

# STA 111 Practice Problems

---

## 1 Summary Statistics

Consider the sample below.

$$-4, 3, 5, 0, 0, 1, 2$$

- 1.1. What is the mode?
- 1.2. What is the IQR?
- 1.3. What is the mean?
- 1.4. What is the median?
- 1.5. What is the range?
- 1.6. What is the variance?

## 2 The Civil Rights Act

The [Civil Rights Act](#), passed in 1964, is a landmark piece of legislation that outlaws discrimination based on race, color, religion, or national origin. It also prohibits racial segregation in schools, employment, and public accommodations. The table below shows the relationship between political party, region, and vote for the voting members of the United States House of Representatives on the Civil Rights Act. Use this data to answer the questions below.

	North		South	
	Aye	Nay	Aye	Nay
Democrat	144	8	8	83
Republican	137	24	0	11

- 2.1. What proportion of Democrats voted *Aye*?
- 2.2. What proportion of Republicans voted *Aye*?
- 2.3. Within each region, which party has a higher proportion of *Aye* votes?
- 2.4. Comment briefly on the relationship in the parts above.
- 2.5. Use a weighted average to adjust the *Aye* proportion for Democrats.
- 2.6. Use a weighted average to adjust the *Aye* proportion for Republicans.
- 2.7. What is going on here? Explain in a few sentences.

## 3 Batting Averages

Suppose two baseball players have the batting averages below.<sup>1</sup>

	Player A		Player B	
	Times at Bat	Hits	Times at Bat	Hits
Against right-hander pitchers	202	45	250	58
Against left-handed pitchers	250	71	108	32

- 3.1. What is the overall batting average of Player A?

---

<sup>1</sup>Knapp, T.R. (1985). "Instances of Simpson's Paradox." *College Mathematics Journal*, 16, 209-211

- 3.2. What is the overall batting average of Player B?
- 3.3. How do the batting averages compare when only examining right-handed pitchers?
- 3.4. How do the batting averages compare when only examining left-handed pitchers?
- 3.5. Use weighting to appropriately adjust the batting average for Player A.
- 3.6. Use weighting to appropriately adjust the batting average for Player B.
- 3.7. What is going on here? Explain in a few concise sentences.

## 4 New Zealand Jury Representation

Policy-makers are interested in determining if ethnic minorities are appropriately represented in the judicial process. In other words, are ethnic minorities in jury pools roughly in proportion to their population? Ian Westbrooke collected the data shown in the table below displaying jury composition in two cities in New Zealand (Rotorura and Nelson). Note here that the Maori are the indigenous people of New Zealand, analogous to Native Americans in the United States. <sup>2</sup>

	Rotorura		Nelson	
	Jury Pool	Population	Jury Pool	Population
Maori	79	8,889	1	1,329
Non-Maori	258	24,009	56	32,658

- 4.1. What percent of the jury pool is Maori in Rotorura?
- 4.2. What percent of the population is Maori in Rotorura?
- 4.3. What percent of the jury pool is Maori in Nelson?
- 4.4. What percent of the population is Maori in Nelson?
- 4.5. Are the Maori overrepresented or underrepresented on juries in Rotorura / Nelson?
- 4.6. Combining the two cities, what is the overall proportion of Maori in the jury pool?
- 4.7. Combining the two cities, what is the overall proportion of Maori in the population?
- 4.8. Comment briefly on the relationship in the parts above.
- 4.9. What is going on here? Explain in a few sentences.

## 5 Death Rates

The table below shows death rates from tuberculosis in Richmond, Virginia and New York, New York in 1910. The death rate in New York is 186 per 100,000 and the death rate in Richmond is 224 per 100,000. <sup>3</sup>

	New York		Richmond	
	Deaths	Population	Deaths	Population
White	8,365	4,675,174	131	80,895
Black	513	91,709	155	46,733

- 5.1. What is going on in this dataset? Answer in a brief paragraph and support your explanation with appropriate numbers.

<sup>2</sup>Westbrooke, I. (1997), *Simpson's Paradox: An Example in a New Zealand Survey of Jury Composition*, Statistics New Zealand

<sup>3</sup>Cohen, M. and Nagel, E. (1934). "An Introduction to Logic and Scientific Method." New York: Harcourt, Brace and Company

## 6 Job Applications

100 Duke graduates are applying for jobs at *Initech Industries*. They apply to one of two positions (sales intern or corporate leadership). There are two types of graduates (non Dean's list and Dean's list graduates).

You examine the data and notice something strange: a higher proportion of the non Dean's list graduates received job offers! Confused, you look at only the corporate leadership applications and discover that a higher proportion of Dean's list graduates were offered jobs compared to non Dean's list students. Finally, you examine the sales intern applications and find that a higher proportion of Dean's list students were offered jobs for this position as well.

	<u>Sales Intern</u>		<u>Corporate Leadership</u>	
	Offered Job	Rejected	Offered Job	Rejected
Not Dean's List				
Dean's List				

- 6.1. What is this surprising phenomenon called?
- 6.2. Fill in the table so that this is an example of the phenomenon above. Note the sum of the table should be 100 since there are 100 applicants.

## 7 Climbing Mt. Cox

**100 climbers** are attempting to reach the summit of Cox Mountain. There are two routes to the summit (easy and hard) and there are two types of climbers (amateur and experienced). You examine the data and notice something surprising - a higher proportion of amateur climbers succeeded in reaching the summit of Mt. Cox compared to experienced climbers! Confused, you examine the hard route, and find a higher proportion of experienced climbers succeeded. Then you examine the easy route and find a higher proportion of experienced climbers succeeded on the easy route as well!

	<u>Easy Route</u>		<u>Hard Route</u>	
	Succeeded	Failed	Succeeded	Failed
Amateur				
Experienced				

- 7.1. What is this surprising phenomenon called?
- 7.2. Fill in the table so that this is an example of the phenomenon above. Note the sum of the table should be 100 since there are 100 climbers.

## 8 Scrambles

Write out an expression for each of the following.

- 8.1. How many ways are there to scramble the letters of DUKE?
  - 8.2. How many ways are there to scramble the letters of DUKEBLUE?
- In my office I have ten textbooks on statistics, four on probability, and six on coding.

- 8.3. Three students come to my office and select a book at random. What is the probability that at least one of them selects a textbook on probability?

- 8.4. Four students come to my office and select a book at random. What is the probability that two of the students get a statistics textbook, one of the students get a probability textbook, and one of the students gets a coding textbook?

## 9 Poker Hands

What is the probability that a five card poker hand dealt from an ordinary deck of 52 cards has...

- 9.1. three face cards (J, Q, K) and two non-face cards?  
9.2. 1 club, 1 spade, 1 diamond, and 2 hearts?  
9.3. a pair of aces?  
9.4. at least one ace?

## 10 Committee Members

The Adventure League has 12 members. Answer the questions below by writing out an appropriate expression.

- 10.1. How many ways are there to choose a subcommittee of four members?  
10.2. How many ways are there to choose a Lead Adventurer, Navigator, and Linguist?

## 11 Committee Selection

A Duke student organization has eight statistics major members and four non-statistics major members.




- 11.1. You want to select a committee of three members. How many ways are there to do this?  
11.2. How many ways are there to select a President, Vice President, Treasurer, and Secretary?  
11.3. How many ways can you form a committee with two statistics majors and two non-statistics majors?

## 12 Indian Food

Tandoor in the Brodhead Center has two types of rice dishes, four meat dishes, and six vegetable dishes. The order of the dishes doesn't matter and you can't order the same dish twice.

- 12.1. If you order a rice, a meat, and a vegetable, how many meals can you eat before you have to repeat a meal?  
12.2. If you order a rice, two meats, and two vegetables, how many meals can you eat before you have to repeat a meal?  
12.3. If you order three vegetable dishes, how many meals can you eat before you have to eat the same meal?  
12.4. Hoping to start a food fight, you select three dishes at random to throw across the Brodhead Center. What is the probability you select all meat dishes?

## 13 Euchered

Euchre is a four-player, team-based card game played with A, K, Q, J, 10, 9 of all four suits (, , , ). You are dealt five cards. Answer the questions below.

- 13.1. What is the probability of being dealt a straight (five cards in a row, for example 10, J, K, Q, A)?  
13.2. What is the probability of being dealt a flush (all five cards the same suit)?  
13.3. What is the probability of four-of-a-kind?

## 14 Anaconda

Suppose you are dealt seven cards from an [ordinary 52 card deck](#).

14.1. What is the probability that four of the seven dealt cards are spades?

You have in your hand 8  $\spadesuit$ , 10  $\heartsuit$ , 2  $\heartsuit$ , Q  $\spadesuit$ , 4  $\heartsuit$ , K  $\clubsuit$ , A  $\heartsuit$ . You randomly choose three to discard.

14.2. What is the probability you select no face cards?

14.3. What is the probability you select all hearts?

## 15 Five Hundred

Five Hundred is a four-player team card game played with 44 cards from the [standard 52 card deck](#). The deck contains 4, 5, 6, 7, 8, 9, 10, J, Q, K, A of all suits. You are dealt five cards (note this isn't how you actually play).

15.1. What is the probability of being dealt a [straight](#) (e.g. {8 $\heartsuit$ , J $\clubsuit$ , 7 $\spadesuit$ , 9 $\diamond$ , 10 $\spadesuit$ } or {K $\diamond$ , J $\clubsuit$ , Q $\spadesuit$ , A $\diamond$ , 10 $\clubsuit$ })?

15.2. What is the probability of being dealt a [flush](#)? A flush is all cards the same suit (e.g. {4 $\heartsuit$ , 9 $\heartsuit$ , J $\heartsuit$ , 7 $\heartsuit$ , A $\heartsuit$ }).

15.3. What is the probability of being dealt [four-of-a-kind](#) (e.g. {7 $\heartsuit$ , 7 $\diamond$ , 5 $\spadesuit$ , 7 $\spadesuit$ , 7 $\clubsuit$ })?

15.4. What is the probability of being dealt a hand with at least one club?

15.5. The deal is a bust if you are dealt no face cards, jacks, or aces. What is the probability your hand is a bust?

## 16 Probability Problems

16.1. A [pawn](#) is placed at random on a regular  $8 \times 8$  chess board. A king of the opposing color is placed at random on one of the remaining squares. What is the probability that the king is in check? Note here "at random" means each position has an equal chance of being selected.

16.2. Draw four cards from an [ordinary deck of cards](#) without replacement. What is the probability that not all of the cards are clubs?

## 17 Chess Problems

A [knight](#) is placed at random on a standard  $8 \times 8$  chessboard and a [rook](#) of the opposing color is placed at random on one of the remaining spaces.

17.1. What is the probability the knight and the rook are attacking each other?

17.2. What is the probability the rook is attacking the knight?

17.3. What is the probability the knight is attacking the rook?

## 18 Duke Undergraduates

At Duke University 40% of female undergraduates and 30% of male undergraduates are involved in Greek life and 52% of undergraduates are female.

18.1. What is the probability that a randomly selected Duke undergraduate is female and involved in Greek life?

18.2. What is the probability that a randomly selected Duke undergraduate is not involved in Greek life?

18.3. What is the probability that a randomly selected Duke undergraduate is female or involved in Greek life or both?

## 19 Food Delivery

Suppose last weekend 40% of Duke undergraduates ordered pizza, 65% ordered Chinese or pizza, and 10% ordered Chinese and pizza. Answer the questions and feel free to use the space below for pictures and scratch work.

	Ordered Chinese	Did not order Chinese	Total
Ordered pizza	0.10	0.30	0.40
Did not order pizza	0.25	0.35	0.60
Total	0.35	0.65	1

- 19.1. What percentage of Duke undergraduates ordered Chinese?
- 19.2. What percentage of Duke undergraduates ordered Chinese and did not order pizza?
- 19.3. Suppose your roommate ordered Chinese. What is the probability he or she ordered pizza?
- 19.4. Suppose your roommate did not order Chinese. What is the probability he or she ordered pizza?
- 19.5. Are the events *Order Pizza* and *Order Chinese* independent? Briefly justify your answer.

## 20 Midterms

You have two midterms this semester, one in STA 111 and one in HIST 101. Suppose the probability you get an A on your STA 111 midterm is 0.70, the probability you get an A on your HIST 101 midterm is 0.50 and the probability you get an A on STA 111 or HIST 101 or both is 0.90.

- 20.1. What is the probability you get an A on both your STA 111 and HIST 101 midterms?
- 20.2. What is the probability you get an A on your STA 111 midterm but do not get an A on your HIST 101 midterm?
- 20.3. What is the probability you get an A on exactly one midterm?
- 20.4. What is the probability you get an A on both midterms?
- 20.5. Are the events **Get an A on your STA 111 midterm** and **Get an A on your HIST 101 midterm** independent? Justify your answer.

## 21 Railroad Crossings

I have to cross two railroad tracks on my way to work. The probability I am stopped by a train on the first crossing is 0.3, the probability I am stopped by a train on the second crossing is 0.2, and the probability I am stopped by at least one train is 0.35.

- 21.1. What is the probability I have to stop for both trains?
- 21.2. What is the probability I have to stop for the first train, but not the second?
- 21.3. What is the probability I have to stop for exactly one train?
- 21.4. Are the events *Stop for Train #1* and *Stop for Train #2* independent?

## 22 Blood Types

Suppose 1% of the American population has AB-negative blood and you take a random sample of size  $n$  from all Americans. How large must  $n$  be if the probability of the sample containing an individual with AB-negative blood must be larger than 0.80?

## 23 Tree Size

Suppose the diameter of cherry trees in inches is normally distributed with mean 13 inches and sd 3 inches.

23.1. What is the probability a randomly selected tree has a diameter between 10 and 14 inches?

23.2. What is the cut-off for the top 20% of tree diameters?

Suppose a tree has a diameter of 17.5 inches.

23.3. What percent of trees have diameters smaller than this tree?

23.4. How many standard deviations is this tree above average?

23.5. In the space below draw a picture of this normal distribution with a clearly labeled and numbered x axis. Indicate on your picture what percent of the trees are within 1, 2, and 3 standard deviations of the mean.

## 24 Duke IQ

Suppose the IQ of Duke statistics students is normally distributed with mean 112 and standard deviation 13.

24.1. What percent of Duke undergraduates have an IQ above 134?

24.2. What IQ is necessary to be in the top 5% of IQ?

24.3. What percent of Duke undergraduates have IQ's between 100 and 115?

24.4. What is the middle 68% of IQ? Provide two numbers here, a lower and an upper bound.

## 25 Heights

Suppose the height of some population is normally distributed with mean 71 inches and standard deviation 4 inches.

25.1. What percent of the population is taller than five feet six inches tall?

25.2. How tall does an adult American man need to be in order to be in the top 35% of height?

25.3. What is the range of the middle 70% of heights (for what range of heights will you be taller than at least 15% of Americans and shorter than at least 15% of Americans)?

25.4. Rogge is 69 inches tall. How many standard deviations is he above or below the average height?

## 26 Normal Problems

Let  $Z$  be a standard normal distribution  $N(0, 1)$ . Find the probabilities below.

26.1.  $P(Z \geq 2.06)$

26.2.  $P(-2.5 \leq Z \leq 1.5)$

26.3.  $P(|Z| \leq 0.5)$

Find the constant  $c$  that makes the below probability statements correct.

26.4.  $\Phi(c) = .4721$

26.5.  $P(c \leq Z) = .0778$

26.6.  $P(|Z| \geq c) = 0.0434$

## 27 Car Trouble

Suppose it is below freezing on 10% of days. When it is below freezing my car only starts 80% of the time, and when it is above freezing, my car starts 95% of the time. Today my car started. What is the chance it is below freezing?

## 28 Domino's

My wife has a 15% chance of receiving a coupon from Domino's via email. If she receives a coupon, there is a 50% chance she orders pizza, but if she doesn't receive a coupon, there is only a 10% chance she orders pizza. I get home from work and see a Domino's pizza box.

28.1. What is the chance my wife received a coupon?

## 29 HIV

Around 2.5% of the population of Togo has HIV. You are using a diagnostic test that is 95% effective in detecting HIV in an infected person, but gives a false positive result in 1% of cases.<sup>4</sup>

29.1. What is the probability of a positive result given that a person has HIV?

29.2. What is the probability of a negative result given that a person has HIV?

29.3. What is the probability of a positive result?

29.4. An individual tests positive. What is the probability they have HIV?

## 30 Electric Scooters

If there is an electric scooter present at an extremely dangerous Durham intersection, the number of accidents is Poisson distributed with mean 2.3 accidents per day and if there is not an electric scooter present, the number of accidents is Poisson distributed with mean 1.1 accidents per day. Suppose electric scooters are present on 15% of days.

30.1. Assume there is an electric scooter present. What is the probability there is at least one accident in a day?

30.2. Assume there is no electric scooter present. What is the probability there are two accidents in a day?

30.3. The Durham Ministry of Transportation reports there were no accidents at the intersection yesterday. What is the probability there was a scooter present?

## 31 Tears in the Rain

The Voight-Kampff test is a diagnostic test used to determine if an individual is an android (an evil robot pretending to be a human). If an individual is an android, the Voight-Kampff gives a positive result with probability 0.85 and if the individual is human, the Voight-Kampff gives a positive result with probability 0.05. Suppose the probability a randomly selected individual from this population is an android is 0.01.

31.1. What is the probability of a positive result on the Voight-Kampff test?

31.2. If an individual receives a positive result from the test, what is the probability that he/she/it is an android?

---

<sup>4</sup>This numbers are not realistic.



## 32 After Graduation

Human Resources at a manufacturing firm finds that 65% of new hires exceed their performance targets. Of those exceeding their targets, 15% were Duke graduates, and of those not exceeding their targets 5% were Duke graduates. What is the probability a new hire who graduated from Duke will exceed his target?

## 33 Prostate Cancer

A 2001 study published by researchers at the University of Illinois at Urbana-Champaign examined 753 men with prostate cancer and compared them to 703 men of the same age without cancer. They found that the prostate cancer risk increased with the lifetime number of female sexual partners. Dr. Karin Rosenblatt hypothesized an infectious agent as a potential cause of prostate cancer.<sup>5</sup>

33.1. Is this an observational study or a designed experiment?

33.2. Can you conclude that a higher number of sexual partners causes prostate cancer? Why or why not?

Identify each of the following as causal links, confounders, or neither.

33.3. Testosterone: Higher testosterone is associated with higher numbers of sexual partners and also increased cancer risk.

- Confounder
- Causal link
- Neither

33.4. Genetics: Some individuals are predisposed to prostate cancer.

- Confounder
- Causal link
- Neither

33.5. Infectious Agent: A sexually transmitted infection causes prostate cancer and a higher number of sexual partners increases your risk of catching the infection.

- Confounder
- Causal link
- Neither

## 34 An Apple A Day

Researchers followed 1,456 Australian women over the age of 70 for 15 years and assessed their apple intake with a periodic questionnaire. The results were idiomatic; individuals who consumed at least an apple a day had a significantly lower cancer mortality relative to those who did not consume an apple a day.<sup>6</sup>

34.1. Was this an observational study or a designed experiment?

34.2. Women who eat at least an apple a day likely differ from those who do not. Explain ways that they differ as well as the relevance for the study above.

34.3. Based on the study above does “an apple a day keep the doctor away”?

<sup>5</sup>Rosenblatt K., et al, Sexual Factors and the Risk of Prostate Cancer *American Journal of Epidemiology*, 153 (12), 2001, 1152–1158)

<sup>6</sup>Hodgson et al (2016). Apple intake is inversely associated with all-cause and disease-specific mortality in elderly women. *British Journal of Nutrition*, 115 (5)

## 35 Experimental Design

Define the terms below and state why each is important in the context of experimental design.

35.1. A **controlled** study:

35.2. A **double-blind** study:

35.3. A **randomized** study:

## 36 Hand Sanitizer

Suppose the amount of hand sanitizer in a particular brand's bottle is normally distributed with mean 14 ounces and standard deviation 3 ounces. You make a bulk order of 100 bottles of hand sanitizer.

36.1. What is the distribution of the total amount of hand sanitizer you ordered?

36.2. What is the approximate probability the total amount of hand sanitizer is greater than 1,450 ounces?

36.3. What is the cut-off for the largest 20% of hand sanitizer?

## 37 Uniform Samples

37.1. A random sample of size  $n = 30$  is taken from a uniform distribution on the interval  $[0, 10]$ . What is the approximate  $\mathbb{P}(4.5 \leq \bar{X}_n \leq 5.5)$ ?

37.2. A random sample of size  $n = 60$  is taken from a uniform distribution on the interval  $[0, 10]$ . What is the approximate  $\mathbb{P}(4.5 \leq \bar{X}_n \leq 5.5)$ ?

37.3. A random sample of size  $n = 90$  is taken from a uniform distribution on the interval  $[0, 10]$ . What is the approximate  $\mathbb{P}(4.5 \leq \bar{X}_n \leq 5.5)$ ?

37.4. Draw pictures of the sampling distribution of  $\bar{X}_n$  from 1, 2, and 3 and comment on the relationship. What is happening here?

## 38 Lightbulb Lifetimes

Suppose the lifetime of a certain brand of light bulb is normally distributed with mean twelve hours and standard deviation 2 hours. You purchase a package containing 50 light bulbs.

38.1. What is the approximate probability that the mean lifetime of the 50 bulbs is greater than 11.5 hours?

38.2. What is the approximate probability that the mean lifetime of the 50 bulbs is between 11.5 and 12.25 hours?

38.3. Consider the total *sum* lifetime of all 50 light bulbs in the package. What is the cut-off for the longest-lasting 10% of packages?

38.4. Again considering the total *sum* lifetime of all 50 light bulbs in the package, what is the middle 50% of lifetimes?

## 39 Traffic

Suppose on my (long) commute to work I have to stop at 40 stoplights and the time I have to wait at each stoplight is a random variable with mean 2 minutes and standard deviation 2 minutes.

39.1. If I leave at 2:05 PM, what is the approximate probability I arrive in time for class at 3:05 PM?

## 40 Grading Quizzes

Suppose the time it takes me to grade one STA 111 quiz is random with expected value three minutes and standard deviation three minutes and that grading times are independent. I have fifty quizzes to grade.

- 40.1. If I begin grading at 6:00, what is the probability I am done in time for the MLB Playoff game that starts at 8:00?
- 40.2. What is the probability I miss more than forty-five minutes of the MLB Playoff game?

## 41 Wire Cracks

You are performing quality control at a firm that manufactures wire. Suppose the number of defects on a given length of wire has a Poisson distribution with mean 9. A random sample of 100 wires is taken and the number of defects on each wire is counted.

- 41.1. What is the probability the average number of defects per wire in the sample is less than 8.5?
- 41.2. What is the probability the average number of defects per wire in the sample is more than 9.3?
- 41.3. What is the probability the average number of defects per wire in the sample is between 8.7 and 9.6?

## 42 Wrapping Paper

You are working at a firm that manufactures wrapping paper. Suppose the number of defects on a given package of wrapping paper has a Poisson distribution with  $\lambda = 4.6$ . You take a random sample of 150 packages and count the number of defects.

- 42.1. What is the probability the average number of defects per package in the sample is less than 4?
- 42.2. What is the probability that average number of defects per package in the sample is greater than 5?
- 42.3. What is the probability the average number of defects per package in the sample is between 4 and 5?