Practice Exercises: Lesson 2.3

STAT 1201
Introduction to Probability and Statistics

ONLINE AND DISTANCE EDUCATION

Exercises

4.1 Area under the curve, Part I. What percent of a standard normal distribution $N(\mu = 0, \sigma = 1)$ is found in each region? Be sure to draw a graph.

(a) $Z < -1.35$ (b) $Z > 1.48$ (c) $-0.4 < Z < 1.5$ (d) $|Z| > 2$

4.2 Area under the curve, Part II. What percent of a standard normal distribution $N(\mu = 0, \sigma = 1)$ is found in each region? Be sure to draw a graph.

(a) $Z > -1.13$ (b) $Z < 0.18$ (c) $Z > 8$ (d) $|Z| < 0.5$

4.3 GRE scores, Part I. Sophia who took the Graduate Record Examination (GRE) scored 160 on the Verbal Reasoning section and 157 on the Quantitative Reasoning section. The mean score for Verbal Reasoning section for all test takers was 151 with a standard deviation of 7, and the mean score for the Quantitative Reasoning was 153 with a standard deviation of 7.67. Suppose that both distributions are nearly normal.

(a) Write down the short-hand for these two normal distributions.
(b) What is Sophia's Z-score on the Verbal Reasoning section? On the Quantitative Reasoning section? Draw a standard normal distribution curve and mark these two Z-scores.
(c) What do these Z-scores tell you?
(d) Relative to others, which section did she do better on?
(e) Find her percentile scores for the two exams.
(f) What percent of the test takers did better than her on the Verbal Reasoning section? On the Quantitative Reasoning section?
(g) Explain why simply comparing raw scores from the two sections could lead to an incorrect conclusion as to which section a student did better on.
(h) If the distributions of the scores on these exams are not nearly normal, would your answers to parts (b) - (f) change? Explain your reasoning.

4.4 Triathlon times, Part I. In triathlons, it is common for racers to be placed into age and gender groups. Friends Leo and Mary both completed the Hermosa Beach Triathlon, where Leo competed in the *Men, Ages 30 - 34* group while Mary competed in the *Women, Ages 25 - 29* group. Leo completed the race in 1:22:28 (4948 seconds), while Mary completed the race in 1:31:53 (5513 seconds). Obviously Leo finished faster, but they are curious about how they did within their respective groups. Can you help them? Here is some information on the performance of their groups:

- The finishing times of the *Men, Ages 30 - 34* group has a mean of 4313 seconds with a standard deviation of 583 seconds.
- The finishing times of the *Women, Ages 25 - 29* group has a mean of 5261 seconds with a standard deviation of 807 seconds.
- The distributions of finishing times for both groups are approximately Normal.

Remember: a better performance corresponds to a faster finish.

(a) Write down the short-hand for these two normal distributions.
(b) What are the Z-scores for Leo's and Mary's finishing times? What do these Z-scores tell you?
(c) Did Leo or Mary rank better in their respective groups? Explain your reasoning.
(d) What percent of the triathletes did Leo finish faster than in his group?
(e) What percent of the triathletes did Mary finish faster than in her group?
(f) If the distributions of finishing times are not nearly normal, would your answers to parts (b) - (e) change? Explain your reasoning.

4.5 GRE scores, Part II. In Exercise 4.3 we saw two distributions for GRE scores: $N(\mu = 151, \sigma = 7)$ for the verbal part of the exam and $N(\mu = 153, \sigma = 7.67)$ for the quantitative part. Use this information to compute each of the following:

(a) The score of a student who scored in the 80^{th} percentile on the Quantitative Reasoning section.
(b) The score of a student who scored worse than 70% of the test takers in the Verbal Reasoning section.

4.6 Triathlon times, Part II. In Exercise 4.4 we saw two distributions for triathlon times: $N(\mu = 4313, \sigma = 583)$ for *Men, Ages 30 - 34* and $N(\mu = 5261, \sigma = 807)$ for the *Women, Ages 25 - 29* group. Times are listed in seconds. Use this information to compute each of the following:

(a) The cutoff time for the fastest 5% of athletes in the men's group, i.e. those who took the shortest 5% of time to finish.

(b) The cutoff time for the slowest 10% of athletes in the women's group.

4.7 LA weather, Part I. The average daily high temperature in June in LA is 77°F with a standard deviation of 5°F. Suppose that the temperatures in June closely follow a normal distribution.

(a) What is the probability of observing an 83°F temperature or higher in LA during a randomly chosen day in June?

(b) How cool are the coldest 10% of the days (days with lowest average high temperature) during June in LA?

4.8 CAPM. The Capital Asset Pricing Model (CAPM) is a financial model that assumes returns on a portfolio are normally distributed. Suppose a portfolio has an average annual return of 14.7% (i.e. an average gain of 14.7%) with a standard deviation of 33%. A return of 0% means the value of the portfolio doesn't change, a negative return means that the portfolio loses money, and a positive return means that the portfolio gains money.

(a) What percent of years does this portfolio lose money, i.e. have a return less than 0%?

(b) What is the cutoff for the highest 15% of annual returns with this portfolio?

4.9 LA weather, Part II. Exercise 4.7 states that average daily high temperature in June in LA is 77°F with a standard deviation of 5°F, and it can be assumed that they to follow a normal distribution. We use the following equation to convert °F (Fahrenheit) to °C (Celsius):

$$C = (F - 32) \times \frac{5}{9}.$$

(a) Write the probability model for the distribution of temperature in °C in June in LA.

(b) What is the probability of observing a 28°C (which roughly corresponds to 83°F) temperature or higher in June in LA? Calculate using the °C model from part (a).

(c) Did you get the same answer or different answers in part (b) of this question and part (a) of Exercise 4.7? Are you surprised? Explain.

(d) Estimate the IQR of the temperatures (in °C) in June in LA.

4.10 Find the SD. Cholesterol levels for women aged 20 to 34 follow an approximately normal distribution with mean 185 milligrams per deciliter (mg/dl). Women with cholesterol levels above 220 mg/dl are considered to have high cholesterol and about 18.5% of women fall into this category. What is the standard deviation of the distribution of cholesterol levels for women aged 20 to 34?

Chapter exercises

4.35 Roulette winnings. In the game of roulette, a wheel is spun and you place bets on where it will stop. One popular bet is that it will stop on a red slot; such a bet has an 18/38 chance of winning. If it stops on red, you double the money you bet. If not, you lose the money you bet. Suppose you play 3 times, each time with a $1 bet. Let Y represent the total amount won or lost. Write a probability model for Y.

4.36 Speeding on the I-5, Part I. The distribution of passenger vehicle speeds traveling on the Interstate 5 Freeway (I-5) in California is nearly normal with a mean of 72.6 miles/hour and a standard deviation of 4.78 miles/hour.[40]

(a) What percent of passenger vehicles travel slower than 80 miles/hour?
(b) What percent of passenger vehicles travel between 60 and 80 miles/hour?
(c) How fast do the fastest 5% of passenger vehicles travel?
(d) The speed limit on this stretch of the I-5 is 70 miles/hour. Approximate what percentage of the passenger vehicles travel above the speed limit on this stretch of the I-5.

4.37 University admissions. Suppose a university announced that it admitted 2,500 students for the following year's freshman class. However, the university has dorm room spots for only 1,786 freshman students. If there is a 70% chance that an admitted student will decide to accept the offer and attend this university, what is the approximate probability that the university will not have enough dormitory room spots for the freshman class?

4.38 Speeding on the I-5, Part II. Exercise 4.36 states that the distribution of speeds of cars traveling on the Interstate 5 Freeway (I-5) in California is nearly normal with a mean of 72.6 miles/hour and a standard deviation of 4.78 miles/hour. The speed limit on this stretch of the I-5 is 70 miles/hour.

(a) A highway patrol officer is hidden on the side of the freeway. What is the probability that 5 cars pass and none are speeding? Assume that the speeds of the cars are independent of each other.
(b) On average, how many cars would the highway patrol officer expect to watch until the first car that is speeding? What is the standard deviation of the number of cars he would expect to watch?

4.39 Auto insurance premiums. Suppose a newspaper article states that the distribution of auto insurance premiums for residents of California is approximately normal with a mean of $1,650. The article also states that 25% of California residents pay more than $1,800.

(a) What is the Z-score that corresponds to the top 25% (or the 75^{th} percentile) of the standard normal distribution?
(b) What is the mean insurance cost? What is the cutoff for the 75th percentile?
(c) Identify the standard deviation of insurance premiums in California.

4.40 SAT scores. SAT scores (out of 1600) are distributed normally with a mean of 1100 and a standard deviation of 200. Suppose a school council awards a certificate of excellence to all students who score at least 1350 on the SAT, and suppose we pick one of the recognized students at random. What is the probability this student's score will be at least 1500? (The material covered in Section 3.2 on conditional probability would be useful for this question.)

4.41 Married women. The American Community Survey estimates that 47.1% of women ages 15 years and over are married.[41]

(a) We randomly select three women between these ages. What is the probability that the third woman selected is the only one who is married?
(b) What is the probability that all three randomly selected women are married?
(c) On average, how many women would you expect to sample before selecting a married woman? What is the standard deviation?
(d) If the proportion of married women was actually 30%, how many women would you expect to sample before selecting a married woman? What is the standard deviation?
(e) Based on your answers to parts (c) and (d), how does decreasing the probability of an event affect the mean and standard deviation of the wait time until success?

[40] S. Johnson and D. Murray. "Empirical Analysis of Truck and Automobile Speeds on Rural Interstates: Impact of Posted Speed Limits". In: *Transportation Research Board 89th Annual Meeting.* 2010.

[41] U.S. Census Bureau, 2010 American Community Survey, Marital Status.

4.42 Survey response rate. Pew Research reported that the typical response rate to their surveys is only 9%. If for a particular survey 15,000 households are contacted, what is the probability that at least 1,500 will agree to respond?[42]

4.43 Overweight baggage. Suppose weights of the checked baggage of airline passengers follow a nearly normal distribution with mean 45 pounds and standard deviation 3.2 pounds. Most airlines charge a fee for baggage that weigh in excess of 50 pounds. Determine what percent of airline passengers incur this fee.

4.44 Heights of 10 year olds, Part I. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with mean 55 inches and standard deviation 6 inches.

(a) What is the probability that a randomly chosen 10 year old is shorter than 48 inches?
(b) What is the probability that a randomly chosen 10 year old is between 60 and 65 inches?
(c) If the tallest 10% of the class is considered "very tall", what is the height cutoff for "very tall"?

4.45 Buying books on Ebay. Suppose you're considering buying your expensive chemistry textbook on Ebay. Looking at past auctions suggests that the prices of this textbook follow an approximately normal distribution with mean \$89 and standard deviation \$15.

(a) What is the probability that a randomly selected auction for this book closes at more than \$100?
(b) Ebay allows you to set your maximum bid price so that if someone outbids you on an auction you can automatically outbid them, up to the maximum bid price you set. If you are only bidding on one auction, what are the advantages and disadvantages of setting a bid price too high or too low? What if you are bidding on multiple auctions?
(c) If you watched 10 auctions, roughly what percentile might you use for a maximum bid cutoff to be somewhat sure that you will win one of these ten auctions? Is it possible to find a cutoff point that will ensure that you win an auction?
(d) If you are willing to track up to ten auctions closely, about what price might you use as your maximum bid price if you want to be somewhat sure that you will buy one of these ten books?

4.46 Heights of 10 year olds, Part II. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with mean 55 inches and standard deviation 6 inches.

(a) The height requirement for *Batman the Ride* at Six Flags Magic Mountain is 54 inches. What percent of 10 year olds cannot go on this ride?
(b) Suppose there are four 10 year olds. What is the chance that at least two of them will be able to ride *Batman the Ride*?
(c) Suppose you work at the park to help them better understand their customers' demographics, and you are counting people as they enter the park. What is the chance that the first 10 year old you see who can ride *Batman the Ride* is the 3rd 10 year old who enters the park?
(d) What is the chance that the fifth 10 year old you see who can ride *Batman the Ride* is the 12th 10 year old who enters the park?

4.47 Heights of 10 year olds, Part III. Heights of 10 year olds, regardless of gender, closely follow a normal distribution with mean 55 inches and standard deviation 6 inches.

(a) What fraction of 10 year olds are taller than 76 inches?
(b) If there are 2,000 10 year olds entering Six Flags Magic Mountain in a single day, then compute the expected number of 10 year olds who are at least 76 inches tall. (You may assume the heights of the 10-year olds are independent.)
(c) Using the binomial distribution, compute the probability that 0 of the 2,000 10 year olds will be at least 76 inches tall.
(d) The number of 10 year olds who enter Six Flags Magic Mountain and are at least 76 inches tall in a given day follows a Poisson distribution with mean equal to the value found in part (b). Use the Poisson distribution to identify the probability no 10 year old will enter the park who is 76 inches or taller.

4.48 Multiple choice quiz. In a multiple choice quiz there are 5 questions and 4 choices for each question (a, b, c, d). Robin has not studied for the quiz at all, and decides to randomly guess the answers. What is the probability that

(a) the first question she gets right is the 3^{rd} question?
(b) she gets exactly 3 or exactly 4 questions right?
(c) she gets the majority of the questions right?

---

[42] Pew Research Center, Assessing the Representativeness of Public Opinion Surveys, May 15, 2012.