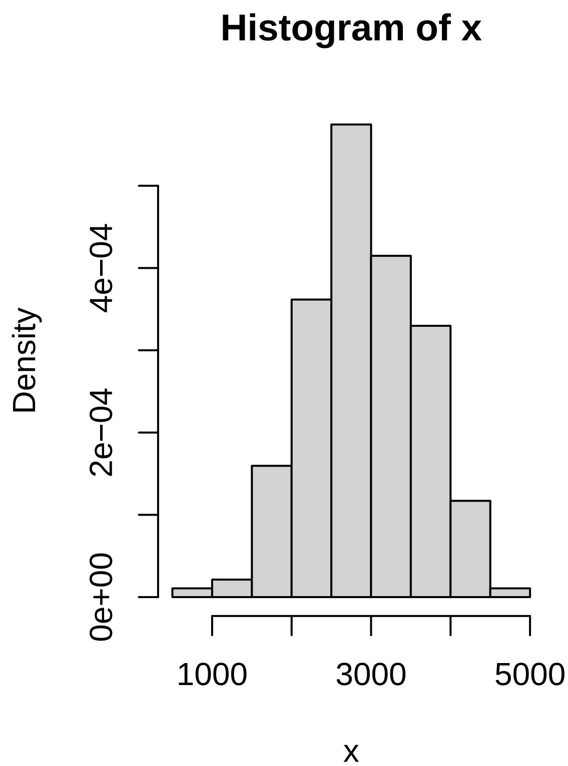
Assignment template

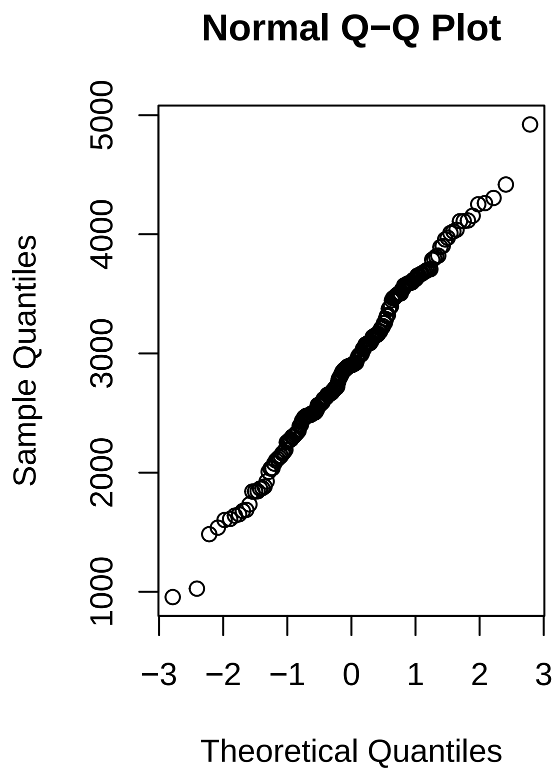
Group 45: Daniel Engbert, Rik Timmer, Koen van der Pool

11 February 2023

Exercise 1: Birthweights

```
data = read.table("birthweight.txt", header=T)
par(mfrow=c(1,2))
x = data$birthweight
# checking normality
hist(x, freq=F)
qqnorm(x)
```






```
shapiro.test(x)
```

```
##
##  Shapiro-Wilk normality test
```

```
##
## data:  x
## W = 0.99595, p-value = 0.8995
```

```
res = t.test(x, conf.level=0.96)
res
```

```
##
##  One Sample t-test
##
## data:  x
## t = 57.269, df = 187, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 96 percent confidence interval:
##  2808.084 3018.501
## sample estimates:
## mean of x
##  2913.293
```

```
sd(x)
```

```
## [1] 697.5002
```

```
tval = qt(0.98, length(x))
m = 50
n = (tval * sd(x) / m) ** 2
n
```

```
## [1] 832.3197
```

```
#sprintf("96% confidence interval: [%.4f, %.4f]", res$conf.int[1], res$conf.int[2])
```

a)

The birthweight data appears to be normal based on a normal-appearing histogram, the straight line in the qqplot, and the shapiro-wilk normality test (having \$p=0.8995 > 0.05\$).

$$m = \frac{t * s}{\sqrt{n}} \rightarrow n = (\frac{t * s}{m})^2$$

The equation above, gives the sample size, \$n\$, needed for the confidence interval to have a width of 100 (meaning \$m=50\$), where t is the t-score for the quantile 0.02 (such that both tails of the distribution have total area 0.04). Our calculations compute \$n = 832.32\$, which rounded indicates the min sample size is $n = 833$ (for a 96% CI).

```
B = 10000
Tstar = numeric(B)

for (i in 1:B){
  # sample with replacement, for a new sample of same length(x)
  Xstar = sample(x, replace=TRUE)
  Tstar[i] = mean(Xstar)
```

```
}

Tstar02 = quantile(Tstar, 0.02)
Tstar04 = quantile(Tstar, 0.98)
Tstar02
```

```
##       2%
## 2808.209
```

```
Tstar04
```

```
##      98%
## 3016.634
```

Our bootstrap CI of 96% is $[2810.32, 3014.50]$ which is approximately consistent with the CI calculated previously (as expected).

b)

Use wilcoxon sign test?

```
t.test(x, mu=2800, alt="g")[[4]]
```

```
## [1] 2829.202      Inf
## attr(,"conf.level")
## [1] 0.95
```

The output of the confidence interval here denotes that we can say with 95% confidence that the true mean is between 2829.2015472, ∞

c)

Exercise 2: Cholesterol

```
##    Before After8weeks
## 1    6.42        5.75
## 2    6.76        6.13
## 3    6.56        5.71
## 4    4.80        4.15
## 5    8.43        7.67
## 6    7.49        7.05
## 7    8.05        7.10
## 8    5.05        4.67
## 9    5.77        5.33
## 10   3.91        3.66
## 11   6.77        5.96
## 12   6.44        5.64
## 13   6.17        5.51
## 14   7.67        6.96
## 15   7.34        6.82
## 16   6.85        6.29
## 17   5.13        4.45
```

```
## 18   5.73        5.17
```

Exercise 3: Diet

```r
data = read.table("diet.txt", header=T)
data
```

```
##    person gender age height preweight diet weight6weeks
## 1       1      0  22    159        58    1         54.2
## 2       2      0  46    192        60    1         54.0
## 3       3      0  55    170        64    1         63.3
## 4       4      0  33    171        64    1         61.1
## 5       5      0  50    170        65    1         62.2
## 6       6      0  50    201        66    1         64.0
## 7       7      0  37    174        67    1         65.0
## 8       8      0  28    176        69    1         60.5
## 9       9      0  28    165        70    1         68.1
## 10     10      0  45    165        70    1         66.9
## 11     11      0  60    173        72    1         70.5
## 12     12      0  48    156        72    1         69.0
## 13     13      0  41    163        72    1         68.4
## 14     14      0  37    167        82    1         81.1
## 15     15      1  39    168        71    1         71.6
## 16     16      1  31    158        72    1         70.9
## 17     17      1  40    173        74    1         69.5
## 18     18      1  50    160        78    1         73.9
## 19     19      1  43    162        80    1         71.0
## 20     20      1  25    165        80    1         77.6
## 21     21      1  52    177        83    1         79.1
## 22     22      1  42    166        85    1         81.5
## 23     23      1  39    166        87    1         81.9
## 24     24      1  40    190        88    1         84.5
## 25     25     NA  41    171        60    2         60.0
## 26     26     NA  32    174       103    2        103.0
## 27     27      0  44    174        58    2         60.1
## 28     28      0  37    172        58    2         56.0
## 29     29      0  41    165        59    2         57.3
## 30     30      0  43    171        61    2         56.7
## 31     31      0  20    169        62    2         55.0
## 32     32      0  51    174        63    2         62.4
## 33     33      0  31    163        63    2         60.3
## 34     34      0  54    173        63    2         59.4
## 35     35      0  50    166        65    2         62.0
## 36     36      0  48    163        66    2         64.0
## 37     37      0  16    165        68    2         63.8
## 38     38      0  37    167        68    2         63.3
## 39     39      0  30    161        76    2         72.7
## 40     40      0  29    169        77    2         77.5
```

```
## 41     41      1  51     191        71    2          66.8
## 42     42      1  38     199        75    2          72.6
## 43     43      1  54     196        75    2          69.2
## 44     44      1  33     190        76    2          72.5
## 45     45      1  45     160        78    2          72.7
## 46     46      1  37     194        78    2          76.3
## 47     47      1  44     163        79    2          73.6
## 48     48      1  40     171        79    2          72.9
## 49     49      1  37     198        79    2          71.1
## 50     50      1  39     180        80    2          81.4
## 51     51      1  31     182        80    2          75.7
## 52     52      0  51     165        60    3          53.0
## 53     53      0  35     169        62    3          56.4
## 54     54      0  21     159        64    3          60.6
## 55     55      0  22     169        65    3          58.2
## 56     56      0  36     160        66    3          58.2
## 57     57      0  20     169        67    3          61.6
## 58     58      0  35     163        67    3          60.2
## 59     59      0  45     155        69    3          61.8
## 60     60      0  58     141        70    3          63.0
## 61     61      0  37     170        70    3          62.7
## 62     62      0  31     170        72    3          71.1
## 63     63      0  35     171        72    3          64.4
## 64     64      0  56     171        73    3          68.9
## 65     65      0  48     153        75    3          68.7
## 66     66      0  41     157        76    3          71.0
## 67     67      1  36     155        71    3          68.5
## 68     68      1  47     179        73    3          72.1
## 69     69      1  29     166        76    3          72.5
## 70     70      1  37     173        78    3          77.5
## 71     71      1  31     177        78    3          75.2
## 72     72      1  26     179        78    3          69.4
## 73     73      1  40     179        79    3          74.5
## 74     74      1  35     183        83    3          80.2
## 75     75      1  49     177        84    3          79.9
## 76     76      1  28     164        85    3          79.7
## 77     77      1  40     167        87    3          77.8
## 78     78      1  51     175        88    3          81.9
```

Exercise 4: Yield of Peas

```r
library(MASS)
data = npk
data
```

```
##    block N P K yield
## 1      1 0 1 1  49.5
## 2      1 1 1 0  62.8
```

```
## 3        1 0 0 0  46.8
## 4        1 1 0 1  57.0
## 5        2 1 0 0  59.8
## 6        2 1 1 1  58.5
## 7        2 0 0 1  55.5
## 8        2 0 1 0  56.0
## 9        3 0 1 0  62.8
## 10       3 1 1 1  55.8
## 11       3 1 0 0  69.5
## 12       3 0 0 1  55.0
## 13       4 1 0 0  62.0
## 14       4 1 1 1  48.8
## 15       4 0 0 1  45.5
## 16       4 0 1 0  44.2
## 17       5 1 1 0  52.0
## 18       5 0 0 0  51.5
## 19       5 1 0 1  49.8
## 20       5 0 1 1  48.8
## 21       6 1 0 1  57.2
## 22       6 1 1 0  59.0
## 23       6 0 1 1  53.2
## 24       6 0 0 0  56.0
```