|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scenario | One Categorical Response | Two Categorical Variables | Paired Differences | Two Quantitative Variables | Quant. Response and Categ. Explanatory (independent samples) |
| Type of plot | Bar plot | Segmented bar plot,  Mosaic plot | Dotplot, histogram, boxplot | Scatterplot | Side-by-sided boxplots,  Stacked dotplots or histograms |
| Summary measure | Proportion | Difference in proportions | Mean Difference | Slope or correlation | Difference in means |
| Parameter notation |  |  |  | or |  |
| Statistic notation |  |  |  | or |  |
| Null hypothesis |  |  |  | or |  |
| Conditions for simulation methods | Independent cases; | Independence (within and between groups); | Independent cases; | Independent cases; Linear form; | Independence (within and between groups); |
| Simulation test (how to generate a null distn)  p-value = proportion of null simulations at or beyond ( direction) the observed statistic | Spin spinner with probability equal to , 𝑛 times or draw with replacement 𝑛 times from a deck of cards created to reflect as probability of success. Plot the proportion of successes. Repeat 1000’s of times. Centered at | Label cards with response values from original data; mix cards together; shuffle into two new groups of sizes and . Plot difference in proportion of successes. Repeat 1000’s of times. Centered at 0. | Shift the original data by adding . Sample with replacement from the shifted data times. Plot sample mean. Repeat 1000’s of times. Centered at (single mean) or 0 (paired mean difference). | Hold the values constant; shuffle new ’s to ’s. Find the regression line for shuffled data; plot the slope or the correlation for the shuffled data. Repeat 1000’s of times.  Centered at 0. | Label cards with response variable values from original data; mix cards together; shuffle into two new groups of sizes and Plot difference in means. Repeat 1000’s of times. Centered at 0. |
| Bootstrap CI (how to generate a boot. distn)  X% CI: | Label 𝑛 cards with the original responses.  Randomly draw with replacement 𝑛 times. Plot the resampled proportion of successes. Repeat 1000’s of times. Centered at . | Label cards with the original responses. Randomly  draw with replacement times from group 1 and times from group 2. Plot the resampled difference in proportion of successes. Repeat 1000’s of times. Centered at | Label cards with the original responses. Randomly draw with replacement times. Plot the resample mean. Repeat 1000’s of times. Centered at | Label cards with the original (response, explanatory) values. Randomly draw with replacement times. Plot the resampled slope or correlation. Repeat 1000’s of times. Centered at or . | Label cards with the original responses. Randomly  draw with replacement times from group 1 and times from group 2. Plot the resampled difference in means. Repeat 1000’s of times. Centered at |
| Theory-based distribution | Standard Normal | Standard Normal | *t-* distribution with df | *t-* distribution with df | *t-* distribution with min of  -1 or -1 df |
| Conditions for theory-based hypothesis tests and confidence intervals | Independent cases;  Number of successes and number of failures in the sample both at least 10. | Independence (within and between groups); Number of successes and number of failures in EACH sample all at least 10. (All four cell counts at least 10.) | Independent cases; with no clear outliers OR with no extreme outliers OR | Linear form; Independent cases; Nearly normal residuals; Variability around the regression line is roughly constant. | Independent cases (within and between groups); In each sample, with no clear outliers OR with no extreme outliers OR |
| Theory-based standardized statistic (test statistic) |  |  |  | is the reported standard error (std. error) of the slope term in the lm() output from R. |  |
| Theory-based confidence interval |  |  |  | is the reported standard error (std. error) of the slope term in the lm() output from R. |  |