Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

***Problem Set 7.1 Solutions***

1.) Give an example of an event with: (and be prepared to share and discuss in the next class)

a) probability = 1

Picking a red skittle out of a bowl of only red skittles

Flipping a fair coin and it lands on either “Heads or Tails”

Nicky will land if he jumps on planet earth

The sun comes up tomorrow morning

b) probability = 0

PG will have school on the weekend

Flipping a fair coin and it lands on neither “Heads or Tails”

A human will life forever

Snow in the summer in Charleston, SC in 2019

c) relative frequency = 1

Neil threw Chai’s pen and it fell on the ground

Picked ten red skittles out of a bowl of 10 red skittles

Nicky jumped and he landed

President Lincoln is assassinated

d) relative frequency = 0

Picked a skittle out of an empty bowl

A human will life forever

A group of people fell asleep at the exact same time on 4/24/19

2.) The data below describes passenger survival from the Titanic. (note, the data excludes crew members). One of the reasons there were so many fatalities was that for aesthetic reasons the ship did not carry enough lifeboats for its capacity. There was only room for a maximum of 52.9% of the boat’s population in the lifeboats, but the survival rate was much less than that.

Source: The Real Reason for the Tragedy of the Titanic. The Wall Street Journal, n.d.Web. 2 Mar. 2013.

This data is organized in a **two- way table** – a table that classifies data based on possible categories for two different variables at the same time, one by rows and one by columns. It also includes the totals for each category and an absolute total.



a) What is the sum of the numbers in the red box? Why?

1316 – because it is the total of all the passengers on the Titanic

b) What does the number 498 in the bottom row represent?

498 represents the total number of survivors

c) What is the relationship between the four numbers in the last column?

The first three rows represent the total number of passengers in each class. The fourth row represents the total number of passengers. “It’s all just totals,” says Neil W.

d) What’s wrong with this question: “What is the probability that any passenger survived?”

Probability is for future events.

e) Does this table give relative frequencies? Explain.

No, but you can use the data from the table to compute relative frequencies.

f) What was the relative frequency of survival?

498/1316 is approximately 0.378 or 38%

g) What was the relative frequency of third-class passengers on the ship?

706/1316 is approximately 0.536 or 53.6%

\*\*\*\*h) What was the relative frequency of third-class survivors?

178/1316 is approximately 0.135 or 13.5%

178/706 is approximately 0.252 or 25.2% (conditional probability)

i) If we only consider survivors, what was the relative frequency of third-class passengers?

178/498 is approximately 0.357 or 35.7%

j) Which is bigger between your answers for *h* and *i*? Why?

I is bigger because only considering survivors which is a smaller group so denominator is smaller hence ratio is bigger.

k) Did class of passenger affect the likelihood of survival? Justify your answer with calculations, and if it did, provide a possible reason why that may have been the case.

Various answers accepted

l) Give one other example of other types of questions about relative frequency that can be answered with this table and answer your question. Show your work.

Various answers accepted

3.) The following graph was published as part of a study on diversity involving teenagers done by Gallup in 2003.



a) If we asked the same questions of a random sample of 400 teenagers, 100 who identify as white and 300 who identify as a race other than white and expected the same proportional responses that Gallup found, how many teens would fit each of the categories below? Use the graph to help you fill in the two-way frequency table. One piece of the chart is given.



b) Why was the row “No Answer” necessary when it wasn’t part of the graph? What observation about the graph makes that row necessary?

If all of the students answered the survey, then the sum of the relative frequencies for each type of student should be 100%. Since 27% + 44% + 25% = 96%, which tells us that 4% of the Nonwhite students chose not to answer the question (or gave an answer that didn’t fit one of the given categories).

4.) The tables you have seen in the previous problems were both two-way frequency tables. A two-way relative frequency table looks very similar, but the cells are filled with numbers between 0 and 1, representing the percentage of outcomes that fall in a certain category.

Let’s take the given frequency table of adult smokers (18 and older) in the US in 2011 and turn it into a relative frequency table:



a) In the frequency table, what does the number in the bottom right corner represent?

The total number of men and women who are smokers and nonsmokers (total number of men and women in the survey)

b) In the relative frequency table, what percentage would represent all responses?

100% or 1

**To create a relative frequency table from a frequency table, divide the value in every cell by the absolute total.**

c) Explain why the instruction above creates a relative frequency table. Then use it to fill in the relative frequency table below, **rounding to the nearest thousandth (third decimal place)**. One cell is done for you.



d) Do the total row and total column still add to the appropriate total in the bottom right cell? Would that always be the case when creating a relative frequency table from a frequency table?

In this case, the total row and total column do still add to the overall total of 100%. That should always be the case when creating a relative frequency table from a frequency table if you account for all decimal places. If you’re rounding the relative frequencies your totals may be slightly off.

e) What does the value 0.105 represent?

0.105 tells us that in 2011, approximately 10.5% of the US population identified as men smokers.

f) What does the value of 0.514 represent?

0.514 tells us that in 2011, approximately 51.4% of people in the US were women (bonus: why not 50%).