5\_hypothesis\_testing.r

abhia

2019-11-13

fifa= read.csv("2\_standard.csv")  
View(fifa)  
  
# We know that the mean of the standardized data should be 0 and the sd should be 1  
print(sd(fifa$Overall))

## [1] 1

print(mean(fifa$Overall))

## [1] 2.149407e-17

# One sample t-test---------------------------  
# Ho : Mean = 0 [Null Hypothesis]  
# H1 : Mean != 0 [Alternate Hypothesis]  
t.test(fifa$Overall, mu = 0)

##   
## One Sample t-test  
##   
## data: fifa$Overall  
## t = 2.9002e-15, df = 18205, p-value = 1  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## -0.01452679 0.01452679  
## sample estimates:  
## mean of x   
## 2.149407e-17

# Ho : Mean =< 0 [Null Hypothesis]  
# H1 : Mean > 0 [Alternate Hypothesis]  
t.test(fifa$Value,alternative = 'greater', mu = 0)

##   
## One Sample t-test  
##   
## data: fifa$Value  
## t = 2.7216e-15, df = 18205, p-value = 0.5  
## alternative hypothesis: true mean is greater than 0  
## 95 percent confidence interval:  
## -0.01219108 Inf  
## sample estimates:  
## mean of x   
## 2.017087e-17

# Two sample t-test---------------------------  
  
# Ho : Mean1 = Mean2 => (Mean1 - Mean2) = 0 [Null Hypothesis]  
# H1 : Mean1 != Mean2 [Alternate Hypothesis]  
t.test(fifa$Overall,fifa$Potential,alternative = 'two.sided', mu = 0)

##   
## Welch Two Sample t-test  
##   
## data: fifa$Overall and fifa$Potential  
## t = 5.9323e-14, df = 36410, p-value = 1  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -0.0205433 0.0205433  
## sample estimates:  
## mean of x mean of y   
## 2.149407e-17 -6.002743e-16

# Ho : (Mean1 - Mean2) =< 0 [Null Hypothesis]  
# H1 : (Mean1 - Mean2) > 0 [Alternate Hypothesis]  
t.test(fifa$Overall,fifa$Potential,alternative = 'greater', mu = 0)

##   
## Welch Two Sample t-test  
##   
## data: fifa$Overall and fifa$Potential  
## t = 5.9323e-14, df = 36410, p-value = 0.5  
## alternative hypothesis: true difference in means is greater than 0  
## 95 percent confidence interval:  
## -0.01724035 Inf  
## sample estimates:  
## mean of x mean of y   
## 2.149407e-17 -6.002743e-16

# Chi square test for independence  
tab = table(fifa$Overall,fifa$Wage)  
chisq.test(fifa$Overall,fifa$Wage,correct = T)

## Warning in chisq.test(fifa$Overall, fifa$Wage, correct = T): Chi-squared  
## approximation may be incorrect

##   
## Pearson's Chi-squared test  
##   
## data: fifa$Overall and fifa$Wage  
## X-squared = 130713, df = 6721, p-value < 2.2e-16