MAT2377C - Assignment 5 - Solutions

Total number of points: 11

Q1. (4 points) Past experience indicates that the breaking strength of yarn used in manufacturing drapery material is normally distributed and that $\sigma = 2$ psi. A random sample of 15 specimens is tested and the average breaking strength is found to be $\bar{x} = 97.5$ psi.

(a) Find a 95% confidence interval on the true mean breaking strength.

(b) Find a 99% confidence interval on the true mean breaking strength.

Solution to Q1:

A 95% confidence interval is

$$\overline{X} \pm z_{.025}\frac{\sigma}{\sqrt{n}} = 97.5 \pm 1.96\left(\frac{2}{\sqrt{15}}\right) = 97.5 \pm 1.012140 = [96.48786, 98.51214].$$

A 99% confidence interval is

$$\overline{X} \pm z_{.005}\frac{\sigma}{\sqrt{n}} = 97.5 \pm 2.576\left(\frac{2}{\sqrt{15}}\right) = 97.5 \pm 1.330241 = [96.16976, 98.83024].$$

Marking scheme for Q1:

1 point for the correct use of confidence interval (i.e. $\overline{X} \pm z_{.025}\frac{\sigma}{\sqrt{n}}$ in the first part), 1 point for the correct answer in each part. Total - 4 points.

Q2. The diameter holes for a cable harness follow a normal distribution with $\sigma = 0.01$ inches. For sample of size 10, an average diameter is 1.5045 inch.

(a) Find a 99% confidence interval on the mean hole diameter.

(b) Repeat this for $n = 100$.

Solution to Q2:

A 99% confidence interval for $n = 10$ is

$$\overline{X} \pm z_{.005}\frac{\sigma}{\sqrt{n}} = 1.5045 \pm 1.96\left(\frac{0.01}{\sqrt{10}}\right) = 1.5045 \pm 0.008146027 = [1.496354, 1.512646].$$

A 99% confidence interval for $n = 100$ is

$$\overline{X} \pm z_{.005}\frac{\sigma}{\sqrt{n}} = 1.5045 \pm 1.96\left(\frac{0.01}{\sqrt{100}}\right) = 1.5045 \pm 0.002576 = [1.501924, 1.507076].$$

Marking scheme for Q2:

This question will not be marked.

Q3. (2 points) An article in a journal describes the effect of delamination on the natural frequency of beams made from composite laminates. The data are as follows:

$$230.66, 233.05, 232.58, 229.48, 232.58, 235.22$$

Assuming that the population is normal, find a 95% confidence interval for the mean.

Solution to Q3:

We have $n = 6$, $\bar{X} = 232.2617$, $S = 1.993935$. A 95% confidence interval is

$$\overline{X} \pm t_{.025,5}\frac{S}{\sqrt{n}} = 232.2617 \pm 2.571\left(\frac{1.993935}{\sqrt{6}}\right) = [230.169, 234.355].$$

Marking scheme for Q3:

1 point for the correct use of CI (i.e. $\overline{X} \pm t_{.005,5}\frac{S}{\sqrt{n}}$), 1 point for the correct answer. Total - 2 points.

Q4. A textile fiber manufacturer is investigating a new drapery yarn, which the company claims has a mean thread elongation of $\mu = 12$ kilograms with standard deviation of $\sigma = 0.5$ kilograms.

(a) What should be the sample size, so that with probability 0.95 we will estimate the mean thread elongation with error at most 0.15 kg?

(b) What should be the sample size, so that with probability 0.95 we will estimate the mean thread elongation with error at most 0.05 kg?

Solution to Q4:

(a)

$$\left(\frac{z_{.025}\,\sigma}{E}\right)^2 = \left(\frac{(1.96)(0.5)}{0.15}\right)^2 = 42.68.$$

Thus (round up) $n = 43$.

(b)

$$\left(\frac{z_{.025}\,\sigma}{E}\right)^2 = \left(\frac{(1.96)(0.5)}{0.05}\right)^2 = 384.16.$$

Thus $n = 385$.

Marking scheme for Q4:

This question will not be marked.

Q5. (1 point) In a sample of $n = 1000$ people, 234 want to vote for Party X. Find the 95% confidence interval for the proportion of people who are going to vote for Party X. Write the answer in a form of a newspaper add.

Solution to Q5:

The confidence interval is

$$\left(\hat{p} - 1.96 * \sqrt{\hat{p}(1-\hat{p})/n}, \hat{p} + 1.96 * \sqrt{\hat{p}(1-\hat{p})/n}\right) = (0.21, 0.26).$$

In the next election we predict that 23.5% people will vote for Party X. The error of this prediction is ±2.5%. Our prediction is valid in 19 out of 20 cases.

Marking scheme for Q5:

1 point for the correct answer. A text will not be marked.

Q6. In this example you have to use `R`. In the solution, included your code and print or write an output. Download data set `election.txt`. The data set is based on voting preferences of 1000 people. "0" represents Party X, "1" represents Party Y. Calculate 95% confidence interval for the proportion of people who are going to vote for Party Y.

Solution to Q6:

After storing the data in R under the name x, I typed

```
how.many=sum(x); n=length(x)
binom.test(how.many,n,conf.level=0.95)
```

The output is

```
data:  how.many and n
number of successes = 502, number of trials = 1000,
95 percent confidence interval:
 0.4705435 0.5334447
```

Marking scheme for Q6:

This question will not be marked.

Q7. (4 points) In this example you have to use `R`. In the solution, included your code and print or write an output

Consider the following data set:

```
170,295,200,165,140,190,195,142,138,148,110,140,103,176,125,126,204,196,98,123,124
152,177,168,175,186,140,147,174,155,195
```

(a) Calculate mean and median;

(b) Discuss normality (use the appropriate tools);

(c) Replace the value 195 with 2000. How does it affect the mean and the median?

Solution to Q7:

```
Data=c(170,295,200,165,140,190,195,142,138,148,110,140,103,176,125,126,204,196,
98,123,124,152,177,168,175,186,140,147,174,155,195);
mean(Data);median(Data);
par(mfrow=c(1,3)); boxplot(Data);hist(Data);qqnorm(Data)
```

(a) mean and median are, respectively, 160.5484, 155.

(b) We plot three graphs. Normality has to be rejected.

(c) For the new data the mean and median are, respectively, 218.7742, 155.

Marking scheme for Q7:

1 point for both correct answers in part a); 1 point for both correct answers in part c); (no partial marks, e.g. if you calculated mean correctly, but the median is wrong, then 0 points); 1 point some plots (one is enough), one point for the correct conclusion. Total - 4 points.

Q8. In this example you have to use `R`. In the solution, included your code and print or write an output. This question is not compulsory, but you can get additional points

(a) Simulate 1000 "6/49" games. Store the results. Check if the numbers are uniformly distributed.

(b) Collect data about recent "6/49" games. Store the results. Check if the numbers are uniformly distributed.

Solution to Q8:

(a)

```
N=1000;
values=NULL;
numbers=c(seq(1,49,by=1));
```

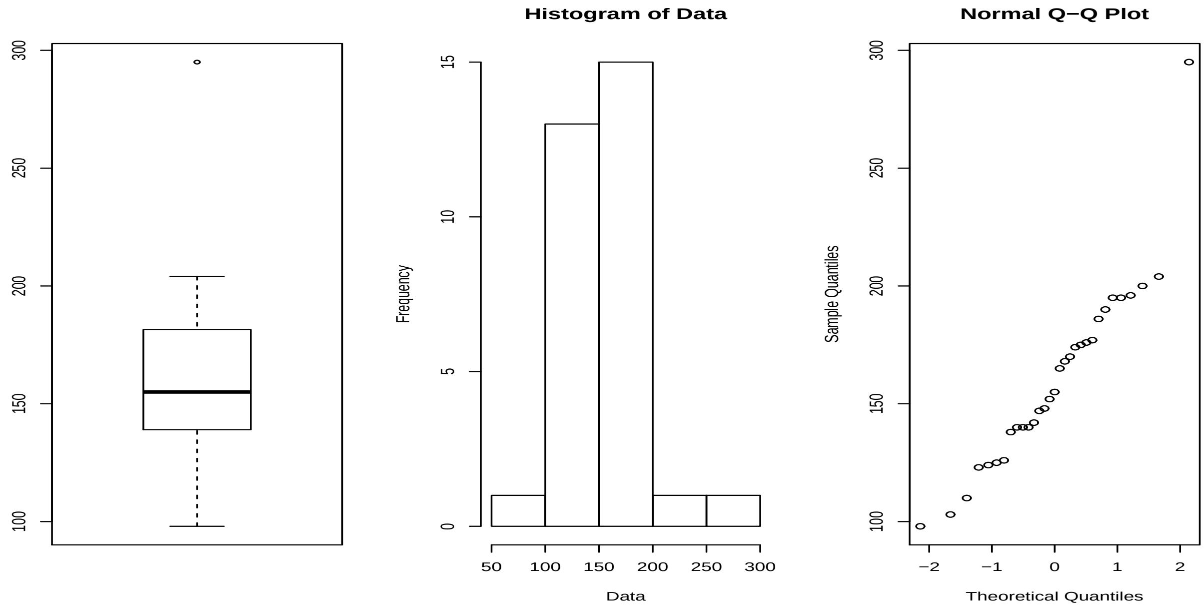

FIGURE 1. Graphs for Question 7

```
for(i in 1:N)
    {
    samples=sample(numbers,replace=FALSE,6)
    values=c(values,samples);
    }
hist(values)
```

The outcome is a vector of size $6N = 6000$ with numbers from 1 to 49. The chance of getting each number is the same, 1/49. Hence, the histogram should be more-or-less flat.

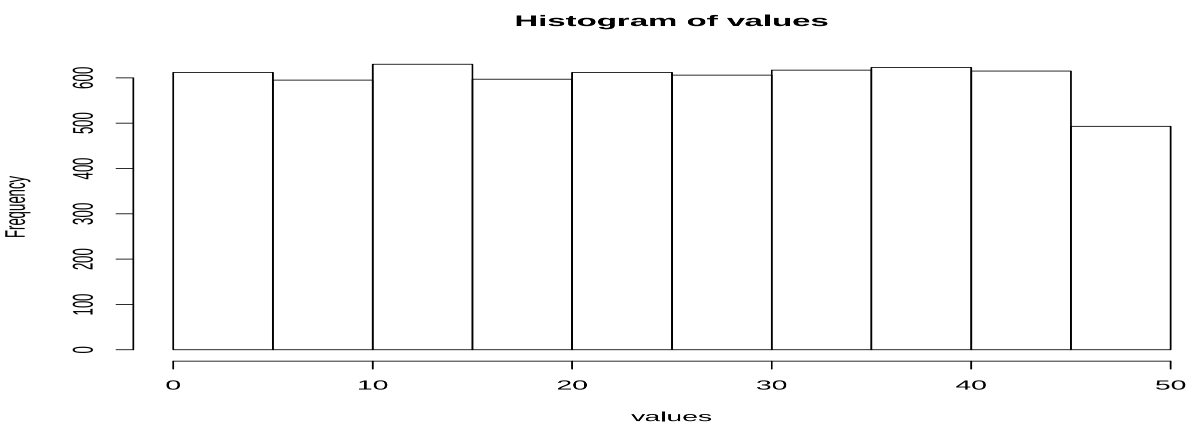

FIGURE 2. Histogram for Question 8a)

Marking scheme for Q8:

Note: this question will be marked separately. There is no "standard" answer for part b), since everyone can have different data.