LAB: INTRO TO STAT ANALYSIS

Zeyi Qian

zeqian@clarku.edu

Office Hours: JC 201, Tuesday 4-5 PM & Thursday 3-4 PM

September 13, 2024

Question 1

- Suppose your Statistics professor reports test grades as z-scores, and you got a score of 2.30 on an exam. Write a sentence explaining what that means.

$$Z = \frac{y - \bar{y}}{Stdev} = \frac{y - \mu}{\sigma}$$

  - Z-score means how many standard deviations a given score is from the mean
  - In this case: grades higher than the mean by 2.3 standard deviations

Question 2: a & b

Suppose the weight of cows is normally distributed with a mean of 1152 pounds and a standard deviation of 84 pounds

- a. How many standard deviations away from the mean would a cow weighing 1000 pounds be?

$$z = \frac{y - \mu}{\sigma} = \frac{1000 - 1152}{84} = -1.81$$

- b. Which would be more **unusual**: a cow weighing 1000 pounds or a cow weighing 1250 pounds?

$$\frac{1000 - 1152}{84} = -1.81 \ V.S. \ \frac{1250 - 1152}{84} = 1.17$$

and $\mid -1.81 \mid > \mid 1.17 \mid$

Question 2: c

- c. Cattle buyers hope that cows will weigh at least 1000 pounds. To see how much over (or under) that goal the cows are, we could subtract 1000 pounds from all the weights. What would the new mean and standard deviation be?

$$mean = 1152 - 1000 = 152, \; stdev = 84$$

  - Subtracting a constant from all values shifts the mean by the same amount but doesn't change the standard deviation

Question 2: d

- d. Suppose such cattle sell at auction for 40 cents a pound. Find the mean and standard deviation of the sales prices (in dollars) for all the cows

$$mean = 1152 \times 0.4 = 460.8, \;\; stdev = 84 \times 0.4 = 33.6$$

  - The standard deviation measures the degree of volatility or dispersion of the data, and when the data is **linearly transformed**, the standard deviation is affected accordingly

Question 3: a

Hens usually begin laying eggs when they are about 6 months old. Young hens tend to lay smaller eggs, often weighing less than the desired minimum weight of **54**grams

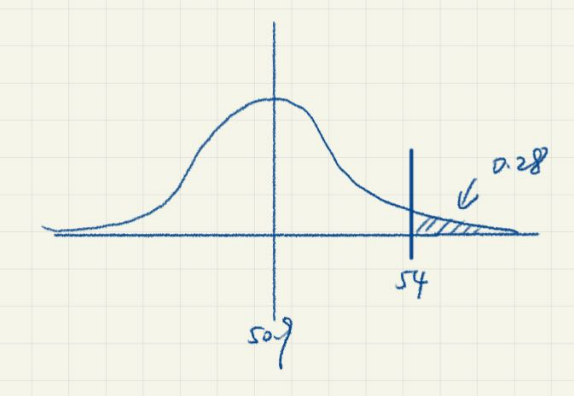
- a. The average weight of the eggs produced by the young hens is **50.9** grams and only **28%**of their eggs exceed the desired minimum weight. If a Normal model is appropriate, what would the standard deviation of the egg weights be?

Question 3: a

- Step 1: Find the Z-score for 28%:

$$P(X > 54) = 0.28, So\ P(X \leq 54) = 1 - 0.28 = 0.72$$

- Use Excel to find the Z score $Z = NORM.S.INV(0.72) = 0.58$



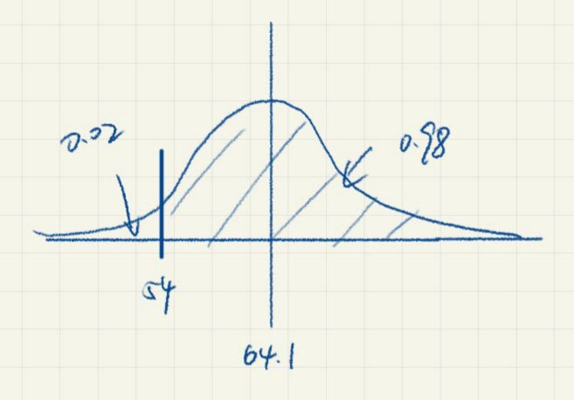
Question 3: a

- Step 2: Use the Z-score formula to solve for the standard deviation

$$\frac{54 - 50.9}{\sigma} = 0.58, \; so \; \sigma = 5.34$$

Question 3: b

- b. By the time these hens have reached the age of 1 year, the eggs they produce average **67.1** grams, and **98%** of them are above the minimum weight. What is the standard deviation for the appropriate Normal model for these older hens?



Question 3: b

- Step 1: Find the Z-score for 98%:

$$P(X > 54) = 0.98, So\ P(X \leq 54) = 1 - 0.98 = 0.02$$

- Use Excel to find the Z score $Z = NORM.S.INV(0.02) = -2.05$
- Step 2: Use the Z-score formula to solve for the standard deviation

$$\frac{54 - 67.1}{\sigma} = -2.05,\ so\ \sigma = 6.39$$

Question 3: c

- c. Are egg sizes more consistent for the younger hens or the older ones? Explain

$$\sigma_{young} = 5.34 < \sigma_{old} = 6.39$$

- The smaller the standard deviation of young hens, the more consistent the egg size

Question 4

- Adult female Dalmatians weigh an average of **50** pounds with a standard deviation of **3.3** pounds. Adult female Boxers weigh an average of **57.5** pounds with a standard deviation of **1.7** pounds. One statistics teacher owns an underweight Dalmatian and an underweight Boxer. The Dalmatian weighs **45** pounds, and the Boxer weighs **52** pounds. Which dog is more underweight?

$$Z_{Dalmatian} = \frac{45 - 50}{3.3} = -1.52$$

$$Z_{Boxer} = \frac{52 - 57.5}{3.3} = -3.24$$

  - Boxer is more underweight ($| -3.24 | > | -1.52 |$)

Question 5

- Based on data collected from its production processes, Crosstiles Inc. determines that the breaking strength of its most popular porcelain tile is normally distributed with a mean of **400** pounds per square inch and a standard deviation of **12.5** pounds per square inch. What percent of its popular porcelain tile will have breaking strengths between **375** and **400** pounds per square inch?

Question 5

- Calculate z score

$$Z = \frac{375 - 400}{12.5} = -2$$

- Use Excel to find the probability: =NORM.S.DIST(-2, TRUE)= 2.2 %

$$50\% - 2.2\% = 47.8\%$$