
Some suggestions for the future development of MINITAB™

Harper Gilmour
Statistics / Public
Health
University of Glasgow

h.gilmour@stats.gla.ac.uk

Minitab responses by:
Gillian Ingram, Trainer
Isaac Newton,
Technical Support
Specialist
Minitab
inewton@minitab.co.uk

“MINITAB and the MINITAB logo are trademarks or registered trademarks of Minitab Inc. in the United States and other countries and are used herein with the owner’s permission.”

For over 20 years Minitab has been developing user-friendly, yet powerful statistical software, used in more than 60 countries and referenced in well over 300 academic textbooks.

Minitab News is a FREE monthly electronic newsletter for MINITAB enthusiasts. You can subscribe or unsubscribe to Minitab News right from our website: <http://www.minitab.com/contacts/minitabnews.asp>



I have been a user of MINITAB™ Statistical Software in its various forms on a range of computing systems since the early 1980s. My earliest memories of using MINITAB are of having to walk several hundred metres from my base in the Mathematics building, cross a busy main road to the basement of the Chemistry building, and queue outside a small room which was equipped with about a dozen interactive terminals with access to MINITAB on a distant mainframe computer. Response times were not quick and it would probably have been less time consuming, for a small data set, to have remained in my office and done the computations by hand!

MINITAB has evolved over the years into a user friendly, Windows based package. Recent versions have incorporated a number of major improvements and additions. For example, the suite of quality control procedures makes MINITAB a leader in this important area of industrial statistics. The general linear models procedure has been revamped and, in particular, now provides a range of methods for follow up multiple comparisons. The addition of the 1 and 2 proportion(s) and 2 variances procedures have made it easier to obtain confidence intervals for population proportions and their differences and for population variances. The power and sample size procedure has obvious applications in study design.

MINITAB is my first choice package both for teaching and for my own collaborative research work. It is an excellent package for use in teaching introductory statistics courses where the main aim is to equip students with practical skills in data analysis. It also has a sufficient coverage of more advanced techniques such as general linear models, logistic regression and multivariate analysis to be useful in teaching more advanced courses and for much collaborative research work involving data analysis. In the department of statistics at Glasgow MINITAB is used in teaching honours students in Statistics in conjunction with Splus and, to a lesser extent, SPSS.

The purpose of this article is not to provide a full review of MINITAB’s capabilities. My main objective in the remainder of this article is to make suggestions for future developments in MINITAB, based on my own experiences and on discussions with colleagues, which I believe would enhance its already strong reputation.

Most introductory courses in Statistics, whether method based or problem based, are very likely to include the same core of statistical methods: presentation of data in tables and diagrams (such as dotplot, boxplot, scatterplot, pie chart, bar chart); numerical summary (using mean, median, standard deviation, quartiles, proportions, percentages); standard errors; appropriate confidence intervals and hypothesis tests for 1, 2 and paired samples contexts for both continuous and categorical variables; one way ANOVA and multiple comparisons. I teach courses, which include these methods, to first year undergraduate students and to postgraduate students doing a Master of Public Health Degree. In both of these courses, MINITAB output is used extensively in lecture notes and MINITAB is used by students in analysis of data practicals. I think there are numerous minor changes which could easily be made to MINITAB which would make these practicals easier for staff to organise and supervise, and improve the students’ educational experience. There are also a number of new features that could be added to avoid the use of locally written macros, or reference to other packages, for simple analyses.

1. Harmonisation of dialog box features

It would make teaching introductory analysis of data practicals in MINITAB much easier if the dialog boxes for basic methods all worked in essentially the same way. For example:

- a) The 2 proportions, 2 sample t and Mann Whitney dialog boxes
 - These should all allow samples of data to be in one column or in two separate columns (Mann Whitney does not allow this at present. Columns need to be unstacked before using this test).

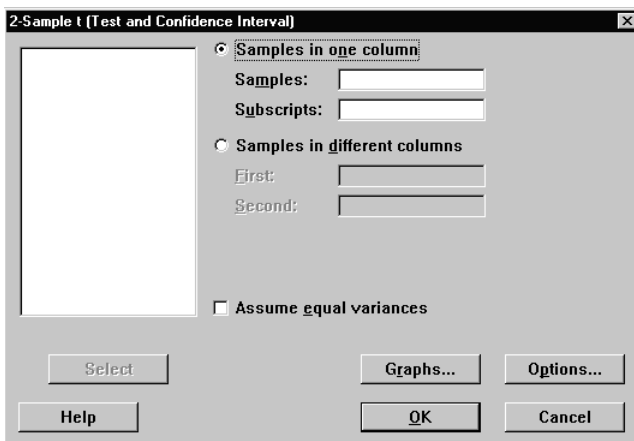


Fig 1 2 Sample t dialog box

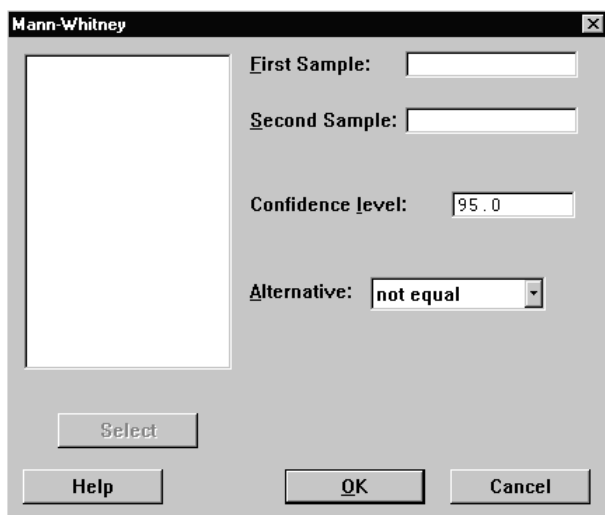


Fig 2 Mann Whitney dialog box

- The terminology “samples” should be replaced by “response”, and “subscripts” should be replaced by “group” or “factor” (This would be more consistent with standard terminology, and is already used in the one way ANOVA procedure).
- There should be check boxes to allow a choice of

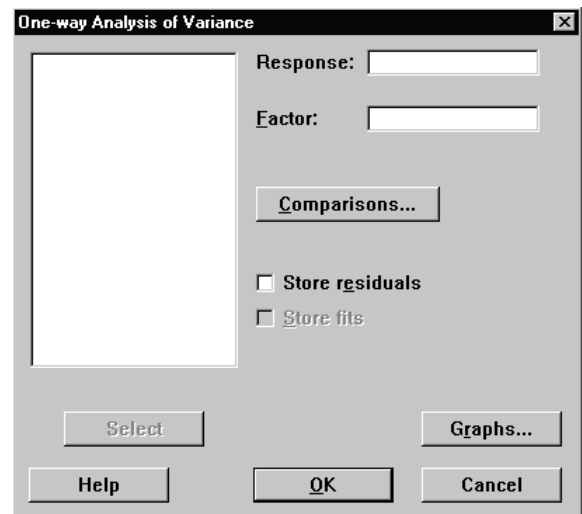


Fig 3 One way Analysis of Variance dialog box

confidence interval (Yes/No), test (Yes/No) to give users maximum control over the output. There are many occasions in teaching when I want students to concentrate on the confidence interval without being distracted by the p value.

Minitab: This is possible to a certain degree in Release 13. With the 1 sample t-test you can choose to have the CI or the CI and test.

- The output for the Mann Whitney test needs to be revamped. At present the p value is displayed in an unhelpful way, and appropriate 1-sided confidence intervals are not constructed when a 1-sided alternative is specified.

Minitab: In Release 13 CI's were changed to one sided for a one-sided test for t-tests to be more statistically accurate. A bonus of the MW output is that you are told if you can reject at $\alpha = 0.05$ which you are not given in the parametric tests.

- The very useful feature of allowing summary data in the 2 proportions dialog box could be extended to allow summary data (mean, sd and sample size) for the 1 and 2 sample t dialog boxes also.
- The number of significant digits used in the 2 proportions dialog box is excessive.

Minitab: 7 decimal places (as was the case with the example I used) is a bit excessive but other customers tell us that some of our session output does not have enough decimal places! This probably comes under the general issue people have of wanting to be able to control the number of decimal places in session output.

b) The 1 proportion, 1 sample t and 1 sample Wilcoxon dialog boxes

- These should be harmonised as above.
- In particular the 1 proportion procedure automatically tests the null hypothesis $P=0.5$ and does not allow this to be deselected. In my experience this test is almost never relevant and causes confusion among inexperienced students. As a general principle I think that the default output for all procedures should be kept to a minimum, with additional output available via options in the dialog boxes.

c) One way ANOVA

- The two options available for one-way ANOVA at present could be combined into a single one way ANOVA dialog box (with options for different data layouts). This would be consistent with changes proposed for t tests earlier.
- It would be extremely useful to have Bonferroni multiple comparisons (the most commonly used in my experience) available here. I know that these are available with the GLM procedure, but students doing introductory courses do not need the complexity of this more advanced procedure.

Minitab: We do offer 4 tests for both one-way ANOVA and GLM.

II. New features to add

a) Basic methods

In my opinion, MINITAB's greatest weakness is its lack of facilities for dealing with simple categorical data. There have been some improvements in this area in recent versions, with the inclusion of the one and two proportion procedures and greater flexibility in dealing with text variables. However the only hypothesis test available in the Tables procedure is the simple chisquare test. It would be extremely useful to have options to select the chisquare test for trend (for use in a $2 \times k$ table where one of the variables has k ordered categories), Fisher's Exact Test (for use in 2×2 tables with small expected values) and the continuity corrected version of the chisquare test for 2×2 tables which is commonly used.

Minitab: The Macro is still available on our website for Fishers Exact test. It is something that is asked for every now and then.

It would also be extremely useful to have a measure of

agreement (such as Kappa) and possibly a measure of the strength of the association between categorical variables. Simple measures of risk used in epidemiological studies (such as relative risk and odds ratio, with confidence intervals) could also be included. A simple goodness of fit test for categorical data would also be useful.

There are no facilities for dealing with paired categorical data at all other than simple tabulation and percentages. It would be extremely useful to have McNemar's Test for 2×2 tables and the tests of symmetry and marginal homogeneity for larger symmetric tables. A paired proportions procedure could be added to the basic statistics menu to provide a confidence interval in this context.

b) More advanced methods

Everyone will have their own 'wish list' of more advanced methods they would like to see added to MINITAB. Top of my list, as a medical statistician, is Cox's Proportional Hazards Model (Cox regression) for Survival Data Analysis. A poll was conducted several years ago among medical statisticians to ask for opinions about the most important advances in medical statistics over the last 30 years. Top of the poll (by a long way) was Cox regression. Survival data analysis features extensively in various courses at Glasgow and at present we have to use other packages in teaching these courses.

Next on my list is stepwise methods for logistic regression, discriminant analysis (and Cox regression). The current versions of these are good at what they do, but since stepwise methods feature in several courses at Glasgow, at present we have to use other packages in teaching these courses.

III. Other issues.

a) Enabling command language

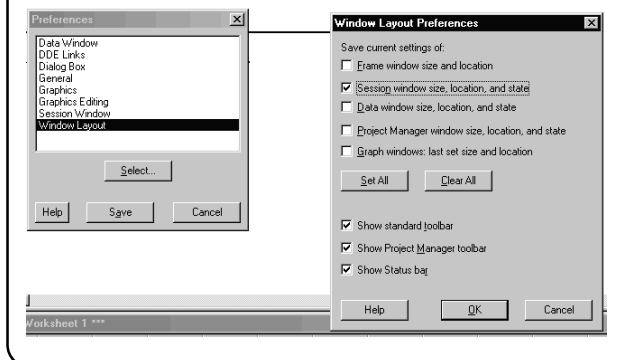
When MINITAB is first loaded onto a PC, the default option is to have the command language disabled, so that there is no prompt in the session window. While I agree that students should be encouraged to use dialog boxes, there are some tasks (such as READING data from a data file or doing arithmetic via the LET command) which are easier to do by typing the command at the MTB> prompt. I have lost count of the number of students who have bought their own copy of the package or downloaded the 30 day trial version from the Minitab website and been completely stumped by this, sometimes wasting several hours trying to discover why their data sets cannot be read. (There is no

indication whatsoever of the underlying problem; MINITAB, apparently, does not respond.) Please change the default.

Minitab: This is easily changed, simply use the preferences to set the new session window layout as default.

Hint from Ewan Crawford on how to do this:

1. Select the session window to make it 'on top'.
2. Go to 'Edit' in the menu and then 'Preferences'
3. Choose 'Window layout' and click on 'select'
4. Tick the 'Session window size, location, and state' box and click on 'OK'
5. Click 'save'.
6. MINITAB should focus on the session window on closing and re-starting.



b) Use of supplementary text

At the start of this academic year when we switched to MINITAB 13, we got two nasty surprises. The first one was that supplementary text was no longer permitted in MINITAB commands. Many of our lecture handouts illustrate MINITAB session window commands with supplementary text for clarity. These now need to be changed. A much more serious problem is that macros containing supplementary text now fail to work. As long term MINITAB users, our department has a large number of macros which had to be checked on all the servers and all the individual PCs they had been copied on to – a non trivial task. It is not clear why this change was made.

Minitab: This is to bring the macro language more into line with programming languages. The use of the # symbol to denote a comment is still allowed.

c) Support for old worksheet formats

The second problem was that support for certain old worksheet formats was discontinued. We have a large collection of data sets for use in teaching many of which were created in 1993 – 95. At present it is only the oldest of these that we have had difficulty opening in

version 13, but many of the others open with obvious problems, especially involving the variable names. If you continue with this policy, we will have to open and resave hundreds of data sets (which again exist in multiple copies on different servers and individual PCs). It could be argued that many of these data sets should be replaced with more recent examples and this is undoubtedly true. However there are many of the old data sets which are still educationally useful and will continue to be for years to come. Please continue support for old worksheet formats.

Minitab: Old worksheets are still supported. However they need to be opened slightly differently, because of the formatting of the worksheets. Simply go to Open Worksheet as normal, and select the file for the older format. Instead of opening the file notice a tick box in the lower right hand corner it says “Merge”. Make sure this is checked, then click open. The file will now merge in to the current worksheet. This is only relevant for worksheets of release 9 or earlier. The session command for this would be

Wopen “K”;
Merge.

For K substitute filename.

Summary

MINITAB is my first choice package both for teaching and for my own collaborative research work. I expect that to continue for the foreseeable future. The modifications and additions I have suggested will, in my opinion, enhance the package and make it more attractive to students. As it stands at present, I find it difficult to recommend to my postgraduate students in Public Health as the one they should continue to use in future, mainly because of the limitations - especially in analysis of categorical data - highlighted above.

[It has been suggested that I should use SPSS in my Public Health teaching partly because of these limitations and partly because SPSS is more commonly used by health professionals outside the university sector. I think there is a huge potential market in this area, but modifications of the package would be required to meet their needs.]

Minitab: From the early 1970’s Minitab’s constant goal has been to solicit feedback from its customers. This has enabled them to create a more useful and robust statistical analysis tool. We continue to solicit feedback from our customers so that we are able to serve them better in the future. Should you wish to communicate your feedback to Minitab, please address your e-mail to inewton@minitab.co.uk.