|  |
| --- |
| SMDM Group Assignment |
| **Case Study- Titan Insurance Company.** |

|  |  |
| --- | --- |
| **Group No:** | 8 |
| **Team Members** | Avinash Kumar S  Meenu Sethuraman  Shiyam Sundar T K  Subashree Bhaskaran |

# **Synopsis:**

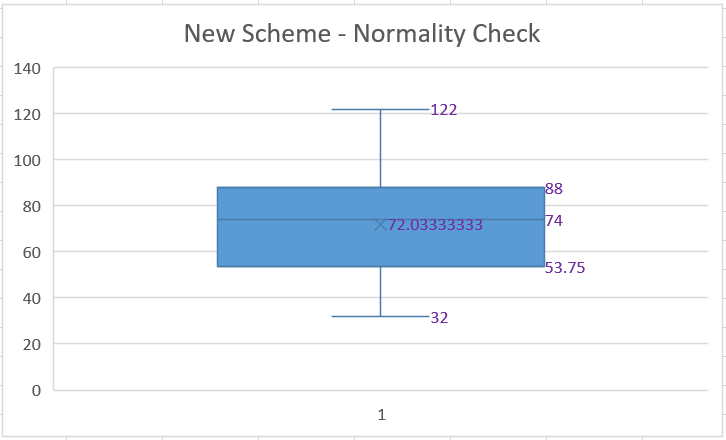
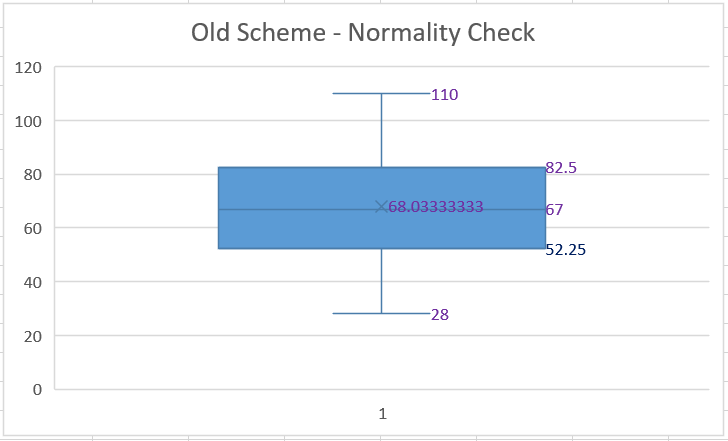
The Titan Insurance Company has introduced a new incentive payment system. We have the data of 30 salespersons’ output for Old Scheme and New Scheme. The data is stable enough, as it is clear from the Case Study, that the fluctuations (due to changeover) are avoided. The company believes the New Scheme is expensive and hence wants to increase sales to compensate for the expenses. We analysed the data provided to help make a decision to continue with the New Scheme or to abandon.

## **Analysis:**

### ***Describe the five per cent significance test you would apply to these data to determine whether new scheme has significantly raised outputs?***

To check for the normality, Old Scheme and New Scheme data are plotted as Box and Whisker Plots individually.

Figure 1: Normality Check using ox and Whisker plots



Skewness in Old Scheme: 0.0401647 Skewness in New Scheme: -0.02355928

* From the Figure 1 above and *Skewness* being close to 0, it can be accounted as a normalized data.
* As the sample Size is 30, by *Central Limit Theorem*, we can assume the data is from normalized population.

Now, t Statistics can be computed for further analysis.

Please note that the data sets (Old Scheme & New Scheme) are from same Salesperson, hence, we can go for a **Paired t-test.**

Let us define the Null & Alternative Hypothesis.

To understand if the New Scheme has significantly raised output, we need to understand if mean difference between these 2 Schemes is “less than 0”. Hence, this becomes the Alternative hypothesis.

|  |
| --- |
| H0: µold - µnew >= 0  H1: µold - µnew < 0 |

Let us run the 5% significant test in Rstudio:

|  |
| --- |
| my\_data = read.csv("TitanInsurace\_OneSample2Tests.csv")  attach(my\_data)  t.test(Old.Scheme, New.Scheme, alternative = "less", paired = TRUE) |

The output is:

|  |
| --- |
| >t.test(Old.Scheme,New.Scheme,alternative = "less", paired = TRUE)  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 0.3681762  sample estimates:  mean of the differences  -4 |

### ***What conclusion does the test lead to?***

* From the above Paired t-test result, the obtained p-value of 0.06529 is > 0.05 (5% significant level), hence, the null hypothesis H0 stays.
* By this, it can be concluded that there is no sufficient evidence available to substantiate the claim that the new scheme has significantly raised outputs.
* Usually, if there is greater than a 5% chance of a result when the null hypothesis is true, then the null hypothesis is retained. This does not necessarily mean that the accepted null

hypothesis is true—only that there is no enough evidence to conclude that it is true. We often use the expression “fail to reject the null hypothesis”

### ***What reservations do you have about this result?***

* From the sample data collected from old scheme and new scheme, it can be seen that the corresponding mean values are 68.03 units and 72.03 units. This approximately signifies that the new scheme has made more output compared to old scheme. But upon testing the null hypothesis, the result arrived is otherwise.
* This difference in result can be attributed to the skewness in the data. Old scheme roughly follows a normal bell curve, but the new scheme doesn’t. It is for this fact that, we use a “T Test”.
* If the sample size be increased with more data, it may, in a way end up proving the alternative hypothesis.
* If the number of sample size is large, we may have precise information as to the value of the mean, but if our sample be small, we have two sources of uncertainty:
  + - The error of random sampling, where the mean of our sample deviates more or less widely from the mean of the population, and
    - The sample is not sufficiently large to determine what the law of distribution of individuals is.

### ***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?***

* The null hypothesis testing above is tested with the condition that the null hypothesis is “TRUE”, with the level of significance at 5%.
* The definition of Type 1 error is the chance of rejecting the null hypothesis when it is true.
* Hence, in this case, the probability or chance of committing a **Type 1 error is 5%.**

***(ii) The probability of a type 2 error?***

* The specified alternative hypothesis of the actual calculated break even point of 5,000 pound sterling is the false null hypothesis.
* In order to calculate the type 2 error, we need to find the for the critical value at α=5%. It can be computed with the formula, , where Tstat = -1.699127, µD = 0, SD = 14.08 and n = 30.
* Based on this, stands out to be -4.36.
* Type 2 error can be denoted as P(Accepting false H0 | µD = -5). Here, µD as -5 is for the reason that True H0 is based on comparison between Old Scheme vs New Scheme, where the actual value is a relation between New Scheme vs Old Scheme. Hence, for the false H0, the sign is change is from 5 to -5.
* Tstat for the false null hypothesis is computed based on the formula , with

= -4.36, µD = -5, SD = 14.08 and n = 30 and that is 0.2457656.

* Based on this, the **probability of Type 2 error** is “1 - P(Rejecting false H0 | µD = -5)”, which is 1-0.5962, which is **0.4038 or 40.38%**.

***(iii) The power of test?***

* + The power of test is defined as probability of correctly rejecting the false Null Hypothesis and that can be denoted as “P(Rejecting false H0 | µD = -5)”.
  + Hence, as calculated above, the **power of test is 0.5962 or 59.62%**.

### ***Are Type 1 and Type 2 errors the same in this case? Should they be equal? Why or why not? If they are to be equated, suggest a way to do so. (Hint: would a change in sample size work?)"***

* Type 1 and Type 2 errors are not the same in this case.
* They should not be equal.
* If the Type 1 and Type 2 errors are made equal, it signifies that, when the null hypothesis is true, the confidence interval gets drastically reduced from the originally employed level of 95%. Alternatively, when the null hypothesis is false, the power of test gets drastically changed. This is because of the fact that, Probability of Type 1 and Type 2 error is inversely related, meaning, with an increase in Type 1 error, the Type 2 error gets decreased. Both the errors may match at one point, but the repercussions are huge as explained above.
* If they are to be equated, this can be done by increasing the sample size or decreasing the standard difference.

1. Increasing the sample size will result in decreasing the standard error. This means the confidence interval gets narrowed down to the mean difference. The increased sample size needed for such instance will have to be calculated based on trial and error method.
2. The other way of equating the two errors, can be done by taking the values in the 5th month hoping that there won’t be significant deviation in the output of sales persons.

In this way, the standard difference can possibly come down. This might influence the probability of Type 1 and Type 2 errors.