

Software Review

A Review of Minitab – Release 11

SIVA GANESH

Department of Statistics, Massey University, Palmerston North, New Zealand

1. Introduction

I have been using Minitab for the past 15 years or so and regard it as my most favoured ‘every day’ statistical software package. Minitab is an excellent tool for both EDA (Exploratory Data Analysis) and Confirmatory Analysis. Its wide use in the University, Industry and Business environments speaks for its popularity in terms of both reliability and ease of use. However, like all good things Minitab also comes with a few ‘black spots’!

The latest version of Minitab, Release 11 for Windows environments, is a 32-bit application, designed to run under Windows 95 or NT operating systems. However, Release 11 is supplied with the necessary 32-bit extensions to run under Windows 3.1x, although some features such as the new ‘tool and status bars’ of Windows 95/NT are dormant. The package comes with one set of disks for installation on any one of Windows 95, NT or 3.1x platforms. This release follows the traditional habit of Minitab Inc. appearing with a range of new features, with a few ‘old’ ones being phased out. The main comfort for an experienced user of Minitab is that the standard features of the Session, Data, Info and History windows, and the pull-down menu structures, remain much the same as that of the previous release - hence, ease of quick transition!

The major change with the Session window is that the output can be pasted into a word processor with formatting intact. The output in the Session window can also be saved in RTF format. The changes to pull-down menus are mainly the re-arrangement of features within each menu plus a few add-ons. Another improvement is, the context-sensitive pop-up menus that can be invoked by right-clicking the mouse in any of the active windows. Windows 95/NT users may also find the new toolbars and status bars useful, although the toolbars which vary window to window, cannot be ‘customized’. Other Windows 95 features are the ability to use long file and column names, and to start Minitab and open the worksheet by double clicking the .MTW file in ‘Windows 95 Explorer’.

2. System Requirements

Minitab will perform satisfactorily with the minimum system requirements quoted (an IBM compatible 16 MHz 386 processor with 8 Mb RAM etc.), though heavy use

of statistical functions and graphics would slow down the processing dramatically. In my opinion, you require at least a 486DX66 processor with 16Mb RAM for use of Minitab at its full-strength.

Although being phased out, one cannot resist mentioning the ‘discomfort’ the device called ‘hardlock’ created among many of the international users. This hardlock must be plugged in the parallel port of the PC in order for Minitab to work. This, of course, is an attempt by Minitab Inc. to stop the illegal copying of the software. It is pleasing to know that Minitab Inc. have decided to discontinue this ‘hardlocking’ process, though many of us still need to put up with it until the next release of Minitab!

2.1. Data and File Management Capabilities

A significant number of enhancements/revisions have been made in Minitab’s *Data and File Management* capabilities:

- Any Minitab worksheet can now be *previewed* before retrieving (as well as other files such as Excel worksheets and text/data files, which was a feature of Release 10.5Xtra).
- Most MINITAB commands now allow *text* (or *alpha*) *data*. A column of text data is often used to specify the levels of a categorical variable (such as using “male” and “female” as factor levels instead of 1 and 2).
- Release 11 also handles *Date/Time data* which can be created, exported or imported. It can read in dates/times from a selection of packages/file formats. In most Minitab commands, date/time columns are treated like numeric columns. Examples include placing date/time labels on x- and/or y-axis of a plot, and subsetting data based on a date/time variable. The ‘millennium’ date formats are handled satisfactorily, though I would advise reading the manual or checking the help facilities before attempting any substantial computations.
- A *calculator* has been introduced by combining (and enhancing) the ‘Mathematical Expressions’ and ‘Functions’ features of the previous release. This can be used for easy calculations and transformations of variables.
- *Improved Worksheet* - Data manipulation capabilities have been improved in that columns can now be inserted and moved in a similar way to typical spreadsheets. Improved Stack/Unstack and Code menus provide facilities for easy stacking (and unstacking) of variables/columns and blocks of columns, and changing ‘text to numeric’ or vice versa or coding ‘numeric to numeric’ or ‘text to text’ or using conversion tables for these activities.
- *Descriptive information* of your worksheet can be viewed and/or saved when saving your Minitab worksheet. This feature is similar to that of many Microsoft products.

- New to Release 11 is the ODBC (Open Database Connectivity) feature that is common among many database packages. This enables Minitab to access data directly from Access, Paradox and the like.
- Minitab now provides *double-precision* accuracy to its worksheet.

It should be noted that many of the above changes will result in incompatibilities with previous releases of Minitab. i.e. Release 11 worksheets can be saved in release 10 format, but Minitab warns you about the likely loss of information. A colleague discovered that Release 11 worksheets saved as ‘portable’ files could not be retrieved by Release 9 of Minitab!

Although Minitab provides facilities to format variables/columns in its worksheet, it lacks the ability to format data to a specified number of ‘significant’ figures. For example, we cannot have a column with values such as 1997, 19.97, 199.7 and 0.1997 - all values become formatted to the default number of decimal places (say, 2 or 4). Although one wouldn’t expect Minitab’s worksheet to behave like a typical spreadsheet, some features such as the above would enhance Minitab’s data handling capabilities.

2.2. Statistical Capabilities

In general, the choice of statistical tools offered by Minitab is ‘ample’, and is well implemented. These span from simple basic statistics and EDA to more advanced GLMs, Time Series, Quality Control and Multivariate techniques.

In release 11, there are numerous new statistical features:

- *Logistic Regression* - similar to linear regression, helps you investigate the relationship between a discrete response variable (binary, ordinal, or nominal) and possible predictors. This was a big gap in Minitab’s previous releases. It should also be mentioned that only a few other packages offer Ordinal Logistic Regression capabilities.
- *Polynomial Regression* - This could only be regarded as a handy shortcut to fitting polynomial (second or third order) regression models as it just computes the 2nd and/or 3rd order terms and sends everything computed to the regular regression command.
- *Correspondence Analysis* - a technique that helps you further explore the relationships in contingency tables with two or more categorical variables. Although the technique has been known since mid-1930’s, it was largely ignored for many years and been popularized only recently. Since then only a few packages (including Minitab Release 11) have a fairly good implementation of the technique.
- *Proportion test* - currently, there are no direct commands to perform hypothesis tests etc. for proportions. However, Release 11 comes with a macro to perform this particular test, but only for single sample cases. I would like to see this

area developed into a genuine menu/dialog of the 'Basic Statistics' command with both single and two sample cases.

- *Statistical Graphs* - Many of MINITAB's statistical analyses now provide options to generate graphs automatically to help you graphically analyze the data. You'll find these new graph options in dialog boxes for many basic statistics and regression analysis commands, as well as analysis of variance and multivariate analysis commands. My only concern is that these dialog boxes do not provide adequate facilities to control the appearance of graphs. For example, you cannot control the appearance of the axes of a regression plot. I would like to see an extended dialog-box, similar to that of the PLOT command, introduced for *Statistical Graphs*.
- Release 11 allows you to use *text variables* as factors in many statistical analyses - from contingency tables to analysis of variance/ designed experiments to quality control. In all these procedures the new text variable replaces the alpha variable type in earlier releases.
- One of the major developments of Minitab in Release 11 is the revision of Statistical Process Control (SPC) facilities and the introduction of *Reliability/Survival Analysis* tools. In some promotional literature Minitab is now described as "Quality control" software! Customizable tests for special causes, lognormal and exponential probability plots, Box-Cox transformations and Gage R & R methods have been added to SPC. Also available are the Distribution ID plot, Overview plot, various Probability plots, and Hazard and Survival plots in the new reliability and reproducibility analyses. Although I am not very familiar with this area of statistics, a 'quality control' colleague of mine informs me that he would like to see extra facilities for storing the output (constants and columns) in the worksheet. He also mentions that the macros (existing and custom written) are considerably slow when executed! The hope is that the many facilities implemented via macros will be built into the package more fully at a later stage.
- *Design of Experiments* - Minitab Inc. claims that they have really simplified the DOE interface to introduce the concept of the design as a data object. It's now easier to design an experiment, fit a model, and plot the response. Release 11 also introduces *High-resolution cube and effects plots* and facilitates the analysis of Taguchi designs. One major drawback with DOE is that it creates an entirely new worksheet, after allowing the user to save the existing worksheet. I would prefer Minitab to allow the DOE worksheet to be 'merged' with the existing one!

Although it is possible to come up with an extensive "wish list", I personally would like to see further developments of facilities in the areas of Multivariate analysis, ANOVA and Non-linear regression. These include, Canonical and Tree-based Discriminant Analyses, Structural Equations and Path Analysis, particular treatment

comparisons using Linear Contrasts (in ANOVA), multiple comparison procedures in ‘Balanced ANOVA’, Least-squares means and comparisons for GLMs, and a significant improvement to Minitab’s curve fitting abilities with Non-linear regression modeling. I do believe that the One-way and Two-way ANOVA commands should be phased out with the above improvements to ‘Balanced ANOVA’ and ‘GLM’.

2.3. Graphical Capabilities

Graphics handling is one of Minitab’s major strengths. Release 11 adds further enhancement by adding in-built association with graphics for many statistical commands (see *Statistical Graphics* above). Additional new features include:

- *Improved graph brushing* - provides the ability to subset, analyze and explore data through brushing. The Data window displays a symbol beside the row number(s) containing the brushed data points, and you can create an indicator column based on brushed points, and specify commands to automatically execute when brushed points change. Although this is a significant improvement, brushing dynamically linked windows/graphs would be an extreme reward for many data analysts. I hope this will be the next revision to ‘brushing’ in Minitab!
- *New probability plot* - generates probability plots using normal, Weibull, log-normal, or exponential distributions. It displays a table of percentile estimates and produces a confidence band for the estimated probabilities.
- Graphic display defaults have been adjusted in a few graphics dialog boxes and commands to implement users’ requests. The default symbol plotted is now a solid circle, and the frame around a Graph window is no longer automatically drawn.

It should be noted that there are a few anomalies between some of the graphing commands. For example, creating histograms for grouped variables (using dialog boxes) works fine with one character group variable, but fails with two or more character group variables. Creation of histograms works satisfactorily with numerical group variables. The PLOT command, on the other hand, works well with character group variables!

Recommendations for further enhancements of the graphics capabilities of Minitab include: 3D plots with interactive rotation, QQ-plots for comparing distributions of two variables/columns and the ability to create/save user defined quantiles, incorporation of mathematical/statistical equations into graphs, either by implementation of such a facility in Minitab, or by allowing a general ‘pasting’ (of equations etc.) into the graph window (when ‘Editing’ graph).

3. Conclusions

Broadly speaking, Release 11 provides a substantial enhancement to Minitab. However, it must be said that there are a number of analyses which are still difficult to do in Minitab. This, of course, is common to many statistical packages! There are a number of improvements that could be suggested, but I have listed only three general ideas below:

- *Storage of statistics* - this has been a major problem with earlier versions of Minitab. Release 10.5Xtra alleviated this to an extent by introducing a Session command “STATS” with several useful sub-commands. I am, however, disappointed to see that this facility has not been improved in Release 11, in particular, it has not even been implemented into the pull-down menu system with easy to use dialog boxes! Another command that requires similar treatment is the ever popular TABLE command. Facilities should be available to not only create “proper” (graphical) two and three dimensional tables with numerous choice of ‘statistics’, but also to save the output/statistics created into the existing worksheet.
- When generating a new release of Minitab, particular attention should be given to making all previous ‘macro only’ commands (e.g. proportion test macro of Release 11) ‘proper’ menu/session commands. Talking of the ‘macro language’, it would be ideal to be able to define macro variables of more than one dimension! - e.g. variables of type $X(I, J)$, $X(I, J, K)$ etc. are not allowed at present. Introduction of such high dimensional variables will greatly reduce not only the efforts of writing a macro but also the execution time.
- *Manuals, Documentation and Help facilities* - Manuals are growing in size (!). Hence they receive the usual “their manual are pretty awful. It’s such a pity, since in most ways the software is so good” on the one hand, and on the other “Minitab provides the most comprehensive documentation and the manuals are easy to follow”. Personally, I hardly use the manuals as the built-in help facilities are very good!

Finally, Minitab has come a long way over the years in terms of ease of use, power and robustness. If you work with *data* and wish to *explore* it statistically, then Minitab is a comprehensive choice.

Acknowledgments

I wish to thank Dr. K Govindaraju of Department of Statistics at Massey University, New Zealand for his fruitful comments.