

Exercise 4

Will Wright

2/6/2019

Applied Statistics and Design of Experiments Exercise 4 February 6, 2018

For these exercises you should work in groups of 3-5 students. Part of homework 5 will involve working with your group to refine and extend your answers to some of these questions.

Data set: 'defects.csv'

The data are from an experiment to compare 4 processing methods for manufacturing steel ball bearings. The 4 process methods were run for one day and a random sample of 1% of the ball bearings from the day was taken from each of the 4 methods. Because the processes produce ball bearings at different rates the sample sizes were not the same for the 4 methods. Each sampled ball bearing had its weight measured to the nearest 0.1 g and the number of surface defects was counted. The variables in the data set are:

Sample: sample number Method: A, B, C, or D Defects: number of defects Weight: weight in g

Question 1.

The target weight for the ball bearings is 10 g. For each of the 4 methods test the null hypothesis that the mean weight is equal to 10. Considering the possibility of inflation of type I error rate due to multiple testing, what do you conclude from these results?

Perform all pairwise comparisons of mean weight for the different methods. Report the p-values from all tests. Considering the possibility of inflation of type I error rate due to multiple testing, what conclusions would you draw from these results?

```
t.test(defectsData$Weight[which(defectsData$Method=="A")], alternative = "two.sided", mu = 10)

##
## One Sample t-test
##
## data: defectsData$Weight[which(defectsData$Method == "A")]
## t = 0.079652, df = 73, p-value = 0.9367
## alternative hypothesis: true mean is not equal to 10
## 95 percent confidence interval:
##  9.935078 10.070327
## sample estimates:
## mean of x
##  10.0027

t.test(defectsData$Weight[which(defectsData$Method=="B")], alternative = "two.sided", mu = 10)

##
## One Sample t-test
##
## data: defectsData$Weight[which(defectsData$Method == "B")]
## t = -0.49565, df = 74, p-value = 0.6216
## alternative hypothesis: true mean is not equal to 10
## 95 percent confidence interval:
##  9.912986 10.052348
## sample estimates:
## mean of x
##  9.982667
```

```
t.test(defectsData$Weight[which(defectsData$Method=="C")], alternative = "two.sided", mu = 10)
```

```
##  
## One Sample t-test  
##  
## data: defectsData$Weight[which(defectsData$Method == "C")]  
## t = 1.4866, df = 63, p-value = 0.1421  
## alternative hypothesis: true mean is not equal to 10  
## 95 percent confidence interval:  
## 9.971497 10.194128  
## sample estimates:  
## mean of x  
## 10.08281
```

```
t.test(defectsData$Weight[which(defectsData$Method=="D")], alternative = "two.sided", mu = 10)
```

```
##  
## One Sample t-test  
##  
## data: defectsData$Weight[which(defectsData$Method == "D")]  
## t = 2.0682, df = 51, p-value = 0.04371  
## alternative hypothesis: true mean is not equal to 10  
## 95 percent confidence interval:  
## 10.00491 10.32971  
## sample estimates:  
## mean of x  
## 10.16731
```

```
t.test(defectsData$Weight[which(defectsData$Method=="A")],  
       defectsData$Weight[which(defectsData$Method=="B")],  
       alternative = "two.sided")
```

```
##  
## Welch Two Sample t-test  
##  
## data: defectsData$Weight[which(defectsData$Method == "A")] and defectsData$Weight[which(defectsData$Method == "B")]  
## t = 0.41119, df = 146.92, p-value = 0.6815  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -0.07625975 0.11633182  
## sample estimates:  
## mean of x mean of y  
## 10.002703 9.982667
```

```
t.test(defectsData$Weight[which(defectsData$Method=="A")],  
       defectsData$Weight[which(defectsData$Method=="C")],  
       alternative = "two.sided")
```

```
##  
## Welch Two Sample t-test  
##  
## data: defectsData$Weight[which(defectsData$Method == "A")] and defectsData$Weight[which(defectsData$Method == "C")]  
## t = -1.2282, df = 105.85, p-value = 0.2221  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -0.20942645 0.04920686
```

```

## sample estimates:
## mean of x mean of y
## 10.00270 10.08281

t.test(defectsData$Weight[which(defectsData$Method=="A")],
       defectsData$Weight[which(defectsData$Method=="D")],
       alternative = "two.sided")

##
## Welch Two Sample t-test
##
## data: defectsData$Weight[which(defectsData$Method == "A")] and defectsData$Weight[which(defectsData$Method == "D")]
## t = -1.8764, df = 69.032, p-value = 0.06482
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.33960503 0.01039505
## sample estimates:
## mean of x mean of y
## 10.00270 10.16731

t.test(defectsData$Weight[which(defectsData$Method=="B")],
       defectsData$Weight[which(defectsData$Method=="C")],
       alternative = "two.sided")

##
## Welch Two Sample t-test
##
## data: defectsData$Weight[which(defectsData$Method == "B")] and defectsData$Weight[which(defectsData$Method == "C")]
## t = -1.5226, df = 108.14, p-value = 0.1308
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.23051478 0.03022311
## sample estimates:
## mean of x mean of y
## 9.982667 10.082812

t.test(defectsData$Weight[which(defectsData$Method=="B")],
       defectsData$Weight[which(defectsData$Method=="D")],
       alternative = "two.sided")

##
## Welch Two Sample t-test
##
## data: defectsData$Weight[which(defectsData$Method == "B")] and defectsData$Weight[which(defectsData$Method == "D")]
## t = -2.0951, df = 70.155, p-value = 0.03977
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.360403604 -0.008878447
## sample estimates:
## mean of x mean of y
## 9.982667 10.167308

t.test(defectsData$Weight[which(defectsData$Method=="C")],
       defectsData$Weight[which(defectsData$Method=="D")],
       alternative = "two.sided")

##

```

```
## Welch Two Sample t-test
##
## data: defectsData$Weight[which(defectsData$Method == "C")] and defectsData$Weight[which(defectsData$Method == "D")]
## t = -0.86028, df = 93.766, p-value = 0.3918
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.2795164 0.1105260
## sample estimates:
## mean of x mean of y
## 10.08281 10.16731
```

The results of testing that each of the methods is not equal to 10 shows that only D is not equal to 10.

The results of testing each method to each other method shows that there is a significant difference in mean weight between B and D. If, however, we apply the Bonferonni principle, then the alpha goes from 0.05 to $0.05/6=0.00833$ and the difference becomes insignificant.

Question 2

Compare the mean weights for the 4 methods using ANOVA. What is your conclusion? How does it compare to the conclusions from the pairwise comparisons?

```
summary(aov(Weight ~ Method, defectsData))
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## Method         3    1.29   0.4299   2.617 0.0515 .
## Residuals    261   42.87   0.1643
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

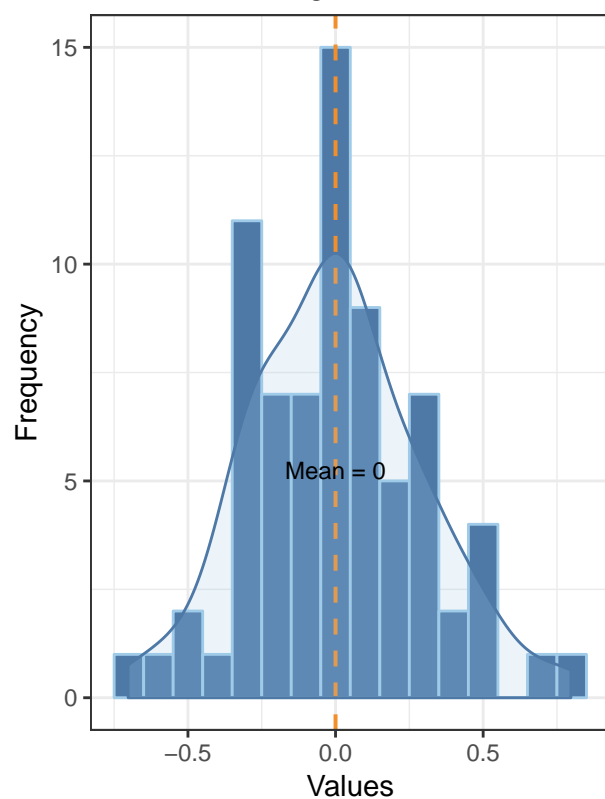
ANOVA shows that we'd fail to reject the null hypothesis of no difference in means between the methods. This is different than the answer in 1 because we're answering a different question. In 1, we were concerned with differences between the method's mean and 10 as well as the differences between each set of methods' means. Here, however, we're answering the question "is there a difference in means between all methods".

Question 3

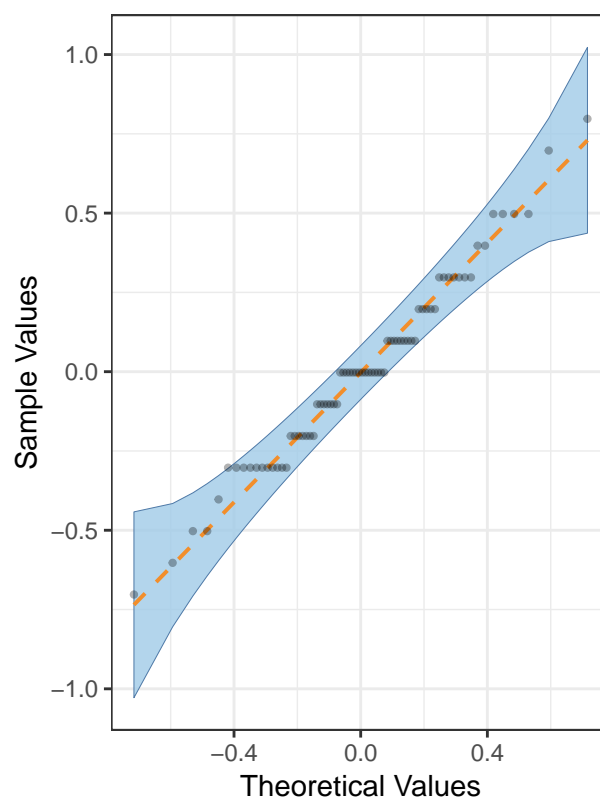
Assess the assumptions of the ANOVA. Are the assumptions met?

```
distribution_visualizer(defectsData$Weight[which(defectsData$Method=="A")]-
                        mean(defectsData$Weight[which(defectsData$Method=="A")]), "A Method Weight Residuals")
```

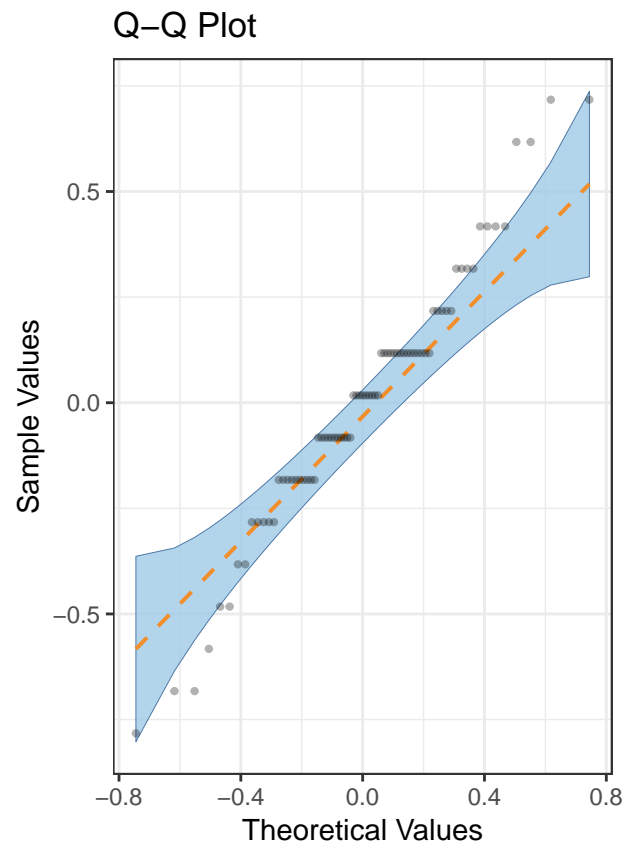
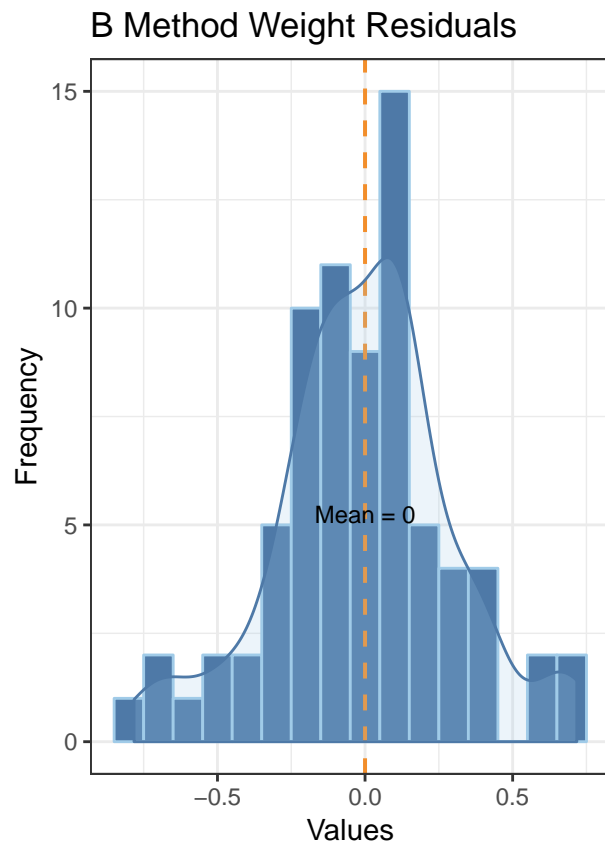
A Method Weight Residuals



Q-Q Plot

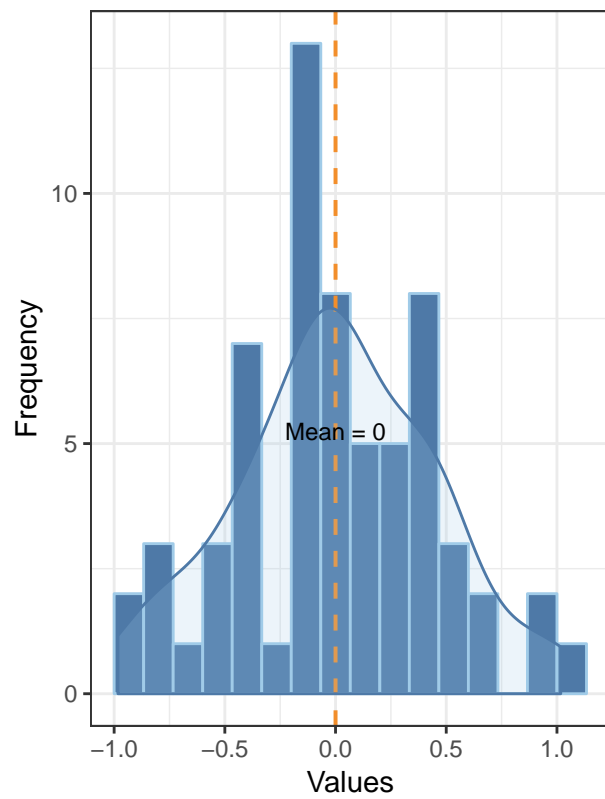


```
distribution_visualizer(defectsData$Weight[which(defectsData$Method=="B")]-
  mean(defectsData$Weight[which(defectsData$Method=="B")]), "B Method Weight Res
```

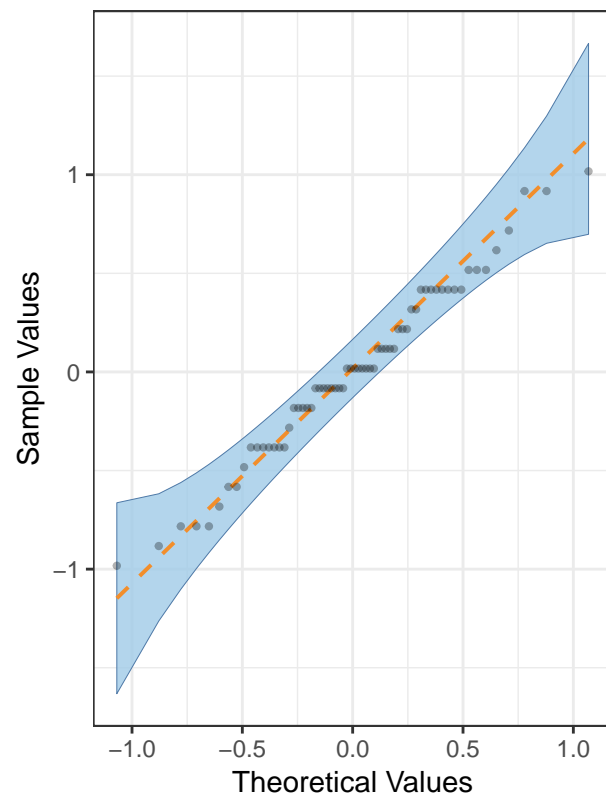


```
distribution_visualizer(defectsData$Weight[which(defectsData$Method=="C")]-
  mean(defectsData$Weight[which(defectsData$Method=="C")]), "C Method Weight Res
```

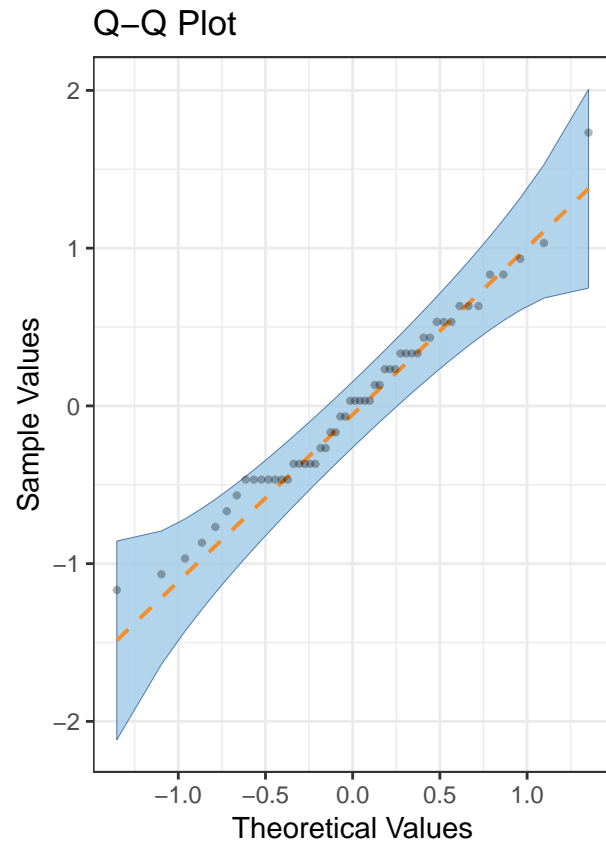
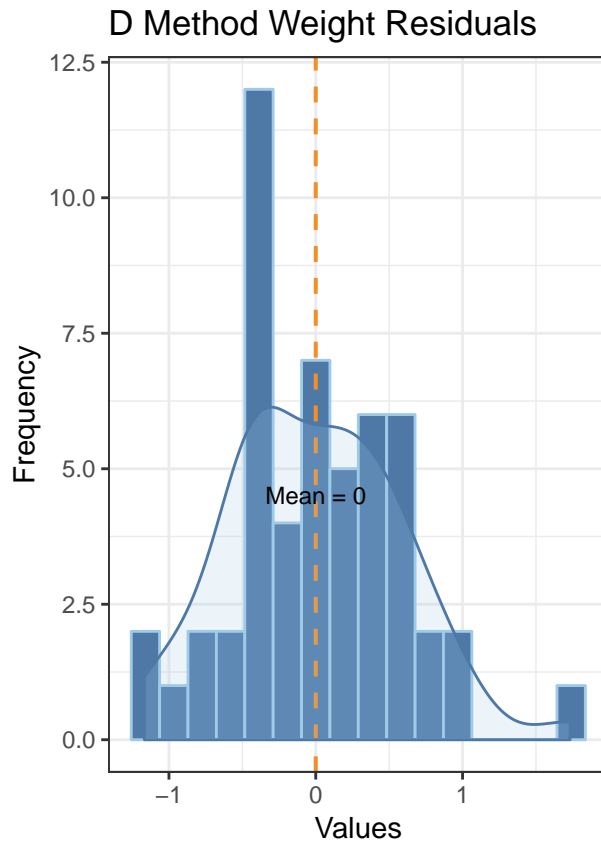
C Method Weight Residuals



Q-Q Plot



```
distribution_visualizer(defectsData$Weight[which(defectsData$Method=="D")]-
  mean(defectsData$Weight[which(defectsData$Method=="D")]), "D Method Weight Res
```



```
var(defectsData$Weight[which(defectsData$Method=="A")])
```

```
## [1] 0.08519807
```

```
var(defectsData$Weight[which(defectsData$Method=="B")])
```

```
## [1] 0.09172252
```

```
var(defectsData$Weight[which(defectsData$Method=="C")])
```

```
## [1] 0.1985888
```

```
var(defectsData$Weight[which(defectsData$Method=="D")])
```

```
## [1] 0.3402828
```

For the assumption of normality, B seems slightly non-normal and skews left, but they seem somewhat normal overall.

For the assumption of equal variances, we can see that they are fairly different, ranging from 0.085 to 0.340.

For the assumption of independence, this seems to be met since random sampling was used.

Question 4

Perform a simulation study to assess the validity of the ANOVA. Choose distributions for weight that seem to be reasonable based on your analysis of the data. Report the estimated type I error rate of the ANOVA F-test.

```
set.seed(0192)
```

```
xbar <- mean(defectsData$Weight)
```

```
sd1 <- sd(defectsData$Weight[which(defectsData$Method=="A")])
```



```

sd2 <- sd(defectsData$Weight[which(defectsData$Method=="B")])
sd3 <- sd(defectsData$Weight[which(defectsData$Method=="C")])
sd4 <- sd(defectsData$Weight[which(defectsData$Method=="D")])
n <- nrow(defectsData)
reps <- 1000
p_vals <- rep(NA, reps)
for(i in 1:reps){
  sample1 <- rnorm(n, xbar, sd1)
  sample2 <- rnorm(n, xbar, sd2)
  sample3 <- rnorm(n, xbar, sd3)
  sample4 <- rnorm(n, xbar, sd4)

  sample_df <- data.frame(data = c(sample1, sample2, sample3, sample4),
                           method = c(rep("A", n), rep("B", n), rep("C", n), rep("D", n)))

  p <- summary(aov(data ~ method, sample_df))[[1]][[5]][[1]]
  p_vals[i] <- p
}
mean(p_vals < 0.05)

```

```
## [1] 0.052
```

Given that we're pretty close to 0.05, ANOVA seems valid.