Midterm

Please complete this midterm using RMarkdown. You may use your notes, the textbook, homework solutions, etc. when working on the problems–the only thing you may not do is discuss the midterm with your classmates or others. If you have any questions, please email me.

Hotel Prices (15 pts)

I am planning a trip to the White Mountains in New Hampshire. The first thing I need to do is book my hotel–the data frame `lodging_prices` contains information on 29 different rentals in Lincoln, a nearby town. The variables are as follows:

- `Nightly_Rate`: the nightly rate as advertised on hotels.com
- `Miles_Away`: the number of miles away from the city center
- `Avg_Rating`: the average rating posted on hotels.com
- `Type`: what type of lodging the property is (`"Hotel"`, `"Condo"`, `"Home"`)
- `Free_Parking`: whether free parking is advertised on hotels.com
- `Internet`: whether wireless internet is advertised on hotels.com
- `Number_Reviews`: the number of reviews posted on hotels.com

(a) Using `ggplot()`, create a histogram of `Nightly_Rate`. Your histogram should have 8 bins (3 pts).

(b) Comment on the distribution of nightly rates (4 pts).

(c) Using `ggplot()`, create a scatterplot of `Avg_Rating` against `Number_Reviews`. Change the color of the points to a color of your choice (3 pts).

(d) Comment on the relationship (4 pts).

(e) Calculate the correlation between `Number_Reviews` and `Avg_Rating` advertising dollars (1 pt).

Moose Tours (18 points)

After I find a hotel, I decide to book a tour with Pemi Valley Moose Tours. According to their website, they have a success rate of 0.97 on their tours–that is, on 97% of their tours, they expect to see at least one moose.

(a) How many tours should the manager expect to run before the tour is unsuccessful (that is, they do not see a moose)?

(b) Pemi Valley Moose Tours plans on carrying out 36 tours in the month of October (one day per week, twice each Saturday and Saturday, no tours on Mondays). What is the probability that more than 4 of their tours are unsuccessful?

(c) What is the probability that all of their tours are successful?

(d) What is the probability that there are exactly 3 tours before the first unsuccessful tour?

(e) What is the probability that there are at least 8 succesful tours before the first unsuccessful tour?

(f) How many tours of the 36 in October should Pemi Valley Moose Tours expect to be successful?

Fall Foliage Activities (15 points)

I've found that other popular activities include driving on the Kancamagus highway to see fall foliage, and hiking in Franconia Notch State Park–the most popular trails in Franconia Notch are Flume Gorge, Artist's Bluff, and the Basin. A survey of a random sample of tourists shows the following:

	Flume Gorge	Artist's Bluff	Basin	None	Total
Drove	80	50	20	10	160
Did not drive	18	8	4	2	32
Total	98	58	24	12	192

(a) What is the probability that a randomly chosen tourist visits Artist's Bluff (2 pts)?

(b) What is the probability that a randomly chosen tourist drove on the Kancamagus Highway (2 pts)?

(c) What is the probability that a randomly chosen tourist drove on the highway and did not do any of the hikes in the state park (2 pts)?

(d) What is the probability that a randomly chosen tourist drove on the highway, given that they hiked Flume Gorge (2 pts)?

(e) What is the probability that a randomly chosen customer did not drive the highway, given that they hiked the Basin (2 pts)?

(f) Are driving on the highway and hiking Flume Gorge independent? Give a mathematical justification (3 pts).

Pancakes for Breakfast (15 points)

A local restaurant, Polly's Pancake Parlor, is very popular–at their busiest, 87.5 customers show up every hour.

(a) What is the probability that exactly 100 customers come in during one hour?

(b) What is the probability that no more than 105 customers during one hour?

(c) The owner takes a random sample of 42 hours during the month of October (because of the fall foliage, October is always a busy time). What is the distribution of the customers arriving each hour?

(d) What is the probability that the average number of customers is between 82 and 86?

(e) What is the probability that the average number of customers is at least 90?

More Pancakes (15 points)

Management at Polly's knows that the average bill is normally distributed \$81, with a standard deviation of \$32.

(a) What is the probability that a single bill is more than \$65?

(b) Service at Polly's is very good, and every customer tips exactly 20%. Thus, the actual cost of the meal will be 20% higher than the bill. If a bill is represented by B, write an expression that represents the actual cost of the meal.

(c) How much should customers expect their meal to cost with tips? What is the standard deviation of cost with tips?

(d) What is the probability that the cost of a meal is between \$55 and \$115?

(e) What is the probability that a meal costs at most \$40?

T-Shirt Sales (8 points)

While on my trip, I visit a store selling t-shirts. The first shirt is always \$17.98, but customers can buy up to four additional boxes for \$12.98 each. The store knows approximately how many customers will buy one, two, three, four, or five shirts; however, some customers come in and do not buy anything. Here is a distribution of their expected sales, S:

	\$0	\$17.98	\$30.96	\$43.94	\$56.92	\$69.90
$P(S)$	0.16	0.12	$P(\$30.96)$	0.22	0.22	0.08

(a) Find the value of $P(\$30.96)$ that will make this distribution a valid distribution (2 pts).

(b) What is $E(S)$ (2 pts)?

(c) What is $Var(S)$ (3 pts)?

(d) What is the probability a customer buys at least two shirts (1 pt)?

Mount Washington (14 points)

Mount Washington, the highest peak in the Northeastern United States, is located in the Presidential Range of the White Mountains. Because of its location, extremely high wind speeds are frequently reported–in a random sample of 40 days in 2021, researchers at the Mount Washington Observatory observed hurricane-force wind gusts 13 times.

(a) Build a 99% confidence interval for the true proportion of days with hurricane-force wind gusts at the top of Mount Washington (9 pts).

(b) What is the margin of error of the confidence interval (3 pts)?

(c) If the researchers choose to report a 90% confidence instead of 99%, will the interval become wider or narrower (2 points)?