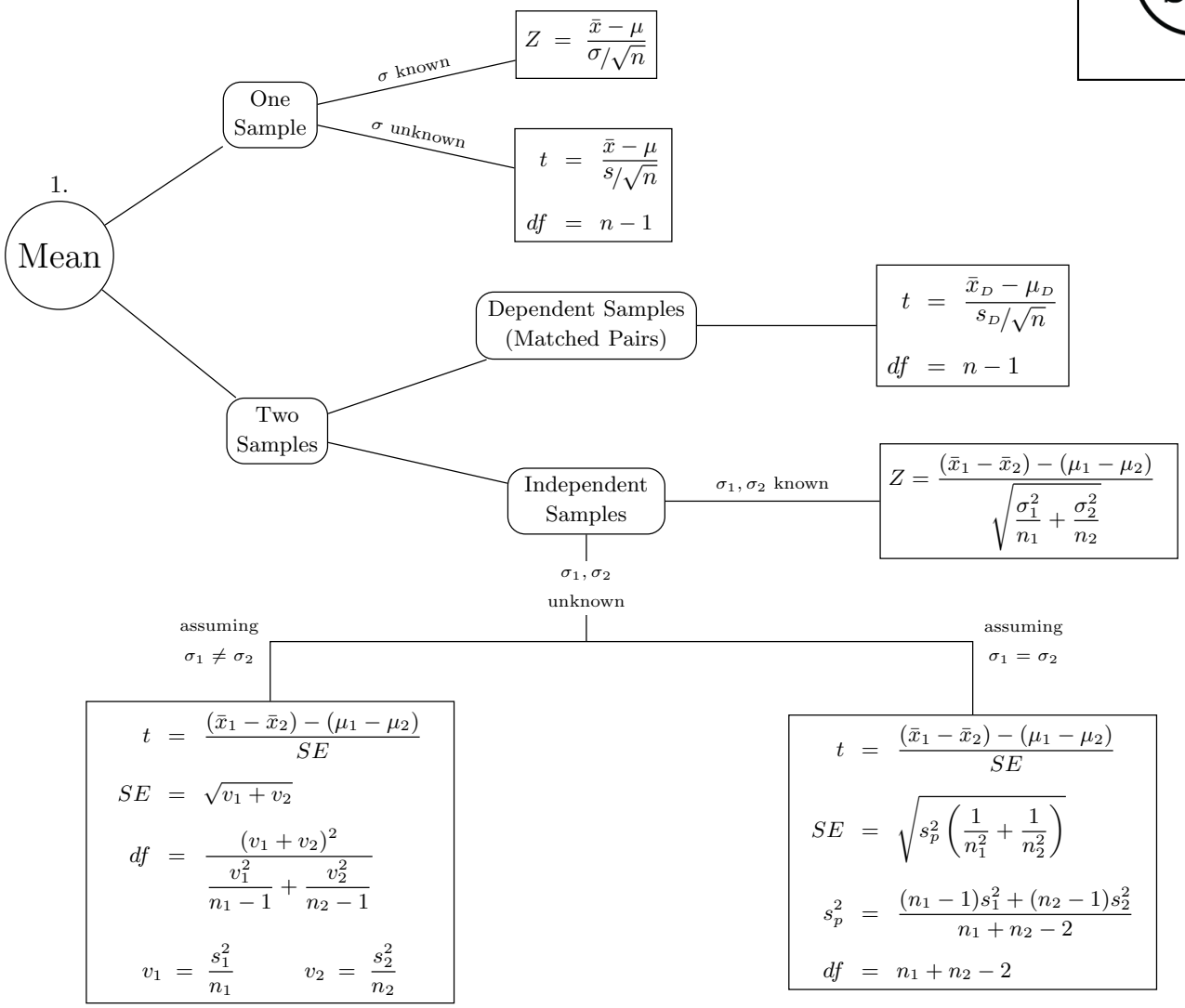
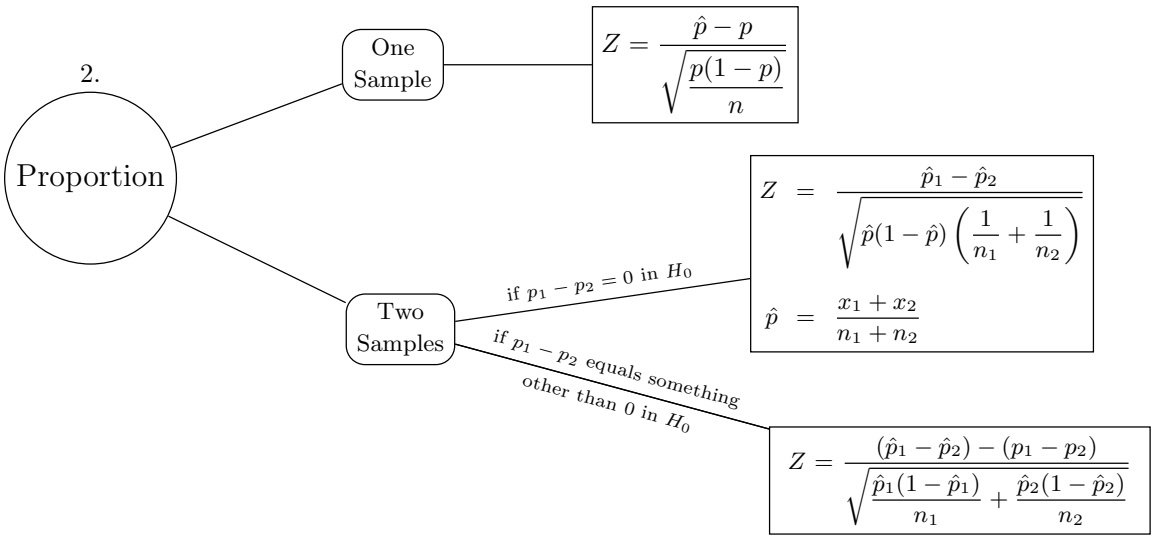
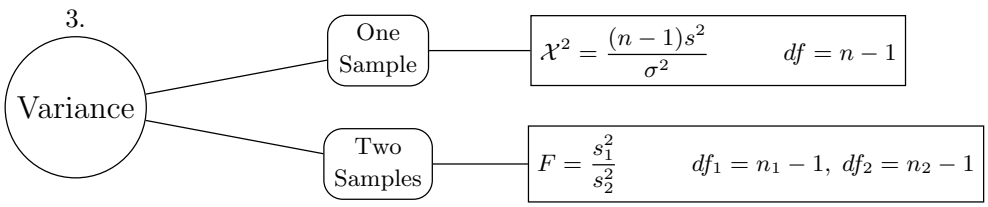
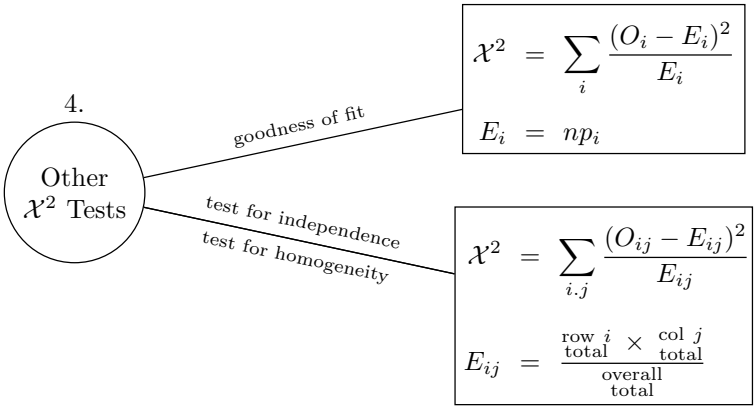
Hypothesis Test Flowchart

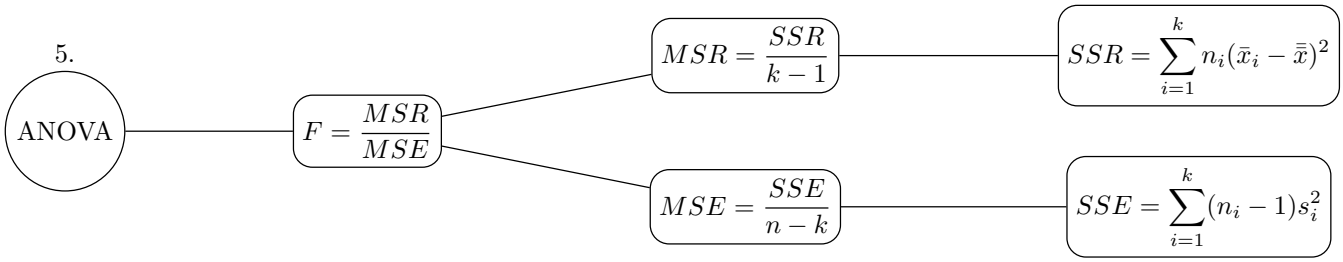
peer tutoring student success center

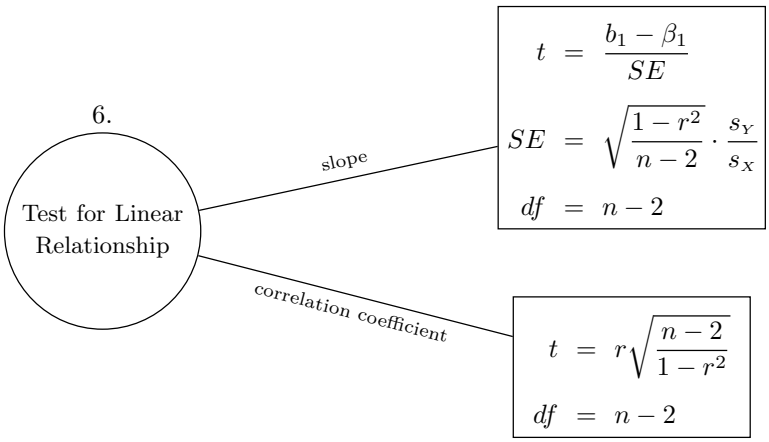












Notes:

0. Many of the statistics classes at UTD teach slight variations to everything presented in this document (e.g. a different formula or two, etc). You should follow what your professor teaches first and foremost.

1. To ensure accuracy in a hypothesis test for one or two population means, the underlying populations should be at least approximately normally distributed, or the sample sizes should be large. Often, a sample size greater than or equal to 30 is considered large.
   - When doing an independent samples t-test comparing two populations means, a common rule for determining if the population variances can be treated as equal or not is to compare the two sample variances: if the larger one is more than twice the smaller one, consider the population variances to be not equal; otherwise, consider them to be equal.

2. To ensure accuracy in a hypothesis test of one population proportion, the sample size should be large enough so that $np \geq 5$ and $n(1-p) \geq 5$. For two populations, both samples should satisfy these inequalities.

3. To ensure accuracy in a hypothesis test for one or two population variances, the underlying populations should be normally distributed. The F test to check if two population variances are equal is generally regarded as being particularly unreliable if the populations are not normal.

4. To ensure accuracy for the chi-square tests, the sample size should be large enough so that all expected counts are greater than or equal to 5.

5. To ensure accuracy in ANOVA, the underlying populations should be at least approximately normal, or the sample sizes should be large; and, they should also be *homoscedastic* (i.e. they should have the same variance).

6. To ensure accuracy in linear regression, the data points in the scatter plot should somewhat follow a straight line, and the residuals should be independent, normally distributed, and homoscedastic.

There are ways of checking to see if the aforementioned requirements for accuracy are met (e.g. histograms for normality, box plots for scedasticity, etc). Textbooks such as the one below discuss this, or you can just search "how to check assumptions for [test]".

Source: Keller, G. (2018). *Statistics for Management and Economics*. Cengage Learning.