**Hypothesis Testing**

For all the hypotheses, mention the following:

1. Null hypothesis
2. Alternate hypothesis
3. Test statistic
4. p value corresponding to the test statistic
5. Insight (explain in less than 100 words)
6. Test hypotheses (95 % conf. level) for two levels of *loan\_status\_1*- *default\_new and current\_new*

There are four continuous variables for which we can run this test.

1. annual\_inc
2. loan\_amnt
3. funded\_amnt
4. DTI

So, we consider the subsets of loan for loan\_status\_1 for default\_new and current\_new cases. Then, we run t.tests on these continuous variables to identify if the mean values for these two categories for the parameter loan\_status\_1 are same.

Test1: For annual income

1. Null Hypothesis: The mean of annual income for the default new and current new are same.
2. Alternate Hypothesis: The mean of annual income for the default new and current new are different.
3. Critical values: Zc = -1.96 , 1.96
4. Test statistic: t-statistic: 9.4737
5. p-value < 2.2e-16
6. Final decision: Safely reject the NULL hypothesis as the test statistic is much higher than critical value.

Test2: For loan\_amnt

1. Null Hypothesis: The mean of loan amount for the default new and current new are same.
2. Alternate Hypothesis: The mean of loan amount for the default new and current new are different.
3. Critical values: Zc = -1.96 , 1.96
4. Test statistic: t-statistic: -16.461
5. p-value: <2.2e-16
6. Final decision: Safely reject the NULL hypothesis as the test statistic is much lower than critical value.

Test3: For funded\_amnt

1. Null Hypothesis: The mean of funded amount for the default new and current new are same.
2. Alternate Hypothesis: The mean of funded amount for the default new and current new are different.
3. Critical values: Zc = -1.96 , 1.96
4. Test statistic: t-statistic: -16.725
5. p-value: <2.2e-16
6. Final decision: Safely reject the NULL hypothesis as the test statistic is much lower than critical value.

Test4: For DTI

1. Null Hypothesis: The mean of DTI for the default new and current new are same.
2. Alternate Hypothesis: The mean of DTI for the default new and current new are different.
3. Critical values: Zc = -1.96 , 1.96
4. Test statistic: t-statistic: -4.1365
5. p-value: 3.736e-05

Final decision: Safely reject the NULL hypothesis as the test statistic is much lower than critical value.

1. Test hypotheses (95 % conf. level) for two levels of int\_rate\_group - high and low

There are four continuous variables for which we can run this test.

* 1. annual\_inc
  2. loan\_amnt
  3. funded\_amnt
  4. DTI

So, we consider the subsets of loan for int\_rate\_grp for high and low cases. Then, we run t.tests on these continuous variables to identify if the mean values for these two categories for the parameter int\_rate\_grp are same.

Test1: For annual income

1. Null Hypothesis: The mean of annual income for the high and low are same.

2. Alternate Hypothesis: The mean of annual income for the high and low are different.

3. Critical values: Zc = -1.96 , 1.96

4. Test statistic: t-statistic: 10.838

5. p-value < 2.2e-16

6. Final decision: Safely reject the NULL hypothesis as the test statistic is much higher than critical value.

Test2: For loan\_amnt

1. Null Hypothesis: The mean of loan amount for the high and low are same.

2. Alternate Hypothesis: The mean of loan amount for the high and low are different.

3. Critical values: Zc = -1.96 , 1.96

4. Test statistic: t-statistic: 25.94

5. p-value: <2.2e-16

6. Final decision: Safely reject the NULL hypothesis as the test statistic is much higher than critical value.

Test3: For funded\_amnt

1. Null Hypothesis: The mean of funded amount for the high and low are same.

2. Alternate Hypothesis: The mean of funded amount for the high and low are different.

3. Critical values: Zc = -1.96 , 1.96

4. Test statistic: t-statistic: 26.265

5. p-value: <2.2e-16

6. Final decision: Safely reject the NULL hypothesis as the test statistic is much higher than critical value.

Test4: For DTI

1. Null Hypothesis: The mean of DTI for the high and low are same.

2. Alternate Hypothesis: The mean of DTI for the high and low are different.

3. Critical values: Zc = -1.96 , 1.96

4. Test statistic: t-statistic: 6.418

5. p-value: 1.709e-10

6. Final decision: Safely reject the NULL hypothesis as the test statistic is much higher than critical value.