Assignment_1

Ted Ladas

1/30/2021

Question 1

a)

Test for independence in a 2-way contingency table with Yate's Continuity correction.

b)

c)

d)

```
rm(list = ls())

# setting up the table and helper variables.
O <- as.table(rbind(c(688, 650),c(21, 59)))
dimnames(O) <- list(smokers = c("Smokers", "Nonsmokers"), lung_cancer=c("Cancer", "Control"))

# Manual Calculation of p-value
alpha <- 0.05
n<-sum(O)
row_n <- rowSums(O)
col_n <- colSums(O)
E <- outer(row_n,col_n)/n # expected cell counts
X2 <- sum((abs(O-E)-0.5)^2/E) # Yate's continuity correction
df <- (nrow(O)-1)*(ncol(O)-1) # df=1 because we have 2x2 matrix

p_manual <- 1 - pchisq(X2,df = df)
# cat('p-value manual calc:', p_manual)


# using the chisq.test method
p_auto <- chisq.test(O)$p.value
cat('\np-value auto   calc:', p_auto)


##
## p-value auto   calc: 2.057117e-05
```

```
diff = round(p_manual - p_auto, 8)
cat('\ncheck for diff between methods ', diff)
```

```
##
## check for diff between methods  0
```

```
if(diff == 0){
  cat('\nReject the null hypothesis under the 5% significance level'%s*%(p_auto<alpha))
  cat('\nNo evidence to reject the null hypothesis under the 5% significance level'%s*%(p_auto>alpha))
}
```

```
##
## Reject the null hypothesis under the 5% significance level
```

e)