myoreos.R

Samuel

Thu Oct 11 14:02:26 2018

# The Art of Oreo Dunking  
# Sam Tenney  
# Finds which type of oreo absorbs more milk   
  
# Read in the data  
oreos <- read.table(text="run,experiment,treatment,weightBefore,weightAfter  
 1,8,TJ,19.69,24.69  
 2,12,DS,20.36,24.43  
 3,11,DS,19.92,24.96  
 4,6,TJ,18.20,23.61  
 5,4,Reg,16.17,23.43  
 6,3,Reg,16.62,22.09  
 7,9,DS,20.31,25.92  
 8,1,Reg,16.27,22.77  
 9,10,DS,20.14,26.00  
 10,7,TJ,18.97,23.54  
 11,5,TJ,18.56,22.96  
 12,2,Reg,16.66,22.15", header = TRUE, sep=",")  
  
# Part d - Provides a summary of the data read in previously  
head(oreos)

## run experiment treatment weightBefore weightAfter  
## 1 1 8 TJ 19.69 24.69  
## 2 2 12 DS 20.36 24.43  
## 3 3 11 DS 19.92 24.96  
## 4 4 6 TJ 18.20 23.61  
## 5 5 4 Reg 16.17 23.43  
## 6 6 3 Reg 16.62 22.09

tail(oreos)

## run experiment treatment weightBefore weightAfter  
## 7 7 9 DS 20.31 25.92  
## 8 8 1 Reg 16.27 22.77  
## 9 9 10 DS 20.14 26.00  
## 10 10 7 TJ 18.97 23.54  
## 11 11 5 TJ 18.56 22.96  
## 12 12 2 Reg 16.66 22.15

str(oreos)

## 'data.frame': 12 obs. of 5 variables:  
## $ run : int 1 2 3 4 5 6 7 8 9 10 ...  
## $ experiment : int 8 12 11 6 4 3 9 1 10 7 ...  
## $ treatment : Factor w/ 3 levels "DS","Reg","TJ": 3 1 1 3 2 2 1 2 1 3 ...  
## $ weightBefore: num 19.7 20.4 19.9 18.2 16.2 ...  
## $ weightAfter : num 24.7 24.4 25 23.6 23.4 ...

# Part e - Subsets the oreo dataset into the Regular and Trader Joe's treatments  
oreosSUB <- subset(oreos, treatment %in% c("Reg", "TJ"))  
oreosSUB$treatment <- droplevels(oreosSUB$treatment)  
str(oreosSUB$treatment)

## Factor w/ 2 levels "Reg","TJ": 2 2 1 1 1 2 2 1

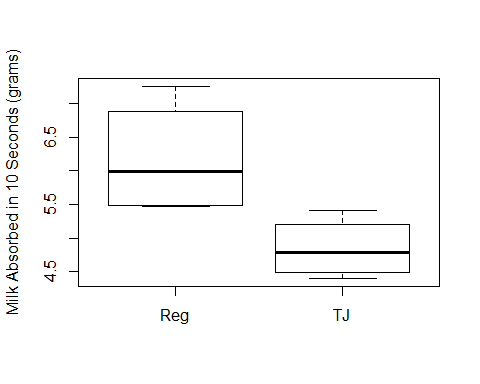
# Part f - Calculate the mean and standard dev. of amount of milk absorbed in 10 seconds for Reg and TJ Oreos  
oreosSUB$milkabsgram <- oreosSUB$weightAfter - oreosSUB$weightBefore  
  
  
#Calculate mean and sd for each cookie type  
aggregate(milkabsgram~treatment, data = oreosSUB, FUN = mean)

## treatment milkabsgram  
## 1 Reg 6.180  
## 2 TJ 4.845

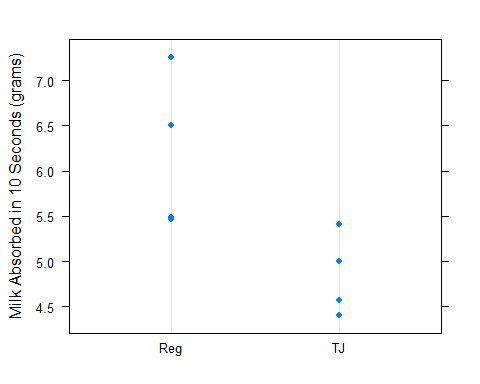
aggregate(milkabsgram~treatment, data = oreosSUB, FUN = sd)

## treatment milkabsgram  
## 1 Reg 0.8658329  
## 2 TJ 0.4534681

# Part g - Boxplots of the amount milk absorbed in 10 seconds for Regular Oreos and Trader Joe's Joe-Joe's  
boxplot(milkabsgram~treatment, data=oreosSUB, ylab="Milk Absorbed in 10 Seconds (grams)")  
  
# Part h - Dotplots of the amount of milk absorbed in 10 seconds  
library(lattice)



dotplot(milkabsgram~treatment, data=oreosSUB, ylab="Milk Absorbed in 10 Seconds (grams)")



# Homework 3;  
oreos$milkabsgram <- oreos$weightAfter - oreos$weightBefore  
  
# Mean for Regular Oreos  
regMean <- mean(oreos$milkabsgram[oreos$treatment == "Reg"])  
regMean

## [1] 6.18

# Standard deviation for Regular Oreos  
Regsd <- sd(oreos$milkabsgram[oreos$treatment == "Reg"])  
Regsd

## [1] 0.8658329

# Sample size for Regular Oreos  
Regsize <- length(oreos$milkabsgram[oreos$treatment == "Reg"])  
Regsize

## [1] 4

# Calculate 95% Confidence Interval for mean amount of milk absorbed for Regular Oreos  
regMean + qt(.975, df=(Regsize-1))\*Regsd/sqrt(Regsize)

## [1] 7.557733

regMean - qt(.975, df=(Regsize-1))\*Regsd/sqrt(Regsize)

## [1] 4.802267

# Mean for Trader Joe's Oreos  
TJmean <- mean(oreos$milkabsgram[oreos$treatment == "TJ"])  
TJmean

## [1] 4.845

# Standard deviation for Trader Joe's Oreos  
TJsd <- sd(oreos$milkabsgram[oreos$treatment == "TJ"])  
TJsd

## [1] 0.4534681

# Sample size for Trader Joe's Oreos  
TJsize <- length(oreos$milkabsgram[oreos$treatment == "TJ"])  
TJsize

## [1] 4

# Calculate 95% Confidence Interval for mean amount of milk absorbed for Trader Joe's Oreos  
TJmean + qt(.975, df=(TJsize-1))\*TJsd/sqrt(TJsize)

## [1] 5.566569

TJmean - qt(.975, df=(TJsize-1))\*TJsd/sqrt(TJsize)

## [1] 4.123431

# Ratio to check if variance is constant  
Regsd / TJsd

## [1] 1.909358

# Part d - Identify variables   
n1 <- Regsize  
n2 <- TJsize  
y14 <- oreosSUB$milkabsgram[oreosSUB$treatment == "Reg"][4]  
y23 <- oreosSUB$milkabsgram[oreosSUB$treatment == "TJ"][3]  
error14 <- y14 - regMean  
error23 <- y23 - TJmean  
  
# Part e - Conduct a two-sample t-test by hand  
sp2 <- ((n1 - 1) \* Regsd^2 + (n2 - 1) \* TJsd^2)/(n1 + n2 - 2)  
teststat <- (regMean - TJmean - 0)/(sqrt(sp2/n1 + sp2/n2))  
teststat

## [1] 2.731753

p\_value <- 2 \* (1 - pt(abs(teststat), df = n1 + n2 - 2))  
p\_value

## [1] 0.03410663

# Part f - Conduct a two-sample t-test using R  
t.test(milkabsgram~treatment, data = oreosSUB, var.equal = TRUE)

##   
## Two Sample t-test  
##   
## data: milkabsgram by treatment  
## t = 2.7318, df = 6, p-value = 0.03411  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 0.1392009 2.5307991  
## sample estimates:  
## mean in group Reg mean in group TJ   
## 6.180 4.845

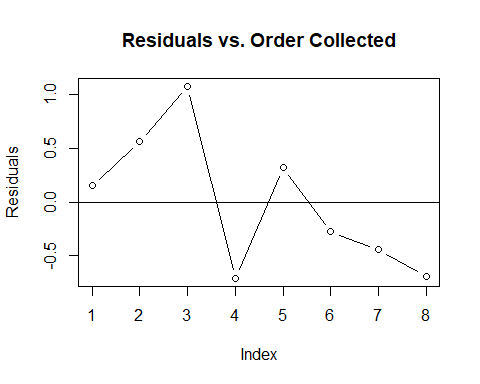
# Part h - Calculate a confidence interval for the difference in mean amount of milk absorbed in 10 seconds  
mean\_diff <- regMean - TJmean  
  
# Homework 4  
  
# Part b - Calculate the residuals and find their mean and standard deviation  
oreosSUB$resids <- resid(aov(milkabsgram~treatment, data=oreosSUB))  
mean(oreosSUB$resids)

## [1] -3.469447e-17

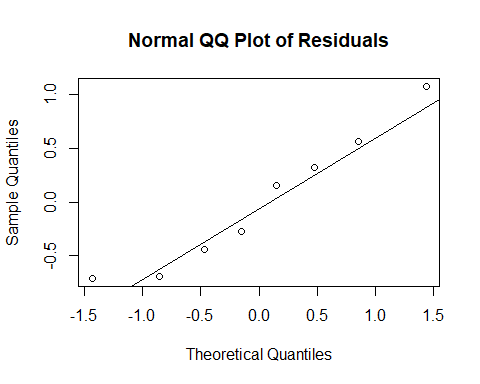
sd(oreosSUB$resids)

## [1] 0.6398549

# Part c - Create a plot of residuals vs the order (index plot)  
plot(oreosSUB$resids, type = "b", main = "Residuals vs. Order Collected", ylab = "Residuals")  
abline(h = 0)



# Part d - Create a normal qqplot  
qqnorm(oreosSUB$resids, main = "Normal QQ Plot of Residuals")  
qqline(oreosSUB$resids)



# Part 6a  
power.t.test(n = 4, sig.level = .05, delta = mean\_diff, sd = sp2)

##   
## Two-sample t test power calculation   
##   
## n = 4  
## delta = 1.335  
## sd = 0.47765  
## sig.level = 0.05  
## power = 0.905521  
## alternative = two.sided  
##   
## NOTE: n is number in \*each\* group

# Part 6b  
power.t.test(sig.level = .05, delta = mean\_diff, power = .80, sd = sp2)

##   
## Two-sample t test power calculation   
##   
## n = 3.305898  
## delta = 1.335  
## sd = 0.47765  
## sig.level = 0.05  
## power = 0.8  
## alternative = two.sided  
##   
## NOTE: n is number in \*each\* group

# Homework 6  
  
# Part b  
# Means for each cookie type  
regMean <- mean(oreos$milkabsgram[oreos$treatment == "Reg"])  
TJmean <- mean(oreos$milkabsgram[oreos$treatment == "TJ"])  
DSMean <- mean(oreos$milkabsgram[oreos$treatment == "DS"])  
  
# Standard deviation for each cookie type  
Regsd <- sd(oreos$milkabsgram[oreos$treatment == "Reg"])  
TJsd <- sd(oreos$milkabsgram[oreos$treatment == "TJ"])  
DSsd <- sd(oreos$milkabsgram[oreos$treatment == "DS"])  
  
# part c  
DSsize <- length(oreos$milkabsgram[oreos$treatment == "DS"])  
  
regMean + qt(.975, df=(Regsize-1))\*Regsd/sqrt(Regsize)

## [1] 7.557733

regMean - qt(.975, df=(Regsize-1))\*Regsd/sqrt(Regsize)

## [1] 4.802267

TJmean + qt(.975, df=(TJsize-1))\*TJsd/sqrt(TJsize)

## [1] 5.566569

TJmean - qt(.975, df=(TJsize-1))\*TJsd/sqrt(TJsize)

## [1] 4.123431

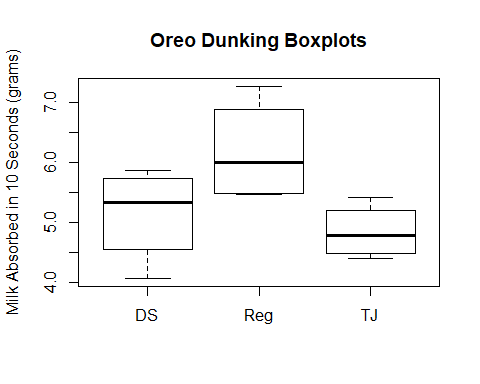
DSMean + qt(.975, df=(DSsize-1))\*DSsd/sqrt(DSsize)

## [1] 6.409363

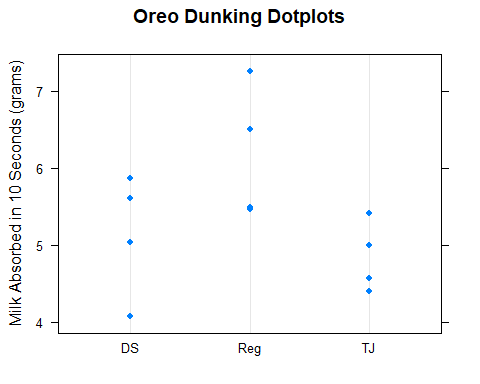
DSMean - qt(.975, df=(DSsize-1))\*DSsd/sqrt(DSsize)

## [1] 3.880637

# part d  
boxplot(milkabsgram~treatment, data=oreos, ylab="Milk Absorbed in 10 Seconds (grams)", main = "Oreo Dunking Boxplots")



# part e  
dotplot(milkabsgram~treatment, data=oreos, ylab="Milk Absorbed in 10 Seconds (grams)", main = "Oreo Dunking Dotplots")



# part 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

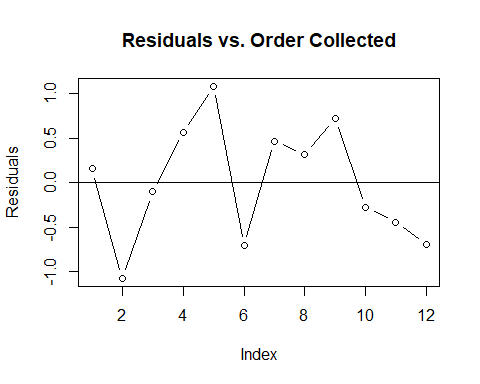
# part j  
TukeyHSD(aov(milkabsgram~treatment, data=oreos))

## 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

# part k  
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  
## [11] -0.445 -0.690

# part l  
plot(oreos$resids, type = "b", main = "Residuals vs. Order Collected", ylab = "Residuals")  
abline(h = 0)



# part m  
qqnorm(oreos$resids, main = "Normal QQ Plot of Residuals")  
qqline(oreos$resids)

