Practical Data Science
(Univariate Stats)

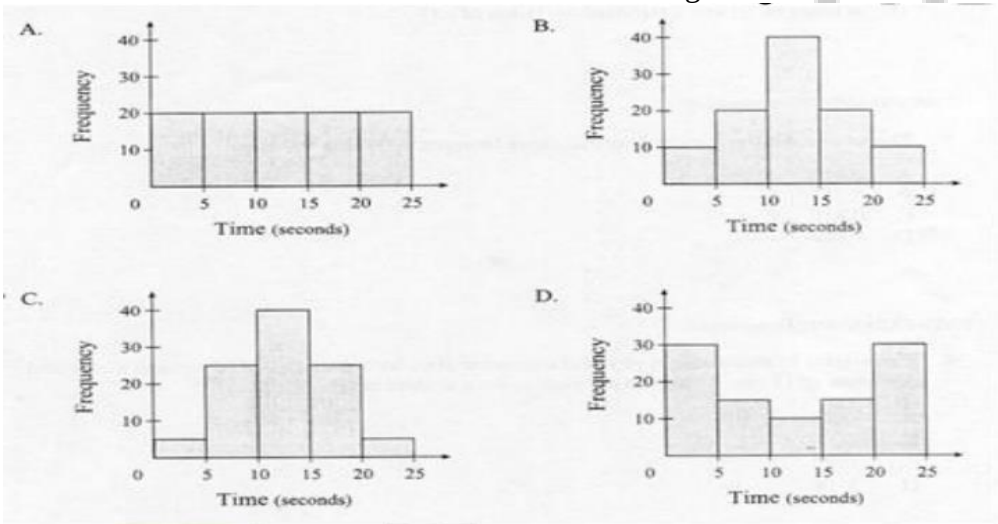
Solve the following problems.

Problem 1: Mean and Standard Deviation

I. A teacher adjusts the marks of an examination by raising each score by 5 percent. What happens to mean and standard deviation?

II. Would you expect the following marks of an examination to have small or large standard deviation? 92 93 92 94 92 91 92

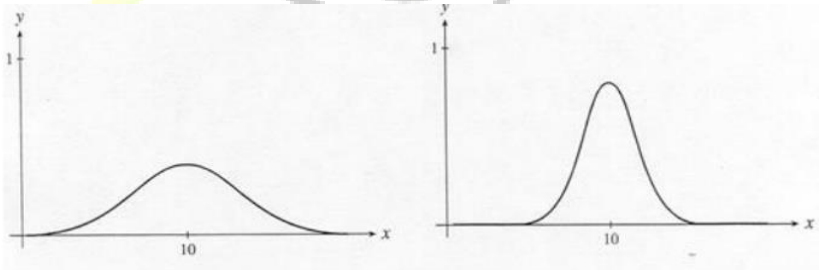
III. Which set of data would have the smallest and largest standard deviation?



IV. If a set of data has a standard deviation of 0, then:

A. the mean of the data must be 0
B. all of the data values are the same
C. the data values collected had a sum of 0
D. the z-score of the mean of the data is equal to 1

V. Which of the two normal distributions graphed in the diagrams below has a higher standard deviation?



Copyright © 2015-16 by Algorithmica

Practical Data Science
(Univariate Stats)

VI. Henry played 24 golf games on the same course during each of two seasons. In the first season, his mean score was 75 with a standard deviation of 2.1. In the second season, his mean score was 74 with a standard deviation of 3.8. Examining the standard deviation of Henry's score for the two seasons, one could conclude which one of the following would be true:

A. scores were more consistent in the first season
B. scores were more consistent in the second season
C. average score was better in the first season
D. average score was better in the second season

Problem 2: Z-Scores

I. The average mark on a test was 58.3 with a standard deviation of 6.7. The zscore of a particular mark was -1.3, what was the mark on test?

II. A teacher marks some exams and finds the mean is 54% and the standard deviation is 8%. The teacher then adjusts the marks by raising the mean to 60% and raising the standard deviation to 9%. The z-scores are kept constant. If the student scored 76% initially, what would be their new mark be?

Problem 3: Comparing among distributions

Paul got a mark of 75 on a math test with a mean of 61 and a standard deviation of 12. He got 72 on a chemistry exam with mean 63 and standard deviation 7. If the marks on both tests were normally distributed, on which test did he do better relative to the class?

Problem 4: Normality Test

Load data "income.csv" available under datasets(github of algorithmica) branch into R data frame. Do the following things:

a) Obtain descriptive statistics for "income", "edu", and "expr". The statistics should include number of observations, min, max, mean, median, std, skewness, kurtosis, quantile(0.25), quantile(0.75).

b) Does income, education and expr of all people follows normal distribution individually?

c) Does income, education and expr of male people follows normal distribution individually?

d) Does income, education and expr of female people follows normal distribution individually?

Copyright © 2015-16 by Algorithmica