|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable(s) of Interest** | **Parameter of Interest** | **Statistic of Interest** | **Descriptive Method(s)** | **Inferential Method(s)** | **Assumption(s) for Inferential Methods** |
| **Single Categorical Variable**  **(Binary – 2 categories)** | True Population Proportion (π) | Sample Proportion ( ) | * Report * Bar chart | * Simulation * Binomial test * CI for π | Check the four conditions for the binomial |
| **Single Categorical Variable**  **(More than two categories)** | True Population Proportion () | Sample proportions | * Report * Stacked bar chart | * Chi-square Goodness of Fit | * Observations are independent * EXPECTED counts should be greater than 5 |
| **Two Categorical Variables (in general)** | True Conditional Population Proportions  () | Sample Proportions | * Report sample proportions * Contingency table * Stacked/Dodged/Filled Bar Plot | * Chi-square test of Independence | * Observations are independent * EXPECTED counts should be greater than 5 |
| **Variables of Interest** | **Parameter of Interest** | **Statistic of Interest** | **Descriptive Methods** | **Inferential Methods** | **Assumptions for Inferential Methods** |
| **Single Numerical Variable** | True Population Mean () | Sample Mean ( ) | * Report measures of center and variation * Dotplot, boxplot, histogram * Describe shape * Identify outliers | * One-sample t-test * CI for | * Either the sample size is fairly large or the data reasonably follow a normal distribution |
| **Comparing Numerical Variable across Two Categories of a Categorical Variable (DEPENDENT samples)** | True Mean Difference  () | Sample Mean Difference  ( ) | * Report measures of center and variation for the differences * Dotplot, boxplot, histogram of the differences * Describe shape, identify outliers | * paired t-test * CI for | * Independent differences * Either the number of pairs is fairly large or the differences reasonably follow a normal distribution |
| **Comparing Numerical Variable across Two Categories of a Categorical Variable (INDEPENDENT samples)** | Difference in True Population Means  () | Difference in Sample Means  ( ) | * Report, and * Side-by-side boxplots, facetted histograms | * Two-sample independent t-test * CI for | * Observations are independent *between* & *within* groups * Either both sample sizes are fairly large or the data from each group reasonably follow a normal distribution |
| **Variables of Interest** | **Parameter of Interest** | **Statistic of Interest** | **Descriptive Methods** | **Inferential Methods** | **Assumptions for Inferential Methods** |
| **Comparing Numerical Variable across 2 or more categories of a Categorical Variable** |  |  | * Group means, group standard dev. * Side-by-side boxplots, facetted histograms | Analysis of Variance (ANOVA)  F-test statistic | * Independence *between* and *within* groups * Equal variances * Normality |
| **Comparing Two Numerical Variables** | Population Slope  () | Sample Slope | * Correlation (r) * Scatterplot * Regression line   ( ) | Linear Regression Analysis  Slope = 0? | * Linearity * Independence * Normality * Equal Variance |