Table of Contents

1 The Assignment

Case Study: Titan Insurance Company

The Titan Insurance Company has just installed a new incentive payment scheme for its lift policy sales force. It wants to have an early view of the success or failure of the new scheme. Indications are that the sales force is selling more policies but sales always vary in an unpredictable pattern from month to month and it is not clear that the scheme has made a significant difference.

Life Insurance companies typically measure the monthly output of a salesperson as the total sum assured for the policies sold by that person during the month. For example, suppose salesperson X has, in the month, sold seven policies for which the sums assured are £1000, £2500, £3000, £5000, £10000, £35000. X's output for the month is the total of these sums assured, £61,500. Titan's new scheme is that the sales force receive low regular salaries but are paid large bonuses related to their output (i.e. to the total sum assured of policies sold by them). The scheme is expensive for the company but they are looking for sales increases which more than compensate. The agreement with the sales force is that if the scheme does not at least break even for the company, it will be abandoned after six months.

The scheme has now been in operation for four months. It has settled down after fluctuations in the first two months due to the changeover. To test the effectiveness of the scheme, Titan have taken a random sample of 30 salespeople measured their output in the penultimate month prior to changeover and then measured it in the fourth month after the changeover (they have deliberately chosen months not too close to the changeover). The outputs of the salespeople are shown in Table 1

Questions:

  a) Describe the five per cent significance test you would apply to these data to determine whether new scheme has significantly raised outputs?
  b) What conclusion does the test lead to?
  c) What reservations have you about this result?
  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:
      a. The probability of a type 1 error?
      b. The probability of a type 2 error?
      c. The power of the test?
  e) What sample size would make the probabilities of type 1 and type 2 errors equal?

Table 1: Titan Insurance Output (in £000)

| Salesperson	Old Scheme	New Scheme
1	57	62
2	103	122
3	59	54
4	75	82
5	84	84
6	73	86
7	35	32
8	110	104
9	44	38
10	82	107
11	67	84
12	64	85
13	78	99
14	53	39
15	41	34
16	39	58
17	80	73
18	87	53
19	73	66
20	65	78
21	28	41
22	62	71
23	49	38
24	84	95
25	63	81
26	77	58
27	67	75
28	101	94
29	91	100
30	50	68

2 The Solution

2.1 Preliminary Analysis

- Sample Size: 30, No. of samples: 2
- Measurement repeated on the same sample after 4 months.
- Five Point Summary and Standard Deviation on both the samples

```
> summary(Mydata)
   Old_Scheme       New_Scheme
 Min.   : 28.00   Min.   : 32.00
 1st Qu.: 54.00   1st Qu.: 55.00
 Median : 67.00   Median : 74.00
 Mean   : 68.03   Mean   : 72.03
 3rd Qu.: 81.50   3rd Qu.: 85.75
 Max.   :110.00   Max.   :122.00
> sd(Old_Scheme)
[1] 20.45598
> sd(New_Scheme)
[1] 24.06239
> var(Old_Scheme)
[1] 418.4471
> var(New_Scheme)
[1] 578.9989
> sd(New_Scheme - Old_Scheme)
[1] 14.08105
```

- Histogram and Box Plot for both the schemes:
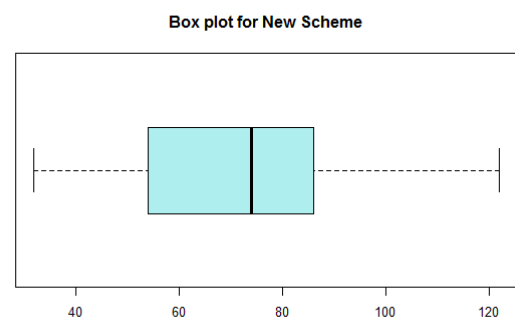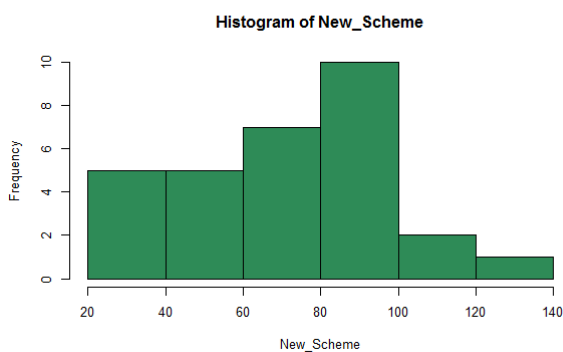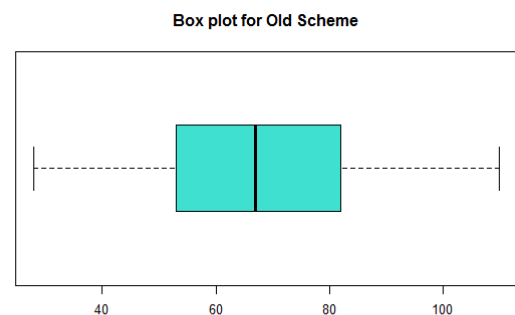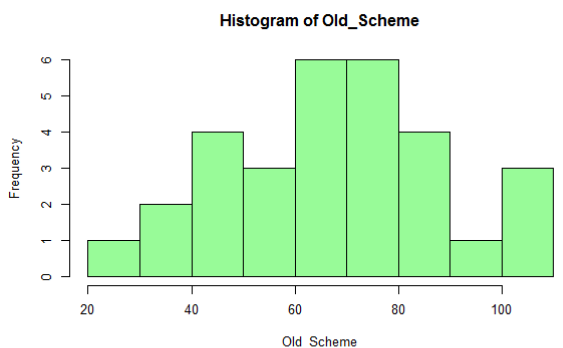
- Observations:
    o Both the samples seems to be normally distributed.
    o Mean and Median Values are not much different.

- o The Old scheme data looks more normally distributed, whereas the New scheme data looks left skewed. there is a dip in the performance.

2.2 Question a and b:

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

- The level of significance (Alpha) = 0.05.
- The sample size , N = 30 which is sufficiently large for a Zstat Test.
- But since the population standard deviation (Sigma) is unknown, we have to use a **Tstat test.**
- Degree of Freedom: Since the sample is the same for both Sampling tests, we have **N-1 degrees of freedom : 29**
- Since the sole purpose of the test is to check whether the New scheme is successful compared to old scheme, we would prefer a **Right tailed T Test.**

- **Hypothesis Formulation**
- Use the T Test: Paired two Sample for Means:

> H_0 : μold $-$ μnew = 0 (New Scheme did not raise the output significantly)
>
> H_1 : μold $-$ μnew < 0 (New Scheme raised the output significantly)

- T Test Equation:

$$t_{stat} = \frac{(\bar{d} - \mu D)/S_d}{\sqrt{n}} \text{ Here } S_d = \sqrt{\sum \frac{(D - \overline{D})^2}{n-1}} \text{ and } d_f = n - 1$$

- The T Test Result from R:

```
> t.test(Old_Scheme,New_Scheme,
+         paired = TRUE,
+         conf.level = 0.95,
+         alternative = "less")

        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
```

2.2.2 What conclusion does the Test lead to ?

- **Conclusion:**
  - In this scenario, the p value is 0.065 which is greater than the 0.05. Hence **accept the null hypothesis** that the new scheme did not raise the output significantly.

2.3 Question c:

2.3.1 What reservations do you have about this result:

- If we compare the means of the two sample distributions, we see that even though visually it seems as if New scheme has fared better, statistically it does not hold.
- The higher mean in the case of New scheme can also be attributed to the higher variance compared to Old Scheme.
- The variance of the second data set, 579.00, is 38% more than the variance of the first data set, 418.45.
- We are unsure of the sampling error present in the data.
- Unpredictable pattern persist even after the bonus scheme direct the management to work on the lift (product portfolio) and to correct the bonus scheme in proportion to revenue generation rather than policy sale.
- The scheme is expensive for the company but they are looking to compensate it by increase in sales. The agreement with the sales force is that if the scheme does not at least break even for the company, it will be abandoned after six months. So, it's better for the management to wait for another two months.
- There is dip in the New Scheme towards end of the cycle, which suggests that the Sales team may have booked all their sales at the earlier part of the period to get their incentives early or may be waiting till the end of quarter (measurement period) to book all their sales at once. Hence aggregation of data for a quarter may provide more clarity on the performance than for a single month.

2.4 Question d:

2.4.1 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:

2.4.1.1 The probability of a type 1 error?

2.4.1.2 The probability of a type 2 error?

2.4.1.3 The power of the test?

- **Type I error (α):** Probability of rejecting the null hypothesis when it is true. The probability of a Type I error in hypothesis testing is predetermined by the **significance level**.

- **Type II error (β) :** Probability of failing to reject null when it is false. Type II error calculation **depends on the population mean which is unknown.**

Question :

- If alternate hypothesis μ(New Scheme) – μ(Old Scheme) = μ_d = £5000 is true. What is the type I and II error.
- Null Hypothesis = μ(New Scheme) – μ(Old Scheme) = μ_d = 0
- First we need to calculate the probability of Type I error which is predetermined by significance level. If the significance level is 0.05,
- then type I error is 0.05. i.e. 5% probability we make Type I error – rejecting null hypothesis when it is true.
- Type II error calculation depends on a particular value of μ. In this case lets assume difference of population μ is £5K. Lets also assume that the significance level for the test is 0.05. Then the calculation is as below :
- This is a right tailed test
- We fail to reject the null hypothesis (commit a type II error) if we get a T-statistic less than **1.699127** for the sample size of 30

```
> # Tcritical at 5% confidence level with 29 Degree of Freedom:
> abs(qt(0.05,29))
[1] 1.699127
```

- To begin with, we need to find the $\overline{D}$ until we fail to reject the null hypothesis. Then we need to identify the probability of calculated $\overline{D}$ appearing when the μ_d = £5000.
- The formula for Tstat is:

$$t_{stat} = \frac{(\bar{d} - \mu D)/S_d}{\sqrt{n}}$$

Hence,

$$1.699 = \frac{\dfrac{\bar{d} - 0}{14.081}}{\sqrt{30}}$$

Or

$$1.699 = \frac{\overline{d}}{2.571}$$

Hence,

$$\bar{d} = 4.37$$

The above calculation states that we will incorrectly fail to reject the null hypothesis as long as we draw samples where the difference of mean is less than £4.37K.

Now given μold – μnew = 5K, we can compute the probability of drawing the sample mean difference lesser than 4.37K as follows:

$$t_{stat} = \frac{(\bar{d} - \mu D)/S_d}{\sqrt{n}}$$

Hence

$$t_{stat} = \frac{4.37 - 5}{2.571}$$

Or,

$$t_{stat} = -0.24504$$

Hence, Probability (p value, or β) =

```
> # Calculate p value for a given T value: -0.24504 and DOF: 29
> pt(-0.24504,29, lower = TRUE)
[1] 0.4040756
```

Hence, Probability of making Type II Error is 40.40%

Power of the Test:

Power of the hypothesis test is 1 minus the probability of Type II error.

Basically the power of a test is the probability that we make the right decision when the null is not correct (i.e. we correctly reject it)

Power of test is 1- beta = 1-0.4041 = 0.5959 . i.e. **59.59%**

2.5 Question e:

2.5.1 What sample size would make the probabilities of type 1 and type 2 errors equal?
Let us assume we need Type 1 error and Type 2 error equal to .05

Assuming sample standard deviation is equal to population standard deviation, we can calculate sample size needed as below :

- Null hypothesis mean difference (M_0) is 0

- Alternative hypothesis mean difference (M_a) is 5

- Sample standard Deviation is 14.081

- Alpha value is 0.05

Beta value is 0.05 i.e. power of test is 0.95 = 95%

Formula is:

$$n = \frac{S^2(t\alpha + t\beta)^2}{(\mu_0 - \mu_a)^2}$$

Sample Size calculation in R:

```
> power.t.test(power = 0.95,
+              delta=5,
+              sd=14.081,
+              sig.level = 0.05,
+              type = "paired",
+              alternative = "one.sided")

    Paired t test power calculation

          n = 87.20281
      delta = 5
         sd = 14.081
  sig.level = 0.05
      power = 0.95
alternative = one.sided

NOTE: n is number of *pairs*, sd is std.dev. of *differences* wit
hin pairs
```

Hence, in order to retain the power, we need to round the value to next whole number. therefore, we may conclude that we need a sample size of 88 to get the Type I and Type II Errors equal.

3 Conclusion:

From the given data, it may be concluded that, statistically there is no significant increase in sales because of the new incentive plan. However, our recommendation is that the test be carried out with a larger sample size covering a longer test period (at least a quarter) to improve the accuracy of test result and negating any one-time effect on the sales performance. Also, the results need to be interpreted and future actions be planned with the understanding of overall business performance.