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: Business Problem

Whether there is any significant difference in the diameter of the cutlet between two units

Step2: Test normality of data

**null hypothesis(H0):Data are normal**

**alternate hypothesis(Ha):data are not normal**

if p-value is > 0.05 => Accept null hypothesis

if p-value is < 0.05 =>Reject null hypothesis

Normality test for Unit1: Shapiro-Wilk normality test

data: Unit1

W = 0.96495, p-value = 0.32

As p-value is 0.32 > 0.05 hence accept null hypothesis(H0)

Normality test for Unit2: Shapiro-Wilk normality test

data: Unit2

W = 0.97273, p-value = 0.5225

As p-value is 0.5225 > 0.05 hence accept null hypothesis(H0)

Step3: Check for Variance

**null hypothesis(H0):Equal variance**

**alternate hypothesis(Ha):Unequal variance**

compare two variances

data: Unit1 and Unit2

p-value = 0.3136

p-value is 0.3136 > 0.05=>P high Ho fly => Accept Ho, Hence we prove variance of Unit1 = variance of Unit2

Step4: **2 Sample t test for compare mean**

H0:Equal means of both units

Ha: Unequal means of both units

Welch Two Sample t-test

data: Unit1 and Unit2

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

p-value is 0.4723 > 0.05=>P high Ho fly => Accept Ho, Hence we prove mean of Unit1 = mean of Unit2.

As per above results we can say that there is similarity between Unit1 and Unit2.

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.

Step1: Business problem

Whether there is any difference in the average Turn Around Time (TAT) of reports of the laboratories on their preferred list

Step2: Test normality of data

Using Shapiro test for testing normality of data

Laboratory1: Shapiro-Wilk normality test

data: Laboratory1

W = 0.99018, p-value = 0.5508

As p-value 0.5508 > 0.05, data follows normal distribution.

Laboratory2: Shapiro-Wilk normality test

data: Laboratory2

W = 0.99363, p-value = 0.8637

As p-value 0.8637> 0.05, data follows normal distribution.

Laboratory3: Shapiro-Wilk normality test

data: Laboratory3

W = 0.98863, p-value = 0.4205

As p-value 0.4205> 0.05, data follows normal distribution.

Laboratory4: Shapiro-Wilk normality test

data: Laboratory4

W = 0.99138, p-value = 0.6619

As p-value 0.6619> 0.05, data follows normal distribution.

Step3: Testing data for equal variance

Using Bartlett test because multiple discrete variable and single continuous variable

Bartlett test of homogeneity of variances

data: values by ind

Bartlett's K-squared = 6.0995, df = 3, p-value = 0.1069

As p-value 0.1069> 0.05, accept H0- data has equal variance

Step4: Anova Test-One way

H0:Average of all laboratory are same

Ha:Average of atleast 1 laboratory are different

Df Sum Sq Mean Sq F value Pr(>F)

ind 3 79979 26660 118.7 <2e-16 \*\*\*

Residuals 476 106905 225

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

The p-value is lower than the usual threshold of 0.05. We can say this as the statistical difference between the groups are marked by a ”\*”. Thus accept Ha:Average of atleast 1 laboratory is different.

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **East** | **West** | **North** | **South** |
| Males | 50 | 142 | 131 | 70 |
| Females | 550 | 351 | 480 | 350 |

Step1: Business Problem

Find if male-female buyer rations are similar across regions.

Step2: Reform the data so that it is easy to operate on

Here, transpose function is used so that the data is easier to understand and to operate.

Step3:Test Normality of data

Shapiro-Wilk normality test

data: Males

W = 0.88076, p-value = 0.3419

As p-value 0.3419> 0.05, data follows normal distribution.

Shapiro-Wilk normality test

data: Females

W = 0.92197, p-value = 0.548

As p-value 0.548> 0.05, data follows normal distribution.

Step-4: Here we perform the Chi-squared test

H0:All averages are same

Ha:atleast 1 are different

Pearson's Chi-squared test

data: data

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

P-value is 0.6603 > 0.05=>P high Ho fly => Accept Ho, hence Average are same

As per results we can say that there is proportion of male and female buying is similar.

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

Step1: Business problem

check whether the defective % varies by centre.

Step-2: Perform Test on Data

As there are multiple x discrete and y discrete, we go with chi-squared test

Pearson's Chi-squared test

data: table(Defective, Country)

X-squared = 3.859, df = 3, p-value = 0.2771

P-value is 0.2771 > 0.05=>P high Ho fly => Accept Ho, hence Average are same

As per results we can say that the % of defectives are all the same at all centres.

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.

Step1: Business problem

*%* of males versus females walking in to the store differ based on day of the week.

Step2: Perform test on data

As we are comparing 2 population with each other, we go with 2-proportion test

H0:Proportion of male vs female in weekdays = Proportion of male vs female in weekends

Ha:Proportion of male vs female in weekdays NOT = Proportion of male vs female in weekends

2-sample test for equality of proportions without continuity correction

data: c(66, 47) out of c(233, 167)

X-squared = 0.0015979, df = 1, p-value = 0.9681

alternative hypothesis: two.sided

95 percent confidence interval:

-0.08761305 0.09126240

sample estimates:

prop 1 prop 2

0.2832618 0.2814371

P-value is 0.9681 > 0.05=>P high Ho fly => Accept Ho, hence Average are same

As per results we can say that the Proportion of male vs female in weekdays = Proportion of male vs female in weekends.