Hypthosis Testing

Sunil Negi

October 9, 2016

## Hypthosis Testing

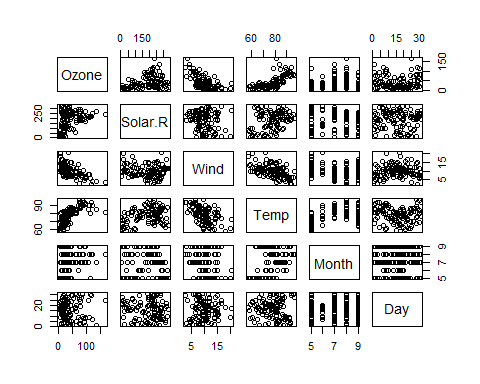
This Assignment is a for Hypthosis Testing using an example of Ozone Layer, with small dataset. The factors affecting ozone layer are Temperature, solar, wind. We are looking to calculate Correlation, Confidence Interval and prove null or alternate hypthosis testing.

setwd("C:/Users/SUNIL.NEGI/Documents")  
mydata<-read.csv("ozone\_data.csv")  
summary(lm(Ozone ~ Temp+Wind+Temp\*Wind, data = mydata))

##   
## Call:  
## lm(formula = Ozone ~ Temp + Wind + Temp \* Wind, data = mydata)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -40.930 -11.193 -3.034 8.193 97.456   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -239.8918 48.6200 -4.934 2.97e-06 \*\*\*  
## Temp 4.0005 0.5935 6.741 8.26e-10 \*\*\*  
## Wind 13.5975 4.2835 3.174 0.001961 \*\*   
## Temp:Wind -0.2173 0.0545 -3.987 0.000123 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 20.37 on 107 degrees of freedom  
## Multiple R-squared: 0.6355, Adjusted R-squared: 0.6253   
## F-statistic: 62.19 on 3 and 107 DF, p-value: < 2.2e-16

# Draws a correlation plot

pairs(mydata)



# Calculate the Pearson's correlation coefficient

cor(mydata)

## Ozone Solar.R Wind Temp Month  
## Ozone 1.000000000 0.34834169 -0.61249658 0.6985414 0.142885168  
## Solar.R 0.348341693 1.00000000 -0.12718345 0.2940876 -0.074066683  
## Wind -0.612496576 -0.12718345 1.00000000 -0.4971897 -0.194495804  
## Temp 0.698541410 0.29408764 -0.49718972 1.0000000 0.403971709  
## Month 0.142885168 -0.07406668 -0.19449580 0.4039717 1.000000000  
## Day -0.005189769 -0.05775380 0.04987102 -0.0965458 -0.009001079  
## Day  
## Ozone -0.005189769  
## Solar.R -0.057753801  
## Wind 0.049871017  
## Temp -0.096545800  
## Month -0.009001079  
## Day 1.000000000

# Confidence Interval

model1=lm(Ozone ~ Temp+Wind+Temp\*Wind, data = mydata)  
confint(model1,conf.level=0.95)

## 2.5 % 97.5 %  
## (Intercept) -336.2751998 -143.5084539  
## Temp 2.8240024 5.1770536  
## Wind 5.1059971 22.0889184  
## Temp:Wind -0.3253122 -0.1092398

# Statistical Inference

t.test(mydata,conf.level=0.90)

##   
## One Sample t-test  
##   
## data: mydata  
## t = 19.614, df = 665, p-value < 2.2e-16  
## alternative hypothesis: true mean is not equal to 0  
## 90 percent confidence interval:  
## 51.57128 61.02722  
## sample estimates:  
## mean of x   
## 56.29925

# just separating Ozone data from the data sheet

newdata <- mydata[c(1)]

Null Hypotheis H0 : Value of Ozone in population is > or = to 50 Alternate Hypotheis Ha: Value of Ozone in population in < 50

# Applying T test

t.test(newdata,alternative = "less", mu= 50)

##   
## One Sample t-test  
##   
## data: newdata  
## t = -2.5015, df = 110, p-value = 0.006919  
## alternative hypothesis: true mean is less than 50  
## 95 percent confidence interval:  
## -Inf 47.33835  
## sample estimates:  
## mean of x   
## 42.0991

Alpha = 1- confidence percentage (example 1-0.95 =0.05 if p < alpha then reject null hypothesis So in our example since p value is < alpha we have to reject null hypothesis Which means Value of Ozone in population is Less than 50