

UNPAIRED T-TEST:

This test is very similar to the paired t-test, but instead of comparing two measurements within our entire population, we use only one measurement but break our population into two natural subgroups, testing whether there is a statistically significant difference between the means of these two subgroups. As in the case of the paired t-test, our primary statistic of concern is the p-value, and again it has the same interpretation.

Exercise 3

Null hypothesis: There is no difference between the price of houses near to and far from the Charles River.

- Open the **Boston Housing Data** file from the CS130 Public folder..
- From the **Analyze** menu, select **Compare Means** then **Independent Samples T-Test**
- Transfer the variable **median** to the **Test Variable(s)** box
- Transfer the variable **Charles** to the **Grouping Variable** box
- Click on **Define Groups** and then enter Nea (short for Near) and Far (see picture below)
- Click Continue and then OK.

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
MEDIAN	Equal variances assumed	8.752	.003	3.996	504	.000	6.3462	1.58795	3.22633	9.46598
	Equal variances not assumed			3.113	36.876	.004	6.3462	2.03841	2.21548	10.47683

Question: Should we accept the null hypothesis?

Correlation Analysis:

Correlation Analysis addresses the following: Is there a statistically significant association between variable X and variable Y?

Null hypothesis tested: There is no statistically significant association between variable X and variable Y.

Level of statistical significance: Set $p=0.05$ so that the probability of rejecting the null hypothesis when it is in fact true is less than 1 in 20.

Remember,

if $p < 0.05$, reject the null hypothesis
if $p \geq 0.05$, accept the null hypothesis

Use scattergrams (scatterplots) to visually display data analyzed with this test.

Question: Is there a statistically significant association between age and cholesterol in the subjects of the Lipid Data study?

Q1: What is the Null Hypothesis?

Q2: Should we accept or reject the Null Hypothesis? Why?

Q3: State your conclusion:

Q4: Question: What is your expected Cholesterol level at age: (a) 25 (b) 50 based on this information?