The Exstatix statistical analysis software package by K. C. Killion for use with Macintosh computers is evaluated. In evaluating the package, the framework developed by C. J. Ansorge et al. (1986) was used. This framework encompasses features such as transportability of files, compatibility of files with other Macintosh software, and ability to merge results with word processing and desk top publishing applications. The review is from the perspective of the educational researcher who wishes to use the package for uncomplicated, quick-and-dirty analysis of relatively small univariate data sets and export the output to reports. Specific features covered in the evaluation include documentation, ease of use, running programs, program output, accuracy, and other features. Strengths of Exstatix are discussed in the areas of statistical decision making, assumption testing, and regression analysis. Weaknesses of the packages include problems with importing and exporting functions, report generation, analysis of variance applications, and add-ons. Sample output of tables and graphs are provided. (TJF)
Exstatix™ Version 1.9.4
by Kevin C. Killian

Select Micro Systems, Inc.
40 Triangle Center - Suite 211
Yorktown Heights, NY 10598 (914) 245-4670

Copyright 1989, Stone House Systems. All Rights Reserved.

Barbara G. Ferrell, Ph.D.
University of Texas Medical Branch

Paper Presented at AERA, Boston, MA
April 17, 1990
Exstatix™: Expandable
Statistical Analysis System for the Macintosh™

A Software Review

Barbara G. Ferrell, Ph.D.
The University of Texas Medical Branch at Galveston

Statistical software packages for the Macintosh™ have proliferated over the last three years. These packages have been described as a “bumper crop” (Levine, 1990) and a “powerful force in statistics” (Seiter, 1989) in the Macintosh-related magazines. The focus of these reviews, however, is usually business application, and the articles are written by programmers or other computer types and not statisticians. The viewpoint of the researcher who will utilize the program in data analysis is always missing.

Because statistical software does not find a ready market in the general population of Mac users, most computer stores carry them only if they have a business focus. Potential users have only the product reviews on which to rely. The criteria used to evaluate a statistics package for a researcher is different from that for business. The purpose of this paper is to review Exstatix™, produced by Select Micro Systems, Inc., from the perspective of an educational researcher.

Evaluation Paradigm

Other statistical software packages for the Macintosh have been evaluated by this reviewer at previous sessions of AEHA (Ferrell, 1989; Ferrell, 1988). Exstatix™ is similar to those previously reviewed in that it might be viewed as a “personal” package; software that is inexpensive enough than an individual would be willing to buy it from his/her own pocket or recommend it to a student to use to analyze thesis data. Exstatix™ is the newest of the software in this category.

In evaluating Exstatix™, the framework developed by Ansorge, Wise and Plake (1986) was used, but other factors (Hamer, 1981) such as transportability of files, compatibility of files with other Mac software, ability to merge results with word processing and desk top publishing applications, etc., were considered. The review is from the perspective of the educational researcher who wishes to use the package for uncomplicated, quick-and-dirty analysis of relatively small univariate data sets and export the output into reports.

General Information

Exstatix™
Version 1.0.4

Vendor:
Select Micro Systems, Inc.
40 Triangle Center, Suite 211
Yorktown Heights, NY 10598
(914) 245-4670
Price: 3349 Suggested Retail
Educational Discounts and "Student Pack" Available

System Requirements: Macintosh 512KE, Plus, SE, Color supported on
Macintosh II. System 4.1 or later.
Supports Imagewriter™, Laserwriter™
and other printers.

Statistics: Descriptive, Crosstabs, Autocorrelation, Correlation,
Regression, t-test, N-way ANOVA, Scheffe, Non-
parametrics, Time Series, Transformations, Box Plots,
Scatter Plots, Bar Chart, Pie Chart, 3-D Scatter & Bar
Charts.

Software Usability

Documentation

Exstatix™ comes with a manual that assumes that the user knows something about the
Macintosh™. Mac novices are referred to the computer manual and no teaching of Mac's
standard features or "mouse ability" are included.

The seasoned statistician is directed to an overview of features designed for "those
experienced in the use of Statistics AND the Macintosh" (Wygonik, Killion & Scott,
1988). The novice statistician is urged to use the extensive tutorial which comprises
the major portion of the manual. The tutorial is more than the usual how-to-do-it. It is
in some sense a mini-applied statistics course, providing the user with guidance as to
when to apply the tests. It is divided into sections dealing with analyses for one variable,
two variables, and three or more variables.

The manual is clearly written and complete. Appendices include sample problems, some
sections on statistical decision making and formulae for the analyses.

Ease of Use

Exstatix™ is menu driven and easy to use. It is a single disk program, so no juggling is
necessary if the user has only one drive, except when data are to be stored on a separate
disk. The program has taken the features of the Macintosh™ into account in its
development and did not just clone PC statistical package features. For example,
variables can be viewed as icons on the screen and the user clicks on the icons to select
variables for analysis.

No on-line help screens are available, but error diagnostics are good. Warnings are
issued when the assumptions of a technique are violated, and they are written in non-
statistical language. For example, to alert the researcher to small expected cell size in a
Chi Square analysis, the following warning is issued:

Warning!
Of 4 cells,
2 have expected values of less than five
Additionally, analyses based on improper use of the data will not be performed. When asked for a correlation matrix using all of the variables, alpha variables were excluded from the analysis.

Running Programs

To test the running of Exstatix™, the data set from previous evaluations of statistical software was used. It was in the attempt to import the data set from another statistics package that a limitation of Exstatix™ that has been cited by other reviewers (Mac World, May, 1989) was demonstrated.

While the Exstatix™ literature states that Exstatix™ files and ASCII text files can be imported or exported, the manual suggests that data files from other Mac applications can be imported if they have been saved in text format. Data files from other statistical programs, the most likely candidates for import, are not usually in text file. Importing was accomplished, however, by copying the file to the clipboard and pasting it into Exstatix™. Variable labels were lost during the process and had to be retyped.

In terms of data analysis, Exstatix™ has a strength in its regression analysis subroutine. Not only does the program permit standard multiple regression analysis, but stepwise, both backward and forward, and subset analyses are available. The linear regression program provides for crossvalidation and a variety of optional tests and output including the regression equation, and nonlinear (log, exponential) and polynomial analyses are available. The regression analysis subroutine in Exstatix™ is the most comprehensive of statistics packages in its price range. (See sample of output in the Appendix.)

Just as the regression subroutine is a strength of Exstatix™, the ANOVA subroutine is a weakness. Other reviews have referred to its ANOVA procedures as "minimal and hard to specify" (Levine, 1990). In attempting to run the ANOVA procedures, real difficulty was encountered in the way the algorithm is set up to handle the data. Two major problems make this subroutine of very little use in most data sets utilized in educational research. First, Exstatix™ ANOVA requires that the data be set up with each level of a factor represented by a column for at least one factor in the data set. If this method is followed, specifying a one-way ANOVA will enable the researcher to obtain the summary table for either an equal or an unequal cell size design. Specifying a two-way ANOVA will give the researcher a repeated measures analysis, because the rows are considered one independent variable while the columns are the second. With the N-way ANOVA algorithm, equal cell sizes are required and the manual specifies that the researcher must set up the data "with each case representing a complete additional experiment" (Wygonik, Killion, & Scott, 1988, p. 107) and the design must be completely balanced. For the most part, this setup makes the ANOVA subroutine entirely unusable to the researcher who rarely has equal group sizes. In addition, using the N-way routine would require making a separate data set for that analysis only, an awkward method that is not necessary with other statistics packages of the same type as Exstatix™.

A good feature of Exstatix™ from the viewpoint of the statistics teacher or the statistical consultant is that it is the only program that encorporates research design and statistical decision making and takes the assumptions of the tests into account. Variables can be designated as independent or dependent, making the user at least think about design. While the usual p values also are printed, the significance level selected remains set until changed by the researcher. Alpha level is preset by the user prior to running the test and statistical decisions based on that level are printed. These two features plus the warnings which are issued when problems such as multicollinearity or small expected
cell size occur make this the only program which pays even lip service to the proper use of statistical findings.

Program Output

Just as the import of data to Exstatix™ was problematic, so was the export of tables produced. When pasted into both desk top publishing and word processing programs, tabs were lost and the font was changed to the default New York font 12 point. Tables, however, are fully editable, so that they need not be pasted to a paint program for adding titles or other information before placing in documents. However, if tables need to be edited once placed into other documents, it is questionable as to whether or not this represents a savings in time. The exported tables were definitely not of a quality which the reviewer would consider “document ready.” Examples of program output are included in the appendix.

A variety of graphics displays are available on Exstatix™. Scatter and box plots may be displayed, and both two and three dimensional bar and scatter plots are possible. The 3-D options may be used to rotate the scatterplot, a simple version of some of the exploratory data analysis programs available for the Mac. Quality of output is adequate. Examples of graphics displays are in the appendix.

Accuracy

Exstatix™ uses Standard Apple Numeric Environment (SANE) arithmetic and provisional means algorithms for accuracy. No information regarding the number of significant digits computed was provided in the manual. Exstatix™ computes a variety of statistics during idle time, thus speeding up subsequent analyses. It can handle up to 100 variables with as many as 10^10 cases per variable.

Other Features

Exstatix™ has been marketed as “the expandable statistical analysis system” based on its capability to add new functions and commands in the form of “custom packets.” Two types of packets may be added, function packets which enable the user to add to the Functions menu, and command packets which place new commands on the menu. This feature would enable the user to add new procedures or to customize those already available for a special purpose. The ability to expand Exstatix™, however, requires knowledge of Pascal or C. The software developers do not intend to add packets, but instead will rely on outside sources for these. While it is possible that good packets may be developed to go along with the basic program, it seems that those with advanced knowledge of Pascal or C probably would not select this program in the first place, but would work with one of the higher end statistical packages which are available.

Summary

Based on the review of Exstatix™ the following is noted:

**Strengths**  
- statistical decision making  
- assumption testing  
- regression

**Weaknesses**  
- importing/exporting  
- reports  
- ANOVA  
- addons
RATING

DOCUMENTATION

EASE OF USE

RUNNING PROGRAMS

PROGRAM OUTPUT

OTHER

OVERALL RATING

Rating based on a maximum of $\Sigma \Sigma \Sigma \Sigma \Sigma$.

References


Appendix
Sample Output

Descriptive*

* Tables have been edited for ease of interpretation. For actual output see display.

Science Project Sample Output from Exstatix™

Summary statistics for all numeric variables:

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>St. Deviation</th>
<th>St. Error</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case Number</td>
<td>18.50</td>
<td>10.54</td>
<td>1.76</td>
<td>18.50</td>
<td>35.00</td>
</tr>
<tr>
<td>Boating Experience</td>
<td>21.86</td>
<td>11.17</td>
<td>1.86</td>
<td>18.50</td>
<td>44.00</td>
</tr>
<tr>
<td>Seasickness Experience</td>
<td>22.50</td>
<td>22.88</td>
<td>3.81</td>
<td>5.50</td>
<td>113.00</td>
</tr>
<tr>
<td>GEFT</td>
<td>13.67</td>
<td>5.43</td>
<td>0.91</td>
<td>16.00</td>
<td>17.00</td>
</tr>
<tr>
<td>Age In Years</td>
<td>44.06</td>
<td>13.60</td>
<td>2.27</td>
<td>45.00</td>
<td>50.00</td>
</tr>
<tr>
<td>Index</td>
<td>1.79</td>
<td>3.84</td>
<td>0.64</td>
<td>0.71</td>
<td>22.60</td>
</tr>
</tbody>
</table>

Basic statistics for GEFT:

Active cases: 36
Missing values: 0
Mean: 13.6667
Sum: 492.0000
Sum of squares: 756.0000
Variance: 29.4857
Std deviation: 5.4301
Std error: 0.9050
Second moment: 28.6667
Third moment: -207.1296
Fourth moment: 2725.1481
Skewness: -1.3495
Kurtosis (norm=3): 3.3162
Median: 16.0000
Minimum: 1.0000
Maximum: 18.0000
Range: 17.0000
**Frequency Distributions and Crosstabs**

**Frequency table for Sex:**

Active cases included: 36

<table>
<thead>
<tr>
<th></th>
<th>Count</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>16</td>
<td>44.4</td>
</tr>
<tr>
<td>Male</td>
<td>20</td>
<td>55.6</td>
</tr>
</tbody>
</table>

**Crosstab of:**
- **Field Ind/Dep**
- **Sex**

Cases included: 36

<table>
<thead>
<tr>
<th>Field Ind/Dep</th>
<th>Dependent</th>
<th>Independent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female Count</td>
<td>2</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>% across</td>
<td>12.5</td>
<td>87.5</td>
<td>100.0</td>
</tr>
<tr>
<td>% down</td>
<td>28.6</td>
<td>48.3</td>
<td>44.4</td>
</tr>
<tr>
<td>% total</td>
<td>5.6</td>
<td>38.9</td>
<td>44.4</td>
</tr>
<tr>
<td>index</td>
<td>64</td>
<td>109</td>
<td>100</td>
</tr>
<tr>
<td>expected</td>
<td>3.1</td>
<td>12.9</td>
<td>16.0</td>
</tr>
<tr>
<td>deviation</td>
<td>-1.1</td>
<td>1.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Male Count</td>
<td>5</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>% across</td>
<td>25.0</td>
<td>75.0</td>
<td>100.0</td>
</tr>
<tr>
<td>% down</td>
<td>71.4</td>
<td>51.7</td>
<td>55.6</td>
</tr>
<tr>
<td>% total</td>
<td>13.9</td>
<td>41.7</td>
<td>55.6</td>
</tr>
<tr>
<td>index</td>
<td>12.9</td>
<td>93</td>
<td>100</td>
</tr>
<tr>
<td>expected</td>
<td>3.9</td>
<td>16.1</td>
<td>20.0</td>
</tr>
<tr>
<td>deviation</td>
<td>1.1</td>
<td>-1.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Total Count</td>
<td>7</td>
<td>29</td>
<td>36</td>
</tr>
<tr>
<td>% across</td>
<td>19.4</td>
<td>80.6</td>
<td>100.0</td>
</tr>
<tr>
<td>% down</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>% total</td>
<td>19.4</td>
<td>80.6</td>
<td>100.0</td>
</tr>
<tr>
<td>index</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>expected</td>
<td>7.0</td>
<td>79.0</td>
<td>36.0</td>
</tr>
<tr>
<td>deviation</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Degrees of freedom: 1

Chi-square: 0.887
Yates corrected chi-square: 0.268
Association (phi): 0.025
Cramer's V: 0.157
Contingency coefficient: 0.155

Warning! Of 4 cells, 2 have expected values of less than five
## Correlation and Regression

### Table of correlations:

<table>
<thead>
<tr>
<th></th>
<th>Case Number</th>
<th>Boating Experience</th>
<th>Seasickness Experience</th>
<th>GEFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case Number</td>
<td>1.000</td>
<td>-0.455Δ</td>
<td>0.192</td>
<td>-0.178</td>
</tr>
<tr>
<td>Boating Experience</td>
<td>-0.455Δ</td>
<td>1.000</td>
<td>-0.178</td>
<td>0.195</td>
</tr>
<tr>
<td>Seasickness Experience</td>
<td>0.192</td>
<td>-0.178</td>
<td>1.000</td>
<td>-0.355Δ</td>
</tr>
<tr>
<td>GEFT</td>
<td>-0.178</td>
<td>0.195</td>
<td>-0.355Δ</td>
<td>1.000</td>
</tr>
<tr>
<td>Age In Years</td>
<td>0.221</td>
<td>-0.010</td>
<td>0.149</td>
<td>-0.362Δ</td>
</tr>
<tr>
<td>Index</td>
<td>0.304</td>
<td>-0.398Δ</td>
<td>0.846Δ</td>
<td>-0.305</td>
</tr>
</tbody>
</table>

Δ denotes correlations significantly different from zero, at the 95% level.

### Table of covariances:

<table>
<thead>
<tr>
<th></th>
<th>Case Number</th>
<th>Boating Experience</th>
<th>Seasickness Experience</th>
<th>GEFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case Number</td>
<td>111.000</td>
<td>-53.586</td>
<td>46.229</td>
<td>-10.171</td>
</tr>
<tr>
<td>Boating Experience</td>
<td>-53.586</td>
<td>124.866</td>
<td>-45.500</td>
<td>11.810</td>
</tr>
<tr>
<td>Seasickness Experience</td>
<td>46.229</td>
<td>-45.500</td>
<td>523.571</td>
<td>-44.057</td>
</tr>
<tr>
<td>GEFT</td>
<td>-10.171</td>
<td>11.810</td>
<td>-44.057</td>
<td>29.486</td>
</tr>
<tr>
<td>Age In Years</td>
<td>31.600</td>
<td>-1.478</td>
<td>46.400</td>
<td>-26.724</td>
</tr>
<tr>
<td>Index</td>
<td>12.282</td>
<td>-17.050</td>
<td>74.260</td>
<td>-6.357</td>
</tr>
</tbody>
</table>

### Table of partial correlations:

(Each entry shows correlation of variables of that row and column, after partialling out the effects of other variables in the table.)

<table>
<thead>
<tr>
<th></th>
<th>Case Number</th>
<th>Boating Experience</th>
<th>Seasickness Experience</th>
<th>GEFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case Number</td>
<td>1.000</td>
<td>-0.589</td>
<td>0.040</td>
<td>0.013</td>
</tr>
<tr>
<td>Boating Experience</td>
<td>-0.589</td>
<td>1.000</td>
<td>0.377</td>
<td>0.219</td>
</tr>
<tr>
<td>Seasickness Experience</td>
<td>0.040</td>
<td>0.377</td>
<td>1.000</td>
<td>-0.316</td>
</tr>
<tr>
<td>GEFT</td>
<td>0.013</td>
<td>0.256</td>
<td>-0.295</td>
<td>1.000</td>
</tr>
<tr>
<td>Age In Years</td>
<td>0.203</td>
<td>0.296</td>
<td>-0.295</td>
<td>-0.375</td>
</tr>
<tr>
<td>Index</td>
<td>0.009</td>
<td>-0.471</td>
<td>0.861</td>
<td>0.192</td>
</tr>
</tbody>
</table>
Compare ranks of:

Age In Years
GEFT

Test used: Spearman Rank Correlation
Degrees of freedom: 34
Spearman R: -0.4531
T-statistic: -2.9635

Test used: Kendall Rank Correlation
S: -203
Kendall tau: -0.3456

Regression of: Seasickness Experience
Independent variables:
   Boating Experience
   GEFT
   Sex
Numeric equivalents used for alpha items:
   1 = Female
   2 = Male

<table>
<thead>
<tr>
<th>Case Number</th>
<th>Age In Years</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.203</td>
<td>0.009</td>
</tr>
</tbody>
</table>

| Boating Experience | 0.296 | -0.471 |
| GEFT               | -0.295 | 0.861 |
| Age In Years       | 1.000 | 0.349 |
| Index              | 0.349 | 1.000 |

Multiple-R | R-square
Unadjusted: | 0.9021 | 0.8138
Adjusted:   | 0.8847 | 0.7828

Sample size = 36
Standard Error of Estimate = 10.6647
F = 26.224
p = 0.0000

Is the F-value statistically significant?
Required critical value of F for 95% level of significance is 2.679.
The achieved F-value IS significant at the 95% level.
### Seasickness Experience Regression Analysis

<table>
<thead>
<tr>
<th></th>
<th>Regression Coefficient</th>
<th>Standard Error of Coefficient</th>
<th>T Value</th>
<th>Beta Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boating Experience</td>
<td>0.493270</td>
<td>0.179057</td>
<td>2.7548</td>
<td>0.240890</td>
</tr>
<tr>
<td>GEFT</td>
<td>-0.744881</td>
<td>0.368961</td>
<td>-2.0189</td>
<td>-0.176769</td>
</tr>
<tr>
<td>Sex</td>
<td>-8.790018</td>
<td>3.717654</td>
<td>-2.3644</td>
<td>-0.193594</td>
</tr>
<tr>
<td>Age In Years</td>
<td>-0.186420</td>
<td>0.149927</td>
<td>-1.2434</td>
<td>-0.110803</td>
</tr>
<tr>
<td>Index</td>
<td>5.504803</td>
<td>0.541939</td>
<td>10.15/6</td>
<td>0.923356</td>
</tr>
<tr>
<td>Intercept</td>
<td>33.943135</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Seasickness Experience = 33.94313456 + 0.49326966 * Boating Experience - 0.74488146 * GEFT - 8.79001776 * Sex - 0.18641991 * Age In Years + 5.5048033 * Index

Tests on residuals:
- Available cases: 36
- Cases with error > 0: 16
- Cases with error < 0: 20

- Standard error of estimate (standard deviation): 10.6647
- Mean absolute error: 7.60096893
- Greatest error: 23.32015644
- Durbin-Watson: 2.294
- Von Neumann ratio: 2.360

### ANOVA of Current Regression

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>Degrees of Freedom</th>
<th>Mean Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>14912.92451</td>
<td>5</td>
<td>2982.58490</td>
</tr>
<tr>
<td>Residuals</td>
<td>3412.07549</td>
<td>30</td>
<td>113.73585</td>
</tr>
<tr>
<td>Total</td>
<td>18325.00000</td>
<td>35</td>
<td>523.57143</td>
</tr>
</tbody>
</table>

### T-Test

(Independent values; variances assumed equal)
- Degrees of freedom: 62
- T-statistic: 4.8488

Is the DIFFERENCE statistically significant?
- Required t value for 95% level of significance is 1.999
- The difference IS significant at the 95% level.

Is the DIRECTION of the difference statistically significant?
- Required t value for 95% level of significance is 1.670
- The direction of the difference IS significant at the 95% level.
Esxtatix™ Software Review

**Scatterplot**

GEFT

Seasickness Experience

R square = 0.146

**3-D Bar Chart**

GEFT

Sex