

# **NOTE A: COMMENTS ON TEACHING/USING EXCEL**

## **REASONS WHY MANY TEACHERS WILL NOT USE EXCEL IN THEIR COURSES.**

Highly limited graphics and visual displays of data.<sup>1</sup>

Excel outputs often are incorrect, and sometimes misleading on the accuracy of the output values.<sup>2</sup>

Inadequate or primitive data management. Data analysis tools should dynamically link output with data.<sup>3</sup>

Very limited set of statistical functions and routines

Inability to process large data sets, and large numbers of variables.<sup>4</sup>

Excel does not have a number of commonly used methods that have come into frequent usage since the 1980's and which are now taught in contemporary statistics courses.

Excel is not extensive enough to be able to be used as the computing tool in a laboratory-based, quality-oriented statistics curriculum for business or engineering students.<sup>5</sup>

Excel data entry and computed outputs are only a small part of the business problem solving process.

It is too much to try to teach both Excel and statistics from scratch in the program time and in a typical classroom usually allocated to the statistics course.<sup>6</sup>

## **ADVANTAGES IN USING EXCEL IN AN INTRODUCTORY COURSE.**

One is that the memorization of formulas and the use of hand calculators can be dropped. This is often proposed as an argument for Excel. However in the real

---

<sup>1</sup> Excel 2007 offers more graphic capabilities., but when compared to graphics in publications and on the internet, it is still limited.

<sup>2</sup> Thus study shows that accuracies in Excel 2007 now are limited mainly by the IEEE-754 floating double.

<sup>3</sup> Excel was always able to dynamically link output to the current inputs. The real issue is with commercial stat software that uses codes to link computations to hidden coded data sets.

<sup>4</sup> Excel 2007 has the capacity to hold 1.718E+10 data cells per worksheet. This would be a very, very, very large data set.

<sup>5</sup> With modern lap-tops, De Leve (2001 & 2004) has found that it is fully able to work in such an environment

<sup>6</sup> See Bell(2000). He has been doing it for years.

world of classroom quizzes, homework and examinations, the student still has to remember formulas.

The emphasis can be on statistical reasoning, how to choose statistical methods, how to interpret data and the building of logical, clean, and defensible arguments.

Excel works very well in illustrating how variations (random) in a value affect some outcome parameter.

## **PROBLEMS WITH USING EXCEL IN THE CLASSROOM**

Dawson (2003) points out that, “I cannot consider telling students ‘do not use any of these graphs, or any of these ones, or any of those except the two dimension versions, and reset all the defaults on this one each time you use it....’ as cost free. It either damages the instructor’s credibility or the credibility of the other functions of the software. He goes on to say, “If students have difficulty using a package properly, then I have difficulty using it to teach them. Excel simply doesn’t do many of the things that are required for serious data analysis.” This is some of the basic reasons why add-ins at the teaching level are so extensively used.

Another problem in using Excel in statistics courses is that the teacher cannot assess (i.e. grade) the student’s understanding, comprehension and abilities without timed, in-class quizzes involving number crunching of data. To deal with quizzes (and examinations) within a limited time period, a student has to learn the formulas, how to use printed distribution tables and how to do hand-calculator numerical calculations. As a result Excel gets put into the background, where it just becomes incidental to the class.

Another consideration is, “What will the students use after the course is over”. An assumption here is that each and every student in an introductory statistics class will eventually do some form of research, and will use the software tools used in research. This approach is a disadvantage to those students who will enter the business world, who will be using Excel and who will not be doing any scientific research. Students may also find that the company they work for, does not have SAS, SPSS, etc or other programs that were used at his school and which are not available off the company’s network. The company also will not allow personal software on the company’s computers. However he will find Excel available in just about all companies.

Excel has certain advantages that other competing software does not have. The spreadsheet capability and the built in Macro and VBA programming capability gives Excel a great capability for problem solving, “what-if” explorations, analyzing data and building complex relationships. It easily interfaces with financial problems and financial data.

Excel also has the capability for doing more advanced statistical modeling by the user, such as doing Structural Equation Modeling. ShROUT (2002) has written a module in Excel that solves the smaller structural models quite nicely. It is the use of the Solver routine applied to the maximum likelihood equation. She is expanding it to include solution of the combined structural-measurement model. The primary reason was to show how the solution process is done.

Love and Hildebrand (2002) discuss the role of Excel in teaching statistics in schools of business. This comes from the “Making Statistics More Effective in Schools of Business (MSMESM)” conferences held every year. The issues regarding Excel are directly from their article.

“Any discussion of teaching materials must of course address software. A particularly relevant topic for recent attendees has been the appropriate role for Microsoft Excel in the core business statistics course. Numerous contributors including Nash (1997) and Allen (1999) have provided resources and prompted arguments for and against spreadsheets for statistics, often leading to heated discussions. The motivations for using spreadsheet software in the business statistics classroom are clear. The software is ubiquitous, if not unavoidable. The spreadsheet is familiar to many students, used frequently by other disciplines such as accounting and finance, and comes at a negligible cost to many schools. Several textbooks are available that use Excel as the primary software tool for performing data analysis.

Yet the existing statistical function set and data analysis toolbox module are unreliable, and do not lend themselves to effective graphical analysis. Several authors have criticized the accuracy of Excel’s statistical procedures, including McCullough and Wilson (1999) and Knuesel (1998). Most textbook authors and instructors using Excel as the primary software tool incorporate add-ins. This produces greater reliability of results, and often succeeds in allowing instructors to get students to look at pictures before numbers, but at the cost of losing the main argument for Excel – that the package will be available “as-is” in the workplace. The market clearly demands the incorporation of Excel output in texts, yet few instructors are happy with the tool. The 1999 conference called for a more united response to the “Excel problem” from the statistical community at large, work that continues today.” (Love and Hildebrand 2002).

For some background on Excel in business, see Barton and Nowack (1998) and the following discussion by Bisgaard. The industry side of business is important, and six sigma improvements are an important aspect. Excel can be one of the effective tools here, and it fits in to what Hahn (1999) says about being able to measure, analyze, improve and control. However as Hillmer (1996) indicates, the use of Excel here would be a very small part of the problem solving process. Love and Hildebrand (2002) refer to schools of business teacher conferences where the role of Excel as part of the instruction has been the source of very heated discussions.

Excel has not kept pace with the increased scope of introductory statistics. Many of the current textbooks include Excel, but when the material gets beyond Excel, Excel gets dropped, and the focus is on Minitab and other statistical problem solving tools. Excel will fit the simple stuff at the beginning of the course, but can’t be used for the rest of the course. So why make the effort to teach Excel?

## **SOME COMMENTS FROM STATISTICS TEACHERS**

**Myles Gartland [myles.gartland@ROCKHURST.EDU]**

I currently teach undergraduate Applied Business Statistics for aspiring managers (every business student has to take it). Almost all of the basic books come with support for Excel, Minitab and SPSS. However, there seems to be a range of opinion on whether to

use MS-Excel in the class or a professional stats package like SPSS. Here are the arguments I hear for both.

For Excel: It is the dominant software package that business managers will have on their desktop "in the real world". Also, most of the other quantitative business disciplines teach with MS-Excel. Finally, because it is the only package that most managers will ever have access to, it increases the odds that a manager would generate statistics in their job.

For a Professional Stats package (like SPSS): Not only does it make most statistical calculations easier (allowing to spend more time on interpretation), it can do quite significantly more detailed and advanced statistics. Finally, it gives students a second software package for them add to their resume. Also, Excel has several statistical limitations.

Over the years, I have used one or the other, and in some semesters I have used both.

**Prof. Humberto Barreto, [barretoh@wabash.edu](mailto:barretoh@wabash.edu)**

I've taught undergraduate statistics and econometrics with a colleague for many years.

We use Excel because of student familiarity with the software and our ability to control the environment via Visual Basic macros. This enables us to develop workbooks that demonstrate via simulation such difficult concepts as a sampling distribution or P value. The ability to run a Monte Carlo simulation with the click of a button is one big plus for using Excel.

I am well aware of the computational deficiencies of Excel and the arguments against its use. Perhaps the disclaimer that we use in our add-ins would make our position clear:

**WARNING:**

This software was written and designed for teaching purposes. It has been tested on several examples, but not on a wide variety of datasets. For mission critical projects always check the results with full-fledged statistical programs.

In other words, Excel for teaching has been a good choice for us.

**Michael Granaas, Assoc. Prof. Dept. of Psychology [mgranaas@usd.edu](mailto:mgranaas@usd.edu)**

I'm going to argue against excel because of efficiency. I can't provide a lot of detail, but I am aware of someone who does large quantities of basic statistics using excel, means, medians, charts for display--basic stuff. As far as I can tell it takes them much longer to work their way through excel to get their results than it would if they were using SPSS.

If it takes longer to conduct an analysis with excel than with SPSS then it is in the employer's best interest to buy SPSS and to have employees who know how to use it.

**Lise DeShea [[DeShea@uky.edu](mailto:DeShea@uky.edu)]**

The argument FOR excel that is made because business folks will have Excel on their machine anyway ... is bogus. One might use the same argument for saying that since

WordPad is a part of Windows XP ... one does not need Word ... or Word Perfect. Sure, if you want to type a note to your pal ... WordPad is fine ... but if you have to deal with decently complicated documents ... WordPad is not going to do.

Everything done in Excel is a plugin of one sort or another ... even if you install Excel as part of the Office package ... you have to specially install the Analysis tool to have some statistical functions. And, there are 3rd party tools too like ... <http://www.analyse-it.com/> ... I'm not pushing it ... just saying it exists.

But then again, think about the students in APstat ... that have to contend ONLY with a TI83! ...

The problems with Excel tools ... either built in or 3rd party ... are many. Here are a few.

1. Data management capabilities
2. Graphing options
3. Range of statistical procedures available

etc.

I think if I were in a business program ... I would want my profs to help me see that it's important to learn good tools for doing different things and Excel is NOT the way to go if one is serious about statistical work.

**Dennis Roberts: [dmr@psu.edu](mailto:dmr@psu.edu)**

The argument FOR excel that is made because business folks will have Excel on their machine anyway ... is bogus. One might use the same argument for saying that since WordPad is a part of Windows XP ... one does not need Word ... or Word Perfect. Sure, if you want to type a note to your pal ... WordPad is fine ... but if you have to deal with decently complicated documents ... WordPad is not going to do.

Everything done in Excel is a plugin of one sort or another ... even if you install Excel as part of the Office package ... you have to specially install the Analysis tool to have some statistical functions. And, there are 3rd party tools too like ... <http://www.analyse-it.com/> ... I'm not pushing it ... just saying it exists.

The problems with Excel tools ... either built in or 3rd party ... are many. Here are a few.

1. Data management capabilities
2. Graphing options
3. Range of statistical procedures available etc.

I think if I were in a business program ... I would want my profs to help me see that it's important to learn good tools for doing different things and Excel is NOT the way to go if one is serious about statistical work.

I think when we focus on ONE specific analysis method, we miss the point about Excel COMPARED to regular statistical packages. What if we want to do a correlation? I am sure Excel and it's plugins will do fine. But, the broader issues are:

1. How can you manage your data with Excel ... data editing ... data transformations? and the like? Most regular stat packages have evolved to

make them good interfaces for data management. What if you need to recode a column of data where codes might be alpha characters and you need numerical values? Can Excel do this easily if at all?

2. What about the set of graphical techniques available ... from simple things like dot plots and stem and leaf diagrams ... to pseudo 3 dimensional plots...? My experience with Excel is that it is very limited indeed when it comes to graphics. What sort of graphic EDITING capabilities does Excel have? How can you highlight say ... M and F data cases in a scatter plot between height and weight?
3. There just is no argument that the set of analytic procedures in Excel and its plugins ... is highly limited.
4. What about support for the product? If there are technical problems, can one contact the Excel Statistical Company and get help? No package is perfect but if we rank the main statistical packages and then put Excel on that chart, it will be so far at the bottom as to make ANY of the main packages look majestic by comparison.

Finally, again, I think what we should be emphasizing with students AND others who are or will be doing rather regular statistical analysis ... we should help them see that there are tools that were designed specifically to do these tasks. Why settle for less?

As for cost ... one does NOT have to purchase SPSS or SAS ... to have a good package. In addition, there are packages even online that probably can outperform Excel in its best-case scenario. I will list two links that might provide useful information about things available online and also through the regular channels.

**Peter L. Flom, PhD, Assistant Director, Statistics and Data Analysis Core, Center for Drug Use and HIV Research, National Development and Research Institutes**

IMHO, even SPSS for Windows makes statistical computations considerably too easy to do WRONG. Point, click, bing. Or maybe BANG.

Even for something "simple" like taking the mean, there are many ways to go wrong. WHAT did you take the mean of? WHY? Might a median have been better? Were there outliers? Were there wrong values? Was the data cleaned?

At least, if you write syntax for something, you have to think a LITTLE while doing it. Programs, which make it easier to avoid thought, are a menace.

**Dale E. Berger, Ph.D., Professor of Psychology, Claremont Graduate University**

I have traced most of my own problems with Excel and other programs back to the nut that holds the mouse. The darned programs do what we tell them to do, not necessarily what we want them to do. That is perhaps the biggest problem our students face when they begin to use sophisticated programs.

**Derek Christie: derek.christie@wintec.ac.nz**

Excel is the ideal tool if you want to use a simulation in your teaching or as a student activity because it can produce results with applet like graphs instantaneously and as often as you want. I find that one problem students have is that they usually only see one of the multitude of possible realizations of a sample from the population – the single data set that they are given. Often students have the feeling that what they see from their data is somehow it – the real thing, and they are usually surprised at just how variable things are in real life. Using Excel’s random numbers and graphs (and possibly on screen controls), it is easy for teachers or students to produce applet type displays, which change or refresh with the press of a key or the click of a button.

To use the CI example, generate 20 samples in Excel with their confidence intervals and plot all the CI’s on one graph. Each time you press F9 another 20 CI’s will appear. On the average one of the 20 CI’s won’t cover the true mean. Students often find the variety of CI’s (their size and position) unexpected because when they do a CI calculation they only ever see one interval and they imagine that that is it.

I have my students (trainee technicians) produce their own Excel CI simulation as a follow up from a practical session in class. Email me [Derek.Christie@wintec.ac.nz](mailto:Derek.Christie@wintec.ac.nz) if you want an example.

**Peter C. Bell, Richard Ivey School of Business, University of Western Ontario, “Teaching Business Statistics with Microsoft Excel”, *INFORMS Transactions on Education*, Vol. 1, No. 1, <http://ite.informs.org/Vol1No1/Bell/>**

“This article discusses the use of Excel in the introductory course in statistics in the business school. An attempt is made to identify how business school programs differ, and how the use of Excel to work cases or examples can contribute to a business statistics course that is in keeping with the school’s and the instructor’s learning objectives.

“There are possibly more different ways to teach statistics to business students than any other subject in the business school curriculum. The business statistics course can range from a highly quantitative presentation of distribution theory, to a course that might be described as “strategic analysis with data”. The observed variations in pedagogy and content reflect differences among schools and faculty, but also reflect a varying response to changing technology. This article discusses some issues that are important in designing the introductory business statistics course, and presents some ideas for using spreadsheets that may be appropriate in teaching statistics in the business school.

•

• *Left out material, see his full article*

•

“The emphasis on real problems suggests the use of software to facilitate performing the statistical calculations, and here the instructor has a choice between using special purpose statistical software (such as SPSS or Minitab) or the general purpose Excel spreadsheet. The case in favor of using Excel includes the following points:

- Excel is very widely available in industry and is now quite standard,
- there is a low set-up cost to using Excel: most students will have used Excel before the course,
- there is both an immediate and a lasting value to Excel: your students will be able to use their Excel skills in other courses, and will almost certainly have access to Excel when they start work,
- improving students' Excel skills supports program objectives of improving students' general problem analysis skills, and
- Excel has instant credibility with students, who recognize Excel as a very widely used and powerful "number-crunching" device.

“There are also some advantages to Excel in terms of statistical calculations:

- The Excel statistical functionality is seamlessly integrated with the other Excel functions: it is simply a part of the spreadsheet.
- Excel is very good for doing basic data analysis, including graphing and producing tables,
- the Excel basic statistical functions (including average, standard deviation, median, slope, intercept) are easily accessible through the “function wizard” and generally produce the answers statisticians expect,
- the “Analysis Toolpak” adds a great deal of more advanced statistical functionality,
- simple (i.e. one-independent variable) regression is easy if done with “Trendlines” although this results in very limited statistics ( $R^2$  and regression coefficients only).

“It is, however, important to recognize some of the weaknesses of Excel:

- Some (a few) students will have little prior Excel experience,
- Excel makes analysis so easy that some students will do too much analysis,
- it is easier to crunch numbers in Excel than it is to think about why this is being done, or what value it is,
- crunching numbers in Excel is much easier than thinking about softer or more managerial issues,
- Excel is poor at handling mathematical formulas, or processing symbols,
- the function formulas are hidden when using Excel, and
- the “Analysis Toolpak” is quite often cumbersome (e.g. Histogram, Multiple Regression),

*Bell responded in an Email:*

“My text "Statistics for Business with Lotus 1-2-3" was published in 1989. The second edition "Statistics for Business with Spreadsheets" updated to Excel was published in 1992. We still use this text (now as a hypertext).

“I can't understand why anyone teaching Stats in a business School would not use a whole lot of Excel.

“Best .....Peter”

**Dennis Roberts, dmr@psu.edu**

One of the add on products for Excel is Winstat ...<http://www.winstat.com/>

They list it's cost as \$99. Another that I mentioned is Analyze It ... <http://www.analyze-it.com/>. A 1 year student license for the General package is \$30 ... but has to be renewed after a year. SigmaXL is available here ... <http://www.sigmaxl.com/> ... for \$199. Stat Tools is here ... <http://www.palisade.com/html/stattools.asp> ... for \$350. An academic version of XLfit is here ...

<<http://www.idbs.com/xlfit4/index.asp?source=GoogleAdWordsXLfit4-Excelstatistics-1>>for \$300.

Now, in almost every university ... the full package of Minitab Rel 14 is available for under \$100. If one can't obtain it there, then see <http://www.e-academy.com/>.

I just cite ONE major package ...

I would rather spend LESS on a real statistical package than MORE on something that is an add on.

Just my opinion.

**Robert Dawson**

Well, Excel isn't alone in doing this; MINITAB (my weapon of choice)will cheerfully sort one column of a paired data set too, and I've made said mistake many a time. I haven't used SPSS much but I recall it being possible there as well.

I understand that there is software out there (and I've probably even experimented with some of it) that by default stores data in structures that enforce the relevant model with some firmness. (ViSta comes to mind but when I tried to download a copy just now it wouldn't run on my machine so I am still unsure] I would suppose that this would make it harder to experiment casually to determine what model was appropriate - though there are good counterarguments that such poking around can lead to data snooping and ought to be hard.

One built-in weakness -from the pedagogical/newbie viewpoint- of Excel is that its data storage model is \*quite\* \*properly\* more unstructured than a data analysis tool's ought to be, (properly, given that Excel is not a data analysis tool). MINITAB and other products with a "[pages of] row[s] of columns of cells" structure are closer to reflecting the real structure of data sets but don't quite get there.

One might use the analogy that Excel is like one of the scripting languages that does little or no type checking but simply tries its best to convert (Perl, some Basics). MINITAB is like classic Fortran, with typing but not a lot else. Some of the older "bundle of programs" packages seem almost like shell scripting. And my impression is that the newer wave of software that I haven't used much is much closer to the modern object-oriented languages. This last is probably a good thing...for most purposes. (Remember, assembler is not dead yet, just a niche tool!)

## **Jay Gould**

At the rate this group is going, the entire pot will melt, spilling hot metal all over the floor! First it was teaching intro stat – possible or not? Now we're reviewing the limitations of Excel (yes, they exist). Where are they in relation to the students' upper limits?

In Excel 98, when I sort one column of a table, it reminds me that adjacent columns contain data, and do I really want to proceed. And yes, my favorite db, FileMaker, would never consider getting that close to that mistake.

In any serious stat situation ('mission critical,' if you will), you must work with validated software. Has your Excel spread sheet been validated? I doubt it, a lot. Mine aren't, either. Then you really should ask, has your stat software been validated? Then the clincher: how do you know? I sure hope that the last place Martin worked at, not to mention his present one, was able to answer the last question satisfactorily.

As for intro stat for MBA wanna be's, how much validation is enough?

Cheers,

## **Dale E. Berger, Ph.D., Professor of Psychology, Claremont Graduate University**

I have traced most of my own problems with Excel and other programs back to the nut that holds the mouse. The darned programs do what we tell them to do, not necessarily what we want them to do. That is perhaps the biggest problem our students face when they begin to use sophisticated programs.

## **Robert Dawson**

Well, that is so; but they do make their preferences clear. I had a bicycle once that had been in a bit of a prang. It would go straight forward if you kept a constant pressure on the left handlebar but otherwise it would turn left in about a 10m arc, towards oncoming traffic.

One day it got stolen. I kept wondering whether the thief figured it out in time or ended up as a hood ornament.

With good software - and much of it is - the novice should be able to assume that the most direct way of doing something standard is the right way because it was made so by somebody who knew what he or she was doing. Used as a spreadsheet, Excel may pass that test - used as a stats package it fails abominably.

## **Derek Christie [Derek.Christie@WINTEC.AC.NZ]**

Applets are nice but so often they aren't quite what you want or what you want isn't there.

Excel is the ideal tool if you want to use a simulation in your teaching or as a student activity because it can produce results with applet like graphs instantaneously and as often as you want. I find that one problem students have is that they usually only see one of the multitude of possible realizations of a sample from the population – the single data set that they are given. Often students have the feeling that what they see from their data

is somehow it – the real thing, and they are usually surprised at just how variable things are in real life. Using Excel's random numbers and graphs (and possibly on screen controls), it is easy for teachers or students to produce applet type displays which change or refresh with the press of a key or the click of a button.

To use the CI example, generate 20 samples in Excel with their confidence intervals and plot all the CI's on one graph. Each time you press F9 another 20 CI's will appear. On the average one of the 20 CI's won't cover the true mean. Students often find the variety of CI's (their size and position) unexpected because when they do a CI calculation they only ever see one interval and they imagine that that is it.

I have my students (trainee technicians) produce their own Excel CI simulation as a follow up from a practical session in class. Email me [Derek.Christie@wintec.ac.nz](mailto:Derek.Christie@wintec.ac.nz) if you want an example.

Cheers

**Frank Bokhorst, University of Cape Town**

Excel does give quite a nice friendly "desktop" for the students to work with.

At UCT we use Excel very extensively for teaching basic numerical skills, numerical literacy, and basic statistical concepts to large numbers of students. We use Excel with some very complex Visual Basic macros to provide online "tutorials" that allow students to perform statistical calculations, which are immediately checked, feedback is given, marks allocated, and stored automatically. The students' spreadsheet, with all the work done is submitted to a database, and may be checked off-line. We do this mostly in the place of face-to face tutorials, simply because the classes are so large (e.g. 400 students in PSY206F). Often the quality of the feedback a student gets is better than if a tutor were to mark the work by hand. We can check not only calculations, but any object, e.g. a graph to see if it has an appropriate title, if the axes are correct, etc.

Excel also has a specific advantage in this respect because it integrates well into networked data systems. See Deacon, Jaftha, Horwitz (2004) Customizing Microsoft Office to develop a tutorial learning environment. British Journal of Educational Technology, 19(1): 24-45.

Some details about an Excel tutorial that I used, in conjunction with a Java applet to collect data from a computer-based experiment, is described in a conference paper by Deacon & Jaftha (2004): Abstract available at <http://emerge2004.net> (the full text can be obtained from me personally)

For the tutorials I design for my classes, I do not use the built in Excel statistical analysis tools, but require the spreadsheet formulas and functions to do the same "by hand".

So, for e.g. to do a t-test, you calculate the mean, standard dev., and then the appropriate std.error, to calculate t, and then use the Excel function for the t distribution to get the probability.

Personally, I am emotionally biased against all Microsoft products, but in this respect, I simply must agree that there is a place for Excel in my teaching, and so I use it!

**Martin Holt, MSc, Medical Statistician, Thornton, Leicester, UK**

**theholts@care4free.net**

Having had a go at Derek Christie's simulation in Excel, yesterday, I would like to echo his comments and those of Frank's just now. Derek's produces 20 CI's on a graph in front of you, so that you can see how many do not encompass the population mean. Then, at a touch of a button, it will work out another series that replaces the first, but is different, and so on. It was fun and instructional (is that a word ?) to do. Most of Derek's functions were 'done by hand' (echoing Frank's comments). Does anyone know where you can get hold of more of these ?

Excel is a spreadsheet. The press so far has not been good. Is that because some of Excel's statistical functions are just plain wrong (that was my impression with Excel97, but what about now ? [I have Excel 2000}). Is that why Frank says that most of his work is 'by hand' ? or is he emphasizing the teaching aspect ?

PS. Thanks to Robert J. MacG. Dawson for his answer to my dof question.

Best Regards,

**Dennis Roberts, [dmr@psu.edu](mailto:dmr@psu.edu),**

**<http://www.personal.psu.edu/users/d/m/dmr/droberts.htm>**

We need to separate the notion of working through a problem by hand ...from letting a computer package go through these same steps ... from actually having the package do the entire simulation for us.

No one is really going to do 20 CI intervals by hand ... even if the data are fairly easy. So, what if you want to mimic the STEPS you have to go through but be able to repeat this process ... which CIs in the long run represent ... over and over?

Now, one can say that in the mimic process, one has to do the actual steps in every calculation ... so if step one is to find the mean, then:

1. One has to physically add up the scores
2. Then divide by n

If one has to find the SD from sample data ... then one has to:

1. Find the mean from above ...
2. Subtract the mean from every score
3. Square the deviations found in 2
4. Take an average by dividing by n-1

And on and on

OR ...

One could mimic those steps by saying:

1. You have to find the mean of each sample
2. You have to find an SD based on the sample data in each case
3. You have to find the SE of the mean for each case

4. You have to add and subtract (to keep it simple say 2) 2 SE units from the sample mean ... to find the end points of the CI
5. You have to evaluate each interval to see if the designated mu value is inside or outside
6. You need to convert that volume of captures to a % ... to see how close it is to say 95%

As long as one knows how to do a CI ... even in one case ... I don't see where going through all the steps in the first case above ... is necessary ... whether you do it in Excel or any other package.

What IS important is the process listed from 1 to 6 ... knowing what you need to find ... to result in n CIs that you can then check to see how often the CI captures the parameter.

In Minitab, one can do this easily ... as I have shown before ... and have Minitab actually evaluate each interval and sum up how many of say 1000 captured the parameter.

If you want to do this repeatedly ... then you make this little routine a macro ... and save ... and then student 1 can run it once ... see what happens ... re-run it ... see what happens again ... until you are blue in the face.

So, like pressing the "do again" button in Excel ... running the macro is the same simple task ... say the macro you have created is called "CIDemo" MTB> exec 'CIDemo' and it's done

OR, you can generate sufficient #s of samples from the git go ... so that doing it over and over again is not necessary.

Personally, I think doing 20 at a clip ... will NOT be sufficient to hone in on 95% ... unless you keep tally across many batches of 20 ... But, if you had it do say 1000 intervals ... or 10000 intervals ... which Minitab (and I am assuming Excel too) can do easily ... then one pass is sufficient to drive the point home.