**Hypothesis Testing**

Problem Statement:-

You are working for a consumer finance company. When the company receives a loan application, the company has to make a decision for loan approval based on the applicant’s profile. Two types of risks are associated with the bank’s decision -

* If the applicant is likely to repay the loan, then not approving the loan to the person results in a loss of business to the company.
* If the applicant is not likely to repay the loan i.e. default, then approving the loan to the person results in a financial loss to the company.

In this case study, we consider only consumers whose loan application is approved. Here, our aim is to understand how consumer attributes and loan attributes influencing the tendency of defaulting.

Initial Insight:-

The below table is given the average of Continuous Variables, belonging to two level categorical Variables:-

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Loan Status | Annual Income | Loan Amount | Funded Amount | DTI (Debt to Income Ratio) |
| Current\_New | A1 = 64304.69 | L1 = 17089.21 | F1 = 15186.56 | D1 = 14.70268 |
| Default\_New | A2 = 55840.83 | L2 = 11798.37 | F2 = 10944.01 | D2 = 14.04377 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Interest Rate Group | Annual Income | Loan Amount | Funded Amount | DTI (Debt to Income Ratio) |
| High | A3 = 64960.48 | L3 = 18487.693 | F3 = 15738.033 | D3 = 14.45543 |
| Low | A4 = 51512.86 | L4 = 8482.883 | F4 = 8231.533 | D4 = 13.27978 |

1. Confirm at 95% Confidence level if the average number of Annual Income of each Applicant differ significantly for Loan Status of Current\_New and Default\_New.
2. Null hypothesis: A1 equal to A2
3. Alternate hypothesis:A1 not equal to A2
4. Test statistic: t.test t = 10.639
5. p value corresponding to the test statistic: P Value < 2.2e-16
6. Insight (Final decision): Since the p-value is less than alpha 2.2e-16 or is less than 0.0001 we can reject the null Hypothesis and say that there is significant difference between means of A1 and A2 with 95% Confidence.
7. Confirm at 95% Confidence level if the average number of Loan Amount of each Applicant differ significantly for Loan Status of Current\_New and Default\_New.
8. Null hypothesis: L1 equal to L2
9. Alternate hypothesis: L1 not equal to L2
10. Test statistic: t.test, t = 19.407
11. p value corresponding to the test statistic: P Value < 2.2e-16
12. Insight (Final decision): Since the p-value is less than alpha 2.2e-16 or is less than 0.0001 we can reject the null Hypothesis and say that there is significant difference between means of L1 and L2 with 95% Confidence.
13. Confirm at 95% Confidence level if the average number of Funded Amount of each Applicant differ significantly for Loan Status of Current\_New and Default\_New.
14. Null hypothesis: F1 equal to F2
15. Alternate hypothesis: F1 not equal to F2
16. Test statistic: t.test, t = 19.016
17. p value corresponding to the test statistic: P Value < 2.2e-16
18. Insight (Final decision): Since the p-value is less than alpha 2.2e-16 or is less than 0.0001 we can reject the null Hypothesis and say that there is significant difference between means of F1 and F2 with 95% Confidence.
19. Confirm at 95% Confidence level if the average number of Debt To Income Ratio of each Applicant differ significantly for Loan Status of Current\_New and Default\_New.
20. Null hypothesis: D1 equal to D2
21. Alternate hypothesis: D1 not equal to D2
22. Test statistic: t.test, t = 3.093
23. p value corresponding to the test statistic: P Value = 0.002016
24. Insight (Final decision): Since the p-value is 0.002016 which is less than alpha 0.025 we can reject the null Hypothesis and say that there is significant difference between means of D1 and D2 with 95% Confidence.
25. Confirm at 95% Confidence level if the average number of Annual Income of each Applicant differ significantly with Interest Rate Group of High and Low.
26. Null hypothesis: A3 equal to A4
27. Alternate hypothesis: A3 not equal to F4
28. Test statistic: t.test, t = 12.034
29. p value corresponding to the test statistic: P Value < 2.2e-16
30. Insight (Final decision): Since the p-value is less than alpha 2.2e-16 or is less than 0.0001 we can reject the null Hypothesis and say that there is significant difference between means of A3 and A4 with 95% Confidence.
31. Confirm at 95% Confidence level if the average number of Loan Amount of each Applicant differ significantly with Interest Rate Group of High and Low.
32. Null hypothesis: L3 equal to L4
33. Alternate hypothesis: L3 not equal to L4
34. Test statistic: t.test, t = 29.92
35. p value corresponding to the test statistic: P Value < 2.2e-16
36. Insight (Final decision): Since the p-value is less than alpha 2.2e-16 or is less than 0.0001 we can reject the null Hypothesis and say that there is significant difference between means of L3 and L4 with 95% Confidence.
37. Confirm at 95% Confidence level if the average number of Funded Amount of each Applicant differ significantly with Interest Rate Group of High and Low.
38. Null hypothesis: F3 equal to F4
39. Alternate hypothesis: F3 not equal to F4
40. Test statistic: t.test, t = 26.684
41. p value corresponding to the test statistic: P Value < 2.2e-16
42. Insight (Final decision): Since the p-value is less than alpha 2.2e-16 or is less than 0.0001 we can reject the null Hypothesis and say that there is significant difference between means of F3 and F4 with 95% Confidence.
43. Confirm at 95% Confidence level if the average number of Debt To Income Ratio of each Applicant differ significantly with Interest Rate Group of High and Low.
44. Null hypothesis: D3 equal to D4
45. Alternate hypothesis: D3 not equal to D4
46. Test statistic: t.test, t = 3.9043
47. p value corresponding to the test statistic: P Value = 9.795e-05
48. Insight (Final decision): Since the p-value is 9.795e-05 which is greater than alpha value of 0.05 we cannot reject the null Hypothesis.

Conclusion:

We see that the applicant who default have a significantly different average income ,Loan amount and funded amount, and DTI, So If we give loan to applicant with the annual income of 64304 we would be having less changes of default, Also since there is significant difference between Interest Rate of High and Low.