AP Statistics

**The Significance Tests Summary**

* **Z-Test**
* What it tests: Population mean, but you have to have the population standard deviation for this test, which is as likely as the tooth fairy, (Santa Claus is real though) **so just don’t do this test.**
* **T-Test**
  + What it tests: A sample against hypothesized mean.
  + CYA:
    - random sample
    - independence OR *(population 10x the sample)*
    - normal distribution or sample size thirty or larger.
  + Test statistic: 
* **Matched Pairs T-test**
  + What it tests: The difference in sample data taken from the same population at different times. In other words the samples are comparative, usually a before after. A t-test is done on the differences.
  + CYA
    - random sample
    - Independence OR (population 10x sample)
    - normal distribution (for the differences) or large enough sample size.
  + Test Statistics: same as above with  usually being 0
* **One Proportion Z-test.** 
  + What it tests: A sample proportion against a hypothesized proportion.
  + CYA:
    - SRS, n(p-hat)>10, n(q-hat)>10
    - Independence OR *(population 10x the sample)*
  + Test Statistic: 
* **Two-Sample T-test**
  + What it tests: Whether or not two population means are likely the same.
  + CYA:
    - TWO random samples
    - two independent samples OR (both populations 10X the sample)
    - two normal populations or sample size large enough
  + Test Statistic: 
* **Two Proportion Z-test**
  + What it Tests: Whether or not two population proportions are likely the same
  + CYA:
    - 2 SRS’s 
    - 2 independent samples OR 2 populations that are 10x the sample
  + Test Statistic (In Sever’s perfect world we do this): 
  + Test Statistic (pooled……what the calculator and most of the world does):  where phatc=phat (combined)

=

**THE CHI-SQUARE TESTS. They are all very similar with subtle differences.**

* Chi-Square Test for Goodness of Fit
  + What it tests: Are the observed counts for a statistic the same as the expected counts? (m&m colors) (this test may not be on older calculators… but it is easy)
* Chi-Square test for independence:
  + What it tests: Is there a relationship between two categorical variables? (uses matrices for observed and expected values)
* Chi-Square test for homogeneity.
  + What it tests: Is there a difference in the distribution of a categorical variable for all of the populations or treatments. Done exactly the same was as independence. (key word: different groups)

For all of the chi-Square tests:

* + CYA:
    - SRS
    - independent observations OR *(population 10x the sample)*
    - all expected counts are at least 5.
  + Test statistic:  (with n-1 (n= categories) degrees of freedom.)
* Be careful with homogeneity and independence. , NEVER SAY ONE CATEGORY CAUSES CHANGES IN THE OTHER. THE Ho IS ALWAYS THAT THERE IS NOT AN ASSOCIATION BETWEEN CATEGORICAL VARIABLES. DON’T SAY ONE EFFECTS THE OTHER ONE.

Also on chi-square tests **do not** say expected values are consistent with observed values, say are roughly equivalent to or better yet, use actual values.

REMEMBER BS.

**LINEAR REGRESSION TESTS.**

* + What it tests: Whether or not there is a linear relationship between x and y.
  + CYA: (LINER)
    - Linear: check scatterplot to see if overall pattern is linear.
    - Independence: how was data produced
    - Normal: Make a histogram or normal plot of the residuals.
    - Equal variance: Make sure the residual plot has *roughly* the same amount of data above and below the y=0 line.
    - Random: Was the data random or a randomized experiment.
  + Test Statistic:  where  (you never have to calculate SE by hand)

With n-2 degrees of freedom.