Exam 2 - Review Exercises

**Acid Rain.** Based on long-term investigation, researchers have suggested that the acidity (pH) of rainfall in the Shenandoan Mountains can be described by the Normal Model N(4.9, 0.6).

a. What percent of storms produce rainfall with pH over 6?
   - Have a normal distribution, with mean and standard deviation given
   - Asked for a percent
   - Need to find
     \[ z = \frac{y - \mu}{\sigma} \]
   - Area in table corresponding to a z-score of
   - To find area to the RIGHT, need to
     \[ \text{Area} = \]
   - Or percent =

b. What percent of storms produce rainfall with pH under 4?
   - Need to find z-score then area in table less than corresponding z-score
     \[ z = \frac{y - \mu}{\sigma} \]
   - Area in table corresponding to a z-score of
   - Percent =

c. The lower the pH, the more acidic the rain. What is the pH level for the most acidic 20% of all storms?
   - What value marks off the
   - Need to find z-score first, then work backwards to find corresponding pH level.
   - Need to look up area of to the left, work backwards to find z-score
   - 0.2000 is between 0.1977 and 0.2005 – but closest to
   - Plug z-score into formula, solve for y
     \[ z\sigma + \mu = y \]

d. What is the pH level for the least acidic 5% of all storms?
   - What value marks off the
   - Need to find z-score first, then work backwards to find corresponding pH level.
   - If area to the right is 0.0500, then area to the left is
   - Need to look up area of to the left, work backwards to find z-score
   - 0.9500 is directly between 0.9495 and 0.9505 – need to average corresponding z-scores
   - \[ z = \]
   - Plug z-score into formula, solve for y
     \[ z\sigma + \mu = y \]
Exercise 8.24
The regression of Duration of a roller coaster ride on the height of its initial Drop has $R^2 = 12.4\%$.

a. What is the correlation between Drop and Duration?
   - $R^2$ is the square of the
     \[ r = \]

b. What would you predict about the Duration of the ride on a coaster whose initial Drop was 1 standard deviation below the mean Drop?
   - Need to use the linear equation for standardized variables
     \[ Z_{\text{duration}} = \]
     \[ Z_{\text{duration}} = \]

   We would expect a ride whose initial Drop is 1 standard deviation below the mean to have a Duration

c. What would you predict about the Duration of the ride on a coaster whose initial Drop was 3 standard deviation above the mean Drop?
   - Need to use the linear equation for standardized variables
     \[ Z_{\text{duration}} = \]
     \[ Z_{\text{duration}} = \]

   We would expect a ride whose initial Drop is 3 standard deviations above the mean to have a Duration

Exercise 8.39
The SAT is a test often used as part of an application to college. SAT scores are between 200 and 800. Tests are given in both Math and Verbal areas. Doing the SAT-Math problems also involves the ability to read and understand the questions, but can a person’s verbal score be used to predict the math score?

Verbal and math SAT scores of a high school graduating class are displayed in the scatterplot, with the regression line added.

a. Describe the relationship.

b. Are there any students whose scores do not seem to fit the overall pattern?
   - The student with the Verbal Score and Math Score

   \[ \sum (x - \bar{x})(y - \bar{y}) = 1,076,459.3 \quad s_x = 99.5 \quad s_y = 98.1 \quad n = 162 \]
These verbal scores averaged 596.3, with a standard deviation of 99.5, and the math scores averaged 612.2, with a standard deviation of 98.1. Write the equation of the regression line.

\[
\hat{y} = b_0 + b_1 x
\]

b_1 =

b_0 =

Equation:

math =

e. Interpret the slope of the line.
- Every point of verbal score adds an average of to the predicted math score.

f. Predict the math score of a student with a verbal score of 500.
- Plug 500 into the equation in part (d)

math =
math =

We would expect a student with a verbal score of 500 to score a 547

g. Every year some student scores a perfect 1600. Based on this model, what would that student’s residual be for her math score?
- If a student scored 1600, then their verbal score would be an 800.
- Plug 800 into the model to see what the predicted math score is.
- Use this predicted math score in the residual equation, with 800 as the observed math score to find the residual.

math =
math =
residual =
residual =
residual =

A student who scores a 1600 on the SAT would have a residual of

Car Sales A car dealer investigated the association between the number of TV commercials he ran each week and the number of cars he sold the following weekend. He found the correlation to be \( r = 0.56 \). During the time he collected the data he ran an average of 12.4 commercials a week with a standard deviation of 1.8, and sold an average of 30.5 cars with a standard deviation of 4.2. Next weekend he is planning a sale, hoping to sell 40 cars. Create a linear model to estimate the number of commercials he should run this week. Write a sentence explaining your recommendation.

- Want to estimate the number of commercials, so commercials will be the response, or y-variable
- Using number of cars the dealer hopes to sell as the predictor, or x-variable
- Need to find the values for slope and intercept to create our equation
\[ \text{slope: } b_1 = r \frac{s_y}{s_x} \]

Slope =

\[ \text{intercept: } b_0 = \bar{y} - b_1 \bar{x} \]

Intercept =

- Equation of model:
  
  Commercials =

- Want to use the model to predict the number of commercials needed to sell 40 cars next weekend.

  Commercials =
  Commercials =
  Commercials =

If the salesman wants to sell 40 cars next weekend, he should show

Exercise 6.56

A Cornell University researcher measured the mouth volumes of 31 men and 30 women. She found a mean of 66 cc for men (SD=17 cc) and a mean of 54 cc for women (SD = 14.5 cc). The man with the largest mouth had a mouth volume of 111.2 cc. The woman with the largest mouth had a mouth volume of 95.8 cc.

a. Which had the more extraordinarily large mouth?
   - Have 2 values from 2 different normal distributions and want to find which is more extreme
   - Need to find \[ Z = \frac{y - \mu}{\sigma} \]
   - Man: \( y = \)
     \[ z = \]
     \[ z= \]
   - Woman: \( y = \)
     \[ z = \]
     \[ z= \]

so the woman has the more extraordinarily large mouth.

b. If the distribution of mouth volumes is nearly Normal, what percentage of men and of women should have even larger mouths than these?
   - Have z-scores, need to find percentage that are larger
   - For men, look up z-score 2.66 in table
   - \( z=2.66 \) corresponds to an area of \( \) in the table
   - This is the area to the left of 2.66, we want the area “larger” or to the right
   - Area =
   - \( \) of mouths are larger than the largest male’s in our sample
- For women, look up z-score 2.88 in table
- Z=2.88 corresponds to an area of in the table
- This is the area to the left of 2.88, we want the area “larger” or to the right
- Area =
- of mouths are larger than the largest female’s in our sample

On the exam:
- I will provide z-tables and formulas
- Should know and be able to list assumptions
- Interpretations are important!!!!
  - \( R^2 \)
  - Slope of a line
  - Estimate from a regression line