Quick Review

Here’s a brief summary of the key concepts and skills in probability and probability modeling:

- The Law of Large Numbers says that the more times we try something, the closer the results will come to theoretical perfection.
  - Don’t mistakenly misinterpret the Law of Large Numbers as the “Law of Averages.” There’s no such thing.
- Basic rules of probability can handle most situations:
  - To find the probability that an event OR another event happens, add their probabilities and subtract the probability that both happen.
  - To find the probability that an event AND another independent event both happen, multiply probabilities.
  - Conditional probabilities tell you how likely one event is to happen, knowing that another event has happened.
  - Mutually exclusive events (also called “disjoint”) cannot both happen at the same time.
  - Two events are independent if the occurrence of one doesn’t change the probability that the other happens.

- A probability model for a random variable describes the theoretical distribution of outcomes.
  - The mean of a random variable is its expected value.
  - For sums or differences of independent random variables, variances add.
  - To estimate probabilities involving quantitative variables, you may be able to use a Normal model—but only if the distribution of the variable is unimodal and symmetric.
  - To estimate the probability you’ll get your first success on a certain trial, use a Geometric model.
  - To estimate the probability you’ll get a certain number of successes in a specified number of independent trials, use a Binomial model.

Ready? Here are some opportunities to check your understanding of these ideas.

REVIEW EXERCISES

1. **Quality control.** A consumer organization estimates that 29% of new cars have a cosmetic defect, such as a scratch or a dent, when they are delivered to car dealers. This same organization believes that 7% have a functional defect—something that does not work properly—and that 2% of new cars have both kinds of problems.
   a) If you buy a new car, what’s the probability that it has some kind of defect?
   b) What’s the probability it has a cosmetic defect but no functional defect?
   c) If you notice a dent on a new car, what’s the probability it has a functional defect?
   d) Are the two kinds of defects disjoint events? Explain.
   e) Do you think the two kinds of defects are independent events? Explain.

2. **Workers.** A company’s human resources officer reports a breakdown of employees by job type and sex shown in the table.

<table>
<thead>
<tr>
<th>Job Type</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Management</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Supervision</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>Production</td>
<td>45</td>
<td>72</td>
</tr>
</tbody>
</table>

   a) What’s the probability that a worker selected at random is
      i) female?
      ii) female or a production worker?
      iii) female, if the person works in production?
      iv) a production worker, if the person is female?
   b) Do these data suggest that job type is independent of being male or female? Explain.

3. **Airfares.** Each year a company must send 3 officials to a meeting in China and 5 officials to a meeting in France. Airline ticket prices vary from time to time, but the company purchases all tickets for a country at the same price. Past experience has shown that tickets to China have a mean price of $1000, with a standard deviation of $150, while the mean airfare to France is $500, with a standard deviation of $100.
   a) Define random variables and use them to express the total amount the company will have to spend to send these delegations to the two meetings.
   b) Find the mean and standard deviation of this total cost.
   c) Find the mean and standard deviation of the difference in price of a ticket to China and a ticket to France.
   d) Do you need to make any assumptions in calculating these means? How about the standard deviations?
4. **Bipolar.** Psychiatrists estimate that about 1 in 100 adults suffers from bipolar disorder. What’s the probability that in a city of 10,000 there are more than 200 people with this condition? Be sure to verify that a Normal model can be used here.

5. **A game.** To play a game, you must pay $5 for each play. There is a 10% chance you will win $5, a 40% chance you will win $7, and a 50% chance you will win only $3.
   a) What are the mean and standard deviation of your net winnings?
   b) You play twice. Assuming the plays are independent events, what are the mean and standard deviation of your total winnings?

6. **Emergency switch.** Safety engineers must determine whether industrial workers can operate a machine’s emergency shutoff device. Among a group of test subjects, 66% were successful with their left hands, 82% with their right hands, and 51% with either hand.
   a) What percent of these workers could not operate the switch with either hand?
   b) Are success with right and left hands independent events? Explain.
   c) Are success with right and left hands mutually exclusive? Explain.

7. **Twins.** In the United States, the probability of having twins (usually about 1 in 90 births) rises to about 1 in 10 for women who have been taking the fertility drug Clomid. Among a group of 10 pregnant women, what’s the probability that
   a) at least one will have twins if none were taking a fertility drug?
   b) at least one will have twins if all were taking Clomid?
   c) at least one will have twins if half were taking Clomid?

8. **Deductible.** A car owner may buy insurance that will pay the full price of repairing the car after an at-fault accident, or save $12 a year by getting a policy with a $500 deductible. Her insurance company says that about 0.5% of drivers in her area have an at-fault auto accident during any given year. Based on this information, should she buy the policy with the deductible or not? How does the value of her car influence this decision?

9. **More twins.** A group of 5 women became pregnant while undergoing fertility treatments with the drug Clomid, discussed in Exercise 7. What’s the probability that
   a) none will have twins?
   b) exactly 1 will have twins?
   c) at least 3 will have twins?

10. **At fault.** The car insurance company in Exercise 8 believes that about 0.5% of drivers have an at-fault accident during a given year. Suppose the company insures 1355 drivers in that city.
    a) What are the mean and standard deviation of the number who may have at-fault accidents?
    b) Can you describe the distribution of these accidents with a Normal model? Explain.

11. **Twins, part III.** At a large fertility clinic, 152 women became pregnant while taking Clomid. (See Exercise 7.)
    a) What are the mean and standard deviation of the number of twin births we might expect?
    b) Can we use a Normal model in this situation? Explain.
    c) What’s the probability that no more than 10 of the women have twins?

12. **Child’s play.** In a board game you determine the number of spaces you may move by spinning a spinner and rolling a die. The spinner has three regions: Half of the spinner is marked “5,” and the other half is equally divided between “10” and “20.” The six faces of the die show 0, 1, 2, 3, and 4 spots. When it’s your turn, you spin and roll, adding the numbers together to determine how far you may move.
    a) Create a probability model for the outcome on the spinner.
    b) Find the mean and standard deviation of the spinner results.
    c) Create a probability model for the outcome on the die.
    d) Find the mean and standard deviation of the die results.
    e) Find the mean and standard deviation of the number of spaces you get to move.

13. **Language.** Neurological research has shown that in about 80% of people, language abilities reside in the brain’s left side. Another 10% display right-brain language centers, and the remaining 10% have two-sided language control. (The latter two groups are mainly left-handers; Science News, 161 no. 24 [2002].)
    a) Assume that a freshman composition class contains 25 randomly selected people. What’s the probability that no more than 15 of them have left-brain language control?
    b) In a randomly chosen group of 5 of these students, what’s the probability that no one has two-sided language control?
    c) In the entire freshman class of 1200 students, how many would you expect to find of each type?
    d) What are the mean and standard deviation of the number of these freshmen who might be right-brained in language abilities?
    e) If an assumption of Normality is justified, use the 68–95–99.7 Rule to describe how many students in the freshman class might have right-brain language control.

14. **Play again.** If you land in a “penalty zone” on the game board described in Exercise 12, your move will be determined by subtracting the roll of the die from the result on the spinner. Now what are the mean and standard deviation of the number of spots you may move?

15. **Beanstalks.** In some cities tall people who want to meet and socialize with other tall people can join Beanstalk Clubs. To qualify, a man must be over 6’2” tall, and a woman over 5’10”. According to the National Health Survey, heights of adults may have a Normal model with mean heights of 69.1” for men and 64.0” for women. The respective standard deviations are 2.8” and 2.5”.
16. **Stocks.** Since the stock market began in 1872, stock prices have risen in about 75% of the years. Assuming that market performance is independent from year to year, what’s the probability that
a) the market will rise for 3 consecutive years?
b) the market will rise 3 years out of the next 5?
c) the market will fall during at least 1 of the next 5 years?
d) the market will rise during a majority of years over the next decade?

17. **Multiple choice.** A multiple choice test has 50 questions, with 4 answer choices each. You must get at least 30 correct to pass the test, and the questions are very difficult.
   a) Are you likely to be able to pass by guessing on every question? Explain.
   b) Suppose, after studying for a while, you believe you have raised your chances of getting each question right to 70%. How likely are you to pass now?
   c) Assuming you are operating at the 70% level and the instructor arranges questions randomly, what’s the probability that the third question is the first one you get right?

18. **Stock strategy.** Many investment advisors argue that after stocks have declined in value for 2 consecutive years, people should invest heavily because the market rarely declines 3 years in a row.
   a) Since the stock market began in 1872, there have been two consecutive losing years eight times. In six of those cases, the market rose during the following year. Does this confirm the advice?
   b) Overall, stocks have risen in value during 95 of the 130 years since the market began in 1872. How is this fact relevant in assessing the statistical reasoning of the advisors?

19. **Insurance.** A 65-year-old woman takes out a $10,000 term life insurance policy. The company charges an annual premium of $500. Estimate the company’s expected profit on such policies if mortality tables indicate that only 2.6% of women age 65 die within a year.

20. **Teen smoking.** The Centers for Disease Control say that about 30% of high-school students smoke tobacco (down from a high of 38% in 1997). Suppose you randomly select high-school students to survey them on their attitudes toward scenes of smoking in the movies. What’s the probability that
a) none of the first 4 students you interview is a smoker?
b) the first smoker is the sixth person you choose?
c) there are no more than 2 smokers among 10 people you choose?

21. **Passing stats.** Molly’s college offers two sections of Statistics 101. From what she has heard about the two professors listed, Molly estimates that her chances of passing the course are 0.80 if she gets Professor Scedastic and 0.60 if she gets Professor Kurtosis. The registrar uses a lottery to randomly assign the 120 enrolled students based on the number of available seats in each class.
   a) What’s the probability that Molly will pass Statistics?
b) At the end of the semester, we find out that Molly failed. What’s the probability that she got Professor Kurtosis?

22. **Teen smoking II.** Suppose that, as reported by the Centers for Disease Control, about 30% of high school students smoke tobacco. You randomly select 120 high school students to survey them on their attitudes toward scenes of smoking in the movies.
   a) What’s the expected number of smokers?
b) What’s the standard deviation of the number of smokers?
c) The number of smokers among 120 randomly selected students will vary from group to group. Explain why that number can be described with a Normal model.
d) Using the 68–95–99.7 Rule, create and interpret a model for the number of smokers among your group of 120 students.

23. **Random variables.** Given independent random variables with means and standard deviations as shown, find the mean and standard deviation of each of these variables:
   a) $X + 50$
   b) $10Y$
   c) $X + 0.5Y$
   d) $X − Y$
   e) $X_1 + X_2$

24. **Merger.** Explain why the facts you know about variances of independent random variables might encourage two small insurance companies to merge. (Hint: Think about the expected amount and potential variability in payouts for the separate and the merged companies.)

25. **Youth survey.** According to a recent Gallup survey, 93% of teens use the Internet, but there are differences in how teen boys and girls say they use computers. The telephone poll found that 77% of boys had played computer games in the past week, compared with 65% of girls. On the other hand, 76% of girls said they had e-mailed friends in the past week, compared with only 65% of boys.
a) For boys, the cited percentages are 77% playing computer games and 65% using e-mail. That total is 142%, so there is obviously a mistake in the report. No? Explain.
b) Based on these results, do you think playing games and using e-mail are mutually exclusive? Explain.
c) Do you think whether a child e-mails friends is independent of being a boy or a girl? Explain.
d) Suppose that in fact 93% of the teens in your area do use the Internet. You want to interview a few who do not, so you start contacting teenagers at random. What is the probability that it takes you 5 interviews until you find the first person who does not use the Internet?

26. Meals. A college student on a seven-day meal plan reports that the amount of money he spends daily on food varies with a mean of $13.50 and a standard deviation of $7.
   a) What are the mean and standard deviation of the amount he might spend in two consecutive days?
b) What assumption did you make in order to find that standard deviation? Are there any reasons you might question that assumption?
c) Estimate his average weekly food costs, and the standard deviation.
d) Do you think it likely he might spend less than $50 in a week? Explain, including any assumptions you make in your analysis.

27. Travel to Kyrgyzstan. Your pocket copy of Kyrgyzstan on 4237 ± 360 Som a Day claims that you can expect to spend about 4237 som each day with a standard deviation of 360 som. How well can you estimate your expenses for the trip?
   a) Your budget allows you to spend 90,000 som. To the nearest day, how long can you afford to stay in Kyrgyzstan, on average?
b) What’s the standard deviation of your expenses for a trip of that duration?
c) You doubt that your total expenses will exceed your expectations by more than two standard deviations. How much extra money should you bring? On average, how much of a “cushion” will you have per day?

28. Picking melons. Two stores sell watermelons. At the first store the melons weigh an average of 22 pounds, with a standard deviation of 2.5 pounds. At the second store the melons are smaller, with a mean of 18 pounds and a standard deviation of 2 pounds. You select a melon at random at each store.
   a) What’s the mean difference in weights of the melons?
b) What’s the standard deviation of the difference in weights?
c) If a Normal model can be used to describe the difference in weights, what’s the probability that the melon you got at the first store is heavier?

29. Home, sweet home. According to the 2000 Census, 66% of U.S. households own the home they live in. A mayoral candidate conducts a survey of 820 randomly selected homes in your city and finds only 523 owned by the current residents. The candidate then attacks the incumbent mayor, saying that there is an unusually low level of homeownership in the city. Do you agree? Explain.

30. Buying melons. The first store in Exercise 28 sells watermelons for 32 cents a pound. The second store is having a sale on watermelons—only 25 cents a pound. Find the mean and standard deviation of the difference in the price you may pay for melons randomly selected at each store.

31. Who’s the boss? The 2000 Census revealed that 26% of all firms in the United States are owned by women. You call some firms doing business locally, assuming that the national percentage is true in your area.
   a) What’s the probability that the first 3 you call are all owned by women?
b) What’s the probability that none of your first 4 calls finds a firm that is owned by a woman?
c) Suppose none of your first 5 calls found a firm owned by a woman. What’s the probability that your next call does?

32. Jerseys. A Statistics professor comes home to find that all four of his children got white team shirts from soccer camp this year. He concludes that this year, unlike other years, the camp must not be using a variety of colors. But then he finds out that in each child’s age group there are 4 teams, only 1 of which wears white shirts. Each child just happened to get on the white team at random.
   a) Why was he so surprised? If each age group uses the same 4 colors, what’s the probability that all four kids would get the same-color shirt?
b) What’s the probability that all 4 would get white shirts?
c) We lied. Actually, in the oldest child’s group there are 6 teams instead of the 4 teams in each of the other three groups. How does this change the probability you calculated in part b?

33. When to stop? In Exercise 27 of the Review Exercises for Part III, we posed this question:
   You play a game that involves rolling a die. You can roll as many times as you want, and your score is the total for all the rolls. But . . . if you roll a 6, your score is 0 and your turn is over. What might be a good strategy for a game like this?

You attempted to devise a good strategy by simulating several plays to see what might happen. Let’s try calculating a strategy.
a) On what roll would you expect to get a 6 for the first time?
b) So, roll one time less than that. Assuming all those rolls were not 6’s, what’s your expected score?
c) What’s the probability that you can roll that many times without getting a 6?

34. Plan B. Here’s another attempt at developing a good strategy for the dice game in Exercise 33. Instead of stopping after a certain number of rolls, you could decide to stop when your score reaches a certain number of points.
a) How many points would you expect a roll to add to your score?
b) In terms of your current score, how many points would you expect a roll to subtract from your score?
c) Based on your answers in parts a and b, at what score will another roll “break even”?
d) Describe the strategy this result suggests.

35. Technology on campus. Every 5 years the Conference Board of the Mathematical Sciences surveys college math departments. In 2000 the board reported that 51% of all undergraduates taking Calculus I were in classes that used graphing calculators and 31% were in classes that used computer assignments. Suppose that 16% used both calculators and computers.
a) What percent used neither kind of technology?
b) What percent used calculators but not computers?
c) What percent of the calculator users had computer assignments?
d) Based on this survey, do calculator and computer use appear to be independent events? Explain.

36. Dogs. A census by the county dog control officer found that 18% of homes kept one dog as a pet, 4% had two dogs, and 1% had three or more. If a salesman visits two homes selected at random, what’s the probability he encounters
a) no dogs?
b) some dogs?
c) dogs in each home?
d) more than one dog in each home?

37. Socks. In your sock drawer you have 4 blue socks, 5 grey socks, and 3 black ones. Half asleep one morning, you grab 2 socks at random and put them on. Find the probability you end up wearing
a) 2 blue socks.
b) no grey socks.
c) at least 1 black sock.
da) a green sock.
e) matching socks.

38. Coins. A coin is to be tossed 36 times.
a) What are the mean and standard deviation of the number of heads?
b) Suppose the resulting number of heads is unusual, two standard deviations above the mean. How many “extra” heads were observed?
c) If the coin were tossed 100 times, would you still consider the same number of extra heads unusual? Explain.
d) In the 100 tosses, how many extra heads would you need to observe in order to say the results were unusual?
e) Explain how these results refute the “Law of Averages” but confirm the Law of Large Numbers.

39. The Drake equation. In 1961 astronomer Frank Drake developed an equation to try to estimate the number of extraterrestrial civilizations in our galaxy that might be able to communicate with us via radio transmissions. Now largely accepted by the scientific community, the Drake equation has helped spur efforts by radio astronomers to search for extraterrestrial intelligence. Here is the equation:

\[
N_c = N \cdot f_p \cdot n_e \cdot f_i \cdot f_l \cdot f_c
\]

OK, it looks a little messy, but here’s what it means:

<table>
<thead>
<tr>
<th>Factor</th>
<th>What It Represents</th>
<th>Possible Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N)</td>
<td>Number of stars in the Milky Way Galaxy</td>
<td>200–400 billion</td>
</tr>
<tr>
<td>(f_p)</td>
<td>Probability that a star has planets</td>
<td>20%–50%</td>
</tr>
<tr>
<td>(n_e)</td>
<td>Number of planets in a solar system capable of sustaining earth-type life</td>
<td>1? 2?</td>
</tr>
<tr>
<td>(f_l)</td>
<td>Probability that life develops on a planet with a suitable environment</td>
<td>1%–100%</td>
</tr>
<tr>
<td>(f_i)</td>
<td>Probability that life evolves intelligence</td>
<td>50%?</td>
</tr>
<tr>
<td>(f_c)</td>
<td>Probability that intelligent life develops radio communication</td>
<td>10%–20%</td>
</tr>
<tr>
<td>(f_L)</td>
<td>Fraction of the planet’s life for which the civilization survives</td>
<td>(\frac{1}{1,000,000})</td>
</tr>
<tr>
<td>(N_c)</td>
<td>Number of extraterrestrial civilizations in our galaxy with which we could communicate</td>
<td>?</td>
</tr>
</tbody>
</table>

So, how many EIs are out there? That depends; values chosen for the many factors in the equation depend on ever-evolving scientific knowledge and one’s personal guesses. But now, some questions.
a) What quantity is calculated by the first product, \(N \cdot f_p\)?
b) What quantity is calculated by the product, \(N \cdot f_p \cdot n_e \cdot f_i\)?
c) What probability is calculated by the product \(f_l \cdot f_i\)?
d) Which of the factors in the formula are conditional probabilities? Restate each in a way that makes the condition clear.

Note: A quick Internet search will find you a site where you can play with the Drake equation yourself.

40. Recalls. In a car rental company’s fleet, 70% of the cars are American brands, 20% are Japanese, and the rest are German. The company notes that manufacturers’ recalls seem to affect 2% of the American cars, but only 1% of the others.
a) What’s the probability that a randomly chosen car is recalled?
b) What’s the probability that a recalled car is American?
41. **Pregnant?** Suppose that 70% of the women who suspect they may be pregnant and purchase an in-home pregnancy test are actually pregnant. Further suppose that the test is 98% accurate. What’s the probability that a woman whose test indicates that she is pregnant actually is?

42. **Door prize.** You are among 100 people attending a charity fundraiser at which a large-screen TV will be given away as a door prize. To determine who wins, 99 white balls and 1 red ball have been placed in a box and thoroughly mixed. The guests will line up and, one at a time, pick a ball from the box. Whoever gets the red ball wins the TV, but if the ball is white, it is returned to the box. If none of the 100 guests gets the red ball, the TV will be auctioned off for additional benefit of the charity.

   a) What’s the probability that the first person in line wins the TV?
   b) You are the third person in line. What’s the probability that you win the TV?
   c) What’s the probability that the charity gets to sell the TV because no one wins?
   d) Suppose you get to pick your spot in line. Where would you want to be in order to maximize your chances of winning?
   e) After hearing some protest about the plan, the organizers decide to award the prize by not returning the white balls to the box, thus ensuring that 1 of the 100 people will draw the red ball and win the TV. Now what position in line would you choose in order to maximize your chances?