VERSION 1.0

MARCH 12, 2017

GREAT LAKES

INSTITUTE OF MANAGEMENT, CHENNAI

THE TITAN INSURANCE COMPANY

CASE STUDY

PRESENTED BY: SMDM GROUP 13

Kishore Kumar Alajangi
Srikant Deverajan
Kuchibhotla V S R Krishna Murthy
Hari Prasad. A
Sundar Kumar

THE TITAN INSURANCE COMPANY

A. DESCRIBE THE FIVE PER CENT SIGNIFICANCE TEST YOU WOULD APPLY TO THESE DATA TO DETERMINE WHETHER NEW SCHEME HAS SIGNIFICANTLY RAISED OUTPUTS?

We shall use Paired t-test for this case as we are interested in evaluating the effectiveness of new scheme that has been Implemented in Titan. We are to infer whether the average output has increased or not. Hence this is a case of two population means where you have two samples in which observations in one sample can be paired with observations in the other sample i.e. Before-and-after execution of the scheme.

B. WHAT CONCLUSION DOES THE TEST LEAD TO?

We have used R as the tool to arrive at a decision;

Since we have to find whether the new scheme has significantly raised output we can formulate the hypotheses as follows:

$$\text{Null Hypothesis->} \quad H_0 : \mu_1 = \mu_2$$

$$\text{Alternate Hypothesis->} \quad H_1 : \mu_1 < \mu_2$$

Following is the R code for the same:

Calculating the paired t-test for the 2 samples

<<R-Code>>

```
paired.t.test=t.test(Old.Scheme, New.Scheme, paired = T, alternative = "less")
```

<<R-Output>>

```
Paired t-test

data:  Old.Scheme and New.Scheme
t = -1.5559, df = 29, p-value = 0.06529
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
-Inf 368.1762
sample estimates:
mean of the differences -4000
```

Inference:

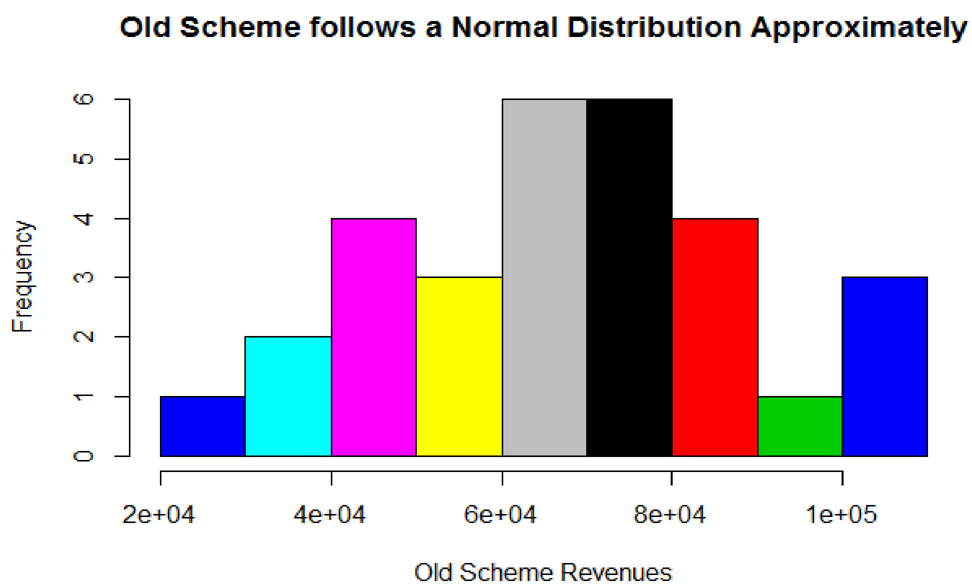
The p-value 0.06529 is greater that 0.05 significant level of the test, hence we accept the Null Hypothesis with 95% confidence.

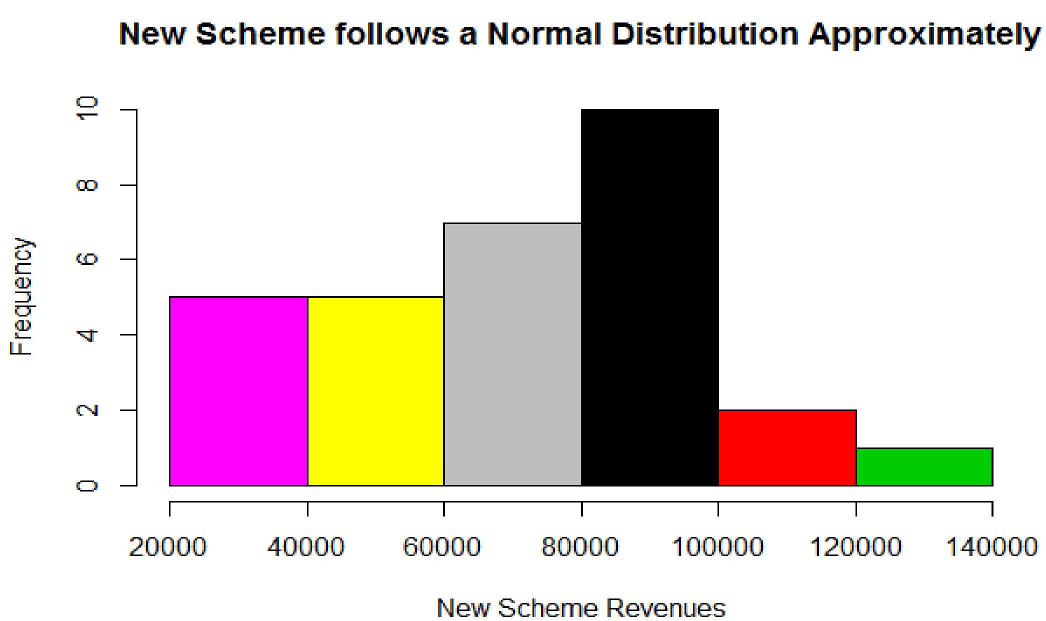
This implies that there is no significant increase in the average output after implementation of the new scheme. Hence the Titan Insurance company can stick to the older scheme which would be less expensive than the new scheme.

C. WHAT RESERVATIONS HAVE YOU ABOUT THIS RESULT?

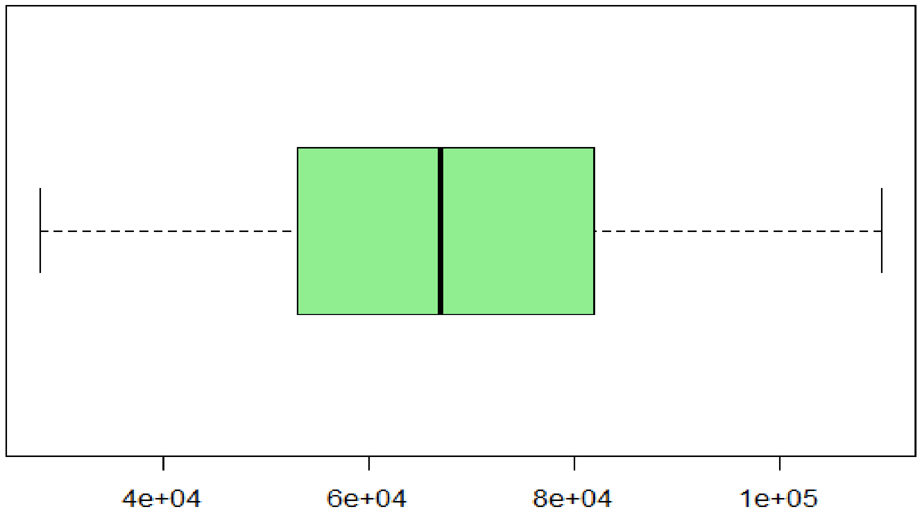
The following reservations are there about the result while performing the paired t-test

- The p-value of 0.0625 implies that there is a weak evidence of rejecting the alternate hypothesis.
- The dependent variable must be continuous (interval/ratio).
- The observations are independent of one another.
- The dependent variable should be approximately normally distributed.

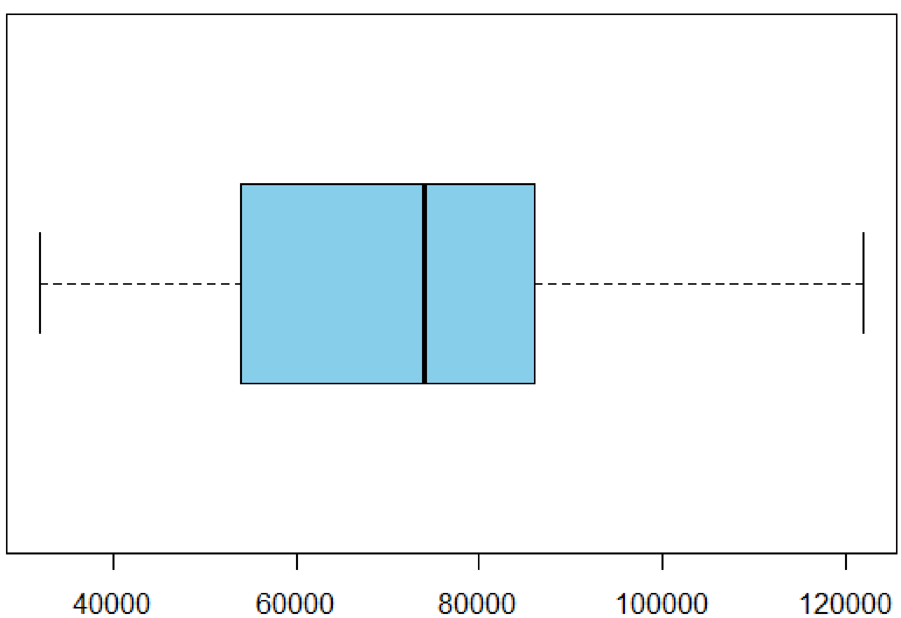




Old Scheme has no outliers



New Scheme has no outliers



D. SUPPOSE IT HAS BEEN CALCULATED THAT IN ORDER FOR TITAN TO BREAK EVEN, THE AVERAGE OUTPUT MUST INCREASE BY £5000. IF THIS FIGURE IS ALTERNATIVE HYPOTHESIS, WHAT IS:

(I) THE PROBABILITY OF A TYPE 1 ERROR?

As we have taken a significance level of 5%, hence the probability of Type 1 error is 5%

(II) THE POWER OF THE TEST?

The power of the test is 0.4592 i.e. 45.92%

<<R CODE>>

```
# finding out Std deviation for the samples
SD1=sd(Old.Scheme)
SD2=sd(New.Scheme)
# Calculating pooled SD
pooledSD=sqrt((SD1^2+SD2^2)/(2))
# Calculating the power of the test
# As the question states that the avg mean has to increase by 5000 in the new scheme to
  achieve  breakeven, Hence 5000 needs to be added to 4000(mean diff of existing sample)
power.t.test(n=30,delta=9000,sd=pooledSD,power = NULL,sig.level = 0.05, type
"two.sample", alternative = "one.sided")
```

<<R OUTPUT>>

```
n = 30
delta = 9000
 sd = 22332.11
sig.level = 0.05
power = 0.4592797
alternative = one.sided
```

(III) THE PROBABILITY OF TYPE 2 ERROR?

The Probability of type 2 error is (1-power) of the test

Power=1-0.4592

= 0.5408

=54.08%

E. WHAT SAMPLE SIZE WOULD MAKE THE PROBABILITIES OF TYPE 1 AND TYPE 2 ERRORS EQUAL?

The number of sample when the Type 1 is equal to Type 2 error is 133.948~134 Nos.

<<R CODE>>

```
# calculating the no. of sample when the Type 1 and Type 2 error is equal
power.t.test(n=NULL,delta = 9000,sd = pooledSD,power = 0.95,sig.level = 0.05, type = "two.sample"
        ,alternative = "one.sided")
```

<<R OUTPUT>>

```
n = 133.948
delta = 9000
sd = 22332.11
sig.level = 0.05
power = 0.95
alternative = one.sided
```

PS: Refer Annexure I for entire program of R

ANNEXURE I

R-CODE

```
# Assignment on Titan Insurance company
#set working directory
setwd("F:/1.00 Home/4.00 PERSONAL/4.00 GreatLakes/1.00 1st Residency 18-22/R Classes")
#Read the data from the source
mydata=read.csv("Titan.csv",header = T)
attach(mydata) # attaching mydata to R DB
mydata
# Calculating the paired t-test for the 2 samples
Paired.t.test=t.test(Old.Scheme, New.Scheme, paired = T, alternative = "less")
# Displaying the result
Paired.t.test
# finding out Std deviation for the samples
SD1=sd(Old.Scheme)
SD1
SD2=sd(New.Scheme)
SD2
# Calculating pooled SD
pooledSD=sqrt((SD1^2+SD2^2)/(2))
pooledSD
# Calculating the power of the test
# As the question states that the avg mean has to increase by 5000 in the new scheme to achieve breakeven, Hence 5000 needs to be added to 4000(mean diff of existing sample)
power.t.test(n=30,delta=9000,sd=pooledSD,power = NULL,sig.level = 0.05, type = "two.sample" ,alternative = "one.sided")
```

```
# Calculating the probability of Type 2 error (1-power)
1-0.459
# Probability of type 2 error = 0.5408
# calculating the no. of sample when the Type 1 and Type 2 error is equal
power.t.test(n=NULL,delta = 9000,sd = pooledSD,power = 0.95,sig.level = 0.05, type = "two.sample",alternative = "one.sided")
# Histogram to check out the normality of old scheme
hist(Old.Scheme, main = "Old Scheme follows a Normal Distribution Approximately", col = c(20:50), ylab = "Frequency", xlab = "Old Scheme")
# Histogram to check out the normality of New Scheme
hist(New.Scheme, main = "New Scheme follows a Normal Distribution Approximately", col = c(30:60), ylab = "Frequency", xlab = "Old Scheme")
#Boxplot to check out the presence of outliers in Old scheme
boxplot(Old.Scheme,main="Old Scheme has No Outliers",col="light green")
#Boxplot to check out the presence of outliers in New Scheme
boxplot(New.Scheme,main="New Scheme has No Outliers",col="Sky Blue")
```

R-OUTPUT

```
> mydata
   Old.Scheme New.Scheme
1       57000      62000
2      103000     122000
3       59000      54000
4       75000      82000
5       84000      84000
6       73000      86000
7       35000      32000
8      110000     104000
9       44000      38000
10      82000     107000
11      67000      84000
12      64000      85000
13      78000      99000
14      53000      39000
15      41000      34000
16      39000      58000
17      80000      73000
18      87000      53000
19      73000      66000
20      65000      78000
21      28000      41000
22      62000      71000
23      49000      38000
24      84000      95000
25      63000      81000
26      77000      58000
27      67000      75000
28     101000      94000
29      91000     100000
30      50000      68000
> Paired.t.test=t.test(Old.Scheme, New.Scheme, paired = T, alternative = "les
s")
> Paired.t.test

        Paired t-test

data:  Old.Scheme and New.Scheme
t = -1.5559, df = 29, p-value = 0.06529
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
     -Inf 368.1762
sample estimates:
mean of the differences
                  -4000

> SD1=sd(Old.Scheme)
> SD1
[1] 20455.98
> SD2=sd(New.Scheme)
> SD2
[1] 24062.39
> pooledSD=sqrt((SD1^2+SD2^2)/(2))
```

```
> pooledSD
[1] 22332.11
> power.t.test(n=30,delta=9000,sd=pooledSD,power = NULL,sig.level = 0.05, typ
e = "two.sample"
+               ,alternative = "one.sided")

     Two-sample t test power calculation

              n = 30
          delta = 9000
             sd = 22332.11
      sig.level = 0.05
          power = 0.4592797
    alternative = one.sided

NOTE: n is number in *each* group

> 1-0.459
[1] 0.541
> power.t.test(n=NULL,delta = 9000,sd = pooledSD,power = 0.95,sig.level = 0.0
5, type = "two.sample"
+               ,alternative = "one.sided")

     Two-sample t test power calculation

              n = 133.948
          delta = 9000
             sd = 22332.11
      sig.level = 0.05
          power = 0.95
    alternative = one.sided

NOTE: n is number in *each* group

> hist(Old.Scheme,
+      main = "Old Scheme follows a Normal Distribution Approximately",
+      col = c(20:50), ylab = "Frequency", xlab = "Old Scheme")
> hist(New.Scheme,
+      main = "New Scheme follows a Normal Distribution Approximately",
+      col = c(30:60), ylab = "Frequency", xlab = "Old Scheme")
> boxplot(Old.Scheme,main="Old Scheme has No Outliers",col="light green")
> boxplot(New.Scheme,main="New Scheme has No Outliers",col="Sky Blue")
```