HYPOTHESIS TESTING

1. A F&B manager wants to determine whether there is any significant difference in the diameter of the cutlet between two units. A randomly selected sample of cutlets was collected from both units and measured? Analyze the data and draw inferences at 5% significance level. Please state the assumptions and tests that you carried out to check validity of the assumptions.

Step1: H0: mu1=mu2

Ha: mu1!=0mu2

Two Sample t-test

x<-c(6.809,6.4376, 6.9157, 7.3012, 7.4488,7.3871,

6.8755,

7.0621,

6.684,

6.8236,

7.393,

7.5169,

6.9246,

6.9256,

6.5797,

6.8394,

6.597,

7.2705,

7.2828,

7.3495,

6.9438,

7.156,

6.5341,

7.2854,

6.9952,

6.8568,

7.2163,

6.6801,

6.9431,

7.0852,

6.7794,

7.2783,

7.1561,

7.3943,

6.9405)

y<-c(6.7703,

7.5093,

6.73,

6.7878,

7.1522,

6.811,

7.2212,

6.6606,

7.2402,

7.0503,

6.881,

7.4059,

6.7652,

6.038,

7.1581,

7.024,

6.6672,

7.4314,

7.307,

6.7478,

6.8889,

7.422,

6.5217,

7.1688,

6.7594,

6.9399,

7.0133,

6.9182,

6.3346,

7.5459,

7.0992,

7.118,

6.6965,

6.578,

7.3875)

t.test(x,y,alternative = "two.sided")

data: x and y

t = 0.72287, df = 66.029, p-value = 0.4723

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-0.09654633 0.20613490

sample estimates:

mean of x mean of y

7.019091 6.964297

Conclusion: P=0.47>alpha Implies there is no significant difference the length of cutlets. In other words, accept null hypothesis

1. A hospital wants to determine whether there is any difference in the average Turn Around Time (TAT) of reports of the laboratories on their preferred list. They collected a random sample and recorded TAT for reports of 4 laboratories. TAT is defined as sample collected to report dispatch.

Analyze the data and determine whether there is any difference in average TAT among the different laboratories at 5% significance level.

library(nortest): “Error in installing”

# Hypothesis Part

# Ho : There is no significant difference between the average TAT values in the 4 libraries

# H1 : There is significant difference between the average TAT values in the 4 libraries

ad.test(labtat$Laboratory.1);ad.test(labtat$Laboratory.2);ad.test(labtat$Laboratory.3);ad.test(labtat$Laboratory.4)

1. Sales of products in four different regions is tabulated for males and females. Find if male-female buyer rations are similar across regions.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Observed Values | East | West | North | South |
| Males | 50 | 142 | 131 | 70 |
| Females | 435 | 1523 | 1356 | 750 |

R-code:

#3 SAles of different regions

BuyerRatio<-read.csv("BuyerRatio.csv")

#H0; Male-Female proportions are similar

#Ha; Male-Female proportions are not similar

chisq.test(BuyerRatio[,2:5])

Result:

chisq.test(BuyerRatio[,2:5])

Pearson's Chi-squared test

data: BuyerRatio[, 2:5]

X-squared = 1.5959, df = 3, p-value = 0.6603

P>0.05

Accept Null Hypothesis

i.e. male-female buyer rations are similar across regions

1. TeleCall uses 4 centers around the globe to process customer order forms. They audit a certain % of the customer order forms. Any error in order form renders it defective and has to be reworked before processing. The manager wants to check whether the defective % varies by centre. Please analyze the data at *5%* significance level and help the manager draw appropriate inferences

H0 =Defective % does not varies

Ha= Defective % varies

attach(`Costomer+OrderForm`)

summary(`Costomer+OrderForm`)

df=as.data.frame(cbind(Phillippines,Indonesia,Malta,India),rep(c("phillippines","Indonesia","Malta","India"),c(300,300,300,300)))

colnames(df)=c("Type","Location")

head(df)

df\_tab1=table(df$Type,df$Location)

#1 is for defective & 2 is for error free

chisq.test(t(df\_tab1))

result

Pearson's Chi-squared test with Yates' continuity correction

data: t(df\_tab1)

X-squared = 0.18564, df = 1, p-value = 0.6666

Warning msg: CHI Squared approximation may be incorrect

p-Value>0.05 Accept null hypothesis i.e Defective % does not varies

1. Fantaloons Sales managers commented that % of males versus females walking in to the store differ based on day of the week. Analyze the data and determine whether there is evidence at 5 % significance level to support this hypothesis."

Hypothesis Testing:---

# Ho : males versus females walking in to the store doesnot differ based on day of the week.

# Ha : males versus females walking in to the store differ based on day of the week.

head(Faltoons)

nrow(Faltoons)

attach(Faltoons)

df\_falt=as.data.frame(cbind(Weekdays,Weekend),rep(c("Weekdays","Weekend"),c(400,400)))

colnames(df\_falt)=c("Gender","Day")

head(df\_falt)

#1 for female, 2 for male

df\_tab2=table(df\_falt)

chisq.test(df\_tab2)

R-TEST

Pearson's Chi-squared test with Yates' continuity correction

data: df\_tab2

X-squared = 2.2781e-30, df = 1, p-value = 1

p>0.05; Accept Null Hypothesis