

Statistics in Excel

P. Cook – DNP Faculty Development 4/30/19

Control Charts for Quality Improvement

1. Format data with “month” (or day, quarter, etc.) as a column, and results for that month in a second column. It’s OK to have more than one entry for the same month.
2. Use Pivot Table to aggregate results – month (or whatever) as rows, and average value as cells
3. Insert a Pivot Chart based on the table to show a line graph
4. Calculate a mean and standard deviation from the original data:
 - a. Formula for the mean: **=average(xx2...xx100)** where “xx2” and “xx100” are the first and last cells in the column of data values
 - b. Formula for the standard deviation: **=stdev(xx2..xx100)**
5. Determine the values for the average + 1.96*SD, and average – 1.96*SD. Create lines at those levels on the Pivot Chart (they are your upper and lower control limits)
6. Right click the graph to add a trend line if you’d like one

Chi-Square or Fisher’s Exact Test from Frequency Data

1. Calculate frequencies that are the raw numbers with and without the desired outcome, pre vs. post intervention. It should make a 2x2 table like this:

	Pre-Intervention	Post-Intervention
Number with Desired Result	100	150
Number without Desired Result	10	5

2. Go to **VassarStats.com**, and select “frequency data,” then “option 2.”
3. Input the raw numbers (*not percentages*) into the 2x2 table, and click “compute.”
4. Look for results including odds ratio, chi-square with p-value, Fisher’s exact test p-value, etc.

T-Test from Raw Data

1. All data should be in a single file, with a grouping variable (e.g., “Pre-Post” as a column, with values of “Pre” or “Post” for any given row, and the numeric result in a second column). It’s OK to have multiple entries labeled Pre or Post.
2. Sort the dataset by the grouping variable, so that all “pre” results are at the top followed by the “post” results (or vice-versa).
3. Use Pivot Table to aggregate results – Pre/Post as the row, and average value as the cell. Or else just use the formula for the mean (see above) to calculate the averages for pre and for post.
4. Use the t-test formulas to get a p-value and test statistic result:
 - a. Formula for the p-value: **=ttest(xx2..xx100,xx101..xx200,2,2)** where xx2 and xx100 are the first and last cells of the “pre” values, xx101 and xx200 are the first and last cells of the “post” values, and the other numbers are just 2s (one tells Excel to do a 2-tailed ttest and the other tells Excel to assume homogeneity of variance).
 - b. Formula for the actual t-test statistic: **=tinv(xx00, count(xx2..xx200)-2)**, where xx00 is the cell where you typed the formula for the p-value, and the second element is the degrees of freedom, defined as the total number of values minus two.

Correlations

1. The formula for a correlation in Excel is: **=correl(xx2..xx50,yy2..yy50)**, where xx and yy are the columns of data for the 2 variables that you want to correlate. Correlations don’t have p-values.