Statistics 101B HW2

Earle Aguilar
804501476
01-20-2017

2.10

a.

$$\frac{11.5}{SE} = 1.88$$
$$SE = 6.12$$

b. This is a two-sided t-test.
c. We fail to reject the null hypothesis because

$$0.0723 > 0.05$$

.
d. We know:

$$t_{0.975,24} = 2.063$$

. Confidence interval:

$$11.5 \pm 2.063 * 6.12 = [-1.13, 24.13]$$

2.15

a. We can reject the null hypothesis

$$0.001 < 0.05$$

b. This is a two-sided t-test.

$$2 * P(t \leq -3.47) = 0.001$$

c.

$$t = \frac{(50.19 - 52.52) - 2}{2.1277\sqrt{\frac{1}{20} + \frac{1}{20}}}$$
$$t = -6.44$$

Rejection regions:

$$t > 2.02 \quad or \quad t < -2.02$$

We reject the null since

$$-6.44 < -2.02$$

d. We reject the null. We only have to check the lower t value -2.02.
e.

$$Confidence \quad Interval: \quad -2.33 \pm 2.02 * 2.13\sqrt{\frac{1}{20} + \frac{1}{20}}$$
$$Confidence \quad Interval: \quad [-3.69, -0.97]$$

f.

$$p - value \ = 7.105286 * 10^{-08}$$

2.24

a.

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 \neq 0$$

b.

```
machine1 <- c(16.03, 16.01, 16.04, 15.96, 16.05, 15.98, 16.05, 16.02, 16.02, 15.99)

machine2 <- c(16.02, 16.03, 15.97, 16.04, 15.96, 16.02, 16.01, 16.01, 15.99, 16.00)

n <- 10

mean_m1 <- mean(machine1)
mean_m2 <- mean(machine2)

z <- (mean_m1 - mean_m2) / sqrt( (0.015)^2/n + (0.018)^2/n )
z
```

```
## [1] 1.349627
```

Rejection regions:

$$z > 1.96 \quad or \quad z < -1.96$$

We fail to reject the null since

$$1.35 < 1.96$$

c.

```
pnorm(1.35)
```

```
## [1] 0.911492
```

d.

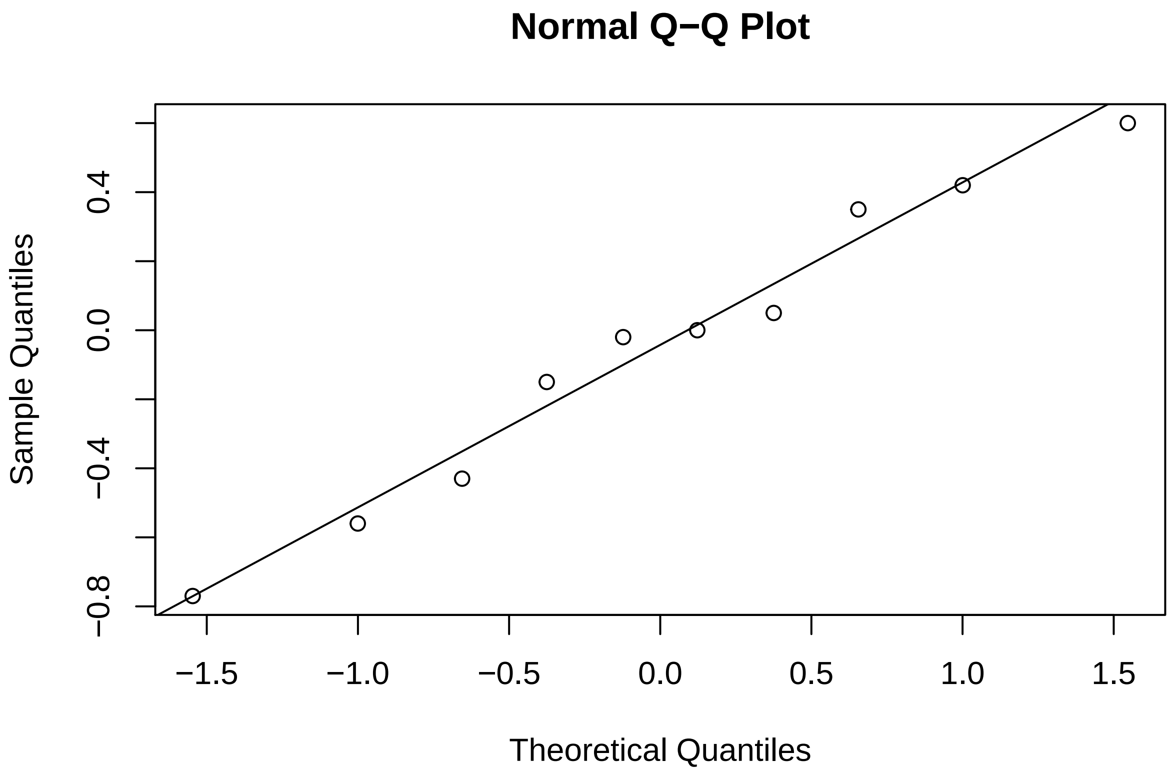
```
lb <- 0.01 - 1.96 * sqrt(((0.015)^2 + (0.018)^2 ) / 10)
ub <- 0.01 + 1.96 * sqrt(((0.015)^2 + (0.018)^2 ) / 10)
c(lb,ub)
```

```
## [1] -0.004522529  0.024522529
```

2.33

a.

```
qqnorm(df$Birth_order1-df$Birth_order2)
qqline(df$Birth_order1-df$Birth_order2)
```



The normal probability plot shows that the distributions of the differences are approximately normal.

b.

```
mean_b1 <- mean(df$Birth_order1)
mean_b2 <- mean(df$Birth_order2)

var_b1 <- var(df$Birth_order1)
var_b2 <- var(df$Birth_order2)

s_pooled <- ( 9*var_b1 + 9*var_b2 ) / (2*10-2)

lb <- (mean_b1-mean_b2) - qt(0.975, 18)*sqrt(s_pooled*(2/10))
up <- (mean_b1-mean_b2) + qt(0.975, 18)*sqrt(s_pooled*(2/10))
c(lb,ub)
```

```
## [1] -0.93349202  0.02452253
```

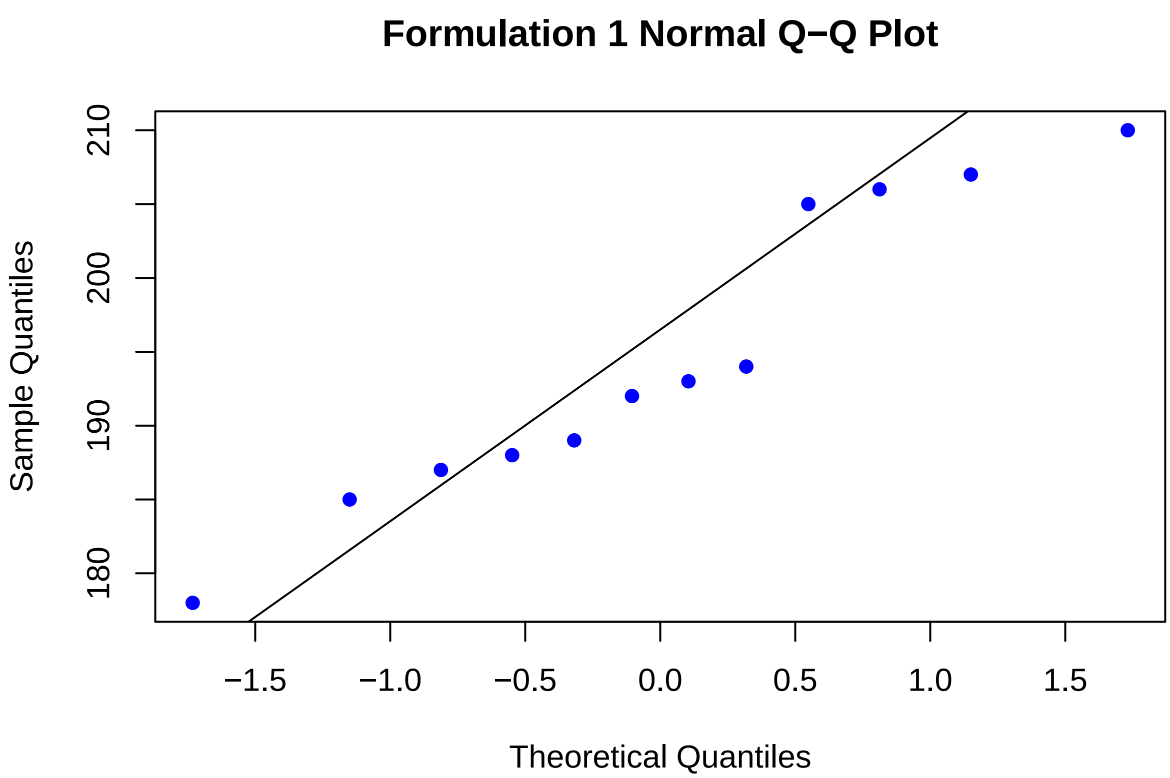
c. 0 is in the 95% confidence interval therefore we cannot reject the null hypothesis that the mean score does not depend on birth order.

2.35

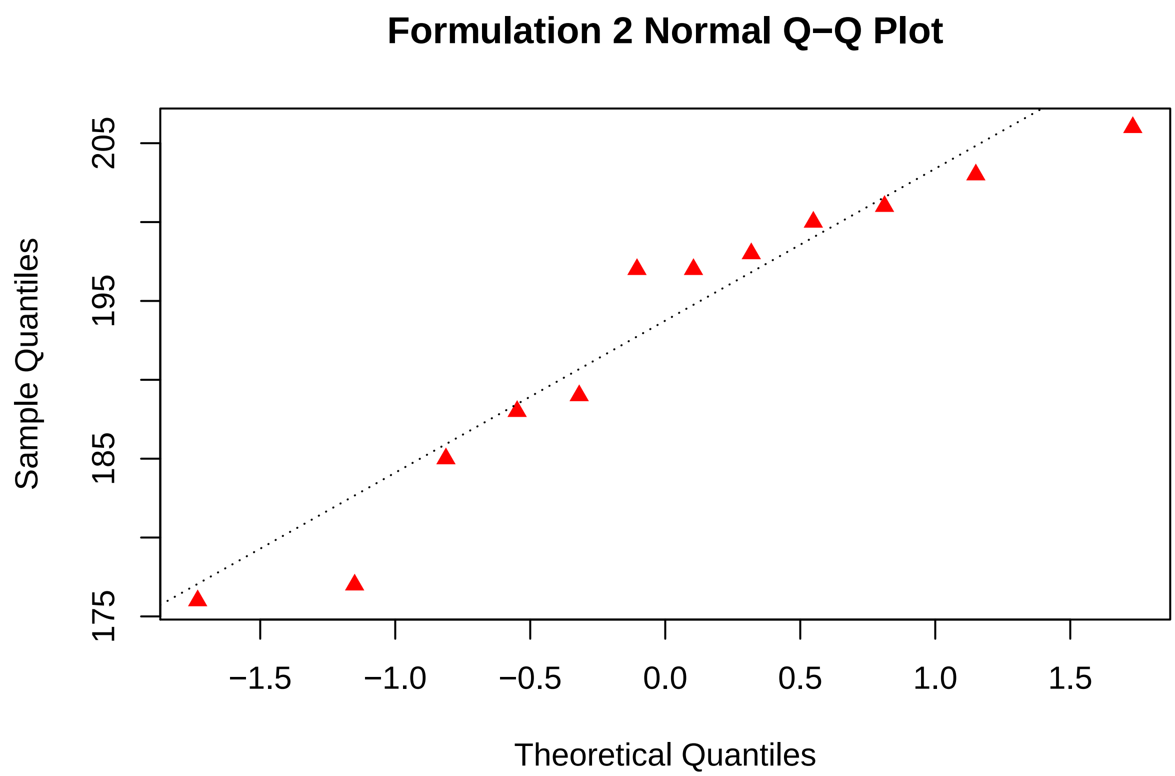
```r
form1 <- c(206, 193, 192, 188, 207, 210, 205, 185, 194, 187, 189, 178)
form2 <- c(177, 176, 198, 197, 185, 188, 206, 200, 189, 201, 197, 203)

mean_f1 <- mean(form1)
mean_f2 <- mean(form2)

var_f1 <- var(form1)
var_f2 <- var(form2)
n <- 12
```

a. Observing all three plots we can see that the assumptions of normality and equal variance are incorrect.

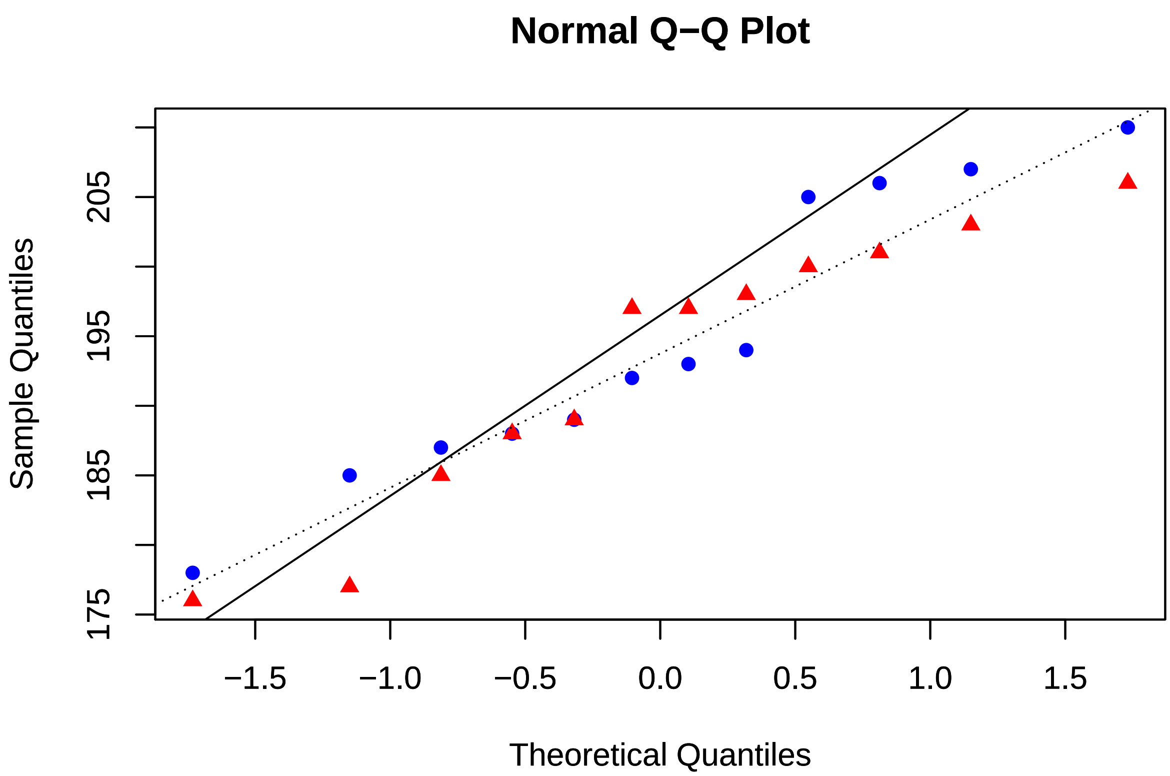
```r
qqnorm(form1, pch = 16, col='blue', main = "Formulation 1 Normal Q-Q Plot")
qqline(form1, lty=1)
```




```r
qqnorm(form2, pch = 17, col='red', main = "Formulation 2 Normal Q-Q Plot")
qqline(form2, lty=3)
```




```
vlmts=range(form1,form2)
qqnorm(form1, ylim=vlmts, pch = 16, col='blue')
qqline(form1, lty=1)
par(new=T)
qqnorm(form2, ylim=vlmts, pch = 17, col='red')
qqline(form2, lty=3)
```



b. The data do not support the claim that the mean deflection temperature under load for formulation 1 exceeds that of formulation 2.

```
a <- ((var_f1+var_f2)/n)^2
b <- (var_f1/n)^2/(n-1) + (var_f2/n)^2/(n-1)
degrees <- a/b

t <- (mean_f1-mean_f2)/sqrt((var_f1+var_f2)/n )
t
```

```
## [1] 0.3448301
```

Rejection Region:

```
qt(0.975, 22)
```

```
## [1] 2.073873
```

since $0.345 < 2.07$ we fail to reject the null. **c.**

```
pt(t, 22)
```

```
## [1] 0.6332515
```

2.47

```
machine1 <- c(16.03, 16.01, 16.04, 15.96, 16.05, 15.98, 16.05, 16.02, 16.02, 15.99)
machine2 <- c(16.02, 16.03, 15.97, 16.04, 15.96, 16.02, 16.01, 16.01, 15.99, 16.00)

n <- 10

mean_m1 <- mean(machine1)
mean_m2 <- mean(machine2)

var_m1 <- var(machine1)
var_m2 <- var(machine2)
```

a.

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 \neq 0$$

b.

```
s_pooled <- ( (n-1)*var_m1 + (n-1)*var_m2 )  / (2*n-2)
t <-  (mean_m1 - mean_m2) / sqrt(s_pooled*(2/n))
```

c.

```
pt(t, 2*n-2)
```

```
## [1] 0.7826281
```

d.

```
d <- (mean_m1 - mean_m2)
lb <- d - qt(0.975, 2*n-2)*sqrt(s_pooled*(2/n))
ub <- d + qt(0.975, 2*n-2)*sqrt(s_pooled*(2/n))
c(lb, ub)
```

```
## [1] -0.01629652  0.03629652
```