Comparing Two Means

Lucy Harlow

12/11/2020

One sample test

Below are some descriptive statistics and plots for the data from a study of reaction time (milliseconds) for a sample of 20 young people on an information processing task:

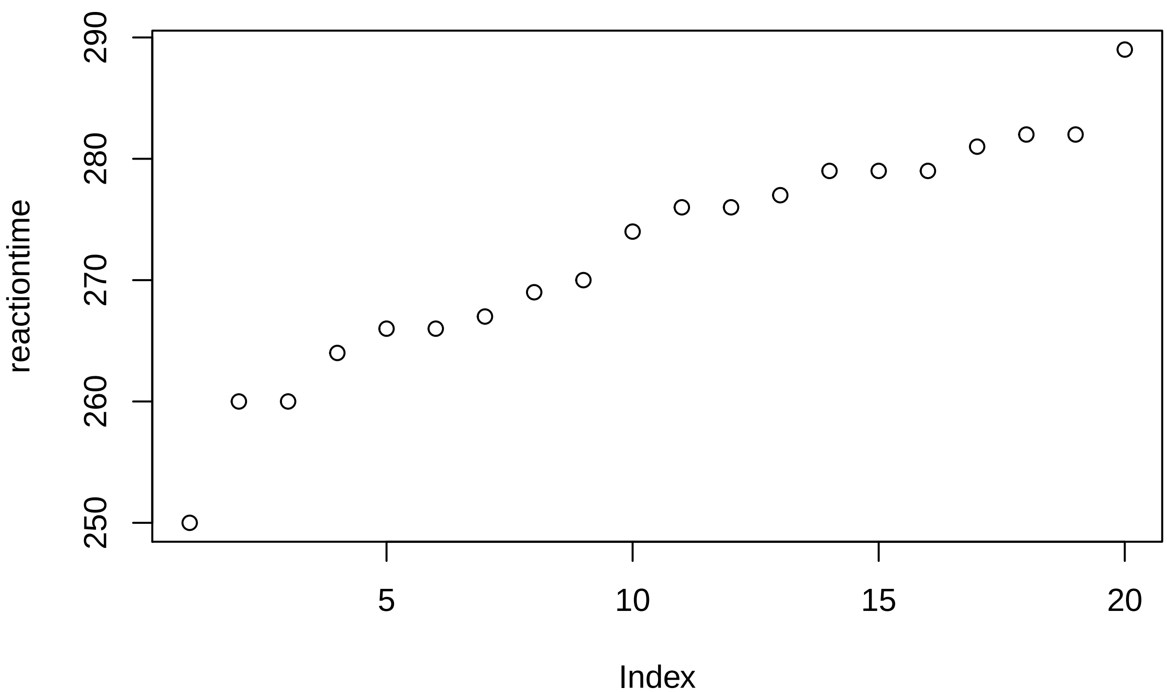
```
reactiontime <- c(250, 260, 260, 264, 266, 266, 267, 269, 270, 274, 276, 276, 277,
                  279, 279, 279, 281, 282, 282, 289)
summary(reactiontime)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   250.0   266.0   275.0   272.3   279.0   289.0
```

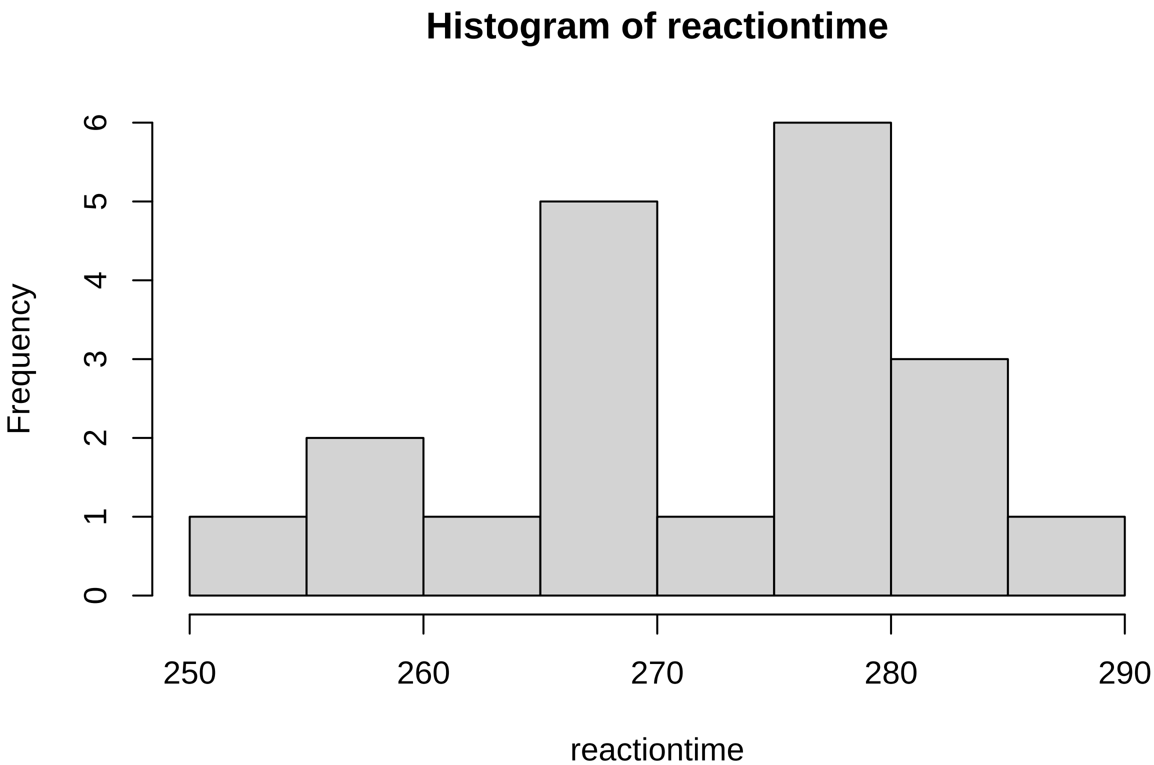
```
sd(reactiontime)
```

```
## [1] 9.520615
```

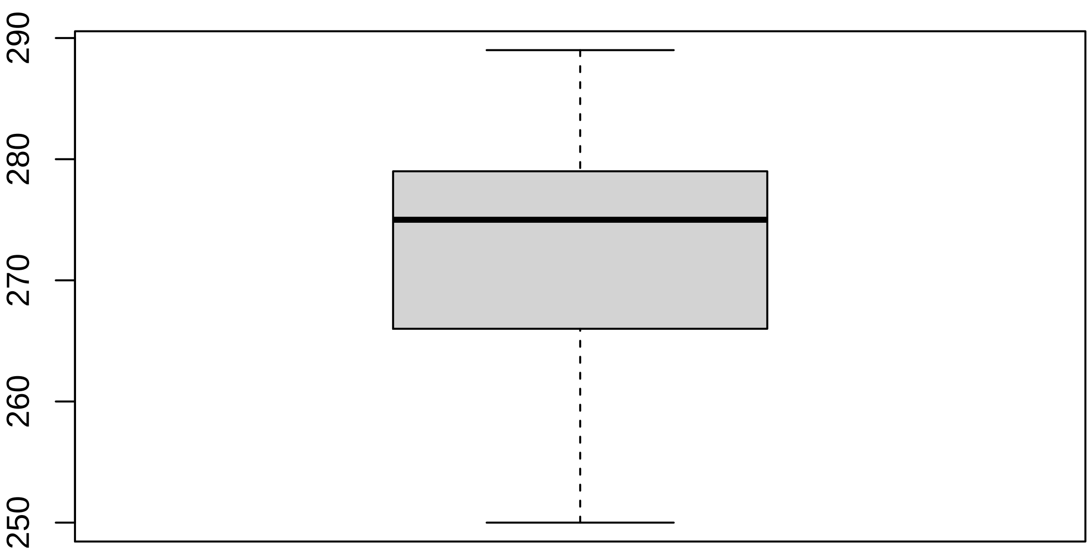
```
plot(reactiontime)
```



```r
hist(reactiontime, breaks = 7)
```




```r
boxplot(reactiontime)
```



A one-sample t-test indicated that in a sample of 20 young people's reaction times on an information processing task, reaction times ($M = 272$, $SD = 9.52$) were significantly different from the prior estimated population mean of 267.5; $t = 2.25$, $p = 0.036$.

```r
results <- t.test(reactiontime, mu = 267.5)
results
```

```
##
##  One Sample t-test
```

```
##
## data:  reactiontime
## t = 2.2547, df = 19, p-value = 0.03615
## alternative hypothesis: true mean is not equal to 267.5
## 95 percent confidence interval:
##  267.8442 276.7558
## sample estimates:
## mean of x
##     272.3
```

Independent Samples Test

Thirty-three students have taken a statistics module. In addition to the lectures, some students are given additional support by Clarence, and the rest are supported by Zebedee. The dataset with the end of module exam grades is in the data file "stats_module_grades.csv".

```r
grades <- read.csv("stats_module_grades.csv")
summary(grades)
```

```
##        ID         grade          tutor          
##  Min.   : 1   Min.   :55.00   Length:33         
##  1st Qu.: 9   1st Qu.:67.00   Class :character  
##  Median :17   Median :72.00   Mode  :character  
##  Mean   :17   Mean   :71.55                     
##  3rd Qu.:25   3rd Qu.:76.00                     
##  Max.   :33   Max.   :90.00
```

```r
grades
```

```
##    ID grade    tutor
## 1   1    65 Clarence
## 2   2    72  Zebedee
## 3   3    66  Zebedee
## 4   4    74 Clarence
## 5   5    73 Clarence
## 6   6    71  Zebedee
## 7   7    66  Zebedee
## 8   8    76  Zebedee
## 9   9    69  Zebedee
## 10 10    79  Zebedee
## 11 11    73  Zebedee
## 12 12    62  Zebedee
## 13 13    83 Clarence
## 14 14    76 Clarence
## 15 15    69  Zebedee
## 16 16    68  Zebedee
## 17 17    65 Clarence
## 18 18    86 Clarence
## 19 19    60  Zebedee
## 20 20    70 Clarence
## 21 21    80 Clarence
## 22 22    73  Zebedee
## 23 23    68  Zebedee
## 24 24    55 Clarence
## 25 25    67  Zebedee
## 26 26    78 Clarence
```

```
## 27 27    78 Clarence
## 28 28    90 Clarence
## 29 29    77 Clarence
## 30 30    74  Zebedee
## 31 31    56  Zebedee
## 32 32    68 Clarence
## 33 33    74  Zebedee
```

You are interested in whether Clarence or Zebedee provides better additional support.

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 \neq \mu_2$

Student's t-test:

```r
t.test(grade ~ tutor, data = grades, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  grade by tutor
## t = 2.1154, df = 31, p-value = 0.04253
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   0.1965873 10.7589683
## sample estimates:
## mean in group Clarence  mean in group Zebedee
##               74.53333               69.05556
```

Welch's t-test:

```r
t.test(grade ~ tutor, data = grades)
```

```
##
##  Welch Two Sample t-test
##
## data:  grade by tutor
## t = 2.0342, df = 23.025, p-value = 0.05361
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.09249349 11.04804904
## sample estimates:
## mean in group Clarence  mean in group Zebedee
##               74.53333               69.05556
```

Mann-Whitney U Test:

```r
wilcox.test(grade ~ tutor, data = grades)
```

```
## Warning in wilcox.test.default(x = c(65L, 74L, 73L, 83L, 76L, 65L, 86L, : cannot
## compute exact p-value with ties
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
```

```
## data:  grade by tutor
## W = 190.5, p-value = 0.04644
## alternative hypothesis: true location shift is not equal to 0
```

Paired Samples Test

In a health psychology experiment, 24 women aged 40-60 are randomly selected to receive an alcohol prevention intervention. Their past-week alcohol consumption (units of alcohol) is measured at baseline (pre-intervention) and again at four week follow-up (post-intervention).

My null hypothesis H_0 is that the intervention will have no effect, and the means of the pre-intervention data and the post-intervention data will be the same. My alternative hypothesis H_a is that the intervention will have an effect, and the means will be different.

```
alcohol <- read.csv("alcohol_experiment.csv")
t.test(alcohol$units_t1, alcohol$units_t2, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  alcohol$units_t1 and alcohol$units_t2
## t = 3.6991, df = 23, p-value = 0.001184
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.671268 5.912065
## sample estimates:
## mean of the differences
##                3.791667
```

NB: a one sample test on the *differences* between each pair of observations is exactly equivalent to a paired samples test. The mean of the differences is equal to the difference of the means.

```
t.test(alcohol$units_t1 - alcohol$units_t2)
```

```
##
##  One Sample t-test
##
## data:  alcohol$units_t1 - alcohol$units_t2
## t = 3.6991, df = 23, p-value = 0.001184
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  1.671268 5.912065
## sample estimates:
## mean of x
##  3.791667
```

A paired t-test assumes that the data are normally distributed, so I check this using both a Q-Q plot and a Shapiro-Wilk test. The null hypothesis in a Shapiro-Wilk test is that the data are normally distributed, and the high p-values of the test for each set of data indicate a failure to reject this null hypothesis:

```
shapiro.test(alcohol$units_t1)
```

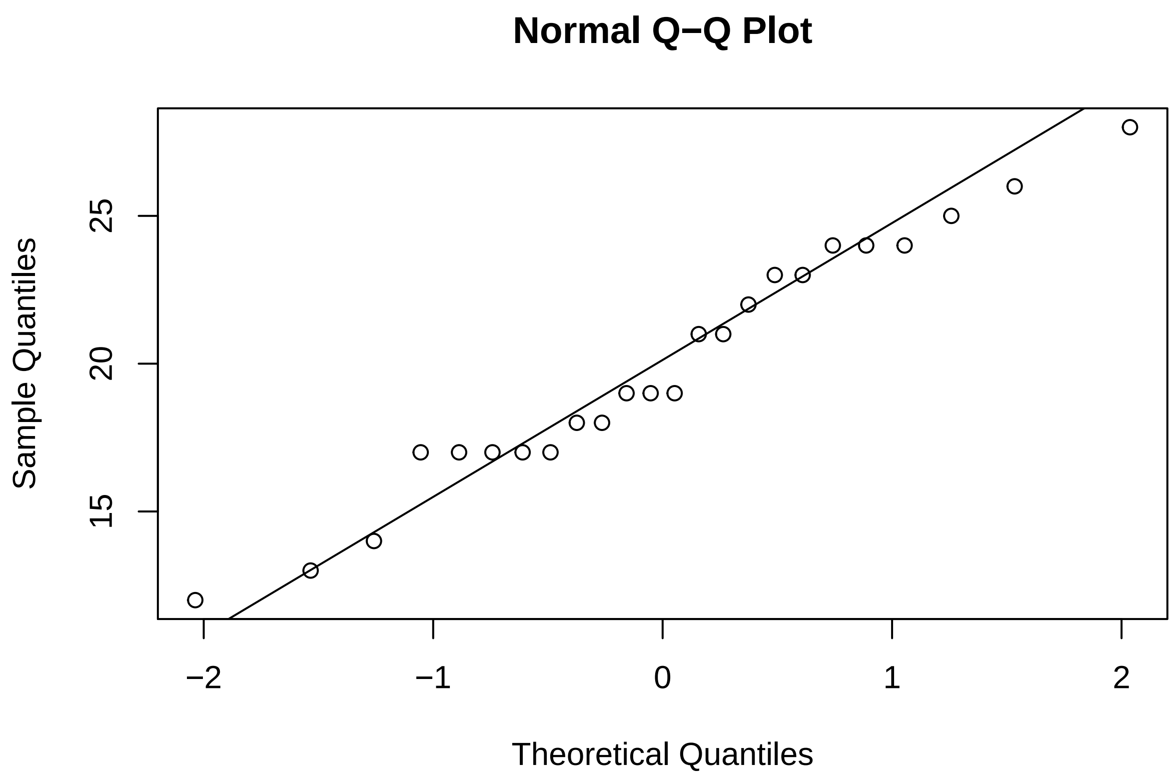
```
##
##  Shapiro-Wilk normality test
##
## data:  alcohol$units_t1
## W = 0.96911, p-value = 0.6451
```

```r
shapiro.test(alcohol$units_t2)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  alcohol$units_t2
## W = 0.98082, p-value = 0.9105
```

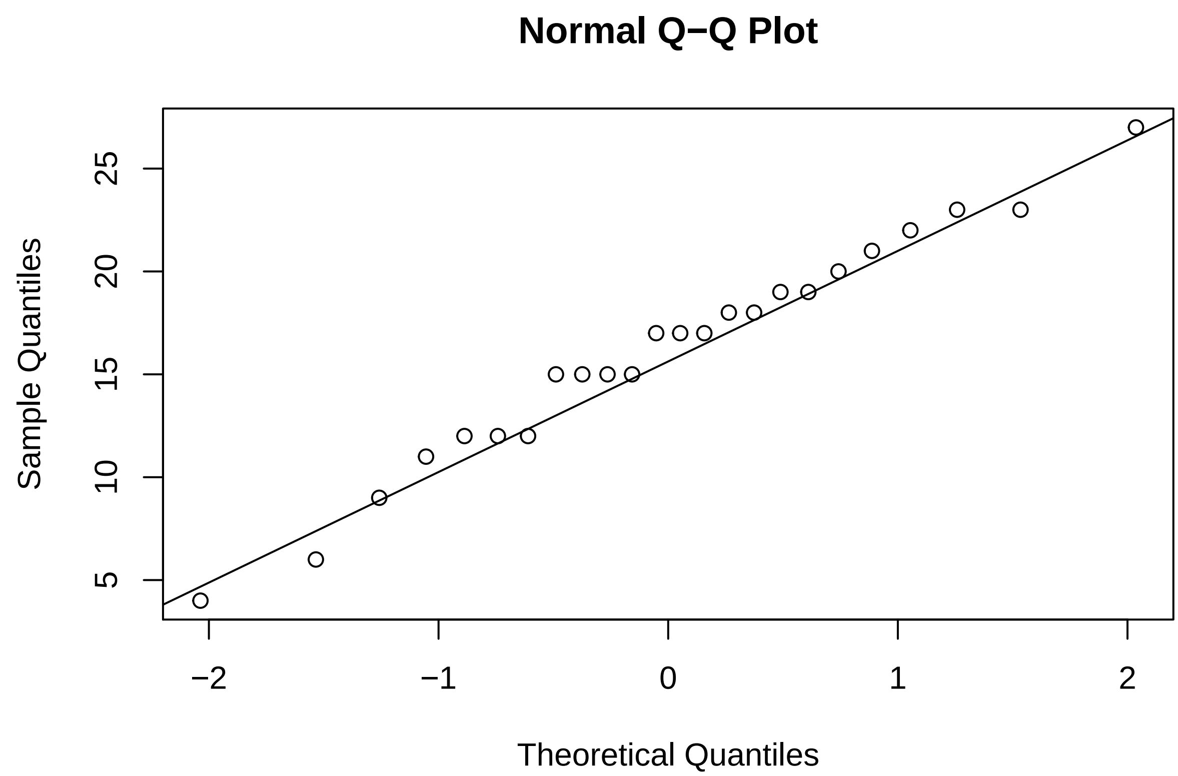
A normal Q-Q plot plots the data (Sample Quantiles) against a standard normal distribution (Theoretical Quantiles). First, I make a Q-Q plot of the pre-intervention data:

```r
qqnorm(alcohol$units_t1)
qqline(alcohol$units_t1)
```



Next, I make a Q-Q plot of the post-intervention data:

```r
qqnorm(alcohol$units_t2)
qqline(alcohol$units_t2)
```



Along with the Shapiro-Wilk test, this suggests that the data are normally distributed.