|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Type of inference** | **Parameter (or question for HT)** | **Conditions needed to use theoretical sampling distribution of test statistic** | **Sample Statistic** | **Confidence Interval formula** | **Theoretical distribution of Test statistic** | **Degrees of freedom** | **Test statistic SE is “standard error”**  **Generally, use software to obtain this. See formulas to the right.** | **Standard Error formula for CI** | **Standard Error for HT** | **Sample size for estimating** |
| One mean |  | Distribution normal or CLT applies, meaning approximately |  |  | t distribution |  |  |  |  | ,  where ME is the chosen margin of error and is an estimate of the population standard deviation. |
| One proportion |  | AND  CI: HT: |  |  | normal distribution | Not relevant |  |  | Where is the value in the null hypothesis | ,  where ME is the chosen margin of error and we use or some other value of if available. |
| Two means |  | In EACH group, distribution normal OR CLT applies, meaning approximately |  |  | t distribution | smaller of   and or Satterthwaite approximation |  |  |  | **Standard Error (degrees of freedom) for two-sample t-procedures:**  In two-sample t procedures, in order to show that the test statistic has an exact t-distribution, we must have that the two population variances are equal. In that case,  If it is not appropriate to assume that the two population variances are equal, then a “conservative” approach (does not overstate our confidence in our answers) is to use the smaller of  .  An adjustment can be made to the degrees of freedom to take into account how different the variances are and how different the sample sizes are.  Most statistical software will use this Satterthwaite approximation as the degrees of freedom for two-sample t procedures. It is derived by a modification of the method of moments method of estimation.  Don’t do this “by hand.”  It is included here because, as you use software, you will see degrees of freedom that do not fit the “simple” method given in this handout and in many applied statistics texts. |
| Two proportions |  | AND  AND  AND |  |  | normal distribution | Not relevant |  | where are the sample proportions from the two separate samples | For testing whether the pop. prop. are equal. Here, is the pooled proportion. |
| Mean of differences from matched pairs data |  | Distribution of differences normal  OR  CLT applies, meaning approximately, where is number of pairs. |  |  | t distribution | where is the number of pairs |  | where the subscripts refer to using the differences | where the subscripts refer to using the differences |
| Correlation coefficient |  | See regression model conditions. Generally linear pattern (rather than a different pattern), same variance across x-values, residuals are independent and have normal distribution. |  |  | t distribution |  |  |  |  |
| Slope coefficient |  | See regression model conditions. Generally linear pattern (rather than a different pattern), same variance across x-values, residuals are independent and have normal distribution. |  |  | t distribution | for simple regression |  | Obtain with technology | Obtain with technology |
| Test of goodness of fit | Do the data fit a particular specified distribution? | Each expected count is at least 5. | chi-squared, degrees of freedom | This investigation is a test. No parameters are estimated with this procedure. | distribution | number of categories minus 1 |  | Not relevant | Not relevant |
| Test of association of two categorical variables | Are the two categorical variables associated? | Each expected count is at least 5. | chi-squared, degrees of freedom | This investigation is a test. No parameters are estimated with this procedure. | distribution | # of rows  # of cols |  | Not relevant | Not relevant |
| Analysis of Variance (ANOVA) for difference of means | Is there a difference in the means of two or more groups? | In EACH group, distribution normal OR CLT applies, meaning approximately.  Variability is similar in all groups (). | F statistic,  and degrees of freedom | This investigation is a test. No parameters are estimated with this procedure. | distribution | # of groups  total sample sizes |  | Not relevant | Not relevant |
| Analysis of Variance (ANOVA) for regression | Is at least one variable in the model useful in predicting the response variable? | For EACH explanatory variable, same conditions as linear model for single explanatory variable. Check with residual analysis, including plots. To start, plot the residuals vs the fitted values. | F statistic,  and degrees of freedom | This investigation is a test. No parameters are estimated with this procedure. | distribution | # of explanatory variables  sample size |  | Not relevant | Not relevant |