

NOTE H: SOME SPECIFIC LISTS OF EXCEL FAULTS

PROBLEMS WITH USING MICROSOFT EXCEL FOR STATISTICS – CRYER’S LIST

One of the frequent references encountered on Excel faults is that of Jonathan D. Cryer. Internet searches relating to Excel faults bring up his web site, and end up as points of questions and discussions on the sci.stat.consult, sci.stat.edu and sci.stat.math newsgroups, and on the EDSTAT-L lists. The following table gives his issues with reference to where it is discussed in the paper.

I consider some of his comments as not being valid; given the situation that Excel was not intended to do advanced statistical analysis or graphics.

Someone who is familiar with the line inputs in Minitab, and in other software programs will find it easier (than in Excel) to enter matrix data (as a name) and enter commands to select appropriate columns as predictor variables. Business applications are all “windows” (except for old COBOL programs) and business users are more comfortable with the point and click method.

Line	Excel General Area	Specific Issue, Fault or Comment	Issue Supported By Illustration or Narrative	Explicit Comment or Illustration by Cryer	Valid Fault	Discussed at:
1	Excel Charts	Portray Numerical Information Visually Without Distortion	No			Excel chart options by menu; allow all kinds of chart variations to properly display data.
2	Excel Charts	Contain No Distracting Elements	Yes	Example: Excel Graphics With False Third Dimension (Taken from JSE)	No	You don't have to select it
3	Excel Charts	Label Axes (Scales) and Tick marks Appropriately	No		No	Excel chart options by menu, allow all kinds of axis and tic mark variations to properly display data.
4	Excel Charts	Have a Descriptive Title and/or Caption Legend	No		No	Excel chart options by menu, allow legends to be removed/relocated and captions changed/relocated

Line	Excel General Area	Specific Issue, Fault or Comment	Issue Supported By Illustration or Narrative	Explicit Comment or Illustration by Cryer	Valid Fault	Discussed at:
5	Excel Charts	Excel does not meet Good Graphic Criteria	No		No	Opinion. He ignores all the other types of charts.
6	Excel Charts	Vast majority of Chart types offered by Excel should never be used	No		No	Opinion. The charts were intended for business applications.
7	Excel Charts	Pyramid Charts	Yes	Does not represent good graphs	No	You don't have to select it
8	Excel Charts	Column, Cone and Cylinder charts	Yes	No redeeming features	No	You don't have to select it
9	Excel Charts	Default Scatterplot	Yes	Data points covered by axis labels	No	Failure to properly set chart options before and after.
10	Excel Charts	Default Scatterplot	Yes	Legend to the right of the graph not useful	No	Failure to properly set chart options before and after.
11	Excel Charts	Default Scatterplot	Yes	No label for the horizontal Axis	No	Failure to properly set chart options before and after.
12	Excel Charts	Default Histogram	Yes	Bin labels impossible to read	No	Failure to properly set chart options before and after.
13	Excel Charts	Default Histogram	Yes	Aspect ratio is poor	No	Failure to properly set chart options before and after.
14	Excel Charts	Default Histogram	Yes	Legend label useless	No	Failure to properly set chart options before and after.
15	Excel Charts	Default Histogram	Yes	Horizontal axis label are useless	No	Failure to properly set chart options before and after.
16	Excel Charts	Default Histogram	Yes	Class intervals bizarre	No	Failure to properly set chart options before and after.
17	Excel Charts	Default Histogram	Yes	Number of digits is atrocious	No	Failure to properly set chart options before and after.

Line	Excel General Area	Specific Issue, Fault or Comment	Issue Supported By Illustration or Narrative	Explicit Comment or Illustration by Cryer	Valid Fault	Discussed at:
18	Excel Help Screens	Two-sample t test:	Yes	Gives 4 sentences from the help screen	Yes	Has errors
19	Excel Help Screens	Confidence Function:	Yes	Misstatement.	Yes	An error.
20	Excel Help Screens	STDEV Function:	Yes	Word "nonbiased" is incorrect	Yes	An error.
21	Excel Help Screens	STDEV Function:	Yes	Standard deviation given here is not unbiased	Yes	Cryer does not correctly state the issue. STDEV is unbiased, STDEVP is biased.
22	Treatment of Missing Data	Excel Does It Incorrectly	No		Yes	Missing data is a complex problem. The main error is in the paired t test. See KBA 829252.
23	Treatment of Missing Data	Excel Does It Inconsistently	No		?	No example of the inconsistency given. Unable to verify.
24	Treatment of Missing Data	Excel Makes Selecting Predictor Variables In Regression Especially Difficult When Data Is Missing	No		?	Opinion. No example given of the difficulty here.
25	Treatment of Missing Data	t-Test: Paired Two Sample for Means	yes	Means are wrong	No	One should never have missing data in a "paired" data test. This is an incorrect application.
26	Treatment of Missing Data	t-Test: Paired Two Sample for Means	yes	Variance is wrong	No	One should never have missing data in a "paired" data test. This is an incorrect application.
27	Treatment of Missing Data	t-Test: Paired Two Sample for Means	yes	Df is wrong	No	One should never have missing data in a "paired" data test. This is an incorrect application.
28	Computing Algorithms for Basic Statistics	Excel Uses Poor Algorithms To Find The Standard Deviation	No		Yes	Excel produces 5 or more correct digits when the COV of the data is greater than 0.00001. Fixed for 2003.
29	Computing Algorithms for Basic Statistics	Excel Defines The First Quartile To Be The Ordered Observation at Position $(n+3)/4$	No		No	Cryer is wrong. This is Gumbel's quartile, an accepted quartile. See note H.

Line	Excel General Area	Specific Issue, Fault or Comment	Issue Supported By Illustration or Narrative	Explicit Comment or Illustration by Cryer	Valid Fault	Discussed at:
30	Computing Algorithms for Basic Statistics	Excel Does Not Treat Tied Observations Correctly When Ranking	No		No	Cryer is wrong. The treatment of tied observations is arbitrary. There is no accepted standard.
31	Computing Algorithms for Basic Statistics	Regression Computations Are Often Erroneous Due To Poor Algorithms	No		Yes	Regression through the origin is wrong. The algorithm is unstable. There errors with certain regressions, but there are easy fixes. Fixed for 2003.
32	Computed numbers.	Displays many more digits than appropriate.	No		No	Formatting cells can set the number of appropriate digits. The number of digits displayed in charts can be set by menu. There is no way to set internally the number of displayed digits based on the internal accuracies of computation.
33	Regression	Does Not Treat Zero-Intercept Models Correctly	No		Yes	The only error is in the correlation coefficients and in the ANOVA table. There is a fix. Fixed for 2003.
34	Regression	Sometimes Gets Negative Sums of Squares	No		Yes	There is an easy fix Fixed for 2003..
35	Regression	Does Not Handle Multicollinearity Correctly	No		?	There is an easy fix when it occurs. Not likely to occur in most regressions. Fixed for 2003.
36	Regression	Displays Normal probability Plots That Are Completely Wrong	No		Yes	Have to generate a new chart.
37	Regression	Makes Variable Selection Very Difficult.	No		No	One has to understand the "Range" concept in Excel first.