Sam Tenney

Section 2

Homework 6

1. **The Art of Oreo Dunking**

    a) # Homework 6
    b)

Table 6.1.b.
Summary Statistics for Oreo Dunking

Cookie Type	Mean (grams)	Standard Deviation (grams)
Regular	6.180	0.866
Trader Joe's	4.845	0.453
Double Stuf	5.145	0.795

Regular Oreos seemed to have a larger mean (6.180 g) and a larger spread (0.866 g) compared to the other two cookies. Trader Joe's had the smallest mean (4.845 g) and the smallest spread (0.453 g).

    c) On average, Regular Oreos absorbed 6.180 grams of milk after being dunked for 10 seconds (95% CI: 4.802g to 7.558g). On average, Trader Joe's Joe Joe's absorbed 4.845 grams of milk after being dunked for 10 seconds (95% CI: 4.123g to 5.567g). On average, Double Stuf Oreos absorbed 5.145 grams of milk after being dunked for 10 seconds (95% CI: 3.881g to 6.409g).
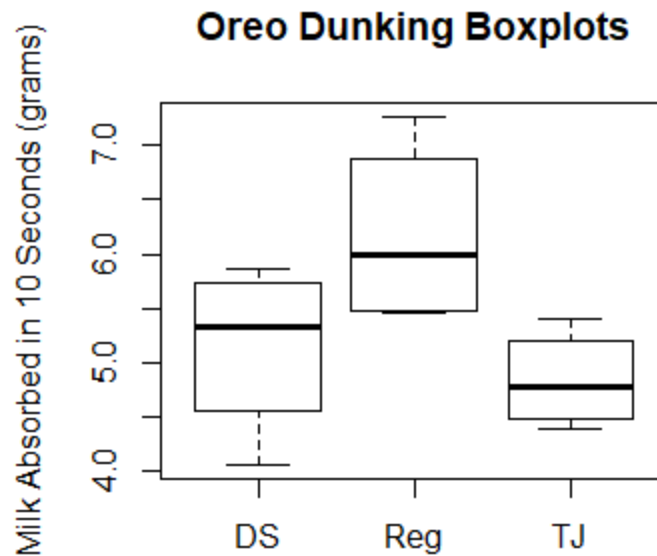
d)

Oreo Dunking Boxplots

Milk Absorbed in 10 Seconds (grams)

DS   Reg   TJ

Figure 6.1.d

None of the boxplots are symmetric. The Double Stuf Oreos are left-skewed, the Trader Joe's are slightly right-skewed, and the regular Oreos are right-skewed. The regular Oreos tend to have a higher amount of milk absorbed in grams than the other two cookie types. The Trader Joe's cookies have the smallest spread. There are no outliers in the data.
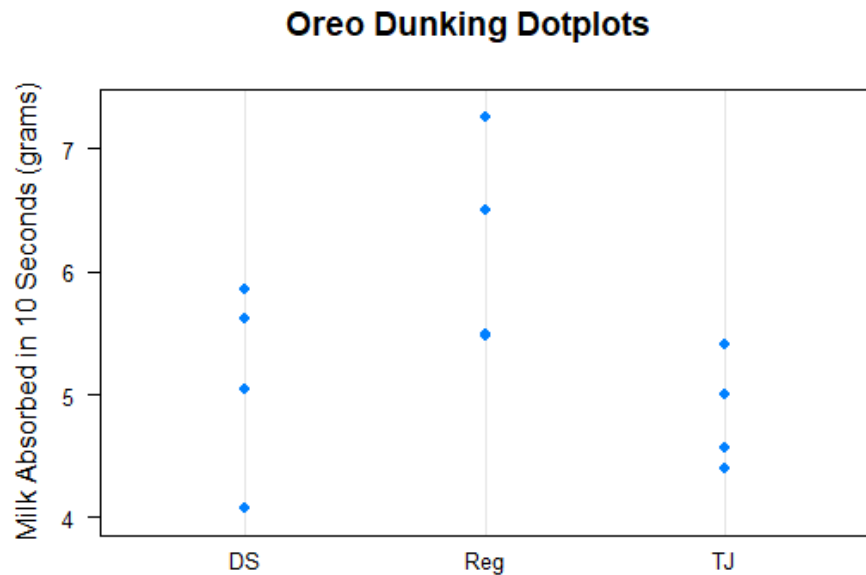
e)

Oreo Dunking Dotplots

Figure 6.1.e

The regular Oreos tend to have a higher amount of milk absorbed in grams according to the dot plot. The Double Stuf could've had one cookie that was lower than usual, as three of them are above five grams while one is at about four grams. Trader Joe's had more consistent results with the smaller spread. Regular Oreos look to have a little bit larger spread than Double Stuf. Regular Oreos and Trader Joe's appear symmetric while Double Stuf appear left-skewed.

f) The ANOVA model is $y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$. The variable y_{ij} is the amount of milk absorbed after 10 seconds by the ith cookie type (Regular, Double Stuf, Trader Joe's) and the jth replicate of that cookie type. The variable μ is the grand mean of the milk absorbed after 10 seconds by all the cookies. The variable α_i is the treatment effect for the ith cookie. The error for the ith cookie type and the jth replicate is represented by ε_{ij} .

g)   The null hypothesis we are testing for is the mean amount of milk absorbed after being dunked for 10 seconds for each cookie type is equal. Our alternative hypothesis is that at least one of the cookie types has a different mean amount of milk absorbed after being dunked for 10 seconds.

h) `anova(aov(milkabsgram~treatment, data=oreos))`

```
## Analysis of Variance Table
## Response: milkabsgram
##              Df      Sum Sq    Mean Sq     F value     Pr(>F)
## treatment   2        3.9246    1.96230     3.7102      0.06681 .
## Residuals   9        4.7600    0.52889
## ---
## Signif. codes:   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

i) The p-value is above our threshold of .05, so we fail to reject the null hypothesis and conclude that the mean amount of milk absorbed for the three cookie types must be equal.

j) The most appropriate approach to ensure our family-wise error rate is no greater than 0.05 is the Tukey HSD approach.

```
TukeyHSD(aov(milkabsgram~treatment, data=oreos))
##    Tukey multiple comparisons of means
##      95% family-wise confidence level
## Fit: aov(formula = milkabsgram ~ treatment, data = oreos)
## $treatment
##            diff        lwr         upr       p adj
## Reg-DS    1.035  -0.4007655  2.4707655  0.1647184
## TJ-DS    -0.300  -1.7357655  1.1357655  0.8322025
## TJ-Reg   -1.335  -2.7707655  0.1007655  0.0678383
```

As found in our results in the one-way ANOVA, none of the p-values for the different tests are below our threshold of .05, so we can fail to reject the null hypothesis as we did in our one-way ANOVA, suggesting that there is no significant difference between the mean amount of milk absorbed (g) after being dunked for 10 seconds in each cookie type.

k) Calculated residuals:

```
oreos$resids <- resid(aov(milkabsgram~treatment, data=oreos))
oreos$resids

##  [1]   0.155 -1.075 -0.105  0.565  1.080 -0.710  0.465  0.320
0.715 -0.275 -0.445 -0.690
```
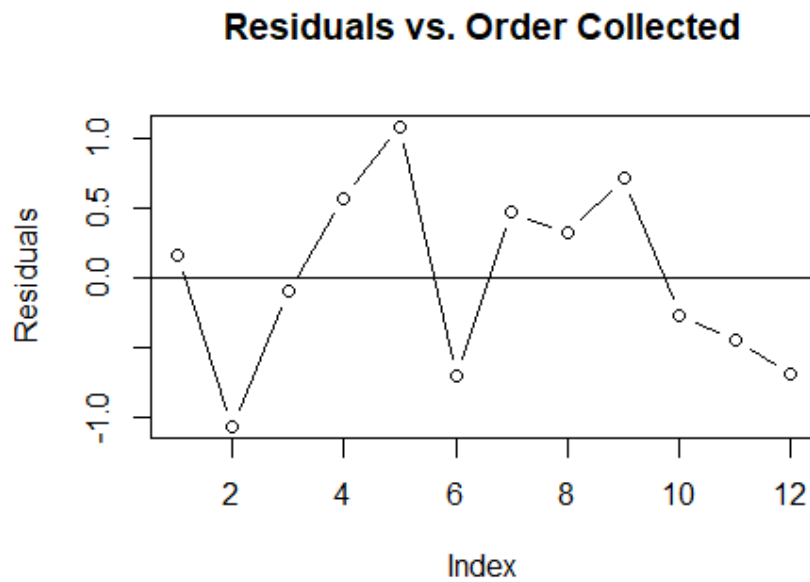
l)

Residuals vs. Order Collected

Figure 6.1.l

There are no obvious patterns to suggest that the observations are not independent. We can assume that our observations are independent.
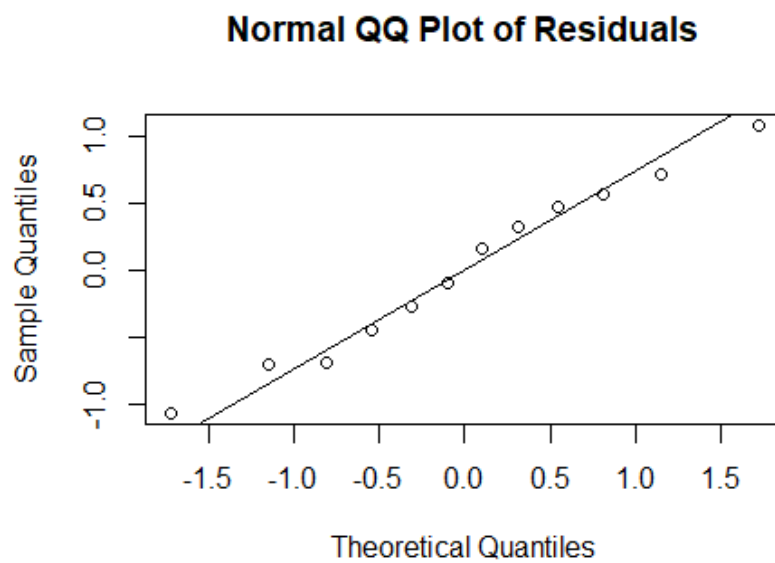
m)

Normal QQ Plot of Residuals

Figure 6.1.m

The residuals form a straight line, so we can assume that the residuals are normally distributed.

n) 0.866 / 0.453 = 1.91
Our standard deviation ratio of 1.91 is less than 2, so the equal variance assumption is satisfied.

o) We want to compare three types of cookies (Regular Oreos, Trader Joe's Joe-Joe's, Double Stuf Oreos) to see which one absorbs the most milk after being dunked for 10 seconds in milk. Since we are comparing means between three different treatments, we used a one-way ANOVA. Our null hypothesis is that the mean amount of milk absorbed between all three cookies is the same, and our alternative hypothesis is that at least one cookie has a different mean amount of milk absorbed. We sampled 4 cookies of each of the three factor levels to run our analysis. We found that on average, Regular Oreos absorbed 6.180 grams of milk after being dunked for 10 seconds (95% CI: 4.802g to 7.558g), this being the most of the three cookie types. Our ANOVA returned a p-value of 0.06 which is greater than our significance level of 0.05, so we fail to reject the null hypothesis and conclude that there is no difference in mean amount of milk absorbed between all the cookie types.

2. Pairwise Comparisons:

Golden - Regular
Golden - Double-Stuff,
Golden - Thins
Regular - Double-Stuff
Regular - Thins
Double Stuff – Thins

The probability at least one test inappropriately rejects H_0 when H_0 is true and when using $\alpha = .05$ is 0.265.
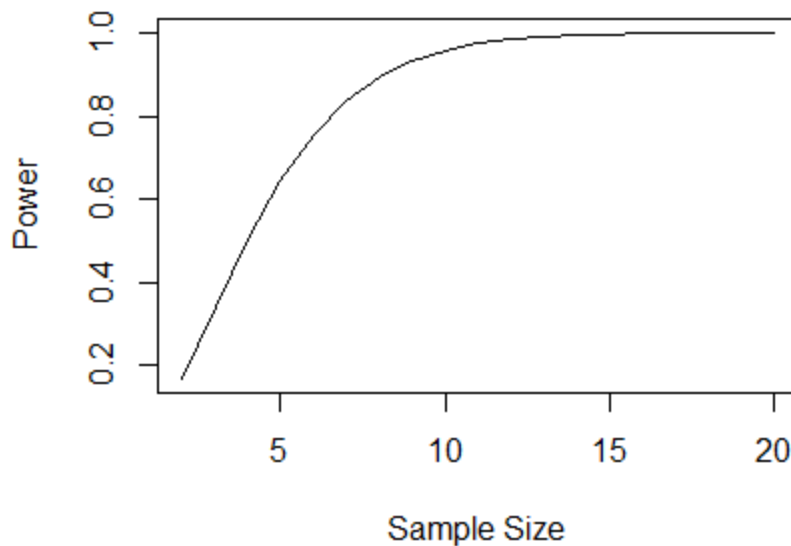
3.

a)
```
n.options = seq(2,20,by=1)
res.power = NA
for(i in 1:length(n.options)) {
  res = power.anova.test(groups=4,
                         between.var=var(c(22, 23, 29, 30)),
                         within.var = 25,
                         sig.level=0.05,
                         n=n.options[i])
  res.power[i] = res$power
}
plot(n.options,
     res.power,
     type="l",
     xlab="Sample Size",
     ylab="Power",
     main="Power Curve for Cookie Type Study")
```

Power Curve for Cookie Type Study

b) ```
Part b
power.anova.test(groups=4,
                 between.var=var(c(22, 23, 29, 30)),
                 within.var = 25,
                 sig.level=0.05,
                 power = 0.85)
Balanced one-way analysis of variance power calculation
##
groups = 4
n = 7.209869
between.var = 16.66667
within.var = 25
sig.level = 0.05
power = 0.85
```

The smallest value for the group size (n) is 7.21 that gives 85% power.

c)

```
# Part c
n.options = seq(2,20,by=1)
res.power = NA

for(i in 1:length(n.options)) {
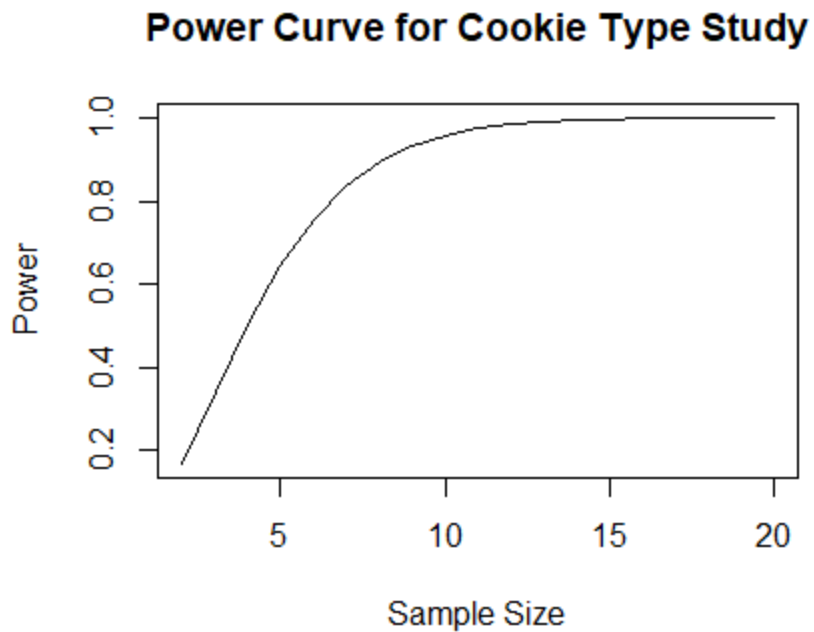  res = power.anova.test(groups=4,
                         between.var=var(c(30, 29, 23, 22)),
                         within.var = 25,
```

```
                                sig.level=0.05,
                                n=n.options[i])
    res.power[i] = res$power
}
plot(n.options,
     res.power,
     type="l",
     xlab="Sample Size",
     ylab="Power",
     main="Power Curve for Cookie Type Study")
```

Power Curve for Cookie Type Study

The power curve remains the same.

d)

```
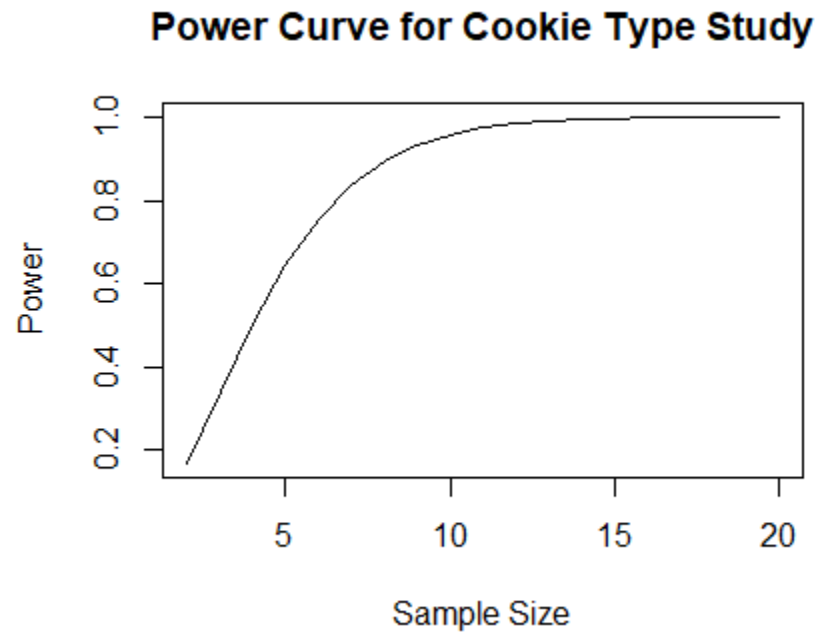# Part d
n.options = seq(2,20,by=1)
res.power = NA

for(i in 1:length(n.options)) {
    res = power.anova.test(groups=4,
                           between.var=var(c(21, 26, 26, 31)),
                           within.var = 25,
                           sig.level=0.05,
                           n=n.options[i])
    res.power[i] = res$power
}
plot(n.options,
     res.power,
     type="l",
```

```
xlab="Sample Size",
ylab="Power",
main="Power Curve for Cookie Type Study")
```

Power Curve for Cookie Type Study

The power curve remains the same.

e) All the curves for parts (a), (c), and (d) are the same because the variance of the four treatment groups remains the same even though the numbers are different. The vary the same about the grand mean.