Project # 3: Power of a z-test. Two proportions. Chi-squared.

Sunny Kharel, William Easterby, Gabrielle Toutin

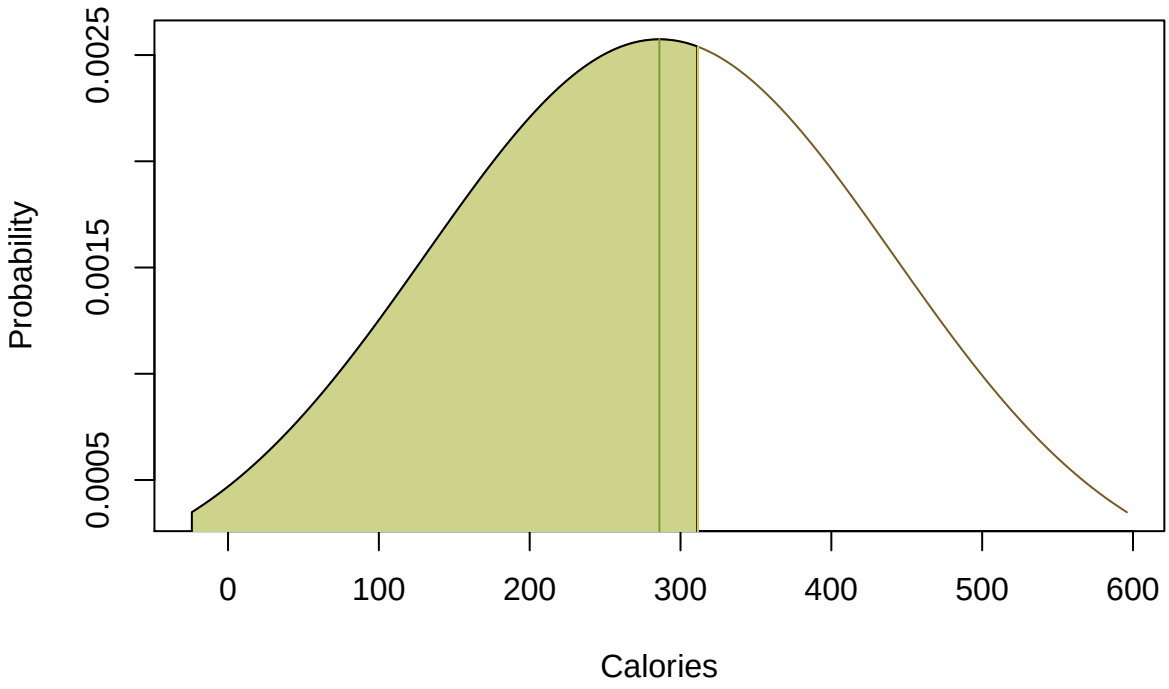
11/23/2020

Problem 3.1: The operating characteristic curve: Power of a z-test.

(i)

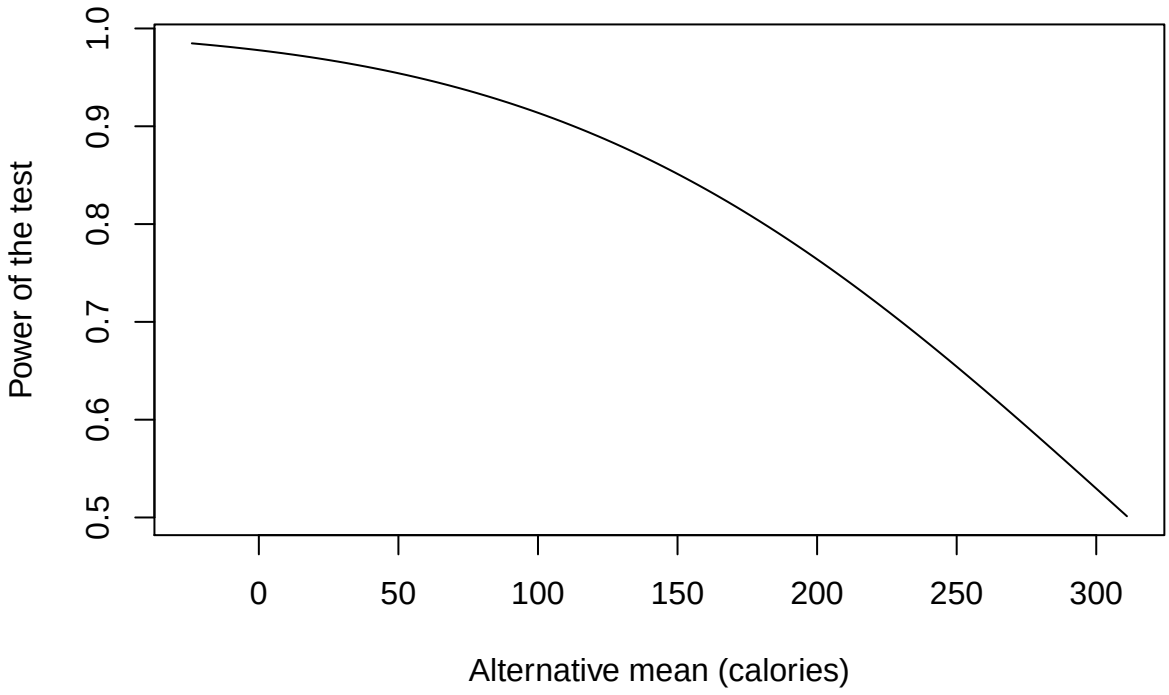
$H_0 : \mu = 286$ vs. $H_a : \mu < 286$
Rejection region:
$\mu \pm z^* \frac{\sigma}{\sqrt{n}}$

```
alpha = 0.05
n = 100
sigma = 155
nullmu = 286

domain = (nullmu-2*(sigma)):(nullmu+2*(sigma))  # plot 2 standard deviations
range = dnorm(domain,nullmu,sigma)
plot(domain,range,"l",col="#775B24",xlab = "Calories",ylab = "Probability")
lines(nullmu,dnorm(nullmu,nullmu,sigma),type="h",col="#7D9D33")

rr_upper = nullmu + qnorm(1-alpha)*sigma/sqrt(n)
lines(rr_upper,dnorm(rr_upper,nullmu,sigma),type="h",col="#DCC949")

x = (nullmu-2*(sigma)):rr_upper
y = dnorm(x,nullmu,sigma)
polygon(c(x,rev(x)),c(y,vector(mode="integer",length=length(x))),col="#CED38C")

#redo things that got deleted
lines(rr_upper,dnorm(rr_upper,nullmu,sigma),type="h",col="#DCC949")
lines(nullmu,dnorm(nullmu,nullmu,sigma),type="h",col="#7D9D33")
```



(ii)

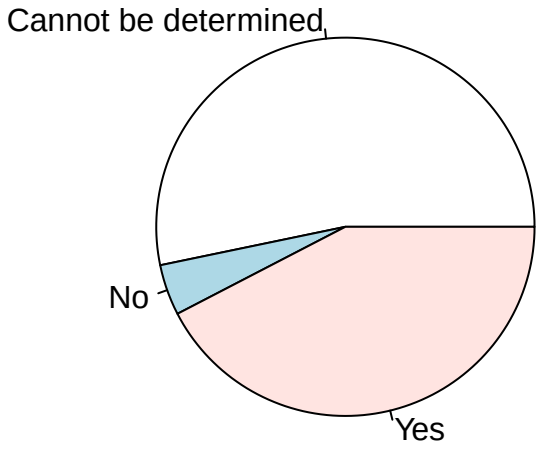
```
ivegotthepower <- function(mua){
  power = pnorm(rr_upper,mua,sigma)
  return(power)
}

plot(x,ivegotthepower(x),"l",xlab="Alternative mean (calories)",ylab = "Power of the test")
```



Problem 3.2: Our logic survey

(i)

```
#load the data
read.csv('logic-puzzle-abridged.csv') -> logicData
colnames(logicData)<-c("time", "answer", "section")
logicResponse <- logicData$answer
pie(table(logicResponse))
```




```
#separate section responses
morning<-logicResponse[logicData$section=="11am"]
afternoon<-logicResponse[logicData$section=="12noon"]

#get proportion of correct responses
am_correct <- as.numeric(table(morning)['Yes'])
am_total <- length(morning)

noon_correct <- as.numeric(table(afternoon)['Yes'])
noon_total <- length(afternoon)

#get p-value
p_hat <- (am_correct + noon_correct)/(am_total + noon_total)
z_score <- ((am_correct/am_total) - (noon_correct/noon_total))/sqrt(p_hat*(1-p_hat)*((1/am_total)+(1/noo

print(sprintf("The z-score is %f", z_score))
```

```
## [1] "The z-score is -0.018503"
```

```
p_value <- 2*pnorm(z_score)
print(sprintf("The p-value is %f", p_value))
```

```
## [1] "The p-value is 0.985237"
```

With a p-value of 0.985, we fail to reject the null hypothesis.

(ii)

```
p_hat = (noon_correct + am_correct)/(noon_total+am_total)
p_not = 0.20

z_score = (p_hat - p_not) / sqrt((p_not*(1-p_not))/(noon_total+am_total))
print(sprintf("The z-score is %f", z_score))
```

```
## [1] "The z-score is 5.369246"
```

```
p_value <- 1-pnorm(z_score)
print(sprintf("The p-value is %e", p_value))
```

```
## [1] "The p-value is 3.953322e-08"
```

With this p-value, we reject the null.

Problem 3.3: Pizza and ice cream

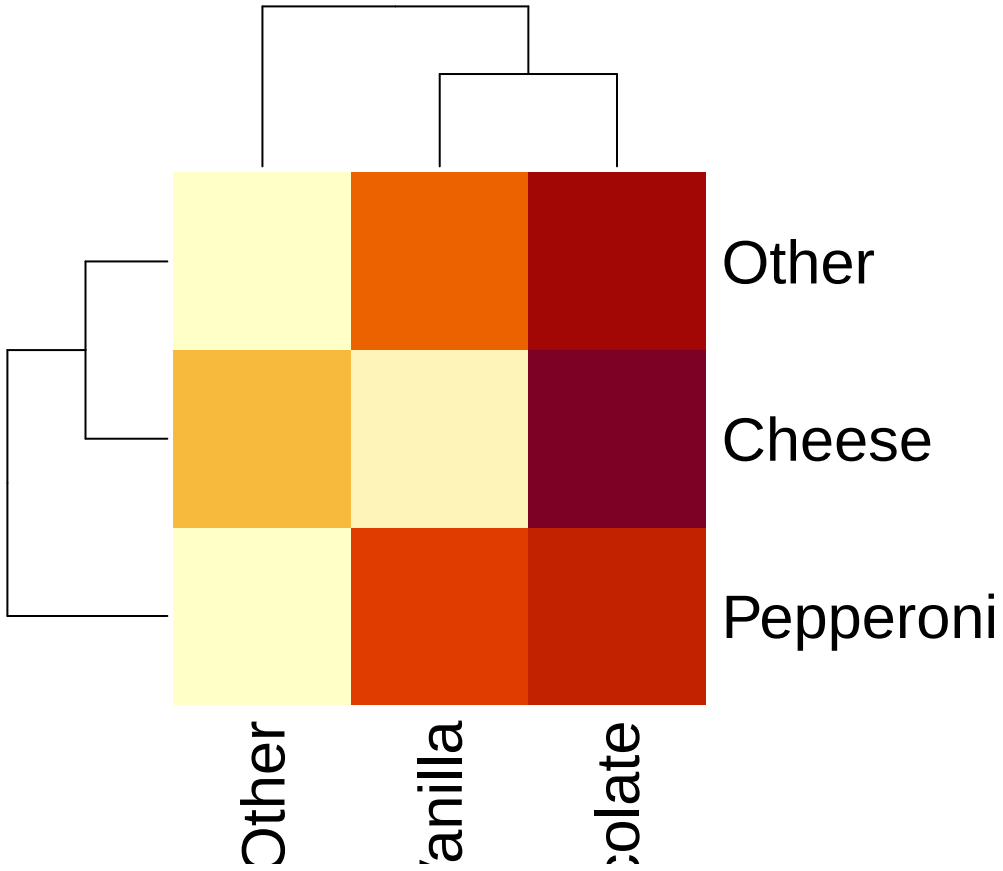
(i)

```
#load data
surveyData <- read.csv("pizza-ice-cream-abridged.csv")
colnames(surveyData) <- c("time", "ice_cream", "pizza")

#separate
pizza <- surveyData$pizza
ice_cream <- surveyData$ice_cream

#clean pizza
pizza <- replace(pizza, grepl("Cheese", pizza), "Cheese")
pizza <- replace(pizza, grepl("Pepperoni", pizza), "Pepperoni")
pizza <- replace(pizza, grepl("Other", pizza), "Other")

#clean ice cream
ice_cream <- replace(ice_cream, grepl("Vanilla", ice_cream), "Vanilla")
ice_cream <- replace(ice_cream, grepl("Chocolate", ice_cream), "Chocolate")
ice_cream <- replace(ice_cream, grepl("Other", ice_cream), "Other")

two_way_table <- table(pizza, ice_cream)
print(two_way_table)
```

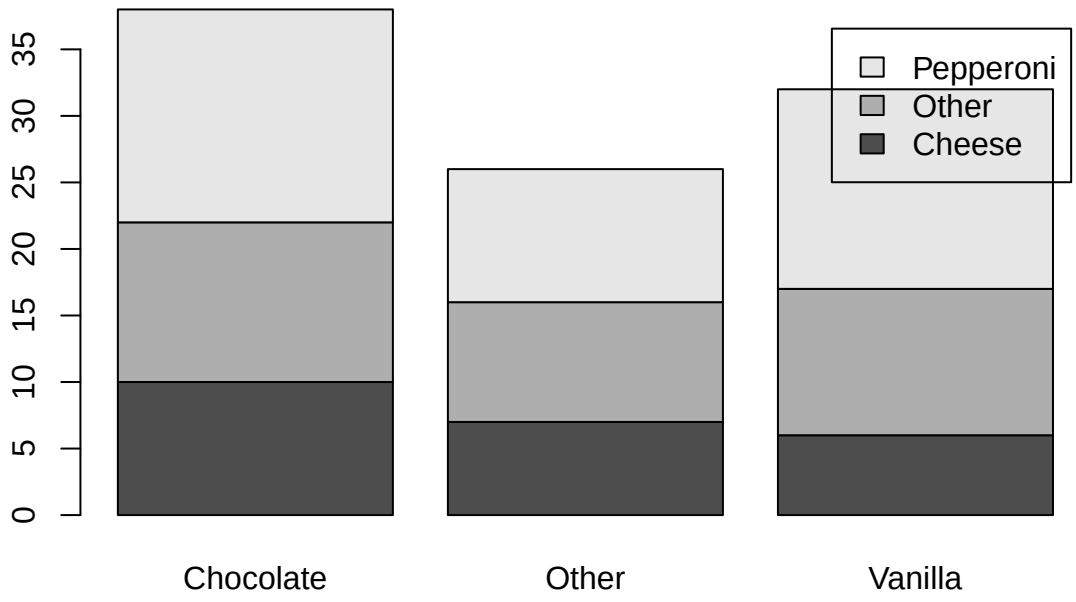
```
##            ice_cream
## pizza       Chocolate Other Vanilla
##   Cheese           10     7       6
##   Other            12     9      11
##   Pepperoni        16    10      15
```

(ii)

```
heatmap(two_way_table)
```




```
barplot(two_way_table, legend.text=TRUE)
```



(iii)

H_o: Favorite pizza and favorite ice cream are independent.
H_o: Favorite pizza and favorite ice cream are dependent.

```
chisq.test(two_way_table)
```

```
##
##  Pearson's Chi-squared test
##
## data:  two_way_table
## X-squared = 0.84729, df = 4, p-value = 0.932
```

With a p-value of 0.932, we fail to reject the null hypothesis.