Exact Binomial Analysis

This platform allows the user to create a graphical representation of binomial data with the appropriate CI lines based on observed defect rates.

Within the dialog, the user will have several options. The columns in the table will be listed to the far left. The user will be able to move columns to three key areas: Sample Size, Defects, and By.

The user needs to specify the organization of the data and has the option of changing Alpha from the default 0.05.

In addition, there is a Sample Size interactive graph which allows the user to select different sample sizes and defects to visualize the change in CI lines. If the user would like to test the script with known working data sets, the user may select a pre-summarized dataset or one broken out into individual data points.

Various actions mimic actions that are standard with JMP platforms.

Sample Size is optional and is limited to 1 numeric column. This column is only used for a data table where the defects have already been summarized.

Defects requires, and is limited to, one column, numeric or character. For already summarized data, this column represents the number of defects observed. For individual row data this column represents the status of each row (Pass or Fail). Should the user be analyzing individual row data, the Defect column must contain "Pass" or "Fail" to code for non-defect or defect results.

By is optional by the user and is unlimited in number of columns added. It represents the hierarchy of data for the defects being tested. For already summarized data all of the hierarchy columns MUST be input into the By variable. For individual row data any combination of the hierarchy columns can be analyzed. In both instances the order of the hierarchy columns affects the final X-axis output (much like multivary charts)

How is Data Organized?:
Already summarized data has the sample size and defect count in each row per category.
Individual row data has one result per row where each row is either a Pass or Fail

Sample Size: Show Graph
This will display a separate, interactive, graph where sample sizes and defect counts can be adjusted to estimate necessary sample sizes to gain resolution.

For either type of data the script will produce a summary table with the following additional columns:

      Proportion
      xx%CI (NAL)
      UCL (NAL)
      LCL (NAL)
      UCL (EBL)
      LCL (EBL)
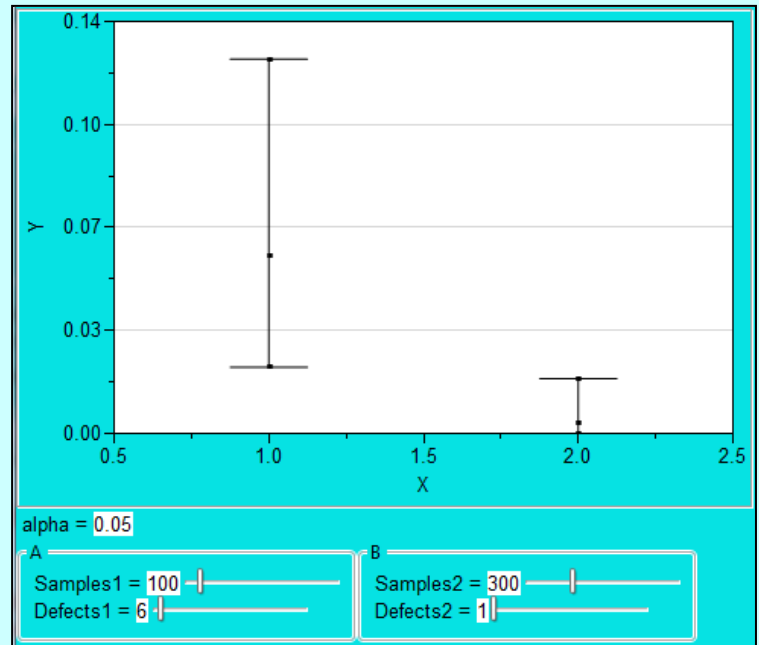   **xx = (1-alpha) * 100
   **NAL = Normal Approximation Limits
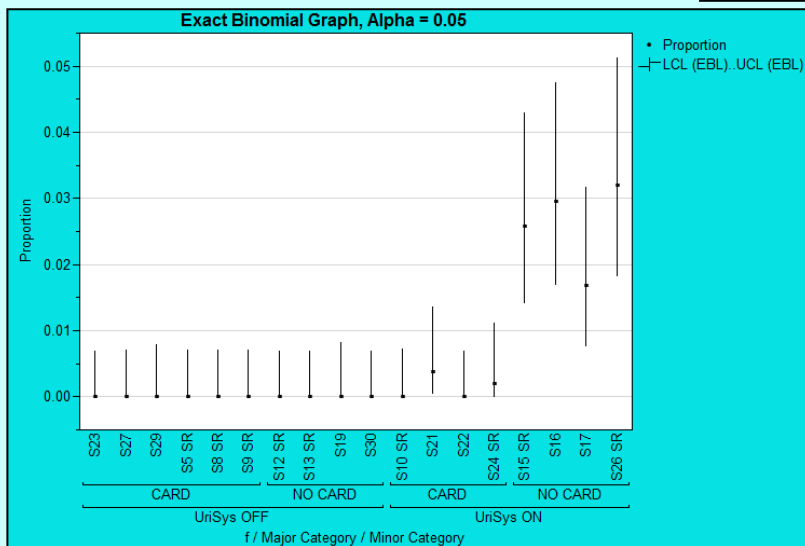   **EBL = Exact Binomial Limits

For already summarized data, the columns will be appended to the summary data. For individual row data the summary will be created as a new table.

Lower confidence limit:
$$1 - \frac{\alpha}{2} = \sum_{k=0}^{d-1} \frac{n!}{k!(n-k)!} p_{LCL}{}^{k} (1 - p_{LCL})^{n-k}$$

Upper confidence limit:
$$\frac{\alpha}{2} = \sum_{k=0}^{d} \frac{n!}{k!(n-k)!} p_{UCL}{}^{k} (1 - p_{UCL})^{n-k}$$

The preceding equations are solved by iterating a value for the confidence limit (pLCL or pUCL) that yields a value of 1-a/2 for the lower confidence limit and a/2 for the upper confidence limit.

The final graph will split the x-axis into the hierarchy created by the order of the columns in "By" with needle charts for the defect proportion +/- the upper and lower exact binomial limits.