Session-11-Assignment-2

Aditya Mulay

July 17, 2018

# Q1. Use the given link below and locate the bank marketing dataset. Data Set Link  
# Perform the below operations:  
Bank <- read.csv("C:/Users/Aditya/Downloads/bank-additional/bank-additional/bank-additional.csv", sep=";")  
# a. Is there any association between Job and default?  
# Ho : There is NO association between Job and default  
chisq.test(Bank$job, Bank$default, simulate.p.value = TRUE)

##   
## Pearson's Chi-squared test with simulated p-value (based on 2000  
## replicates)  
##   
## data: Bank$job and Bank$default  
## X-squared = 224.29, df = NA, p-value = 0.0004998

# Since the P value is less than 0.05,  
# there is association between Job and default at 95% confidence level

# b. Is there any significant difference in duration of last call between people having housing loan or not?  
# Ho : There is NO any significant difference in duration of last call between people having housing loan or not?  
chisq.test(Bank$duration, Bank$loan, simulate.p.value = TRUE)

##   
## Pearson's Chi-squared test with simulated p-value (based on 2000  
## replicates)  
##   
## data: Bank$duration and Bank$loan  
## X-squared = 1644.4, df = NA, p-value = 0.5652

# Since the P value is greater than 0.05,  
# There is NO significant difference in duration of last call between people having housing loan or not at 95% confidence level

# d. Is the employment variation rate consistent across job types?  
# Ho : There is NO employment variation rate consistent across job types  
chisq.test(Bank$emp.var.rate, Bank$job, simulate.p.value = TRUE)

##   
## Pearson's Chi-squared test with simulated p-value (based on 2000  
## replicates)  
##   
## data: Bank$emp.var.rate and Bank$job  
## X-squared = 512.04, df = NA, p-value = 0.0004998

# Since the P value is less than 0.05,  
# there is employment variation rate consistent across job types at 95% confidence level

# e. Is the employment variation rate same across education?  
# Ho : There is NO employment variation rate consistent across education  
chisq.test(Bank$emp.var.rate, Bank$education, simulate.p.value = TRUE)

##   
## Pearson's Chi-squared test with simulated p-value (based on 2000  
## replicates)  
##   
## data: Bank$emp.var.rate and Bank$education  
## X-squared = 193.46, df = NA, p-value = 0.004498

# Since the P value is less than 0.05,  
# there is employment variation rate consistent across education at 95% confidence level