

MAT 2377 – Assignment 3

Deadline: [Thursday Apr 08, 2021 at 3:00 pm](#)

Total = 100 marks

Please complete **all** the questions in this assignment. However, due to our limited TA resources, it is possible that not all the questions will be marked. You will not be informed beforehand which questions will be marked.

Please follow the instructions on course syllabus to submit your assignment online. **Late assignments will not be accepted.**

1. A company manufactures resistors. The mean resistance for these resistors is 1000 ohms and the standard deviation is 200 ohms. We collect 50 resistors at random from the assembly line. Approximate the probability that the mean resistance of these 50 resistors will be 1005 ohms or larger.

2. The lifetime of a 75 watts lightbulb is normally distributed with mean μ hours and variance σ^2 . From a random sample of 20 lightbulbs, the sample mean is 1014 hours and the sample standard deviation is 25 hours.

(a) Compute a 95% confidence interval for μ .

(b) Suppose that $\sigma = 25$ hours. We want to re-construct a 95% confidence interval for μ with the interval length no longer than 9 hours. Determine the required sample size.

3. The mean breaking strength of a certain type of fiber is required to be larger than 200 psi. Past experience has shown that we can assume that the breaking strength is normally distributed with standard deviation $\sigma = 4.5$ psi. A sample of 8 fibers yielded breakage at the following pressures (in psi):

210, 206, 198, 202, 201, 198, 199, 205.

(a) Formulate a null hypothesis versus an alternative hypothesis to verify that this type of fiber is acceptable.

(b) Compute a test statistic to test the hypotheses from part (a).

(c) Based on the value of the test statistic in part (a): (i) give the conclusion at $\alpha = 5\%$; (ii) give the conclusion at $\alpha = 10\%$.

(d) Suppose that a sample of 30 fibers yielded a mean breakage of $\bar{x} = 202.375$ psi. For this new study, compute a test statistic to test the hypotheses from part (a) and give the conclusion at $\alpha = 5\%$.

4. Suppose that the probability density function (p.d.f.) of the life (in weeks) of a certain part is

$$f(x) = \frac{3x^2}{(400)^3}, \quad 0 \leq x < 400.$$

- (a) Compute the probability the a certain part will fail in less than 200 weeks.
- (b) Compute the mean lifetime of a part and the standard deviation of the lifetime of a part.
- (c) Suppose that we select $n = 50$ parts at random. Approximate the probability that the average lifetime for these 50 parts will be less than 275 weeks?

5. Cloud seeding has been studied for many decades as a weather modification procedure. The rainfall in acre-feet from 20 clouds that were selected at random and seeded with silver nitrate. We are displaying the data in an increasing order below :

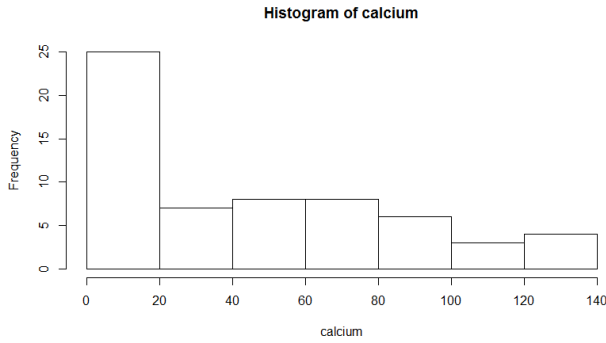
18	18.8	19.8	21.2	21.8	22.3	23.4
24.7	25	26.7	26.9	27.1	27.1	27.9
29.2	30.7	31.6	31.8	31.9	34.8	

- (a) Compute the sample mean and the sample standard deviation.
- (b) Find the first, second, and third sample quartile.
- (c) Are there any outliers in this dataset? (Explain)

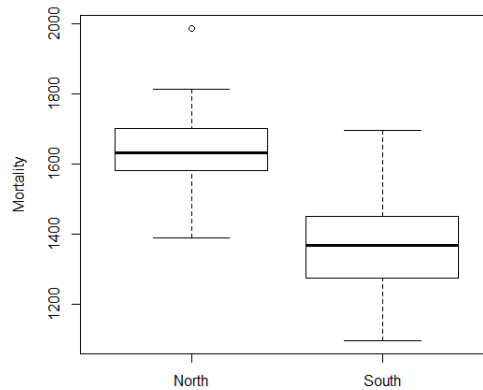
6. In an investigation of environmental causes of disease, data were collected on the annual mortality rate (deaths per 100 000) for males in 61 large towns in England and Wales. In addition, the water hardness was recorded as the calcium concentration (parts per million, ppm) in the drinking water. Below, we provide some descriptive statistics for both variables, i.e. mortality and calcium concentration.

```
> summary(calcium)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  5.00  14.00   39.00   47.18  75.00  138.00
> summary(mortality)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 1096  1379   1555   1524  1668   1987
```

Here are histograms for both variables.



- (a) For each variable (i.e. calcium concentration and mortality) describe its distribution in shape.
- (b) For each variable, are there any outliers in the sample? (Explain.)
- (c) Suppose that the towns in the sample are either South or North of Derby. To describe the mortality according to region, we produce the following side-by-side boxplots.



Answer the following questions based on the above boxplots.

- (i) Which region (North or South) has the town with the largest mortality?
- (ii) Which region (North or South) has the town with the smallest mortality?
- (iii) In terms of the central tendency, which region has a higher mortality?
- (iv) Is the mortality more dispersed in the North or the South?