ANLY510-50-SU2016\_Assignment-1

Dean D'souza

2016-06-04

# Assignment1: Homework Exercise 1

* **A political candidate conducted a poll and found 577 out of 900 respondents said that they would vote for her in a two-person race.**
* **What is the 95% confidence interval for her expected proportion of the vote?**

**Solution:** We can obtatin the confidence interval using the binconf() function in the Hmisc package in R as follows:

## PointEst Lower Upper  
## 0.6411111 0.6092341 0.6717886

# Assignment1: Homework Exercise 2

* **A retailer believes that its new advertising strategy will increase sales. Previously, the mean spending in 15 categories of consumer items in both the 18-34 and 35+ age groups was $90.00 with a standard deviation of $5.**

1. Write the Alternative and Null hypotheses to determine if the mean spending has increased.
2. After the advertising campaign, 250 customers were surveyed. The average spent by these individuals was $86.35 with a standard deviation of $23.90. Do you believe the advertising was successful?
3. For 500 respondents in the 35+ age group, mean and standard deviation was $88.53 and $18.29. Do you believe the advertising was successful?

**Solution:**

1. We write the alternative and null hypothesis as follows:

* H0: The mean spending in 15 categories of consumer items in both the 18-34 and 35+ age group remained the same. HA: The mean spending in 15 categories of consumer items in both the 18-34 and 35+ age group increased.
* or in other words:
* H0: mu=90
* HA: mu>90

1. We first try to obtain the p value as follows:

## [1] -2.41471

## [1] -1.644854

As the test statistic is less than the critical value, at 0.05 significance level, we can reject the claim that the sales increased. Hence, the advertising was not really successful.

1. We try to obtain the p-value as follows:

## [1] -1.797168

## [1] -1.644854

Again, as the test statistic is less than the critical value, at 0.05 significance level, we can reject the claim that the sales increased. Hence, the advertising was not that effective.

# Assignment1: Homework Exercise 3

* **Using the USArrestsCoasts data, is the East Coast states suffer more violent crime than all other states?**
* The USArrests data set is attached to the homework assignment in Moodle
* The East Coast states are marked as such
* Perform one tail t tests for murder, assault, and rape

**Solution:**

We perform the one tail t tests on murder, assault and rape as follows:

#Importing the data  
USArrests <- read.csv("C:/Users/Dean/Desktop/PrinAn2/USArrests\_Coasts.csv", row.names=1)  
  
#Subsettting data for ease of use  
EC<-subset(USArrests,USArrests$East.Coast=="Yes")  
othr<-subset(USArrests,USArrests$East.Coast=="No")  
  
# One tail t test on Murder  
t.test(EC$Murder,othr$Murder, alternative = "less")

##   
## Welch Two Sample t-test  
##   
## data: EC$Murder and othr$Murder  
## t = -0.10323, df = 21.811, p-value = 0.4594  
## alternative hypothesis: true difference in means is less than 0  
## 95 percent confidence interval:  
## -Inf 2.412987  
## sample estimates:  
## mean of x mean of y   
## 7.680000 7.834286

# One tail t test on Assault  
t.test(EC$Assault,othr$Assault, alternative="less")

##   
## Welch Two Sample t-test  
##   
## data: EC$Assault and othr$Assault  
## t = 0.3451, df = 24.496, p-value = 0.6335  
## alternative hypothesis: true difference in means is less than 0  
## 95 percent confidence interval:  
## -Inf 55.33992  
## sample estimates:  
## mean of x mean of y   
## 177.2667 167.9714

# One tail t test on Rape  
t.test(EC$Rape,othr$Rape, alternative="less")

##   
## Welch Two Sample t-test  
##   
## data: EC$Rape and othr$Rape  
## t = -1.846, df = 33.475, p-value = 0.03688  
## alternative hypothesis: true difference in means is less than 0  
## 95 percent confidence interval:  
## -Inf -0.397949  
## sample estimates:  
## mean of x mean of y   
## 17.90 22.66

From the above data we can conclude that the East coast states do suffer from more Violent crime such as Murder and Assault (as can be seen from the high p-values), however in terms of Rape it is significantly less.