

STT 421 Exam 1: Summer 2004 SOLUTIONS

NAME: _____

PID: _____

1. (5 points) Figure 1 contains a plot of a normal density. Circle all of the following which are correct.

- The standard deviation is about 15.
- About 68% of the area is between 20 and 30.
- The standard deviation is about 0.08.
- The standard deviation is about 5.
- The mean is about 25.

Solution: The second statement is true. The standard deviation is about 5, so the interval from 20 to 30 is the interval of values within one standard deviation of the mean, which has area of approximately 0.68. The fourth and fifth statements are true. The other two statements are false. \square

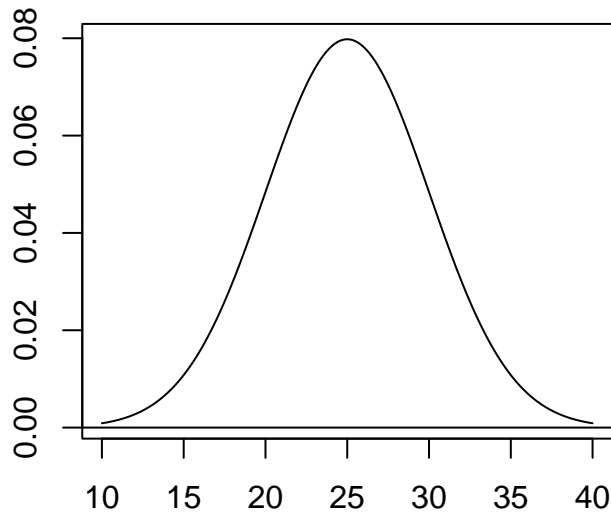


Figure 1: The normal density for Problem 1

2. (4 points) Data were collected on the calorie content of various brands of hot dogs, classified by type (beef, meat, poultry). The data are summarized in the boxplot in Figure 2. Answer the following questions based on the plot. For some questions you'll have to read values from the graph; you will get credit as long as your answer is reasonably close to the actual answer. If it is not possible to answer the question from the boxplot, write **NA** as your answer.

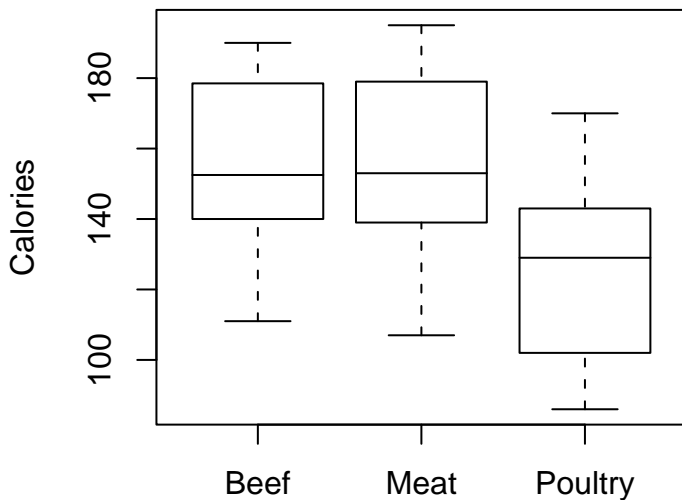


Figure 2: Boxplots of data on calorie content of hot dogs for Problem 2.

- (a) Is the median calorie content for beef hot dogs less than 160?

Solution: Yes, the median is about 154.

- (b) What is the third quartile of calorie content for poultry hot dogs?

Solution: Approximately 142.

- (c) What is the IQR of calorie contents for meat hot dogs?

Solution: Approximately $180 - 140 = 40$.

- (d) Were there more beef, meat, or poultry hot dogs in the data set?

Solution: NA.

3. **(9 points)** As part of a study of the nutrition of London bus drivers, measurements of daily caloric intake were made. The mean and standard deviation of the measurements were 2821 and 436 respectively. Assuming that the daily caloric intake is accurately modeled by a normal density with these values for mean and standard deviation, compute the following quantities.

(a) The proportion of drivers whose daily caloric intake is less than 2300 calories.

Solution: Standardizing yields $(2300 - 2821)/436 \approx -1.19$. Look in the standard normal table to get the area to the left of -1.19 , which is 0.1170. \square

(b) The proportion of drivers whose daily caloric intake is between 3000 and 3200 calories.

Solution: Standardizing yields the interval from $(3000 - 2821)/436 \approx 0.41$ to $(3200 - 2821)/436 \approx 0.87$. The area between these values is $0.8078 - 0.6591 = 0.1487$. \square

(c) The caloric intake which 85% of drivers exceed.

Solution: We want 85% of the area to the right of this value, which means we want 15% of the area to the left. For the standard normal density this leads to -1.04 . So the answer is $(436)(-1.04) + 2821 = 2367.56$. \square

4. (4 points) Figure 3 contains a histogram of the stopping distances of 50 cars (which were traveling at various speeds). Answer the following questions as accurately as you can based on the histogram.

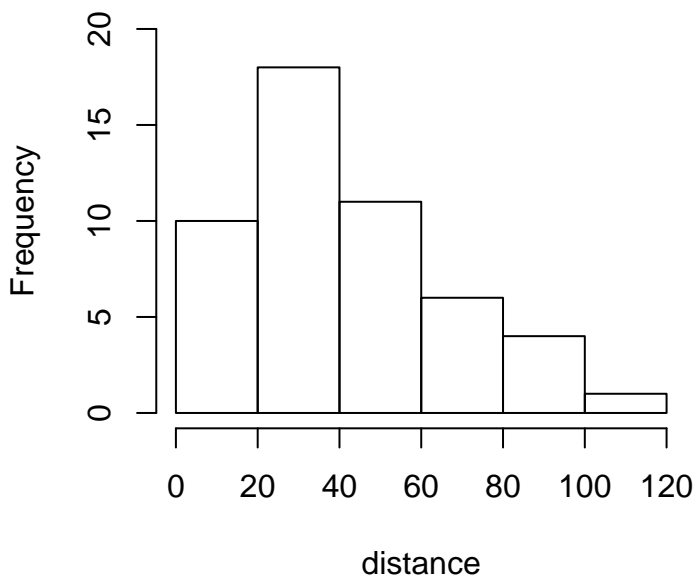


Figure 3: Histogram of stopping distances for Problem 4.

- (a) What proportion of stopping distances were less than 40?

Solution: $28/50 = 0.56$.

- (b) Would a standard normal distribution provide a reasonable model for the data?

Solution: No. The mean and standard deviation are clearly not close to 0 and 1. In addition, the histogram is skewed to the right.

- (c) Is the distribution of the data right-skewed or left-skewed?

Solution: Right-skewed.

- (d) Would you expect the mean of the data to be larger or smaller than the median of the data?

Solution: Larger.

5. (3 points) A list of numbers has the following associated statistics.

- Mean $\bar{x} = 20$.
- Median $M = 22$.
- Maximum value 38.
- Sample size $n = 20$.

(a) The number 20 is added to the original list. Will the mean change?

Solution: No. If you don't understand why, construct a small list of numbers with mean 20, then add another 20 to the list, and recompute the mean. \square

(b) The number 0 is added to the original list. Will the mean change?

Solution: Yes. \square

(c) The number 50 is added to the original list. Which will change more, the mean or the median?

Solution: We're trained to answer that the mean will change more, since the mean is not resistant to outliers. But the answer actually depends on the particular data set. For example, the data set

$$-2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 6, 38, 38, 38, 38, 38, 38, 38, 38, 38, 38$$

has 20 values, with mean 20, median 22, and maximum 38. Adding the value 50 changes the median to 38, but only changes the mean to $450/21 \approx 21.4$. \square