find the expected value to the company of a single policy if a person in this risk group has a 97.25% chance of surviving one year.

15. An insurance company estimates that the probability that an individual in a particular risk group will survive one year is 0.9825. Such a person wishes to buy a $150,000 one-year term life insurance policy. Let *C* denote how much the insurance company charges such a person for such a policy.

    a. Construct the probability distribution of *X*. (Two entries in the table will contain *C*.)
    b. Compute the expected value $E(X)$ of *X*.
    c. Determine the value *C* must have in order for the company to break even on all such policies (that is, to average a net gain of zero per policy on such policies).

d. Determine the value *C* must have in order for the company to average a net gain of $250 per policy on all such policies.

16. An insurance company estimates that the probability that an individual in a particular risk group will survive one year is 0.99. Such a person wishes to buy a $75,000 one-year term life insurance policy. Let *C* denote how much the insurance company charges such a person for such a policy.

    a. Construct the probability distribution of *X*. (Two entries in the table will contain *C*.)
    b. Compute the expected value $E(X)$ of *X*.
    c. Determine the value *C* must have in order for the company to break even on all such policies (that is, to average a net gain of zero per policy on such policies).
    d. Determine the value *C* must have in order for the company to average a net gain of $150 per policy on all such policies.

17. A roulette wheel has 38 slots. Thirty-six slots are numbered from 1 to 36; half of them are red and half are black. The remaining two slots are numbered 0 and 00 and are green. In a $1 bet on red, the bettor pays $1 to play. If the ball lands in a red slot, he receives back the dollar he bet plus an additional dollar. If the ball does not land on red he loses his dollar. Let *X* denote the net gain to the bettor on one play of the game.

    a. Construct the probability distribution of *X*.
    b. Compute the expected value $E(X)$ of *X*, and interpret its meaning in the context of the problem.
    c. Compute the standard deviation of *X*.

18. A roulette wheel has 38 slots. Thirty-six slots are numbered from 1 to 36; the remaining two slots are numbered 0 and 00. Suppose the "number" 00 is considered not to be even, but the number 0 is still even. In a $1 bet on even, the bettor pays $1 to play. If the ball lands in an even numbered slot, he receives back the dollar he bet plus an additional dollar. If the ball does not land on an even numbered slot, he loses his dollar. Let *X* denote the net gain to the bettor on one play of the game.

    a. Construct the probability distribution of *X*.
    b. Compute the expected value $E(X)$ of *X*, and explain why this game is not offered in a casino (where 0 is not considered even).
    c. Compute the standard deviation of *X*.

19. The time, to the nearest whole minute, that a city bus takes to go from one end of its route to the other has the probability distribution shown. As sometimes happens with probabilities computed as empirical relative frequencies, probabilities in the table add up only to a value other than 1.00 because of round-off error.

x	42	43	44	45	46	47
$P(x)$	0.10	0.23	0.34	0.25	0.05	0.02

a. Find the average time the bus takes to drive the length of its route.

b. Find the standard deviation of the length of time the bus takes to drive the length of its route.

20. Tybalt receives in the mail an offer to enter a national sweepstakes. The prizes and chances of winning are listed in the offer as: $5 million, one chance in 65 million; $150,000, one chance in 6.5 million; $5,000, one chance in 650,000; and $1,000, one chance in 65,000. If it costs Tybalt 44 cents to mail his entry, what is the expected value of the sweepstakes to him?

ADDITIONAL EXERCISES

21. The number *X* of nails in a randomly selected 1-pound box has the probability distribution shown. Find the average number of nails per pound.

x	100	101	102
$P(x)$	0.01	0.96	0.03

22. Three fair dice are rolled at once. Let *X* denote the number of dice that land with the same number of dots on top as at least one other die. The probability distribution for *X* is

x	0	u	3
$P(x)$	p	$\frac{15}{36}$	$\frac{1}{36}$

a. Find the missing value *u* of *X*.

b. Find the missing probability *p*.

c. Compute the mean of *X*.

d. Compute the standard deviation of *X*.

23. Two fair dice are rolled at once. Let *X* denote the difference in the number of dots that appear on the top faces of the two dice. Thus for example if a one and a five are rolled, $X = 4$, and if two sixes are rolled, $X = 0$.

a. Construct the probability distribution for *X*.

b. Compute the mean μ of *X*.

c. Compute the standard deviation σ of *X*.

24. A fair coin is tossed repeatedly until either it lands heads or a total of five tosses have been made, whichever comes first. Let *X* denote the number of tosses made.

a. Construct the probability distribution for *X*.

b. Compute the mean μ of *X*.

c. Compute the standard deviation σ of *X*.

25. A manufacturer receives a certain component from a supplier in shipments of 100 units. Two units in each shipment are selected at random and tested. If either one of the units is defective the shipment is rejected. Suppose a shipment has 5 defective units.

a. Construct the probability distribution for the number *X* of defective units in such a sample. (A tree diagram is helpful.)

b. Find the probability that such a shipment will be accepted.

26. Shylock enters a local branch bank at 4:30 p.m. every payday, at which time there are always two tellers on duty. The number *X* of customers in the bank who are either at a teller window or are waiting in a single line for the next available teller has the following probability distribution.

x	0	1	2	3
$P(x)$	0.135	0.192	0.284	0.230

x	4	5	6
$P(x)$	0.103	0.051	0.005

a. What number of customers does Shylock most often see in the bank the moment he enters?

b. What number of customers waiting in line does Shylock most often see the moment he enters?

c. What is the average number of customers who are waiting in line the moment Shylock enters?

27. The owner of a proposed outdoor theater must decide whether to include a cover that will allow shows to be performed in all weather conditions. Based on projected audience sizes and weather conditions, the probability distribution for the revenue *X* per night if the cover is not

installed is

Weather	x	$P(x)$
Clear	\$3000	0.61
Threatening	\$2800	0.17
Light rain	\$1975	0.11
Show-cancelling rain	\$0	0.11

The additional cost of the cover is $410,000. The owner will have it built if this cost can be recovered from the increased revenue the cover affords in the first ten 90-night seasons.

a. Compute the mean revenue per night if the cover is not installed.
b. Use the answer to (a) to compute the projected total revenue per 90-night season if the cover is not installed.
c. Compute the projected total revenue per season when the cover is in place. To do so assume that if the cover were in place the revenue each night of the season would be the same as the revenue on a clear night.
d. Using the answers to (b) and (c), decide whether or not the additional cost of the installation of the cover will be recovered from the increased revenue over the first ten years. Will the owner have the cover installed?

ANSWERS

1. a. no: the sum of the probabilities exceeds 1
   b. no: a negative probability
   c. no: the sum of the probabilities is less than 1

3. a. 0.4
   b. 0.1
   c. 0.9
   d. 79.15
   e. $\sigma^2 = 1.5275$
   f. $\sigma = 1.2359$

5. a. 0.6528
   b. 0.7153

c. μ = 7.8333

d. σ^2 = 5.4866

e. σ = 2.3424

7. a. 0.79

b. 0.60

c. μ = 5.8, σ = 1.2570

9.

x	0	1	2	3
$P(x)$	1/8	3/8	3/8	1/8

11. a.

x	−1	999	499	99
$P(x)$	$\frac{4987}{5000}$	$\frac{1}{5000}$	$\frac{2}{5000}$	$\frac{10}{5000}$

b. −0.4

c. 17.8785

13. 136

15. a.

x	C	$C-150{,}000$
$P(x)$	0.9825	0.0175

b. $C-2625$

c. $C \geq 2625$

d. $C \geq 2875$

17. a.

x	−1	1
$P(x)$	$\frac{20}{38}$	$\frac{18}{38}$

b. $E(X)$ = −0.0526 In many bets the bettor sustains an average loss of about 5.25 cents per bet.

c. 0.9986

19. a. 43.54

    b. 1.2046

21. 101.02

23. a.

x	0	1	2	3	4	5
$P(x)$	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{8}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$

    b. 1.9444

    c. 1.4326

25. a.

x	0	1	2
$P(x)$	0.902	0.096	0.002

    b. 0.902

27. a. 2523.25

    b. 227,092.5

    c. 270,000

    d. The owner will install the cover.