# COMMENTS ON MICRO COMPUTER SOFTWARE

*Y. E. PITTELKOW*

The diversity of possible approaches to the analysis of longitudinal data is described in HAND (1991). This diversity is also present in the types of analyses available in computer packages for analyzing longitudinal data. The purpose of this note is to provide some indication of the range of software that is available on micro computers and which may be used for analyzing longitudinal data. Attention is restricted to software packages, thus excluding libraries of subroutines.

Software for the analysis of longitudinal data may be broadly categorized into three main types as follows:

1. Special purpose software that is tailored for longitudinal data. This type of software is sometimes restrictive in the range of analyses that it performs, but it is often efficient, since it can take account of special structures and restrictions;
2. General model fitting software, where analyses suitable for longitudinal data are performed as a special case of a more general model (eg. MANOVA in SPSS, MGLH in SYSTAT);
3. General purpose software that has no specific facilities for handling longitudinal data, but where it is possible to program a series of steps using available functions and facilities supplied with the software to perform suitable analyses. These steps (sometimes referred to as macros or procedures) may be stored and used repeatedly (eg. GAUSS, MATLAB, S, S-SPLUS, SAS, and X-LISP STAT).

Software systems, or packages, may contain implementations of the three types of software within the single system. This is common amongst the 'bigger' systems such as BMDP, GENSTAT, PSTAT, SAS, and SPSS. Some software that may be classified as either of the first two types, have matrix manipulation and other facilities, which may be used to fit 'non-standard' models, or extend and/or modify supplied analyses.

Attention is focused on software packages which contain software that may be classified as either of the first two types described above. The choice of software to include was arbitrarily based on the availability of manuals and software to the author. It is in no way intended to be a complete listing of software or facilities. CROWDER and HAND (1990) describe some of the software available on mainframe computers. Additional information may be obtained in the books on statistical computing by KENNEDY and GENTLE (1980) and MAINDONALD (1984), plus reviews of statistical packages. Other programs of which the author is aware, that are commonly used for analyzing longitudinal data, include BMDP, GENSTAT, LISREL (general analysis of covariance structure), MASS, and P-STAT. In addition to commercially available software, such as described in this note, there are special-purpose programs, often written by non-commercial organizations and individuals.

The manuals were used to obtain information about the general approach to the analysis of longitudinal data implemented in the package. Details concerning model designs, the handling of incomplete data and the availability of tests for evaluating model assumptions were also collected, and are described for the following packages: EQS, LIMDEP, RATS, SAS, SPSS, STATVIEW, Super ANOVA and SYSTAT.

---

**Software: EQS**
**Structural Equations Program**

BMDP Statistical Software, Inc
1440 Sepuveda Boulevard
Suite 316
Los Angeles, CA 90025
213.479.7799 Tel
213.312.0161 Fax

Mainframes
PC (ver 3)

General approach Estimation of linear structural equation models by minimization of a general weighted function. Includes least-squares, generalized least-squares, maximum likelihood, and re-

weighted least-squares methods. The incorporation of correlation and covariance structures is an integral part of the models. The overall model subsumes a variety of covariance and mean structure models, including multiple regression, path analysis, simultaneous equations, and a wide range of latent variable and factor models. Incorporating models with correlations across time ('waves') is a natural application of structural equation modelling.

Incomplete data Missing data models (missing at random) may be estimated when there are a few patterns of missing data, by separating the data into groups corresponding to the patterns, and estimating the parameters of the models, while restricting the associated parameters in the different groups to be equal.

Tests Goodness-of-fit tests (between the sample and model covariance structures), Wald tests and Lagrange-Multiplier test are available.

Reference

Bentler, P.M. (1989). *EQS Structural Equations Program Manual.* BMDP Statistical Software, Inc . L.A.

---

**Software: LIMDEP**
**Limited Dependent variable Models**

Econometric Software, Inc
43 Maple Avenue
Belport, NY 11713-2010
516.286.7049 Tel
16.286.6848 Fax

Hardware
Mainframes
PC (vers 5)

General approach LIMDEP is a general econometric package, estimating a wide variety of econometric models with an emphasis on limited dependent variable models. Estimation methods include generalized least squares, two stage least squares, and general likelihood maximization (amongst others). Constraints on estimates, weighting, and the incorporation of autocorrelation structures are a standard feature of the estimation procedures. Some specific models for repeated measures (panel data or

combined cross sectional and time series data) are available and are listed below. Facilities for creating procedures are available.

Incomplete data Unbalanced designs are possible. All data do not need to be present to be included in the analysis.

Notes on designs The following features are available for analyzing panel data (the descriptions from the manual are used):

- One and two way fixed and random effects models.
- AR(1) with one way fixed and and random effects.
- Logit models for panel data
- Random coefficients models
- Cross-sectionally correlated regression

Tests General Lagrange multiplier tests available.

References

Green, W. H. (1990) *LIMDEP* . Econometric software, Inc. NY.

Green, W. H. (1991) *Econometric Analysis* . Macmillan Publishing Company. NY.

---

**Software: RATS**
**Regression Analysis Time Series**

Var Econometrics
1800 Sherman Avenue
Belport, NY 11713-2010
516.286.7049 Tel
516.286.6848 Fax

Hardware
Mainframes
PC (vers 3.01)
Mac

General approach Rats is a general econometric package, estimating a wide variety of econometric models with emphasis on time series data. Estimation methods include generalized least squares, two stage least squares, non linear least squares and simultaneous models (and of course other methods, including those suitable for time series). Constraints and weighting are available. The package consists of procedures (instructions) which may be combined to provide estimation procedures not directly available. For example there is an instruction for computing a

residual variance covariance matrix, another for linear regression, and so on.

Notes on designs For balance panel data, fixed and random effects models may be estimated using a series of instructions described in the manual.

Reference

Doan, T. A.(1988) User's Manual RATS version 3.0, Var Econometric, Inc. IL.

---

**Software: SAS**
**Statistical Analysis System**

SAS Institute Inc
SAS Circle Box 8000
Cary, NC 27512-8000

Hardware
Mainframes
(Limited Unix systems)
PC (vers 6.03)

**Procedures:** ANOVA, GLM

General approach General univariate and multivariate analysis of variance and covariance programs. The ANOVA program is designed to handle balanced data sets.

Incomplete data Unbalanced designs in GLM are handled by partitioning the sums of squares four ways. However all data must be present to be included in the analysis if a multivariate approach is taken to analyzing repeated measures. If there is only one dependent variable (ie univariate approach) in GLM then all cases with non missing information are processed.

Notes on designs

- Full factorials or user specified designs available.
- User defined error terms available.
- Random effects may be declared.
- Pre-packaged contrasts available, including polynomial contrasts, with and without equal spacing.
- User specified contrasts possible.
- Weighted sums of squares available.

Tests General multivariate (Hotelling-Lawley's, Pillai's trace, Roy's maximum root and Wilk's lambda) tests and univariate F tests available. The Huynh-Feldt test and Greenhouse-Geisser test are also available. Tests for the homogeneity of dispersion matrices and variances are available.

**Other procedures:**

The procedure NESTED performs random effects analysis of variance and covariance for balanced data from an experiment with a completely nested (hierarchical) structure. Variance components are estimated for unbalanced designs but significance tests are not printed. The procedure NLR fits nonlinear regression models by least squares (or weighted least squares). Alternative general iterative methods may be chosen. The procedure VARCOMP computes estimates of the variance component in a general linear model, especially designed to handle random effects. Maximum likelihood, Restricted Maximum likelihood and MIVQUEO are available.

Reference

SAS Institute Inc. (1988) *SAS/STAT User's Guide,* Release 6.03 Edition. Cary, NC: SAS Institute Inc.

---

## Software: SPSS
## Statistical Package for the Social Sciences

SPSS Inc
444 N Michigan Avenue
Chicago, Illinois 60611
312 329 3500

Hardware
Most Mainframes
PC (vers 4)
Mac (ver 4)

**Procedure:** MANOVA

General approach General univariate and multivariate analysis of variance and covariance program for factorial and nested designs.

Incomplete data Unbalanced designs are handled by partitioning the sums of squares in either of two ways. However all data must be present to be included in the analysis.

Notes on designs

- Full factorials or user specified designs available.
- User defined error terms available.
- Within-subject factors possible.
- Pre-packaged contrasts available, including polynomial contrasts, with and without equal spacing.
- User specified contrasts possible.
- Profile analysis available.

Tests General multivariate (Hotelling-Lawley's, Pillai's trace, Roy's maximum root and Wilk's lambda tests ) and univariate F tests available. The Huynh-Feldt test and Greenhouse-Geisser test are also available. Tests for the homogeneity of dispersion matrices and variances available (eg Box-M, Bartlett-Box's test).

**Other Procedures:** Repeated measures may also be analyzed in the SPSS procedure Reliability. Mainframe versions of SPSSX have limited 'Macro' facilities. Mainframe versions include a general non-linear estimation software routine, with and without constraints. It is possible to perform iteratively reweighted least-squares analysis with this procedure.

References

Norusis, M.J.(1988) *SPSS/PC + Advanced Statistics v2.0.* SPSS Inc

Burns, P. (1984) . *SPSS-6000 MANOVA update manual.* Chicago: Volgelback Computing Center.

## Software: Statview SE+ graphics

Abacus Concepts, Inc
1984 Bonita Avenue,
Berkeley, CA 94704
415.540.1949
415.540.0260 Fax

Hardware

Mac

**Procedure:** ANOVA

Notes on designs

- Repeated measures on up to 15 between subject factors and a single within subject factor. If there are two or more between subject models the design must be balanced. Only full interaction models are estimated.

Reference

Abacus Concepts (1987) *Statview II.* Abacus Concepts, Inc., Berkeley, CA .

---

## Software: SuperANOVA 1.1

Abacus Concepts, Inc
1984 Bonita Avenue,
Berkeley, CA 94704
415.540.1949
415.540.0260 Fax

Hardware

Mac

General approach General multivariate linear models, including regression, ANOVA, MANOVA (and covariance models). Repeated measures data may be analyzed by either a univariate or a multivariate approach. Split plot and nested design data may be analyzed. "Canned" models are available for many standard designs, but there is full flexibility to build additional models.

Incomplete data Unbalanced designs are handled by partitioning the sums of squares in four ways. If some of the repeated measures for some of the factors are missing, the multivariate approach discards the case, but the univariate approach uses all non-missing data.

Notes on designs

- Canned models include models for one, two and three within subject factors and none, one and two between subject factors.
- User specified models.
- Default and user specified error terms.
- Pre-packaged contrasts available, including orthogonal polynomials of equal and unequal spacing.
- User specified contrasts possible.
- The univariate approach may accommodate time dependent variables.

Tests General multivariate (Hotelling-Lawley's, Pillai's trace, Roy's maximum root and Wilk's lambda) tests and univariate F tests available.

- Greenhouse-Geisser epsilon and Huynh-Feldt epsilon tests.

Reference

Gagnon, J., Haycock, J.A., Roth, J. M., Plamondon, J., Caroll, M. and Feldman, D. S. Jr. (1990). *SuperAnova Accessible General Linear Modeling*. Abacus Concepts, Inc. CA.

---

**Software: SYSTAT**
**The System for Statistics**

SYSTAT, Inc
1800 Sherman Avenue
Evanston, IL 60201-3793
708.864.5670 Tel
708.492.3567 Fax

Hardware

PC (vers 5)
Mac(ver 5)

**Procedure: MGLH**

General approach General multivariate general linear hypothesis modelling; includes regression, ANOVA, MANOVA (and covariance models). Repeated measures may be analyzed by either

a univariate or multivariate approach. Parameters for split plot and nested design models, as well as cross-over designs may be estimated.

Incomplete data Unbalanced designs are handled by partitioning the sums of squares, using a 'regression approach'. All data must be present to be included in the analysis. The covariance or correlation may be input directly for user defined methods of handling missing data.

Notes on designs

- Full factorials or user specified designs available.
- User specified within and between subject design matrix available.
- User specified error term.
- Within-subject factors available.
- Pre-packaged contrasts available, including polynomial contrasts, with and without equal spacing.
- User specified contrasts possible.
- Profile analysis available.

Tests General multivariate (Hotelling-Lawley's, Pillai's trace, Roy's maximum root and Wilk's lambda) tests and univariate F tests available.

References

Wilkinson, L. (1975) REGM: A Multivariate general linear hypothesis program for least squares analysis of multivariate data. *Behaviour Research Methods and Instrumentation.*, 7, 485-486.

Wilkinson, L. (1980) REGM: A Multivariate general linear hypothesis program. *American Statistician.*, 34, 182.

Wilkinson, L. (1989) *SYSTAT: The System for Statistics. Evanston*, IL. SYSTAT, Inc.

---

## LIST OF ADDITIONAL ADDRESSES FOR SOFTWARE .

| | |
|---|---|
| BMDP | BMDP Statistical software, Inc. 1440 Sepulveda Boulevard, Los Angeles, California 90025. |
| GAUSS | APTECH Systems Inc, 26250-196th Place South East, Kent, Washington 98042, Tel 206.631.6679 , Fax 206.630.1220 |

GENSTAT Numerical Algorithms Group Ltd, Wilkinson House, Joran Hill Rd., Oxford, United Kingdom OX2 8DR, Tel+44.865.511245 , Fax +44.865.310139.

GLIM Numerical Algorithms Group Ltd. Address as above.

MASS Western Statistical Computing, (WESTAT) Assoc. Pty Ltd., Nedlands, Western Australia, 6009.

MATLAB The MathWorks, Inc. 20 North Main St., Suite 250 Sherborn, MA 01770. 617.653.1415.

P-STAT P-STAT, Inc. P.O. Box 285, Princeton, NJ 08540

S AT&T Software Sales, P.O. Box 25000, Greensboro, North Carolina 26420.

S-PLUS Statistical Sciences, Inc. 1700 Westlake Ave. N., Suite 500, Seattle, WA 98109.

X-LISP-STAT *see:*
Tierney, L. (1990) *Lisp-Stat : An Object-Oriented Environment for Statistical Computing and Dynamic Graphics.* John Wiley & Sons, New York.

## REFERENCES

CROWDER, M.J. and HAND, D.J. (1990) *Analysis of Repeated Measures.* Chapman & Hall. London.

HAND,D.J. (1991) How to analyze your Repeated Measures Data. *Ibid.*

KENNEDY, W.J.,Jr., and GENTLE, J.E. (1980) *Statistical Computing.* Marcel Dekker, N.Y.

MAINDONALD, J. H.(1984) *Statistical Computation.* John Wiley & Sons, N.Y.

Coombs Computing Unit
Australian National University
GPO Box 4, Canberra, ACT 2601