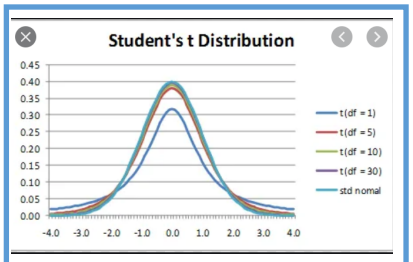
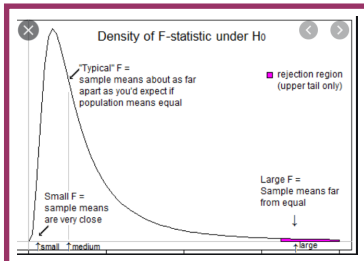
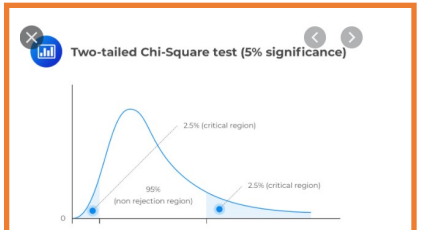
Test	Estimate of the effect or association (co-variation) between X and Y and effect size	Test Statistic	Distribution
	It's important to not just acknowledge the estimate and whether it is statistically significant, we should attend to the size of the effect.	Estimates are translated into test statistics. In one way or another, these are the estimate/the likelihood of sampling error.	The test statistics are compared to the distribution of mathematically derived test statistics that one would expect in a "null world" (where there is no association; it is zero). Those distributions are the source of the p-value, which is the probability of getting a test statistic of this magnitude or larger in a null world. If small (p<.05), we can reject the null. Our finding is "statistically significant."
T-test intvar + dichotomous	The estimate of the effect is the **difference between the means**. Effect size measure is **Cohen's d**. It follows the .1, .3, and .5 and up benchmarks for small, moderate, and large.	**t-statistic** $t = \frac{\text{observed difference between sample means}}{\text{standard error of the difference between the means}}$ or $t = \frac{\bar{X}_1 - \bar{X}_2}{s_{\bar{X}_1 - \bar{X}_2}}$ where $\bar{X}_1$ is the mean for sample 1 $\bar{X}_2$ is the mean for sample 2 $s_{\bar{X}_1 - \bar{X}_2}$ is the standard error of the difference between the means	
ANOVA intvar + multicategorical	The estimate of the effect is the **difference between the means.** Effect size measure is **eta-squared**. This is interpreted as the percent of variation in Y that is explained by the categorical variable. (.1, .3, .5 and up for small, moderate, and large).	**F-value** The calculation of the F is fairly involved (see Urdan). At base, it is the ratio of the variation we see between the groups and the variation we see within groups. $F = \frac{MS_{group}}{MS_{error}}$	
Crosstab categorical + categorical	The effect is observed in the **overall difference between observed frequencies in cells and the expected (by chance) frequencies.** In practice, we compare each group's %s in the outcome categories to see how they differ. We typically **comparing the within row %s.** Effect size measure is **Cramer's V**. (follows the .1, .3, .5 and up benchmarks for small, moderate, and large effects.	**Chi-square** TABLE 14.5 Formula for Calculating χ^2 $\chi^2 = \sum\left(\frac{(O-E)^2}{E}\right)$ where ***O*** is the observed value in each cell and ***E*** is the expected value in each cell	
Correlation/Simple Regression intvar + intvar	The estimate is the Pearson's r correlation coefficient. |. 1| small/weak |.3| moderate |.5 and up| strong $r = \frac{\sum z_x z_y}{n-1}$	**t-value** **T – test for Correlation Coefficient** $t = r\sqrt{\frac{n-2}{1-r^2}}$ where, r = correlation coefficient n = total number of observations Degree of freedom, df = n – 2	Depending on whether we specify a direction in our alternative hypothesis, we use either the one-tailed test or the two-tailed test. You can see that a one-tailed test (directional) doesn't require as large a test statistic in order to reach statistical significance.