

Attach all of your R code at the end of your solution file or in a separate text file.

**Problem 1 (12 points, 1 point for each)**

The following multiple-choice questions are intended to provide practice in methods and reinforce some of the concepts presented in Chapter 1.

1.1 The scores of eight persons on the Stanford–Binet IQ test were:

95    87    96    110    150    104    112    110

The median is:

- (1) 107
- (2) 110
- (3) 112
- (4) 104
- (5) none of the above

1.2 The concentration of DDT, in milligrams per liter, is:

- (1) a nominal variable
- (2) an ordinal variable
- (3) a continuous variable
- (4) a discrete variable

1.3 If the interquartile range is zero, you can conclude that:

- (1) the range must also be zero (Note that the range of the data is the difference between the largest and smallest values)
- (2) the mean is also zero
- (3) at least 50% of the observations have the same value
- (4) all of the observations have the same value
- (5) none of the above is correct

1.4. The species of each insect found in a plot of cropland is:

- (1) a nominal variable
- (2) an ordinal variable
- (3) a continuous variable
- (4) a discrete variable

1.5 A sample of 100 IQ scores produced the following statistics:

mean = 95      lower quartile = 70

median = 100   upper quartile = 120

standard deviation = 30

Which statement(s) is (are) correct?

- (1) Half of the scores are less than 95.
- (2) The middle 50% of scores are between 100 and 120.
- (3) One-quarter of the scores are greater than 120.

1.6 A sample of 100 IQ scores produced the following statistics:

mean = 100      lower quartile = 70

median = 95   upper quartile = 120

standard deviation = 30

Which statement(s) is (are) correct?

- (1) Half of the scores are less than 100.
- (2) The middle 50% of scores are between 70 and 120.
- (3) One-quarter of the scores are greater than 100.

1.7 A sample of pounds lost in a given week by individual members of a weight reducing clinic produced the following statistics:

mean = 5 pounds      first quartile = 2 pounds

median = 7 pounds      third quartile = 8.5 pounds

standard deviation = 2 pounds

Identify the correct statement:

- (1) One-fourth of the members lost less than 2 pounds.
- (2) The middle 50% of the members lost between 2 and 8.5 pounds.
- (3) Both (1) and (2) are correct.
- (4) Neither (1) and (2) is correct.

1.8 A measurable characteristic of a population is:

- (1) a parameter
- (2) a statistic
- (3) a sample
- (4) an experiment

1.9 What is the primary characteristic of a set of data for which the standard deviation is zero?

- (1) All values of the variable appear with equal frequency.
- (2) All values of the variable have the same value.
- (3) The mean of the values is also zero.
- (4) All of (1), (2), and (3) are correct.
- (5) Neither of (1), (2) or (3) is correct.

1.10 A subset of a population is:

- (1) a parameter
- (2) a population
- (3) a statistic
- (4) a sample
- (5) none of the above

1.11 The median is a better measure of central tendency than the mean if:

- (1) the variable is discrete
- (2) the distribution is skewed
- (3) the variable is continuous
- (4) the distribution is symmetric
- (5) none of the above is correct

1.12 A small sample of automobile owners at Texas A & M University produced the following number of parking tickets during a particular year: 4, 0, 3, 2, 5, 1, 2, 1, 0. The mean number of tickets (rounded to the nearest tenth) is:

- (1) 1.7
- (2) 2.0
- (3) 2.5

(4) 3.0

(5) none of the above

### Problem 2 (10 points)

On ten days, a bank had 18, 15, 13, 12, 8, 3, 7, 14, 16, and 3 bad checks. Find the sample mean, sample median, sample variance, and sample standard deviation, of the number of bad checks. You need to show the details about how you calculate them.

### Problem 3 (24 points, 6 points for each part)

Data file anl.csv contains the times in days from remission induction to relapse for 50 patients with acute nonlymphoblastic leukemia who were treated on a common protocol at university and private institutions in the Pacific Northwest. This is a portion of a larger study reported by Glucksberg et al. (1981). The data file only contains one variable “days”.

- (1) Draw a histogram using R function hist with default settings. Comments on the features of your plot. You must use appropriate main title and labels for x-axis and y-axis for the plot.
- (2) Draw a boxplot using R function boxplot. Comments on the features of your plot. You must use appropriate main title and labels for x-axis and y-axis for the plot.
- (3) Use R to calculate the following summary statistics: sample mean, sample median, sample variance, sample standard deviation, sample median, sample lower quartile, and sample upper quartile. You only need to present your final answer here. You do not need to present any details about your calculations.
- (4) Calculate the following values used in the boxplot: interquartile, step, lower inner fence, upper inner fence, lower outer fence, and upper outer fence. You need to show the details about how you calculate them based on summary the statistics obtained from (3).