Guided Practice A.54 After an introductory statistics course, 78% of students can successfully construct tree diagrams. Of those who can construct tree diagrams, 97% passed, while only 57% of those students who could not construct tree diagrams passed. (a) Organize this information into a tree diagram. (b) What is the probability that a randomly selected student passed? (c) Compute the probability a student is able to construct a tree diagram if it is known that she passed.[38]

A.3 Random variables

Example A.55 Two books are assigned for a statistics class: a textbook and its corresponding study guide. The university bookstore determined 20% of enrolled students do not buy either book, 55% buy the textbook only, and 25% buy both books, and these percentages are relatively constant from one term to another. If there are 100 students enrolled, how many books should the bookstore expect to sell to this class?

Around 20 students will not buy either book (0 books total), about 55 will buy one book (55 books total), and approximately 25 will buy two books (totaling 50 books for these 25 students). The bookstore should expect to sell about 105 books for this class.

Guided Practice A.56 Would you be surprised if the bookstore sold slightly more or less than 105 books?[39]

Example A.57 The textbook costs \$137 and the study guide \$33. How much revenue should the bookstore expect from this class of 100 students?

About 55 students will just buy a textbook, providing revenue of

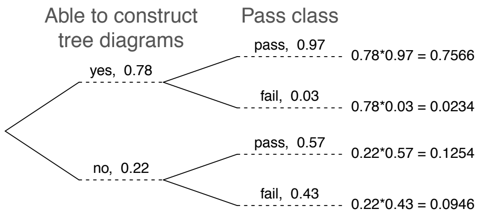
$$\$137 \times 55 = \$7,535$$

The roughly 25 students who buy both the textbook and the study guide would pay a total of

$$(\$137 + \$33) \times 25 = \$170 \times 25 = \$4,250$$

Thus, the bookstore should expect to generate about $\$7,535 + \$4,250 = \$11,785$ from these 100 students for this one class. However, there might be some *sampling variability* so the actual amount may differ by a little bit.

[38] (a) The tree diagram is shown to the right. (b) Identify which two joint probabilities represent students who passed, and add them: $P(\text{passed}) = 0.7566 + 0.1254 = 0.8820$. (c) $P(\text{construct tree diagram} \mid \text{passed}) = \frac{0.7566}{0.8820} = 0.8578$.



[39] If they sell a little more or a little less, this should not be a surprise. Hopefully Chapter 1 helped make clear that there is natural variability in observed data. For example, if we would flip a coin 100 times, it will not usually come up heads exactly half the time, but it will probably be close.

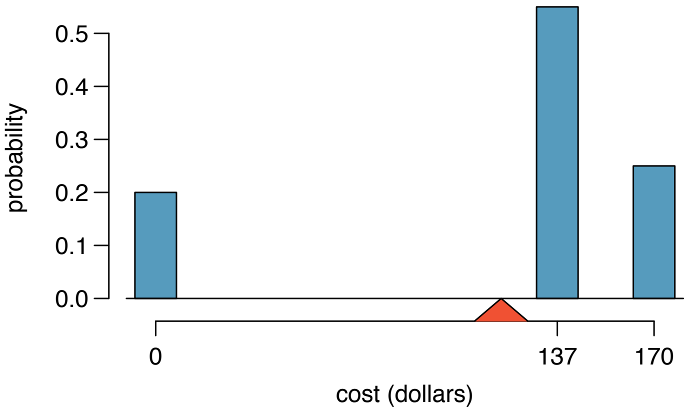


Figure A.19: Probability distribution for the bookstore's revenue from a single student. The distribution balances on a triangle representing the average revenue per student.

i	1	2	3	Total
x_i	\$0	\$137	\$170	–
$P(X = x_i)$	0.20	0.55	0.25	1.00

Table A.20: The probability distribution for the random variable X, representing the bookstore's revenue from a single student.

● **Example A.58** What is the average revenue per student for this course?

The expected total revenue is \$11,785, and there are 100 students. Therefore the expected revenue per student is $\$11,785/100 = \$117.85$.

A.3.1 Expectation

We call a variable or process with a numerical outcome a **random variable**, and we usually represent this random variable with a capital letter such as X, Y, or Z. The amount of money a single student will spend on her statistics books is a random variable, and we represent it by X.

Random variable
A random process or variable with a numerical outcome.

The possible outcomes of X are labeled with a corresponding lower case letter x and subscripts. For example, we write $x_1 = \$0$, $x_2 = \$137$, and $x_3 = \$170$, which occur with probabilities 0.20, 0.55, and 0.25. The distribution of X is summarized in Figure A.19 and Table A.20.

$E(X)$
Expected value of X

We computed the average outcome of X as \$117.85 in Example A.58. We call this average the **expected value** of X, denoted by $E(X)$. The expected value of a random variable is computed by adding each outcome weighted by its probability:

$$\begin{aligned} E(X) &= 0 \times P(X = 0) + 137 \times P(X = 137) + 170 \times P(X = 170) \\ &= 0 \times 0.20 + 137 \times 0.55 + 170 \times 0.25 = 117.85 \end{aligned}$$

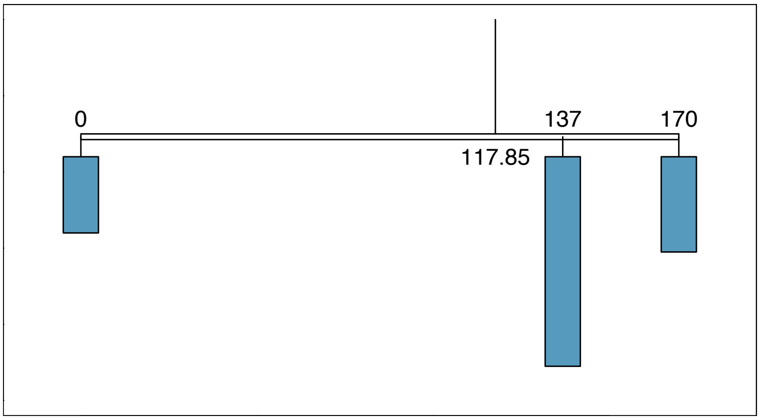


Figure A.21: A weight system representing the probability distribution for X. The string holds the distribution at the mean to keep the system balanced.

Expected value of a Discrete Random Variable

If X takes outcomes x_1, ..., x_k with probabilities $P(X = x_1)$, ..., $P(X = x_k)$, the expected value of X is the sum of each outcome multiplied by its corresponding probability:

$$\begin{aligned} E(X) &= x_1 \times P(X = x_1) + \cdots + x_k \times P(X = x_k) \\ &= \sum_{i=1}^{k} x_i P(X = x_i) \qquad \text{(A.59)} \end{aligned}$$

The Greek letter μ may be used in place of the notation $E(X)$.

The expected value for a random variable represents the average outcome. For example, $E(X) = 117.85$ represents the average amount the bookstore expects to make from a single student, which we could also write as $\mu = 117.85$.

It is also possible to compute the expected value of a continuous random variable. However, it requires a little calculus and we save it for a later class.[40]

In physics, the expectation holds the same meaning as the center of gravity. The distribution can be represented by a series of weights at each outcome, and the mean represents the balancing point. This is represented in Figures A.19 and A.21. The idea of a center of gravity also expands to continuous probability distributions. Figure A.22 shows a continuous probability distribution balanced atop a wedge placed at the mean.

A.3.2 Variability in random variables

Suppose you ran the university bookstore. Besides how much revenue you expect to generate, you might also want to know the volatility (variability) in your revenue.

The variance and standard deviation can be used to describe the variability of a random variable. Section 1.6.4 introduced a method for finding the variance and standard deviation for a data set. We first computed deviations from the mean ($x_i - \mu$), squared those deviations, and took an average to get the variance. In the case of a random variable, we again compute squared deviations. However, we take their sum weighted by their corresponding probabilities, just like we did for the expectation. This weighted sum of squared deviations equals the variance, and

[40] $\mu = \int x f(x) dx$ where $f(x)$ represents a function for the density curve.

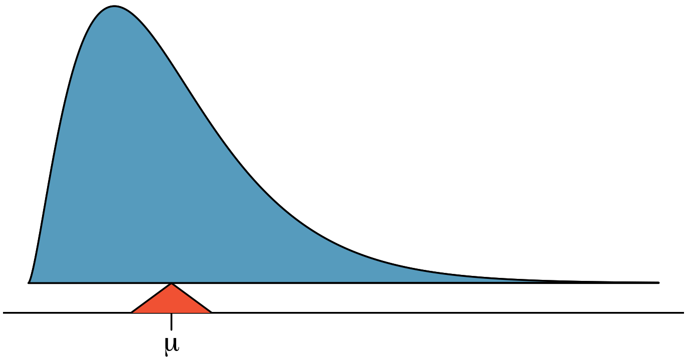


Figure A.22: A continuous distribution can also be balanced at its mean.

we calculate the standard deviation by taking the square root of the variance, just as we did in Section 1.6.4.

$Var(X)$
Variance of X

General variance formula

If X takes outcomes x_1, ..., x_k with probabilities $P(X = x_1)$, ..., $P(X = x_k)$ and expected value $\mu = E(X)$, then the variance of X, denoted by $Var(X)$ or the symbol σ^2, is

$$\begin{aligned} \sigma^2 &= (x_1 - \mu)^2 \times P(X = x_1) + \cdots \\ &\qquad \cdots + (x_k - \mu)^2 \times P(X = x_k) \\ &= \sum_{j=1}^{k} (x_j - \mu)^2 P(X = x_j) \end{aligned} \tag{A.60}$$

The standard deviation of X, labeled σ, is the square root of the variance.

● **Example A.61** Compute the expected value, variance, and standard deviation of X, the revenue of a single statistics student for the bookstore.

It is useful to construct a table that holds computations for each outcome separately, then add up the results.

i	1	2	3	Total
x_i	\$0	\$137	\$170	
$P(X = x_i)$	0.20	0.55	0.25	
$x_i \times P(X = x_i)$	0	75.35	42.50	117.85

Thus, the expected value is $\mu = 117.85$, which we computed earlier. The variance can be constructed by extending this table:

i	1	2	3	Total
x_i	\$0	\$137	\$170	
$P(X = x_i)$	0.20	0.55	0.25	
$x_i \times P(X = x_i)$	0	75.35	42.50	117.85
$x_i - \mu$	-117.85	19.15	52.15	
$(x_i - \mu)^2$	13888.62	366.72	2719.62	
$(x_i - \mu)^2 \times P(X = x_i)$	2777.7	201.7	679.9	3659.3

The variance of X is $\sigma^2 = 3659.3$, which means the standard deviation is $\sigma = \sqrt{3659.3} = \$60.49$.

Guided Practice A.62 The bookstore also offers a chemistry textbook for \$159 and a book supplement for \$41. From past experience, they know about 25% of chemistry students just buy the textbook while 60% buy both the textbook and supplement.[41]

(a) What proportion of students don't buy either book? Assume no students buy the supplement without the textbook.

(b) Let Y represent the revenue from a single student. Write out the probability distribution of Y, i.e. a table for each outcome and its associated probability.

(c) Compute the expected revenue from a single chemistry student.

(d) Find the standard deviation to describe the variability associated with the revenue from a single student.

A.3.3 Linear combinations of random variables

So far, we have thought of each variable as being a complete story in and of itself. Sometimes it is more appropriate to use a combination of variables. For instance, the amount of time a person spends commuting to work each week can be broken down into several daily commutes. Similarly, the total gain or loss in a stock portfolio is the sum of the gains and losses in its components.

Example A.63 John travels to work five days a week. We will use X_1 to represent his travel time on Monday, X_2 to represent his travel time on Tuesday, and so on. Write an equation using X_1, ..., X_5 that represents his travel time for the week, denoted by W.

His total weekly travel time is the sum of the five daily values:

$$W = X_1 + X_2 + X_3 + X_4 + X_5$$

Breaking the weekly travel time W into pieces provides a framework for understanding each source of randomness and is useful for modeling W.

[41] (a) 100% - 25% - 60% = 15% of students do not buy any books for the class. Part (b) is represented by the first two lines in the table below. The expectation for part (c) is given as the total on the line $y_i \times P(Y = y_i)$. The result of part (d) is the square-root of the variance listed on in the total on the last line: $\sigma = \sqrt{Var(Y)} = \$69.28$.

i (scenario)	1 (`noBook`)	2 (`textbook`)	3 (`both`)	Total
y_i	0.00	159.00	200.00	
$P(Y = y_i)$	0.15	0.25	0.60	
$y_i \times P(Y = y_i)$	0.00	39.75	120.00	$E(Y) = 159.75$
$y_i - E(Y)$	-159.75	-0.75	40.25	
$(y_i - E(Y))^2$	25520.06	0.56	1620.06	
$(y_i - E(Y))^2 \times P(Y)$	3828.0	0.1	972.0	$Var(Y) \approx 4800$

● **Example A.64** It takes John an average of 18 minutes each day to commute to work. What would you expect his average commute time to be for the week?

We were told that the average (i.e. expected value) of the commute time is 18 minutes per day: $E(X_i) = 18$. To get the expected time for the sum of the five days, we can add up the expected time for each individual day:

$$\begin{aligned} E(W) &= E(X_1 + X_2 + X_3 + X_4 + X_5) \\ &= E(X_1) + E(X_2) + E(X_3) + E(X_4) + E(X_5) \\ &= 18 + 18 + 18 + 18 + 18 = 90 \text{ minutes} \end{aligned}$$

The expectation of the total time is equal to the sum of the expected individual times. More generally, the expectation of a sum of random variables is always the sum of the expectation for each random variable.

⊙ **Guided Practice A.65** Elena is selling a TV at a cash auction and also intends to buy a toaster oven in the auction. If X represents the profit for selling the TV and Y represents the cost of the toaster oven, write an equation that represents the net change in Elena's cash.[42]

⊙ **Guided Practice A.66** Based on past auctions, Elena figures she should expect to make about \$175 on the TV and pay about \$23 for the toaster oven. In total, how much should she expect to make or spend?[43]

⊙ **Guided Practice A.67** Would you be surprised if John's weekly commute wasn't exactly 90 minutes or if Elena didn't make exactly \$152? Explain.[44]

Two important concepts concerning combinations of random variables have so far been introduced. First, a final value can sometimes be described as the sum of its parts in an equation. Second, intuition suggests that putting the individual average values into this equation gives the average value we would expect in total. This second point needs clarification – it is guaranteed to be true in what are called *linear combinations of random variables*.

A **linear combination** of two random variables X and Y is a fancy phrase to describe a combination

$$aX + bY$$

where a and b are some fixed and known numbers. For John's commute time, there were five random variables – one for each work day – and each random variable could be written as having a fixed coefficient of 1:

$$1X_1 + 1X_2 + 1X_3 + 1X_4 + 1X_5$$

For Elena's net gain or loss, the X random variable had a coefficient of +1 and the Y random variable had a coefficient of -1.

When considering the average of a linear combination of random variables, it is safe to plug in the mean of each random variable and then compute the final result. For a few examples of nonlinear combinations of random variables – cases where we cannot simply plug in the means – see the footnote.[45]

[42] She will make X dollars on the TV but spend Y dollars on the toaster oven: $X - Y$.

[43] $E(X - Y) = E(X) - E(Y) = 175 - 23 = \$152$. She should expect to make about \$152.

[44] No, since there is probably some variability. For example, the traffic will vary from one day to next, and auction prices will vary depending on the quality of the merchandise and the interest of the attendees.

[45] If X and Y are random variables, consider the following combinations: X^{1+Y}, $X \times Y$, X/Y. In such cases, plugging in the average value for each random variable and computing the result will not generally lead to an accurate average value for the end result.

Linear combinations of random variables and the average result

If X and Y are random variables, then a linear combination of the random variables is given by

$$aX + bY \tag{A.68}$$

where a and b are some fixed numbers. To compute the average value of a linear combination of random variables, plug in the average of each individual random variable and compute the result:

$$a \times E(X) + b \times E(Y)$$

Recall that the expected value is the same as the mean, e.g. $E(X) = \mu_X$.

● **Example A.69** Leonard has invested \$6000 in Google Inc. (stock ticker: GOOG) and \$2000 in Exxon Mobil Corp. (XOM). If X represents the change in Google's stock next month and Y represents the change in Exxon Mobil stock next month, write an equation that describes how much money will be made or lost in Leonard's stocks for the month.

For simplicity, we will suppose X and Y are not in percents but are in decimal form (e.g. if Google's stock increases 1%, then $X = 0.01$; or if it loses 1%, then $X = -0.01$). Then we can write an equation for Leonard's gain as

$$\$6000 \times X + \$2000 \times Y$$

If we plug in the change in the stock value for X and Y, this equation gives the change in value of Leonard's stock portfolio for the month. A positive value represents a gain, and a negative value represents a loss.

⊙ **Guided Practice A.70** Suppose Google and Exxon Mobil stocks have recently been rising 2.1% and 0.4% per month, respectively. Compute the expected change in Leonard's stock portfolio for next month.[46]

⊙ **Guided Practice A.71** You should have found that Leonard expects a positive gain in Guided Practice A.70. However, would you be surprised if he actually had a loss this month?[47]

A.3.4 Variability in linear combinations of random variables

Quantifying the average outcome from a linear combination of random variables is helpful, but it is also important to have some sense of the uncertainty associated with the total outcome of that combination of random variables. The expected net gain or loss of Leonard's stock portfolio was considered in Guided Practice A.70. However, there was no quantitative discussion of the volatility of this portfolio. For instance, while the average monthly gain might be about \$134 according to the data, that gain is not guaranteed. Figure A.23 shows the monthly changes in a portfolio like Leonard's during the 36 months from 2009 to 2011. The gains and losses vary widely, and quantifying these fluctuations is important when investing in stocks.

Just as we have done in many previous cases, we use the variance and standard deviation to describe the uncertainty associated with Leonard's monthly returns. To do so, the variances of each stock's monthly return will be useful, and these are shown in Table A.24. The stocks' returns are nearly independent.

[46] $E(\$6000 \times X + \$2000 \times Y) = \$6000 \times 0.021 + \$2000 \times 0.004 = \$134$.

[47] No. While stocks tend to rise over time, they are often volatile in the short term.

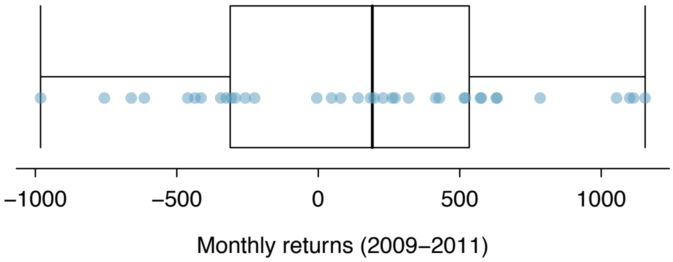


Figure A.23: The change in a portfolio like Leonard's for the 36 months from 2009 to 2011, where \$6000 is in Google's stock and \$2000 is in Exxon Mobil's.

	Mean ($\bar{x}$)	Standard deviation (s)	Variance (s^2)
GOOG	0.0210	0.0846	0.0072
XOM	0.0038	0.0519	0.0027

Table A.24: The mean, standard deviation, and variance of the GOOG and XOM stocks. These statistics were estimated from historical stock data, so notation used for sample statistics has been used.

Here we use an equation from probability theory to describe the uncertainty of Leonard's monthly returns; we leave the proof of this method to a dedicated probability course. The variance of a linear combination of random variables can be computed by plugging in the variances of the individual random variables and squaring the coefficients of the random variables:

$$Var(aX + bY) = a^2 \times Var(X) + b^2 \times Var(Y)$$

It is important to note that this equality assumes the random variables are independent; if independence doesn't hold, then more advanced methods are necessary. This equation can be used to compute the variance of Leonard's monthly return:

$$\begin{aligned} Var(6000 \times X + 2000 \times Y) &= 6000^2 \times Var(X) + 2000^2 \times Var(Y) \\ &= 36,000,000 \times 0.0072 + 4,000,000 \times 0.0027 \\ &= 270,000 \end{aligned}$$

The standard deviation is computed as the square root of the variance: $\sqrt{270,000} = \$520$. While an average monthly return of \$134 on an \$8000 investment is nothing to scoff at, the monthly returns are so volatile that Leonard should not expect this income to be very stable.

Variability of linear combinations of random variables

The variance of a linear combination of random variables may be computed by squaring the constants, substituting in the variances for the random variables, and computing the result:

$$Var(aX + bY) = a^2 \times Var(X) + b^2 \times Var(Y)$$

This equation is valid as long as the random variables are independent of each other. The standard deviation of the linear combination may be found by taking the square root of the variance.

Example A.72 Suppose John's daily commute has a standard deviation of 4 minutes. What is the uncertainty in his total commute time for the week?

The expression for John's commute time was

$$X_1 + X_2 + X_3 + X_4 + X_5$$

Each coefficient is 1, and the variance of each day's time is $4^2 = 16$. Thus, the variance of the total weekly commute time is

$$\begin{aligned} \text{variance} &= 1^2 \times 16 + 1^2 \times 16 + 1^2 \times 16 + 1^2 \times 16 + 1^2 \times 16 = 5 \times 16 = 80 \\ \text{standard deviation} &= \sqrt{\text{variance}} = \sqrt{80} = 8.94 \end{aligned}$$

The standard deviation for John's weekly work commute time is about 9 minutes.

Guided Practice A.73 The computation in Example A.72 relied on an important assumption: the commute time for each day is independent of the time on other days of that week. Do you think this is valid? Explain.[48]

Guided Practice A.74 Consider Elena's two auctions from Guided Practice A.65 on page 320. Suppose these auctions are approximately independent and the variability in auction prices associated with the TV and toaster oven can be described using standard deviations of \$25 and \$8. Compute the standard deviation of Elena's net gain.[49]

Consider again Guided Practice A.74. The negative coefficient for Y in the linear combination was eliminated when we squared the coefficients. This generally holds true: negatives in a linear combination will have no impact on the variability computed for a linear combination, but they do impact the expected value computations.

[48] One concern is whether traffic patterns tend to have a weekly cycle (e.g. Fridays may be worse than other days). If that is the case, and John drives, then the assumption is probably not reasonable. However, if John walks to work, then his commute is probably not affected by any weekly traffic cycle.

[49] The equation for Elena can be written as

$$(1) \times X + (-1) \times Y$$

The variances of X and Y are 625 and 64. We square the coefficients and plug in the variances:

$$(1)^2 \times Var(X) + (-1)^2 \times Var(Y) = 1 \times 625 + 1 \times 64 = 689$$

The variance of the linear combination is 689, and the standard deviation is the square root of 689: about \$26.25.