Day16

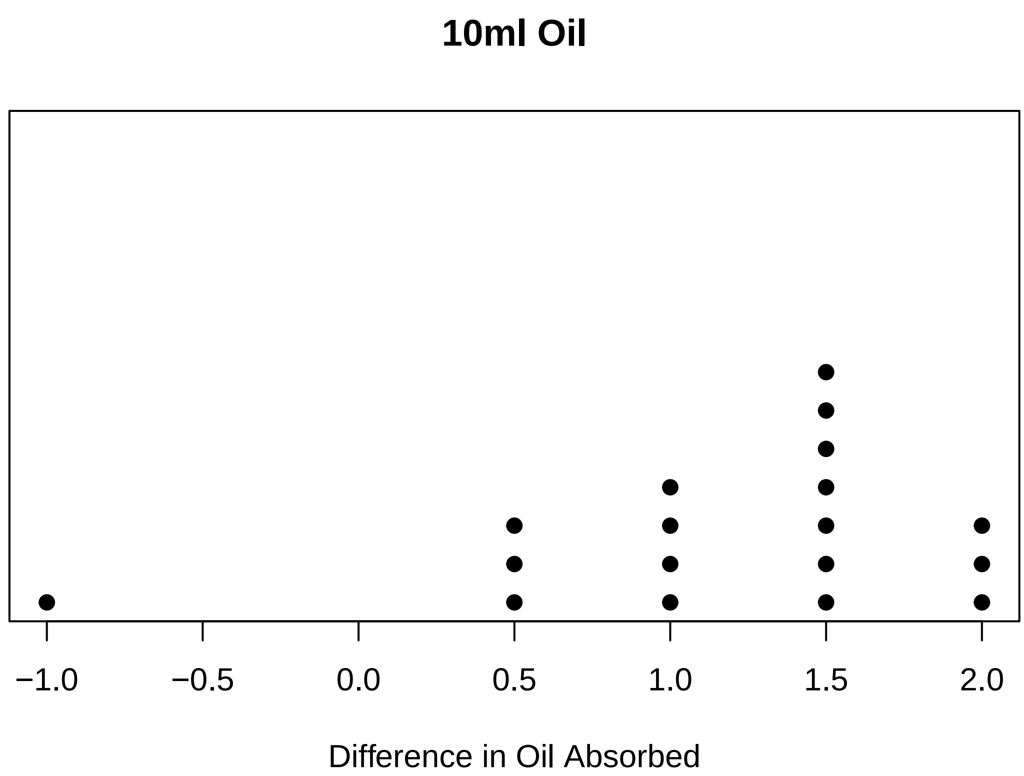
Olivia Wu

2024-04-25

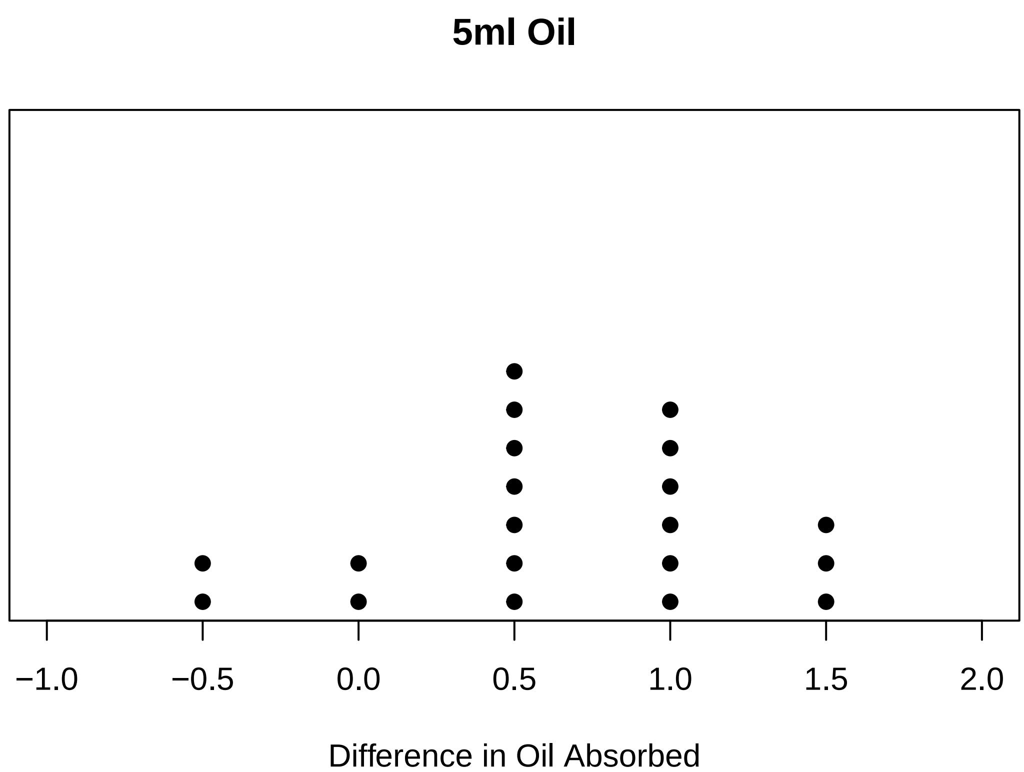
Problem 5.53

a) The differences looke normally distributed, and the standard deviations are close enough (0.847 vs 0.587). The mean of the 10ml oil is about twice as much as the mean of the 5ml oil group (1.325 vs 0.65), and I think there is a difference.

```
OilYes <- OilD[OilD$Oil %in% 10,]
OilNo <- OilD[OilD$Oil %in% 5,]
stripchart(OilYes$diff,method="stack",pch=19,at=0, offset=1, main="10ml Oil", xlab="Difference in Oil Al
```




```
stripchart(OilNo$diff,method="stack",pch=19,at=0, offset=1, main="5ml Oil", xlab="Difference in Oil Abso
```




```r
favstats(OilYes$diff)
```

```
##  min Q1 median    Q3 max  mean       sd  n missing
##   -1  1    1.5 1.625   3 1.325 0.847209 20       0
```

```r
favstats(OilNo$diff)
```

```
##   min  Q1 median Q3 max mean        sd  n missing
##  -0.5 0.5    0.5  1 1.5 0.65 0.5871429 20       0
```

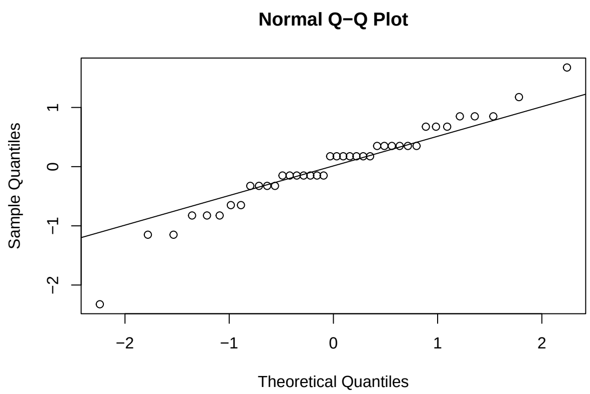
b) The t-test shows a signfiicant difference in group means because the p-value $= 0.006 < 0.05$. 0 is not included in the 95% confidence interval.

```r
t.test(OilYes$diff, OilNo$diff)
```

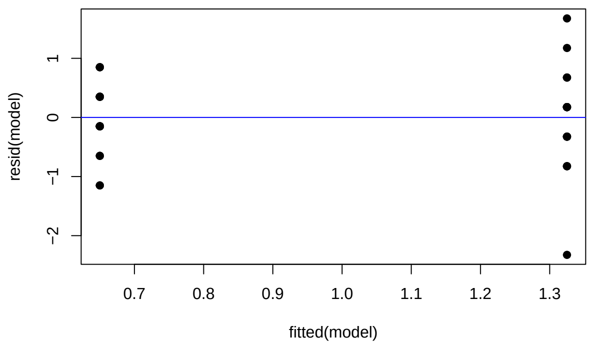
```
##
##  Welch Two Sample t-test
##
## data:  OilYes$diff and OilNo$diff
## t = 2.9286, df = 33.83, p-value = 0.006055
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.2065041 1.1434959
## sample estimates:
## mean of x mean of y
##     1.325     0.650
```

c) The normal probability plot is roughly linear so normality is met. Variance is roughly equal. We saw earlier that Since $p < 0.05$, we reject the null hypothesis. There is a significant difference in mean differences between the 5ml oil group and 10ml oil group.

```
model <- aov(diff~Oil, data=OilD)
qqnorm(model$residuals)
qqline(model$residuals)
```




```
plot(resid(model)~fitted(model),pch=19)
abline(0,0,col="blue")
```




```
summary(model)
```

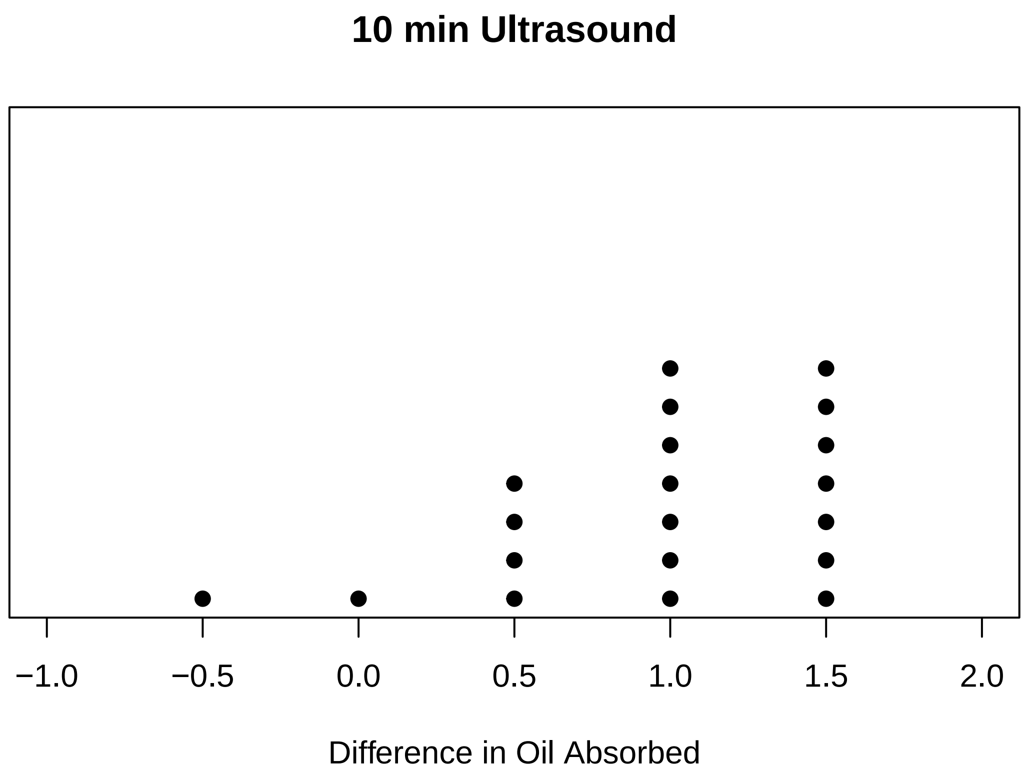
```
##             Df Sum Sq Mean Sq F value  Pr(>F)
## Oil          1  4.556   4.556   8.576 0.00573 **
## Residuals   38 20.188   0.531
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

d) The conclusions are the same. Running an ANOVA for two groups produces the same result as running a two sample t-test.

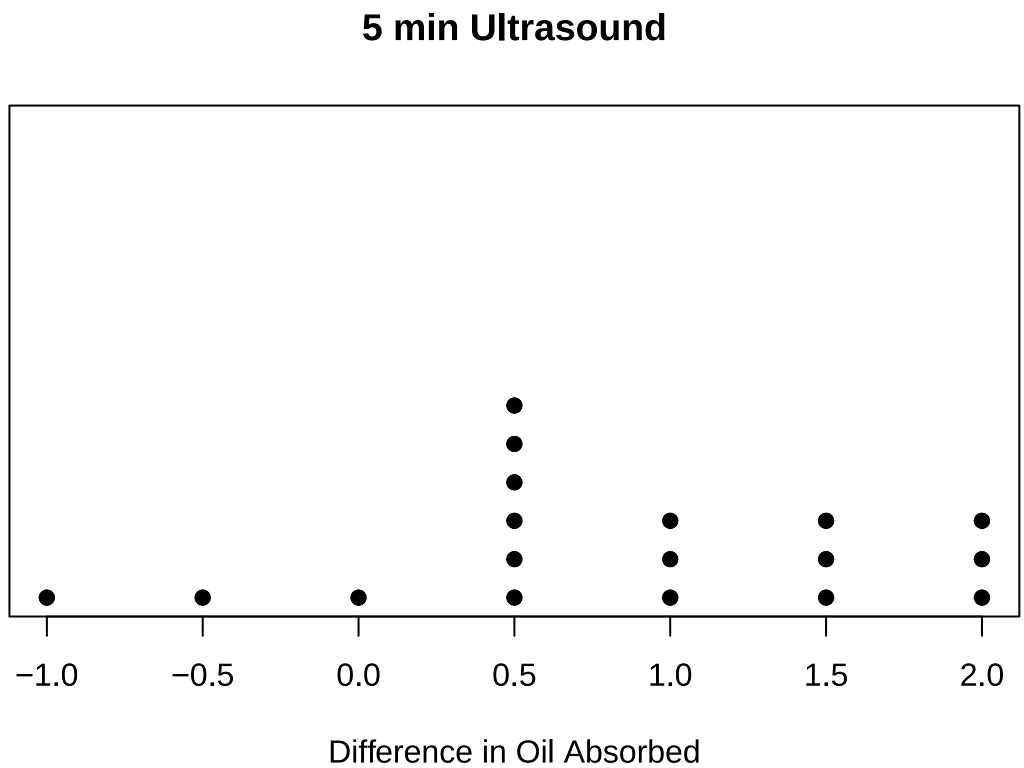
Problem 5.55

a) Neither dotplot looks normal; the 10 minute is skewed left. The means are similar (1.025 vs. 0.95) and there is no large difference. The standard deviations are close enough (0.55 vs 0.99)

```
UltY <- OilD[OilD$Ultra %in% 10,]
UltN <- OilD[OilD$Ultra %in% 5,]
stripchart(UltY$diff,method="stack",pch=19,at=0, offset=1, main="10 min Ultrasound", xlab="Difference ir
```




```
stripchart(UltN$diff,method="stack",pch=19,at=0, offset=1, main="5 min Ultrasound", xlab="Difference in
```




```r
favstats(UltY$diff)
```

```
##   min  Q1 median  Q3 max mean        sd  n missing
##  -0.5 0.5      1 1.5 1.5 0.95 0.5596051 20       0
```

```r
favstats(UltN$diff)
```

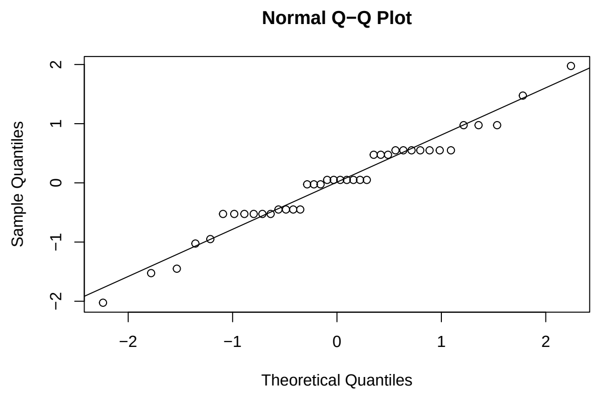
```
##  min  Q1 median    Q3 max  mean        sd  n missing
##   -1 0.5      1 1.625   3 1.025 0.9930681 20       0
```

b) Since $p = 0.77 > 0.05$, we fail to reject the null hypothesis. There is not enough evidence to suggest that the mean differences in oil absorption is different in the 10 min ultrasound and 5 min ultrasound groups.

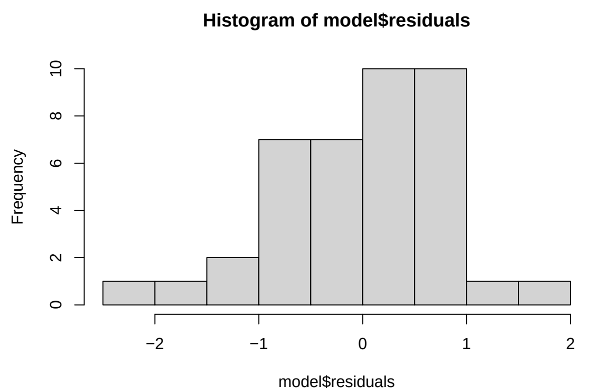
```
##
##  Welch Two Sample t-test
##
## data:  UltY$diff and UltN$diff
## t = -0.29425, df = 29.961, p-value = 0.7706
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.5955757  0.4455757
## sample estimates:
## mean of x mean of y
##     0.950     1.025
```

c) The normal probability plot is roughly linear so the normal condition is met, however the dotplot shows that the variance is not constant. The anova gave $p = 0.77 > 0.05$, so the conclusion is the same as before.

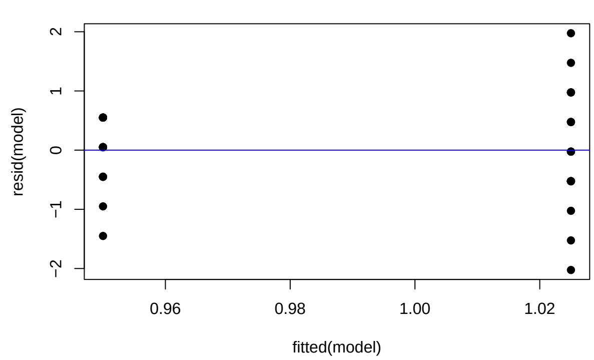
```
model <- aov(diff~Ultra, data=OilD)
qqnorm(model$residuals)
qqline(model$residuals)
```




```
hist(model$residuals)
```




```
plot(resid(model)~fitted(model),pch=19)
abline(0,0,col="blue")
```



```r
summary(model)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Ultra        1  0.056  0.0563   0.087   0.77
## Residuals   38 24.688  0.6497
```

d) The conclusions are the same.

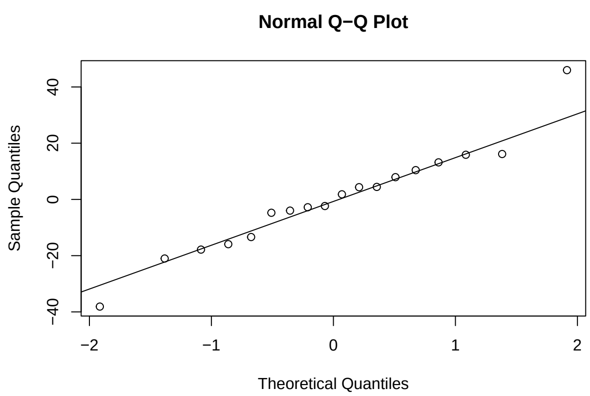
Problem 5.57

a) $H_0 : \mu_1 = \mu_2 = \mu_3$ where μ_i is the average failure load for method i.

$H_a: $at least one of these means is different from the others

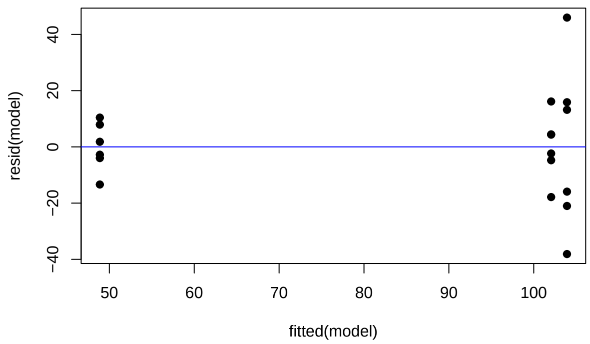
b) The residuals are roughly normally distributed, but the variance is not constant. The standard deviation for the third method is over three times as much as the second method.

```r
model <- aov(FailureLoad~Method, data=Meniscus)
qqnorm(model$residuals)
qqline(model$residuals)
```




```r
hist(model$residuals, breaks=10)
```

```r
plot(resid(model)~fitted(model),pch=19)
abline(0,0,col="blue")
```




```r
sd(Meniscus$FailureLoad[Meniscus$Method==1])
```

```
## [1] 11.36886
```

```r
sd(Meniscus$FailureLoad[Meniscus$Method==2])
```

```
## [1] 8.685716
```

```r
sd(Meniscus$FailureLoad[Meniscus$Method==3])
```

```
## [1] 30.62178
```