Homework 12
Do at least 4.

Soln's at end

SECTION 4.4 Summary

The probability distribution of a random variable X, like a distribution of data, has a **mean** μ_X and a **standard deviation** σ_X.

The **law of large numbers** says that the average of the values of X observed in many trials must approach μ.

The **mean** μ is the balance point of the probability histogram or density curve. If X is discrete with possible values x_i having probabilities p_i, the mean is the average of the values of X, each weighted by its probability:

$$\mu_X = x_1p_1 + x_2p_2 + \cdots + x_kp_k$$

The **variance** σ_X^2 is the average squared deviation of the values of the variable from their mean. For a discrete random variable,

$$\sigma_X^2 = (x_1 - \mu)^2p_1 + (x_2 - \mu)^2p_2 + \cdots + (x_k - \mu)^2p_k$$

The **standard deviation** σ_X is the square root of the variance. The standard deviation measures the variability of the distribution about the mean. It is easiest to interpret for Normal distributions.

The mean and variance of a continuous random variable can be computed from the density curve, but to do so requires more advanced mathematics.

The means and variances of random variables obey the following rules. If a and b are fixed numbers, then

$$\mu_{a+bX} = a + b\mu_X$$
$$\sigma_{a+bX}^2 = b^2\sigma_X^2$$

If X and Y are any two random variables having correlation ρ, then

$$\mu_{X+Y} = \mu_X + \mu_Y$$
$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y$$
$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2 - 2\rho\sigma_X\sigma_Y$$

If X and Y are **independent,** then $\rho = 0$. In this case,

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2$$
$$\sigma_{X-Y}^2 = \sigma_X^2 + \sigma_Y^2$$

To find the standard deviation, take the square root of the variance.

SECTION 4.4 Exercises

For Exercise 4.67, see page 260; for Exercise 4.68, see page 264; for Exercises 4.69 and 4.70, see page 268; and for Exercise 4.71, see page 270.

4.72 What's wrong? In each of the following scenarios, there is something wrong. Describe what is wrong and give a reason for your answer.

(a) If you toss a fair coin three times and get heads all three times then the probability of getting a tail on the next toss is much greater than a half.

(b) If you multiply a random variable by 10, then the mean is multiplied by 10 and the variance is multiplied by 10.

(c) When finding the mean of the sum of two random variables, you need to know the correlation between them.

4.73 Servings of fruits and vegetables. The following table gives the distribution of the number of servings of fruits and vegetables per day in a population.

Number of servings X	0	1	2	3	4	5
Probability	0.3	0.1	0.1	0.2	0.1	0.2

Find the mean and the standard deviation for this random variable.

4.74 Mean of the distribution for the number of aces. In Exercise 4.54 you examined the probability distribution for the number of aces when you are dealt two cards in the game of Texas hold 'em. Let X represent the number of aces in a randomly selected deal of two cards in this game. Here is the probability distribution for the random variable X:

Value of X	0	1	2
Probability	0.8507	0.1448	0.0045

Find μ_X, the mean of the probability distribution of X.

4.75 Mean of the grade distribution. Example 4.22 gives the distribution of grades (A = 4, B = 3, and so on) in English 210 at North Carolina State University as

Value of X	0	1	2	3	4
Probability	0.05	0.04	0.20	0.40	0.31

Find the average (that is, the mean) grade in this course.

4.76 Mean of the distributions of errors. Typographical and spelling errors can be either "nonword errors" or "word errors." A nonword error is not a real word, as when "the" is typed as "teh." A word error is a real word, but not the right word, as when "lose" is typed as "loose." When undergraduates are asked to write a 250-word essay (without spell-checking), the number of nonword errors has the following distribution:

Errors	0	1	2	3	4
Probability	0.1	0.3	0.3	0.2	0.1

The number of word errors has this distribution:

Errors	0	1	2	3
Probability	0.4	0.3	0.2	0.1

What are the mean numbers of nonword errors and word errors in an essay?

4.77 Standard deviation of the number of aces. Refer to Exercise 4.74. Find the standard deviation of the number of aces.

4.78 Standard deviation of the grades. Refer to Exercise 4.75. Find the standard deviation of the grade distribution.

4.79 Suppose that the correlation is zero. Refer to Example 4.37 (page 273).

(a) Recompute the standard deviation for the total of the natural gas bill and the electricity bill assuming that the correlation is zero.

(b) Is this standard deviation larger or smaller that the standard deviation computed in Example 4.37 (page 273)? Explain why.

4.80 Find the mean of the sum. Figure 4.12 (page 259) displays the density curve of the sum $Y = X_1 + X_2$ of two independent random numbers, each uniformly distributed between 0 and 1.

(a) The mean of a continuous random variable is the balance point of its density curve. Use this fact to find the mean of Y from Figure 4.12.

(b) Use the same fact to find the means of X_1 and X_2. (They have the density curve pictured in Figure 4.9, page 254.) Verify that the mean of Y is the sum of the mean of X_1 and the mean of X_2.

4.81 Calcium supplements and calcium in the diet. Refer to Example 4.38 (page 273). Suppose that people who have high intakes of calcium in their diets are more compliant than those who have low intakes. What effect would this have on the calculation of the standard deviation for the total calcium intake? Explain your answer.

4.82 The effect of correlation. Find the mean and standard deviation of the total number of errors (nonword errors plus word errors) in an essay if the error counts have the distributions given in Exercise 4.76 and

(a) the counts of nonword and word errors are independent.

(b) students who make many nonword errors also tend to make many word errors, so that the correlation between the two error counts is 0.5.

4.83 Means and variances of sums. The rules for means and variances allow you to find the mean and variance of a sum of random variables without first finding the distribution of the sum, which is usually much harder to do.

(a) A single toss of a balanced coin has either 0 or 1 head, each with probability 1/2. What are the mean and standard deviation of the number of heads?

(b) Toss a coin four times. Use the rules for means and variances to find the mean and standard deviation of the total number of heads.

(c) Example 4.23 (page 251) finds the distribution of the number of heads in four tosses. Find the mean and standard deviation from this distribution. Your results in parts (b) and (c) should agree.

4.84 CHALLENGE **Toss a four-sided die twice.** Role-playing games like Dungeons & Dragons use many different types of dice. Suppose that a four-sided die has faces marked 1, 2, 3, 4. The intelligence of a character is determined by rolling this die twice and adding 1 to the sum of the spots. The faces are equally likely and the two rolls are independent. What is the average (mean) intelligence for such characters? How spread out are their intelligences, as measured by the standard deviation of the distribution?

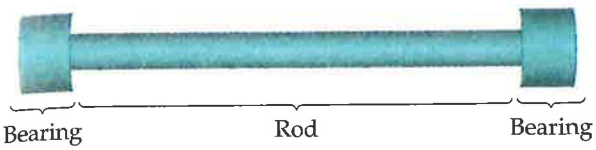

FIGURE 4.15 Sketch of a mechanical assembly, for Exercise 4.85.

4.85 A mechanical assembly. A mechanical assembly (Figure 4.15) consists of a rod with a bearing on each end. The three parts are manufactured independently, and all vary a bit from part to part. The length of the rod has mean 12 centimeters (cm) and standard deviation 0.004 millimeters (mm). The length of a bearing has mean 2 cm and standard deviation 0.001 mm. What are the mean and standard deviation of the total length of the assembly?

4.86 Sums of Normal random variables. Continue your work in the previous exercise. Dimensions of mechanical parts are often roughly Normal. According to the 68–95–99.7 rule, 95% of rods have lengths within $\pm d_1$ of 12 cm and 95% of bearings have lengths within $\pm d_2$ of 2 cm.

(a) What are the values of d_1 and d_2? These are often called the "natural tolerances" of the parts.

(b) Statistical theory says that any sum of independent Normal random variables has a Normal distribution. So the total length of the assembly is roughly Normal. What is the natural tolerance for the total length? It is *not* $d_1 + 2d_2$, because standard deviations don't add.

4.87 Will you assume independence? In which of the following games of chance would you be willing to assume independence of X and Y in making a probability model? Explain your answer in each case.

(a) In blackjack, you are dealt two cards and examine the total points X on the cards (face cards count 10 points). You can choose to be dealt another card and compete based on the total points Y on all three cards.

(b) In craps, the betting is based on successive rolls of two dice. X is the sum of the faces on the first roll, and Y the sum of the faces on the next roll.

4.88 Transform the distribution of heights from centimeters to inches. A report of the National Center for Health Statistics says that the heights of 20-year-old men have mean 176.8 centimeters (cm) and standard deviation 7.2 cm. There are 2.54 centimeters in an inch. What are the mean and standard deviation in inches?

4.89 What happens when the correlation is 1? We know that variances add if the random variables involved are uncorrelated ($\rho = 0$), but not otherwise. The opposite extreme is perfect positive correlation ($\rho = 1$). Show by using the general addition rule for variances that in this case the standard deviations add. That is, $\sigma_{X+Y} = \sigma_X + \sigma_Y$ if $\rho_{XY} = 1$.

4.90 A random variable with given mean and standard deviation. Here is a simple way to create a random variable X that has mean μ and standard deviation σ: X takes only the two values $\mu - \sigma$ and $\mu + \sigma$, each with probability 0.5. Use the definition of the mean and variance for discrete random variables to show that X does have mean μ and standard deviation σ.

Insurance. *The business of selling insurance is based on probability and the law of large numbers. Consumers buy insurance because we all face risks that are unlikely but carry high cost. Think of a fire destroying your home. So we form a group to share the risk: we all pay a small amount, and the insurance policy pays a large amount to those few of us whose homes burn down. The insurance company sells many policies, so it can rely on the law of large numbers. Exercises 4.91 to 4.94 explore aspects of insurance.*

4.91 Fire insurance. An insurance company looks at the records for millions of homeowners and sees that the mean loss from fire in a year is $\mu = \$300$ per person. (Most of us have no loss, but a few lose their homes. The \$300 is the average loss.) The company plans to sell fire insurance for \$300 plus enough to cover its costs and profit. Explain clearly why it would be stupid to sell only 10 policies. Then explain why selling thousands of such policies is a safe business.

4.92 Mean and standard deviation for 10 and for 12 policies. In fact, the insurance company sees that in the entire population of homeowners, the mean loss from fire is $\mu = \$300$ and the standard deviation of the loss is $\sigma = \$400$. What are the mean and standard deviation of the average loss for 10 policies? (Losses on separate policies are independent.) What are the mean and standard deviation of the average loss for 12 policies?

4.93 Life insurance. According to the current Commissioners' Standard Ordinary mortality table, adopted by state insurance regulators in December 2002, a 25-year-old man has these probabilities of dying during the next five years:[22]

Age at death	25	26	27	28	29
Probability	0.00039	0.00044	0.00051	0.00057	0.00060

(a) What is the probability that the man does not die in the next five years?

(b) An online insurance site offers a term insurance policy that will pay \$100,000 if a 25-year-old man dies within the next five years. The cost is \$175 per year. So the insurance company will take in \$875 from this policy if the man does not die within five years. If he does die, the company must pay \$100,000. Its loss depends on how many premiums were paid, as follows:

Age at death	25	26	27	28	29
Loss	\$99,825	\$99,650	\$99,475	\$99,300	\$99,125

What is the insurance company's mean cash intake from such polices?

4.94 Risk for one versus thousands of life insurance policies. It would be quite risky for you to insure the life of a 25-year-old friend under the terms of Exercise 4.93. There is a high probability that your friend would live and you would gain \$875 in premiums. But if he were to die, you would lose almost \$100,000. Explain carefully why selling insurance is not risky for an insurance company that insures many thousands of 25-year-old men.

Solutions

4.72. **(a)** Each toss of the coin is independent (that is, coins have no memory). **(b)** The variance is multiplied by $10^2 = 100$. (The mean and *standard deviation* are multiplied by 10.) **(c)** The correlation does not affect the mean of a sum (although it does affect the variance and standard deviation).

4.73. The mean is

$$\mu_X = (0)(0.3) + (1)(0.1) + (2)(0.1) + (3)(0.2) + (4)(0.1) + (5)(0.2) = 2.3 \text{ servings.}$$

The variance is

$$\sigma_X^2 = (0-2.3)^2(0.3) + (1-2.3)^2(0.1) + (2-2.3)^2(0.1) + (3-2.3)^2(0.2) + (4-2.3)^2(0.1) + (5-2.3)^2(0.2) = 3.61,$$

so the standard deviation is $\sigma_X = \sqrt{3.61} = 1.9$ servings.

4.74. The mean number of aces is $\mu_X = (0)(0.8507) + (1)(0.1448) + (2)(0.0045) = 0.1538$.
Note: *The exact value of the mean is 2/13, because 1/13 of the cards are aces, and two cards have been dealt to us.*

4.75. The average grade is $\mu = (0)(0.05) + (1)(0.04) + (2)(0.20) + (3)(0.40) + (4)(0.31) = 2.88$.

4.76. The means are

$$(0)(0.1) + (1)(0.3) + (2)(0.3) + (3)(0.2) + (4)(0.1) = 1.9 \text{ nonword errors and}$$
$$(0)(0.4) + (1)(0.3) + (2)(0.2) + (3)(0.1) = 1 \text{ word error}$$

4.77. In the solution to Exercise 4.74, we found $\mu_X = 0.1538$ aces, so

$$\sigma_X^2 = (0 - 0.1538)^2(0.8507) + (1 - 0.1538)^2(0.1448) + (2 - 0.1538)^2(0.0045) \doteq 0.1391,$$

and the standard deviation is $\sigma_X \doteq \sqrt{0.1391} \doteq 0.3730$ aces.

4.78. In the solution to Exercise 4.75, we found the average grade was $\mu = 2.88$, so

$$\sigma^2 = (0 - 2.88)^2(0.05) + (1 - 2.88)^2(0.04) + (2 - 2.88)^2(0.2) + (3 - 2.88)^2(0.4) + (4 - 2.88)^2(0.31) = 1.1056,$$

and the standard deviation is $\sigma = \sqrt{1.1056} \doteq 1.0515$.

4.79. **(a)** With $\rho = 0$, the variance is $\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 = (75)^2 + (41)^2 = 7306$, so the standard deviation is $\sigma_{X+Y} = \sqrt{7306} \doteq \$85.48$. **(b)** This is larger; the negative correlation decreased the variance.

4.80. **(a)** The mean of Y is $\mu_Y = 1$—the obvious balance point of the triangle. **(b)** Both X_1 and X_2 have mean $\mu_{X_1} = \mu_{X_2} = 0.5$ and $\mu_Y = \mu_{X_1} + \mu_{X_2}$.

4.81. The situation described in this exercise—"people who have high intakes of calcium in their diets are more compliant than those who have low intakes"—implies a positive correlation between calcium intake and compliance. Because of this, the variance of total calcium intake is greater than the variance we would see if there were no correlation (as the calculations in Example 4.38 demonstrate).

4.82. Let N and W be nonword and word error counts. In Exercise 4.76, we found $\mu_N = 1.9$ errors and $\mu_W = 1$ error. The variances of these distributions are $\sigma_N^2 = 1.29$ and $\sigma_W^2 = 1$, so the standard deviations are $\sigma_N \doteq 1.1358$ errors and $\sigma_W = 1$ error. The mean total error count is $\mu_N + \mu_W = 2.9$ errors for both cases. **(a)** If error counts are independent (so that $\rho = 0$), $\sigma_{N+W}^2 = \sigma_N^2 + \sigma_W^2 = 2.29$ and $\sigma_{N+W} \doteq 1.5133$ errors. (Note that we add the *variances*, not the standard deviations.) **(b)** With $\rho = 0.5$, $\sigma_{N+W}^2 = \sigma_N^2 + \sigma_W^2 + 2\rho\sigma_N\sigma_W \doteq 2.29 + 1.1358 = 3.4258$ and $\sigma_{N+W} \doteq 1.8509$ errors.

4.83. **(a)** The mean for one coin is $\mu_1 = (0)\left(\frac{1}{2}\right) + (1)\left(\frac{1}{2}\right) = 0.5$ and the variance is $\sigma_1^2 = (0 - 0.5)^2\left(\frac{1}{2}\right) + (1 - 0.5)^2\left(\frac{1}{2}\right) = 0.25$, so the standard deviation is $\sigma_1 = 0.5$. **(b)** Multiply μ_1 and σ_1^2 by 4: $\mu_4 = 4\mu_1 = 2$ and $\sigma_4^2 = 4\sigma_1^4 = 1$, so $\sigma_4 = 1$. **(c)** Note that because of the symmetry of the distribution, we do not need to compute the mean to see that $\mu_4 = 2$; this is the obvious balance point of the probability histogram in Figure 4.7. The details of the two computations are

$$\mu_W = (0)(0.0625) + (1)(0.25) + (2)(0.375) + (3)(0.25) + (4)(0.0625) = 2$$

$$\sigma_W^2 = (0 - 2)^2(0.0625) + (1 - 2)^2(0.25) + (2 - 2)^2(0.375) + (3 - 2)^2(0.25) + (4 - 2)^2(0.0625) = 1.$$

4.84. If D is the result of rolling a single four-sided die, then $\mu_D = (1 + 2 + 3 + 4)\left(\frac{1}{4}\right) = 2.5$, and $\sigma_D^2 = [(1 - 2.5)^2 + (2 - 2.5)^2 + (3 - 2.5)^2 + (4 - 2.5)^2]\frac{1}{4} = 1.25$. Then for the sum

$I = D_1 + D_2 + 1$, we have mean intelligence $\mu_I = 2\mu_D + 1 = 6$. The variance of I is $\sigma_I^2 = 2\sigma_D^2 = 2.5$, so $\sigma_I \doteq 1.5811$.

4.85. With R as the rod length and B_1 and B_2 the bearing lengths, we have $\mu_{B_1+R+B_2} = 12 + 2 \cdot 2 = 16$ cm and $\sigma_{B_1+R+B_2} = \sqrt{0.004^2 + 2 \cdot 0.001^2} \doteq 0.004243$ mm.

4.86. **(a)** $d_1 = 2\sigma_R = 0.008$ mm $= 0.0008$ cm and $d_2 = 2\sigma_B = 0.002$ mm $= 0.0002$ cm. **(b)** The natural tolerance of the assembled parts is $2\sigma_{B_1+R+B_2} \doteq 0.008485$ mm $=$ 0.0008485 cm.

4.87. **(a)** Not independent: Knowing the total X of the first two cards tells us something about the total Y for three cards. **(b)** Independent: Separate rolls of the dice should be independent.

4.88. Divide the given values by 2.54: $\mu \doteq 69.6063$ in and $\sigma \doteq 2.8346$ in.

4.89. With $\rho = 1$, we have:

$$\sigma_{X+Y}^2 = \sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y = \sigma_X^2 + \sigma_Y^2 + 2\sigma_X\sigma_Y = (\sigma_X + \sigma_Y)^2$$

And of course, $\sigma_{X+Y} = \sqrt{(\sigma_X + \sigma_Y)^2} = \sigma_X + \sigma_Y$.

4.90. The mean of X is $(\mu - \sigma)(0.5) + (\mu + \sigma)(0.5) = \mu$, and the standard deviation is $\sqrt{(\mu - \sigma - \mu)^2(0.5) + (\mu + \sigma - \mu)^2(0.5)} = \sqrt{\sigma^2} = \sigma$.

4.91. Although the probability of having to pay for a total loss for one or more of the 10 policies is very small, if this were to happen, it would be financially disastrous. On the other hand, for thousands of policies, the law of large numbers says that the average claim on many policies will be close to the mean, so the insurance company can be assured that the premiums they collect will (almost certainly) cover the claims.

4.92. The total loss T for 10 fires has mean $\mu_T = 10 \cdot \$300 = \$3000$, and standard deviation $\sigma_T = \sqrt{10 \cdot \$400^2} = \$400\sqrt{10} \doteq \$1264.91$. The average loss is $T/10$, so $\mu_{T/10} = \frac{1}{10}\mu_T = \$300$ and $\sigma_{T/10} = \frac{1}{10}\sigma_T \doteq \$126.49$.

The total loss T for 12 fires has mean $\mu_T = 12 \cdot \$300 = \$3600$, and standard deviation $\sigma_T = \sqrt{12 \cdot \$400^2} = \$400\sqrt{12} \doteq \$1385.64$. The average loss is $T/12$, so $\mu_{T/12} = \frac{1}{12}\mu_T = \$300$ and $\sigma_{T/12} = \frac{1}{12}\sigma_T \doteq \$115.47$.

Note: *The mean of the average loss is the same regardless of the number of policies, but the standard deviation decreases as the number of policies increases. With thousands of policies, the standard deviation is very small, so the average claim will be close to $300, as was stated in the solution to the previous problem.*

4.93. **(a)** Add up the given probabilities and subtract from 1; this gives P(man does not die in the next five years) $= 0.99749$. **(b)** The distribution of income (or loss) is given below. Multiplying each possible value by its probability gives the mean intake $\mu \doteq \$623.22$.

Age at death	21	22	23	24	25	Survives
Loss or income	−\$99,825	−\$99,650	−\$99,475	−\$99,300	−\$99,125	\$875
Probability	0.00039	0.00044	0.00051	0.00057	0.00060	0.99749

4.94. The mean μ of the company's "winnings" (premiums) and their "losses" (insurance claims) is positive. Even though the company will lose a large amount of money on a small number of policyholders who die, it will gain a small amount on the majority. The law of large numbers says that the average "winnings" minus "losses" should be close to μ, and overall the company will almost certainly show a profit.

4.95. The events "roll a 3" and "roll a 5" are disjoint, so $P(3 \text{ or } 5) = P(3) + P(5) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$.

4.96. The events E (roll is even) and G (roll is greater than 4) are *not* disjoint—specifically, $E \text{ and } G = \{6\}$—so $P(E \text{ or } G) = P(E) + P(G) - P(E \text{ and } G) = \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$.

4.97. Let A be the event "next card is an ace" and B be "two of Slim's four cards are aces." Then, $P(A \mid B) = \frac{2}{48}$ because (other than those in Slim's hand) there are 48 cards, of which 2 are aces.

4.98. Let A_1 = "the next card is a diamond" and A_2 = "the second card is a diamond." We wish to find $P(A_1 \text{ and } A_2)$. There are 27 unseen cards, of which 10 are diamonds, so $P(A_1) = \frac{10}{27}$, and $P(A_2 \mid A_1) = \frac{9}{26}$, so $P(A_1 \text{ and } A_2) = \frac{10}{27} \times \frac{9}{26} = \frac{5}{39} \doteq 0.1282$.

Note: *Technically, we wish to find* $P(A_1 \text{ and } A_2 \mid B)$, *where* B *is the given event (25 cards visible, with 3 diamonds in Slim's hand). We have* $P(A_1 \mid B) = \frac{10}{27}$ *and* $P(A_2 \mid A_1 \text{ and } B) = \frac{9}{26}$, *and compute* $P(A_1 \text{ and } A_2 \mid B) = P(A_1 \mid B) \times P(A_2 \mid A_1 \text{ and } B)$.

4.99. This computation uses the addition rule for disjoint events, which is appropriate for this setting because B (full-time students) is made up of four disjoint groups (those in each of the four age groups).

4.100. With A and B as defined in Example 4.44 (respectively, 15- to 19-year-old students, and full-time students), we want to find

$$P(B \mid A) = \frac{P(B \text{ and } A)}{P(A)} = \frac{0.21}{0.21 + 0.02} \doteq 0.9130$$

For these two calculations, we restrict our attention to different subpopulations of students (that is, different rows of the table given in Example 4.44). For $P(A \mid B)$, we ask what fraction of full-time students (the subpopulation) are aged 15 to 19 years. For $P(B \mid A)$, we ask what fraction of the subpopulation of 15- to 19-year-old students are full-time.