MAT 2379 3X (Spring 2013)
Assignment 4 - Solutions

Part I) Answer the following 4 questions **without** the use of R.

[4] 1)

(a) We have $s/\sqrt{n} = 4.755$, thus the sample standard deviation is $s = 4.755\sqrt{15} = 18.41604$.

(b) A 95% confidence interval for the mean water hardness is

$$\overline{x} \pm t_{0.025,14} \frac{s}{\sqrt{n}} = 103.35 \pm 2.145\,(4.755) = [93.2, 113.5].$$

(c) Since the confidence interval does not fall exactly in one of the three categories, we cannot classify the mean hardness in exactly one of the three categories. We are 95% confident that the mean hardness is either soft or medium hard.

[4] 2) The population mean for the number of tumours per fish is

$$\mu = \sum x\,f(x) = 0.14,$$

and the population variance for the number of tumours per fish is

$$\sigma^2 = \left[\sum x^2\,f(x)\right] - \mu^2 = 0.24 - (0.14)^2 = 0.2204.$$

Let $\overline{X}$ be the mean number of tumours per fish among the $n = 30$ selected fish. Since the sample size is large, we know that $\overline{X}$ has an approximate normal distribution, that is

$$\overline{X} \sim N(\mu, \sigma^2/n) = N(0.14, 0.2204/30).$$

We want

$$P(\overline{X} > 0.35) \approx 1 - \Phi\left(\frac{0.35 - 0.14}{\sqrt{0.2204/30}}\right) = 1 - \Phi(2.45) = 0.0071.$$

[3] 3)

(a) A 95% confidence interval for the average MAO activity level of bulimic patients is

$$\overline{x} \pm t_{0.025,29} \frac{s}{\sqrt{30}} = 4.35 \pm (2.045)\,(2.75/\sqrt{30}) = [3.32, 5.38].$$

(b) We are 95% confident that the mean MAO levels for bulimic patients is below the normal range for the MAO levels.

[5] 4) This is a binomial experiment. The sample proportion for the effectiveness of the new drug is $\hat{p} = 119/170 = 0.7$.

(a) A 95% confidence interval for the rate of effectiveness of the new drug for treating this type of infection is

$$\hat{p} \pm z_{0.025} \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} = [0.631, 0.769],$$

where $z_{0.025} = 1.96$.

(b) Since a larger proportion of the confidence intervals will contain the true proportion, then we expect the 98% confidence interval for the rate of effectiveness, to be longer compared to the 95% confidence interval.

(c) A 98% confidence interval for the rate of effectiveness of the new drug for treating this type of infection is

$$\hat{p} \pm z_{0.01} \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} = [0.618, 0.782],$$

where $z_{0.01} = 2.326$.

Part II) Answer the following 2 questions **with** the use of R.

Remarks:

- You must provide the R commands and output that were used in answering the question.

- The R output alone is not an answer to a question. The R output is used to support your answer.
- Please do not printout your whole R session. Only provide the R commands and output that are necessary to answer the question.

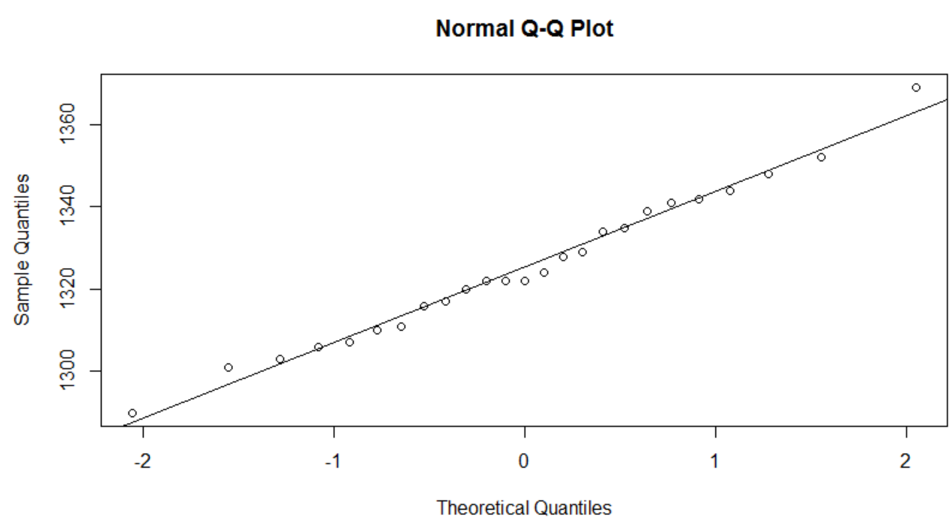
[**4**] 5) We start by importing the data into R and displaying the names of the variables. We assign the values to the variable x.

```
> qqnorm(x)
> abline(mean(x),sd(x))
```

(a) Below are the commands to produce the quantile-quantile plot for the household water use.

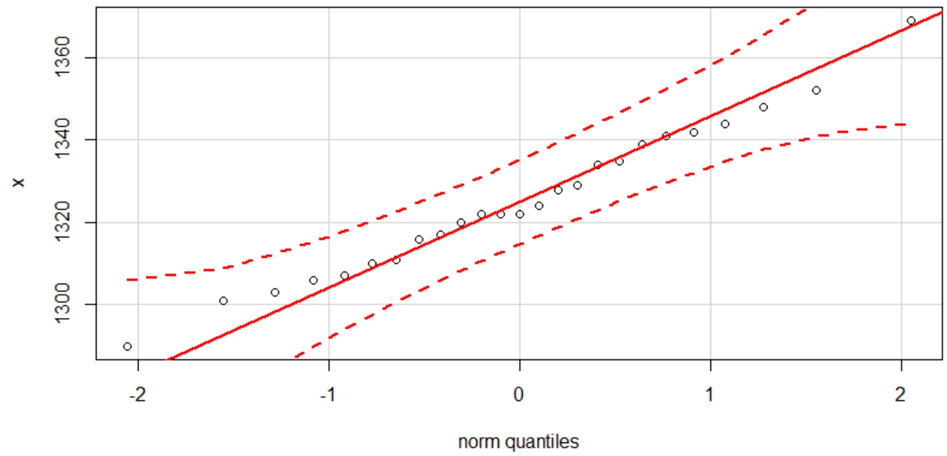
```
> table = read.table(file.choose(),header=TRUE,sep="\t")
> names(table)
[1] "water.use"
> x=table$water.use
```

Since there is a linear tendency in the qq-plot, then it is reasonable to assume that the household water use is normally distributed.



Remark: We could have used the function *qqPlot* from the *car* package to produce the plot. Here are the commands and the plot.

```
> library(car)
> qqPlot(x)
```



(b) We are 90% confident that the mean household water use per day is between 1,319.0 liters to 1,331.6 liters. Based on this confidence interval, we have evidence that the mean household water use is not 1,315 liters per day.

R commands and output:

```
> t.test(x,conf.level=0.9)$conf.int
[1] 1318.998 1331.562
attr(,"conf.level")
[1] 0.9
```

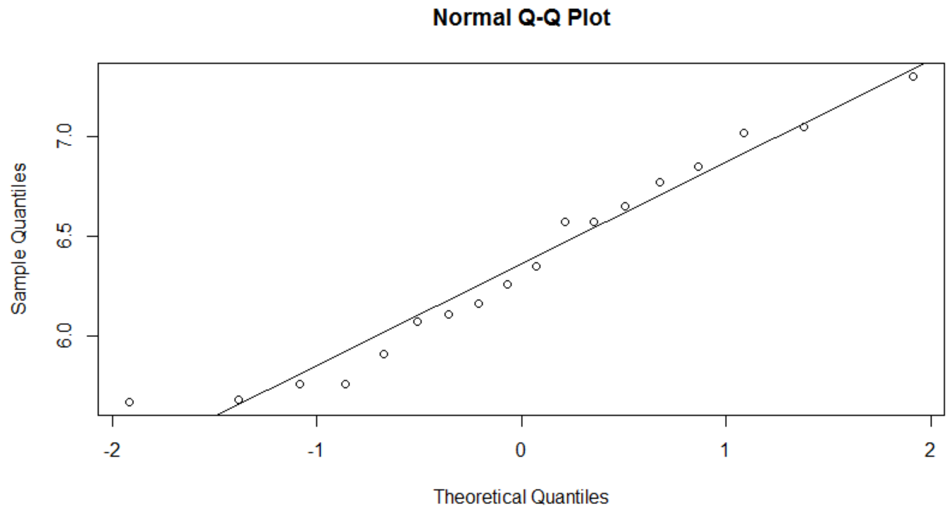
[**4**] 6) We start by importing the data into R and displaying the names of the variables. We assign the values to the variable x.

```
> table = read.table(file.choose(),header=TRUE,sep="\t")
> names(table)
[1] "pH.measurement"
> x=table$pH.measurement
```

(a) Below are the commands to produce the quantile-quantile plot for the household water use.

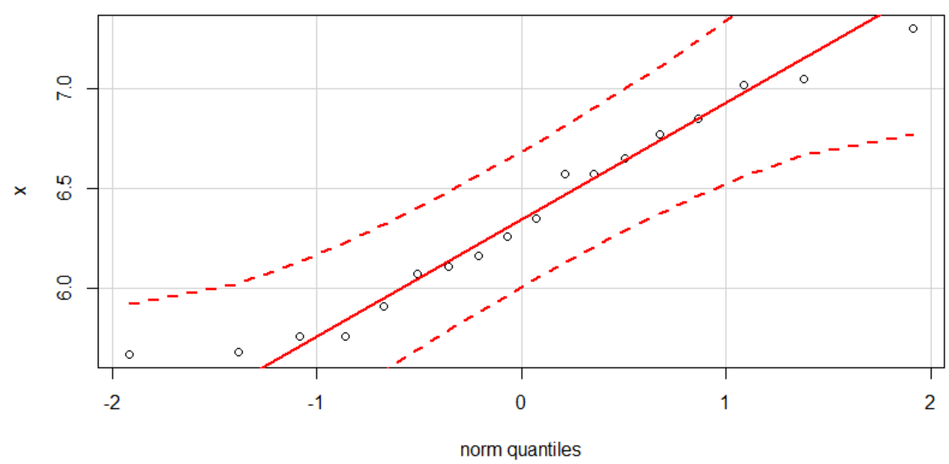
```
> qqnorm(x)
> abline(mean(x),sd(x))
```

Since there is a linear tendency in the qq-plot, then it is reasonable to assume that the household water use is normally distributed.



Remark: We could have used the function *qqPlot* from the *car* package to produce the plot. Here are the commands and the plot.

```
> library(car)
> qqPlot(x)
```



(b) We are 95% confident that the mean pH level is between 6.1 and 6.6.

R commands and output:

```
> t.test(x)$conf.int
[1] 6.107713 6.615620
attr(,"conf.level")
[1] 0.95
```

[/**24**]