

AP Statistics: Assumptions for Inference

You must demonstrate (or assume, if reasonable) that assumptions are met before performing a significance test or constructing a confidence interval.

Work to show:

hypotheses, verification of assumptions, name of procedure, sketch, test statistic with substituted values and result, p-value [alternative if given α -level: comparison of critical value (t^ , z^*) or CI if a 2-tailed test], decision (strength of p-value or comparison to α), and conclusion in context.*

Assumptions for Inference	Conditions that confirm or override them
<i>M e a n s (t - t e s t)</i>	
I will use one-sample t procedures with n-1 degrees of freedom for the mean of a population if:	
1. Individuals/observations are independent. 2. Population is Normal.	1. SRS and sample is less than 10% of the population. 2. ONE of the following: <ul style="list-style-type: none"> • population is stated as normal, OR • sample size is large (>40), OR • sample size is medium (15-40) and plot of sample shows little skewness and no extreme outliers, OR • sample size is small (<15) and plot of sample shows no skewness and outliers. Rough sketch of plot of sample data (histogram, dotplot, stemplot) is unimodal and symmetric which becomes less critical as sample size increases.
I will use one-sample t procedures for matched pairs with n-1 degrees of freedom for the mean of the differences of a population if:	
1. Data are matched. 2. Individuals/observations are independent. 3. Population of differences is Normal.	1. Think about design. 2. SRS and sample is less than 10% of the population. 3. Population of differences is stated as normal OR the sample size of differences <ul style="list-style-type: none"> • large (>40), OR • medium (15-40) and plot of differences shows little skewness and no extreme outliers, OR • small (<15) and plot of differences shows no skewness and outliers. Rough sketch of plot of differences (histogram, dotplot, stemplot) is unimodal and symmetric which becomes less critical as sample size of differences increases.
I will use two-sample t procedures (with df from calculator OR df of n-1 from smaller n) for the difference of two population means if:	
1. Samples are independent of each other. 2. Individuals/observations in each sample are independent. 3. Both populations are Normal.	1. Think about the design. 2. SRS and sample is less than 10% of the population for BOTH samples. 3. For BOTH samples, ONE of the following: <ul style="list-style-type: none"> • populations are stated as normal, OR • sample sizes are large (>40), OR • sample sizes are medium (15-40) and plots of samples show little skewness and no extreme outliers, OR • sample sizes are small (<15) and plots of samples show no skewness and outliers. Rough sketches of BOTH plot of sample data (histogram, dotplot, stemplot) is unimodal and symmetric which becomes less critical as sample size increases.

Assumptions for Inference	Conditions that confirm or override them
<i>Proportions (z-test)</i>	
I will use one-proportion z procedure for a population proportion if:	
1. Individuals/observations are independent. 2. Sample is sufficiently large.	1. SRS and sample is less than 10% of the population. 2. Successes & failures are greater than or equal to 10. <ul style="list-style-type: none"> • For confidence interval: $n\hat{p} \geq 10$ and $n(1 - \hat{p}) \geq 10$ • For significance test: $np_0 \geq 10$ and $n(1 - p_0) \geq 10$ (remember: some texts use ≥ 5)
I will use two-proportion z procedures for the difference of two population proportions if:	
1. Samples are independent of each other. 2. Individuals/observations in each sample are independent. 3. Both samples are sufficiently large.	1. Think about how data were collected. 2. SRS(or random allocation in experiment design) and sample is less than 10% of the population for BOTH samples. 3. For BOTH samples, successes and failures are greater than or equal to 5. <ul style="list-style-type: none"> • For confidence interval: $n_1\hat{p}_1 \geq 5$ and $n_1(1 - \hat{p}_1) \geq 5$ AND $n_2\hat{p}_2 \geq 5$ and $n_2(1 - \hat{p}_2) \geq 5$ • For significance test: using pooled proportion, \hat{p}_p $n_1\hat{p}_p \geq 5$ and $n_1(1 - \hat{p}_p) \geq 5$ AND $n_2\hat{p}_p \geq 5$ and $n_2(1 - \hat{p}_p) \geq 5$
<i>Distributions (Chi-Square)</i>	
I will use a Chi-Square procedure for Goodness of Fit (one sample, one categorical variable with df = n-1) if:	
1. Data are counts 2. Data in sample are independent. 3. Sample is sufficiently large.	1. Are they? 2. SRS and sample is less than 10% of the population. 3. All <u>expected</u> counts are at least 5 OR all <u>expected</u> counts are >1 AND no more than 20% are <5 .
I will use a Chi-Square procedure for Association (Independence) (one sample from one population, classified with regard to 2 categorical variables with df = (r-1)(c-1)) if:	
1. Data are counts 2. Data in sample are independent. 3. Sample is sufficiently large.	1. Are they? 2. SRS and sample is less than 10% of the population. 3. All <u>expected</u> counts are at least 5 OR all <u>expected</u> counts are >1 AND no more than 20% are <5 .
I will use a Chi-Square procedure for Homogeneity of Populations (samples from more than one population classified with regard to a single categorical response variable, with df = (r-1)(c-1)) if:	
1. Data are counts 2. Data in groups are independent. 3. Groups are sufficiently large.	1. Are they? 2. SRS's (or random allocation) and samples are less than 10% of the population. 3. All <u>expected</u> counts are at least 5 OR all <u>expected</u> counts are >1 AND no more than 20% are <5 .
<i>Regression (Slope) (t-test)</i>	
I will use t procedures for the slope of the regression line with df=n-2 (n=# of (x,y) pairs) if:	
1. Form of relationship is linear. 2. Errors are independent and variability of errors is constant. 3. Errors have a normal model.	1. Scatterplot looks approximately linear. 2. Residual plot has not apparent pattern, has a consistent spread. 3. Histogram of residuals is approximately unimodal and symmetric. No strong skewness or extreme outliers.