

Midterm-Solutions

Fall 2017 Mat1371-Descriptive Statistics
Dr.Dr. Alessandro Selvitella

Question 1. We first compute the main "order" statistics. We get:

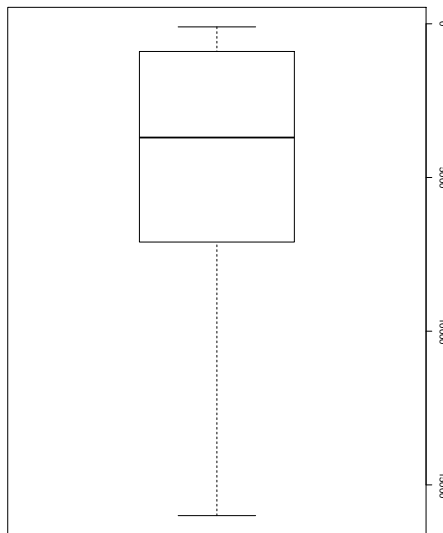
$$\min = 100 \quad q_1 = 900 \quad m = 3700 \quad q_3 = 7200 \quad \max = 16.000.$$

- (a) [6 points] Since

$$q_3 + 1.5 * (q_3 - q_1) = 16650,$$

then there isn't any outlier ($\max = 16000 < 16650$).

- (b) [2 points]



- (c) [2 points] The stem-and-leaves plot (as well as the boxplot) shows that the distribution is strongly positively skewed. Therefore, the mean is going to be most likely greater than the median.

Question 2.

- (a) [2+1/3 points] $\frac{562}{1639} \simeq 34.3\%$.
- (b) [2+1/3 points] $\frac{13}{562} \simeq 2.3\%$.
- (c) [2+1/3 points] $\frac{57}{612} \simeq 9.3\%$.
- (d) [3 points] There is an association between the consumption of coffee and the status of the disease. The more the patient drinks coffee, the less the chances the patient develops cirrhosis. For example 9.3% of the patients which drink zero cups of coffee a day develop cirrhosis. This percentage drops to 2.3% for patients drinking at least three coffee cups a day. Further analysis is needed to verify what the visual analysis of the plot and descriptive statistics suggest.

Question 3.

- (a) [3 points]

$$Pr(X > 9) = 1 - Pr(X < 9) = 1 - Pr(Z < 3) \simeq 1 - 0.9987 = 0.0013 = 0.13\%.$$

- (b) [5 points]

$$Pr(7 < X < 9) = Pr(-1 < Z < 3) = Pr(Z < 3) - Pr(Z < -1) = 0.9987 - 0.1587 = 0.8400 \simeq 84\%.$$

- (c) [2 points]

$$-1.67 = \frac{c - 7.5}{0.5}.$$

Therefore,

$$c = -1.67 * 0.5 + 7.5 = 6.665 \text{ m}.$$

Question 4.

- (a) [2 points] Serum 2.
- (b) [2 points] Serum 1.
- (c) [2 points] Serum 1 (It has the smallest IQR).
- (d) [2 points] Serum 2.
- (e) [1 points] Serum 2 and No Treatment.
- (f) [1 points] Serum 2 and No Treatment.

Question 5.

- (a) [3 points]

$$\min = 35 \text{ cm/s}^2 \quad q_1 = 50 \text{ cm/s}^2 \quad m = 72.5 \text{ cm/s}^2 \quad q_3 = 82.5 \text{ cm/s}^2 \quad \max = 87 \text{ cm/s}^2.$$

- (b) [2 points] $\bar{x} = \frac{531}{8} = 66.375 \text{ cm/s}^2$, $s = \sqrt{37839 - (53)^2/8} = 19.2498 \text{ cm/s}^2$.

- (c) [3 points]

$$z_1 = \frac{35 - 66.375}{19.2498} = -1.63;$$

$$z_2 = \frac{68 - 75.25}{3.6547} = -1.98.$$

Since z_2 is farther from 0 than z_1 , then z_2 is more extreme.

- (d) [2 points] The mean of all the deviations from the mean is always zero. The standard deviation is unchanged, because scaling does not affect it $s = 3.654743 \text{ cm/s}^2$.

TOTAL /50