Doing ANOVA Calculations on the Computer

Alan M Sykes
European Business Management School
University of Wales Swansea
Singleton Park
Swansea, SA2 8PP, Wales, UK
A.M.Sykes@Swansea.ac.uk

Abstract

This paper presents an approach to providing educational software for students in Management Science and Statistics which aids the calculations and understanding required for the analysis of experimental design data.

Keywords ANOVA, MANOVA, experimental design, Latin square, nested designs, mean, sum of squares

1 Introduction

Norman Thomson's excellent contributions on the subject of Analysis of Variance (ANOVA) calculations in APL [1, 2] amply demonstrate the ability of APL, in the hands of an expert APL user, to provide terse code for the analysis of many different experimental designs. This paper looks at the same subject material, from an antipodal viewpoint —that of a student who is neither an expert statistician (yet!) nor a competent programmer. It assumes that computing software is needed for the following reasons:-

- calculators necessitate the use of numerically poor algorithms and their use is labour intensive as data cannot be stored;
- a package (e.g., Statgraphics) or a packaged program does not require the student to really understand what is going on statistically in order to provide the correct figures in the Analysis of Variance table;
- in an educational context it is desirable for computational methods to flow naturally and easily from an understanding of the mathematical formulation of Analysis of Variance found in statistics textbooks, and a firm grasp of concepts such as 'between groups variation' and 'within groups variation'.

Over the last few years, the approach to analysis of variance calculations outlined here has been gradually introduced into the teaching methodology at Swansea. Provided the user understands the basic ideas or notation, calculations are easy, even for complicated designs. However, if understanding is lacking, then the methodology is unlikely to deliver a meaningful analysis. Both of these attributes are important in an educational environment!

The theoretical underpinning of the approach is simple, and is described in detail in Heiberger's book [3] which also includes APL programs. However, the nitty-gritty of applying it in a classroom situation for students who know little APL is, I believe, new. Because we are dealing with users of APL rather than APL programmers, nifty one-liners are avoided. Instead a **small, carefully designed set of utility functions** is constructed to provide a tool-kit of sufficient flexibility to cover **all** ANOVA calculations. This is very much in the spirit of McIntyre's 'pedagogic sequence' [4]. The key function is *AVGE*, which will be described in the next section.

2 Introducing the function *AVGE*

Early on, a student of analysis of variance techniques should realise that ANOVA is all about looking at variation in a sample about different means. For example, consider a small data set of examination marks corresponding to two different groups of students. The data are placed in one vector called *EXAM*. The group structure (the first four values all pertain to students in one group, e.g., female students, and the last four from students in the second group e.g., male students) is specified by the variable *I*.

```
EXAM←53 89 53 95 36 42 32 3
I←1 1 1 1 2 2 2 2
```

The total corrected sum of squares (*TCSS*) measures the amount of variation in the whole data set about the overall average. It is obtained by calculating the differences between the data and the overall average, and summing their squares. Alternatively, we can measure the amount of

APL '95, San Antonio, Texas, USA
© 1995 ACM 0-89791-722-7/95/0006...$3.50

variation within the groups about the group averages (the group structure indicated by the different values of *I*). Calculate the averages for data corresponding to `I=1` and `I=2` respectively, subtract these average values from the original data, and then square and sum to derive the **Within Groups Sums of Squares** (*WGSS*).

All of these operations can be done easily in APL with the help of the function *AVGE*. First of all, *AVGE* with no left argument just returns an array of values equal to the overall average of the right argument:-

```
      AVGE EXAM
50.375 50.375 50.375 50.375 50.375
      50.375 50.375 50.375
```

We emphasize the fact that the function returns an array of the same shape as the data array it operates on as this is crucial for the second use of *AVGE*. Remember that *I* indicates a group structure—the first four values of *EXAM* correspond to `I=1` and so are in one group, and the last four, corresponding to `I=2`, are in another group:-

```
      I AVGE EXAM
72.5 72.5 72.5 72.5 28.25 28.25 28.25
      28.25
```

Consequently *I AVGE EXAM* returns a vector whose first four elements are the average of the first four elements of *EXAM*, and equivalently for the last four. Similarly, if *J* is a vector given by `J←1 2 1 2 1 2 1 2`

```
      J AVGE EXAM
43.5 57.25 43.5 57.25 43.5 57.25 43.5
      57.25
```

where the first, third, fifth and seventh values are all equal to `43.5`, and the second, fourth, sixth and eighth values all equal `57.25`, corresponding to the group structure indicated by *J*.

3 Application to one-way ANOVA

Recall that one-way ANOVA corresponds to breaking down data into deviations of group means from the overall mean, and deviations of the data from the group mean:-

$$(y_{ij}-\bar{y}_{..})=(y_{ij}-\bar{y}_{i.})+(\bar{y}_{i.}-\bar{y}_{..})$$

and, remarkably, if you square and add elements in this identity, you also get equality (thanks to Pythagoras' Theorem):-

$$\sum(y_{ij}-\bar{y}_{..})^2=\sum(y_{ij}-\bar{y}_{i.})^2+\sum(\bar{y}_{i.}-\bar{y}_{..})^2$$

TCSS = *WGSS* + *BGSS*

The names for these objects are the Total (Corrected) Sum of Squares, the Within Groups (or Residual) Sum of Squares, and the Between Groups (or Treatments) Sum of Squares.

Now we can use the small data set of examination results to illustrate how easy it is to calculate the sums of squares as defined above using the function *AVGE*. **APL variables are created with names deliberately chosen to sound like their mathematical counterparts, thereby helping to bridge the gap between textbook and computation.** Thus the data are assigned to a variable *Y* and the mean vector to a variable called *YBAR*.

```
      Y←EXAM
      YBAR←AVGE Y
      YBAR
50.375 50.375 50.375 50.375 50.375
      50.375 50.375 50.375
```

Subtract *YBAR* from *Y* to get the deviations from the overall mean:-

```
      Y-YBAR
2.625 38.625 2.625 44.625 ¯14.375 ¯8.375
      ¯18.375 ¯47.375
```

If these are squared and summed, they yield *TCSS*. This can be done using a simple, but useful function *SSQ* (so that `SSQ 1 2 3` gives the answer `14`).

```
      SSQ Y-YBAR
6355.9
```

To calculate the term *BGSS*, we need to calculate the equivalent of $\bar{y}_{i.}-\bar{y}_{..}$ and then square and sum:-

```
      YIBAR←I AVGE Y
      YIBAR
72.5 72.5 72.5 72.5 28.25 28.25 28.25
      28.25
      SSQ YIBAR-YBAR
3916.1
```

Similarly the Within Groups Sum of Squares *WGSS* can be calculated by using the data *Y* and the group mean data *YIBAR* as directed by its formula:-

```
      SSQ Y-YIBAR
2439.7
```

If the results are allocated to variables during the calculation, APL can be used to validate the results, by checking that `TCSS=WGSS+BGSS`:-

```
      YBAR←AVGE Y
      YIBAR←I AVGE Y
      TCSS←SSQ Y-YBAR
      BGSS←SSQ YIBAR-YBAR
      WGSS←SSQ Y-YIBAR

      TCSS
6355.9
      BGSS
3916.1
      WGSS
2439.7
```

```
      TCSS=BGSS+WGSS
1
```

indicating that the calculations are indeed correct.

To summarise and accentuate the one-to-one correspondence between the mathematical formulation and the APL calculations, we have :-

Overall mean $\bar{y}_{..}$

```
YBAR←AVGE Y
```

Group means $\bar{y}_{i.}$

```
YIBAR←I AVGE Y
```

Total Corrected SS $\sum(y_{ij}-\bar{y}_{..})^2$

```
TCSS←SSQ Y-YBAR
```

Within Groups SS $\sum(y_{ij}-\bar{y}_{i.})^2$

```
WGSS←SSQ Y-YIBAR
```

Between Groups SS $\sum(\bar{y}_{i.}-\bar{y}_{..})^2$

```
BGSS←SSQ YIBAR-YBAR
```

4 Completing the ANOVA calculations in APL

Mean squares and hence F ratios are easy to calculate using the variables derived above. The mean squares can be calculated by dividing by appropriate degrees of freedom:-

```
      WGSS÷6
406.63
      BGSS÷1
3916.1
```

and, if we want to calculate the F ratio without creating further variable names, type:-

```
      (BGSS÷1)÷(WGSS÷6)
9.6308
```

Alternatively, and probably better is:-

```
      BGMS←BGSS÷1
      WGMS←WGSS÷6
      F←BGMS÷WGMS
      F
9.6308
```

Then we can either look this up in *F* tables or calculate the p-value directly using:-

```
      1 6 FTAIL 9.631
0.02103
```

which is less than 0.05 and hence the hypothesis of equal group means across the *I* factor is rejected at the 5% (indeed at the 2.5%) level in favour of the hypothesis of unequal group means. To investigate the group means further, calculate confidence intervals for each group mean or a confidence interval for the difference in the two group means by using the 97.5% (2.5%) percentile of the *t* distribution with six degrees of freedom:-

```
      6 TQUANTILE .975
2.447
      SIGMA←SQRT WGMS
      YIBAR[1]+¯1 1×2.447×SIGMA÷SQRT 4
47.828 97.172
      YIBAR[5]+¯1 1×2.447×SIGMA÷SQRT 4
3.5782 52.922
```

Notice how APL helps you to do two calculations in one using the vector `¯1 1`. If a confidence interval for the difference of the two means is required, then a similar procedure is possible:-

```
      (YIBAR[1]-YIBAR[5])+¯1 1×2.447×
       SIGMA×SQRT (.25+.25)
9.3588 79.141
```

and of course this confidence interval does not include zero, as the hypothesis of equal means has been rejected at the 5% level.

5 Designs involving more than one factor

In Section 2 the variable *J* indicated a second group structure, allowing the analysis to evaluate whether the *J* factor affects the examination results. Because the joint values of *I* and *J* occur more than once (e.g., the first and third values both have *I*=1 and *J*=1), a within-replicates sum of squares can also be evaluated to estimate the inherent variation in the data. Further, an 'interaction sum of squares' representing the interaction between the *I* factor and the *J* factor can also be calculated. To do this, we need to be able to indicate the joint group structure defined by the action of both the variables *I* and *J*. The function *WITH* merely combines the two vectors into columns of a matrix, so that when used as a left argument to *AVGE*, the joint group structure is indicated.

```
      I
1 1 1 1 2 2 2 2
      J
1 2 1 2 1 2 1 2
```

The composite index is now constructed:-

```
      IJ←I WITH J
      IJ
1 1
1 2
1 1
1 2
2 1
2 2
2 1
2 2
```

Clearly, rows one and three are identical as are rows two and four, five and seven, six and eight. Hence using *AVGE* with *IJ* simply returns a vector with appropriate averages of each pair of values:-

```
      IJ AVGE Y
53 92 53 92 34 22.5 34 22.5
```

We are therefore able to calculate sums of squares relevant to the analysis of a two-factor design with replicates. First the sum of squares due to the *J* factor is easy, as it merely follows the calculation already done for *I*.

```
      BIGSS←SSQ YIBAR-YBAR
      YJBAR←J AVGE Y
      BJGSS←SSQ YJBAR-YBAR
      BJGSS
378.12
```

Secondly, the within-replicates residuals measuring the variation inherent in the observations other than that caused by either the *I* or *J* factors, can be calculated from the difference between the original *Y* values and the within replicates average:-

```
      RES←Y-IJ AVGE Y
      SSQ RES
786.5
```

But *TCSS-(BIGSS+BJGSS)* equals 2061.6 which is 1275.1 greater than the within-replicates residual sum of squares, 786.5. This difference is due to the interaction between the two factors. This interaction sum of squares can be calculated directly (with care*) by squaring and summing the quantities

$$y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$$

```
      INT←YIJBAR-YIBAR
      INT←INT-YJBAR
      INT←INT+YBAR
      SSQ INT
1275.1
```

As a result a full two-way ANOVA table with interaction and within replicates sums of squares has been produced very easily.

(*Note: I suggest that students do the calculation in stages to avoid the pitfalls associated with an unwary use of APL working from right to left.)

6 Nested designs

One of the complications which students have to cope with is the variety of different experimental designs. Suppose that in our very small database of examination results the *I* factor indicates two different schools, and the *J* factor indicates two different classes in each of the two schools. The *J* factor is then nested within the *I* factor and, *a priori*, it does not make sense to equate the *J*=1 observations when *I*=1 with the *J*=1 observations when *I*=2. The appropriate sums of squares to calculate are (a) the Between Schools Sum of Squares (easily done, using the *I* variable) and (b) the Between *J* levels within *I* levels or Between Classes within Schools Sum of Squares.

First construct the index that points to the individual classes. This could be a simple vector of the form 1 2 1 2 3 4 3 4, where classes in the second school receive different numbers to those in the first school. Alternatively, as in the two-way ANOVA model, construct an equivalent index from *I* and *J* as already defined:-

```
      IJ←I WITH J
```

Now calculate appropriate averages (the overall average, the school averages and the class averages):-

```
      YBAR←AVGE Y
      YIBAR←I AVGE Y
      YIJBAR←IJ AVGE Y
```

Then the appropriate sums of squares:-

```
      ⍝ Between Schools Sum of Squares
      SSQ YIBAR-YBAR
3916.125
      ⍝ Between Classes Within Schools
      SSQ YIJBAR-YIBAR
1653.25
      ⍝ Within Classes
      SSQ Y-YIJBAR
786.5
```

7 A Latin square design

The function *AVGE* has so far been applied to data in a vector. However, it has been carefully designed so that it may be used on matrix arrays. One could argue that data for a two-way analysis of variance design are best conceived as a matrix, so that the row and column structure indicate the two factors. This is a powerful argument but in order to cope with **all** ANOVA designs it is necessary to use indexing variables (such as *I* and *J* used above) as the next example illustrates. If the user wants to present the data as a two- (or more) dimensional array then this is catered for.

Consider a four by four Latin square design where the different treatment levels are indicated in the usual way by the letters *A,B,C,D* and are:-

```
A  B  C  D          22  31  32  41
D  A  B  C          43  25  36  37
B  C  D  A          32  33  49  19
C  D  A  B          35  46  26  31
```

Use *I* to indicate the row structure, *J* for the column structure, and *K* for the structure given by the letters *A,B,C,D* (note that each letter appears once and only once in each row and column—the essence of the Latin square design).

```
      Y←4 4⍴22 31 32 41 43 25 36 37 32 33
49 19 35 46 26 31
      Y
22 31 32 41
43 25 36 37
32 33 49 19
35 46 26 31
      I←4 4⍴⍳4
      J←4 4⍴4/⍳4
      K←4 4⍴1 2 3 4 4 1 2 3 2 3 4 1 3 4 1
2
```

(These APL statements are used here for the sake of clear exposition to the APL audience—in practice a student might enter all the required data using a spreadsheet or numeric APL editor.)

```
      K
1 2 3 4
4 1 2 3
2 3 4 1
3 4 1 2
```

Appropriate averages can now be calculated as before:-

```
      YBAR←AVGE Y
      YIBAR←I AVGE Y
      YJBAR←J AVGE Y
      YKBAR←K AVGE Y
```

Because of the nature of the Latin square design, appropriate sums of squares for all three factors can be calculated in a way which merely extends the one-way ANOVA methodology. It is helpful, however, if a running-residual calculation is kept so that the final calculation of the residual sum of squares can be checked. This is a major advantage of computer-based calculations over calculators (with which checking is tedious).

```
      TCSS←SSQ Y-YBAR
```

(first step in the residual calculation)

```
      RES←Y-YBAR
      BIGSS←SSQ YIBAR-YBAR
```

(now strip off the *I* effect)

```
      RES←RES-(YIBAR-YBAR)
      BJGSS←SSQ YJBAR-YBAR
```

(now strip off the *J* effect)

```
      RES←RES-(YJBAR-YBAR)
      BKGSS←SSQ YKBAR-YBAR
```

(finally strip off the *K* effect)

```
      RES←RES-(YKBAR-YBAR)
      TCSS
1051.75
      BIGSS
30.25
      BJGSS
32.25
      BKGSS
953.25
      SSQ RES
36
```

This should correspond to the difference between the Total Corrected Sum of Squares and the sum of all the Sum of Squares due to each factor:-

```
      TCSS-(BIGSS+BJGSS+BKGSS)
36
```

One advantage of using the matrix approach is that the residuals naturally inherit the structure of the original data and so can be easily displayed in row and column format. Any glaringly obvious row or column structure in the residuals may then be seen.

```
      RES
 1.75  0.5  ¯2.25  0
¯2.75  0.25 ¯0.25  2.75
 0.5  ¯1     2.5  ¯2
 0.5   0.25  0    ¯0.75
```

8 Calculating orthogonal contrasts

In this section we demonstrate how sums of squares due to orthogonal contrasts can be calculated. For example, in a one-way ANOVA design with *k* levels of treatment, the Between Treatments Sum of Squares (i.e., the Between Groups Sum of Squares) can be broken down into the sum of `k-1` orthogonal contrasts. To demonstrate this, add a further data point to the data used in Sections 2 and 3 and define a different group structure given by *I* again:-

```
      Y
53 89 53 95 36 42 32 3 50
      I
1 1 1 2 2 2 3 3 3
      YIBAR←I AVGE Y
      YBAR←AVGE Y
      SSQ YIBAR-YBAR
2258.666667
```

This is, therefore, the new Between Groups Sum of Squares. Now contrast the first and third groups with the second. This is done by indexing the vector `1 ¯2 1` using the group index variable *I*. (Although we have assumed group vectors *I*, *J* to be positive integers, the function *AVGE* does not require them to be. Here however, it is mandatory!)

```
      C1←1 ¯2 1[I]
      C1
1 1 1 ¯2 ¯2 ¯2 1 1 1
```

Similarly, we can contrast the first and third groups using the contrast *C2*:-

```
      C2←¯1 0 1[I]
      C2
¯1 ¯1 ¯1 0 0 0 1 1 1
```

The fact that these are orthogonal can be checked by multiplying and summing:-

```
      SUM C1×C2
0
```

To calculate the contrast, we multiply either the actual `Y` values or the `YIBAR` values by the contrast and sum:-

```
      C1×Y
53 89 53 ¯190 ¯72 ¯84 32 3 50
      SUM C1×Y
¯66
      SUM C1×YIBAR
¯66
```

Calculating the sum of squares due to the contrast (which will have one degree of freedom) is akin to regressing `Y` on `C1` and calculating the sums of squares due to the regression. This is best done by remembering the standard calculations for Linear Regression in the following form:-

$$S_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

$$S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2$$

from sum of squares due to regression of y on x is given by S_{xy}^2 / S_{xx}. In applying these formulae, the means need not be subtracted because contrasts consist of linear combinations of observations where the sum of the coefficients is zero. Hence the sum of squares S_{xy}^2 / S_{xx} for a contrast is:-

$$\frac{\left(\sum_{i=1}^{n} c_i y_i\right)^2}{\sum_{i=1}^{n} c_i^2}$$

which in APL, using obvious cover functions, becomes:-

```
      (SQ SUM C1×Y)÷SSQ C1
242
```

Thus the sum of squares due to the first contrast is equal to `242`. Similarly, for the second contrast:-

```
      (SQ SUM C2×Y)÷SSQ C2
2016.666667
```

Because the Between Groups Sum of Squares has only two degrees of freedom, these two sums of squares total to the Between Groups Sum of Squares, i.e., `2016.666667+242` equals `2258.666667`

9 Application to multivariate designs

With very slight adaptation the function `AVGE` can be converted into a multivariate equivalent `MAVGE`, which assumes that its right argument is a matrix whose columns correspond to different measurements on the same object. The use of the function `SSQ` is replaced by `SSQP` which calculates not only sums of squares for each of the columns of the right argument, but also all cross-products, i.e., the sum of squares and products matrix. Consider a multivariate equivalent of the example of Sections 2 and 3, consisting of examination results from eight individuals sitting the same three examinations. Suppose `I` indicates the sex of the examinees, and a Multivariate Analysis of Variance (MANOVA) is required for the `Y` data:-

```
      77  27   5
      74  33  64
      76 100  37
      25  99  73
      76  66   8
      64  89  28
      44  77  48
      24  28  36
```

A `MEAN` function operates, as one would expect on the columns to give column means:-

```
      MEAN Y
57.5 64.875 37.375
```

However, the function `MAVGE` always produces an array of the same shape as it uses:-

```
      MAVGE Y
57.5 64.875 37.375
57.5 64.875 37.375
57.5 64.875 37.375
57.5 64.875 37.375
57.5 64.875 37.375
57.5 64.875 37.375
57.5 64.875 37.375
57.5 64.875 37.375
      I
1 1 1 1 2 2 2 2
      I MAVGE Y
63 64.75 44.75
63 64.75 44.75
63 64.75 44.75
63 64.75 44.75
52 65    30
52 65    30
52 65    30
52 65    30
```

Hence the calculations of Section 3 can easily be repeated in the multivariate context. Each derived sum of squares in

Section 3 is now replaced by a Sum of Squares and Products matrix:-

```
      YBAR←MAVGE Y
      YIBAR←I MAVGE Y

      TCSSP←SSQP Y-YBAR
      BGSSP←SSQP YIBAR-YBAR
      WGSSP←SSQP  Y-YIBAR

      TCSSP
 3740   ¯474.5    ¯2058.5
 ¯474.5 6938.875   1500.375
¯2058.5 1500.375   4091.875

      BGSSP
242   ¯5.5   324.5
 ¯5.5  0.125  ¯7.375
324.5 ¯7.375 435.125

      WGSSP
 3498 ¯469     ¯2383
 ¯469 6938.75   1507.75
¯2383 1507.75   3656.75
```

Just as `TCSS=BGSS+WGSS`, so we have an equality between the sum of squares and products matrices:-

```
      TCSSP=WGSSP+BGSSP
1 1 1
1 1 1
1 1 1
```

The computations involving these matrices required to test the hypothesis of equality of mean vectors between the two groups (and other more complicated hypotheses) are outside the remit of this paper but present no problems provided that good eigenvalue-eigenvector routines are available.

10 Function listings

The most important functions are listed below. Other cover functions mentioned are easily constructed. More detailed listings of functions mentioned, such as `FTAIL`, `TQUANTILE` are available as are the Dyalog workspaces in which they were written.

```
      ∇ R←MEAN X
[1]     R←(+⌿X)÷⍬ρρX
      ∇

      ∇ R←SSQ X
[1]     R←+/,X×X
      ∇

      ∇ R←SSQP X
[1]     R←(⍉X)+.×X
      ∇

      ∇ R←{A}AVGE X;O;I;D;N
[1]     D←ρX ◇ →DOIF 0=⎕NC'A' ◇ A←(ρX)ρ1
[2]     X←,X ◇ →DOIF(×/ρA)=×/D ◇ A←,A
[3]     ⍝ just catenate A if a matrix
          of right dimension
[4]     →DOIF((ρX)≠1↑ρA)◇R←⍳0◇
          'Incompatible dimensions' ◇ →0
[5]     →DOIF 2=≡A ◇ A←,A
[6]     →DOIF 1=ρρA ◇ A←((ρA),1)ρA
[7]     A←A[O←⍋A;] ◇ X←X[O]
[8]     I←1,1↓∨/A≠¯1⊖A
[9]     N←I/⍳ρI ◇ N←(1↓N,1+ρI)-N
[10]    R←(+\X)[+\N] ◇ R←(R-0,¯1↓R)÷N
[11]    R←Dρ(N/R)[⍋O]
      ∇

      ∇ R←{A}MAVGE X;O;I;D;N
[1]     D←ρX ◇ →DOIF 0=⎕NC'A'◇A←(ρX)ρ1
[2]     →DOIF((1↑ρX)≠1↑ρA)◇R←⍳0 ◇
          'Incompatible dimensions' ◇ →0
[3]     →DOIF 1=ρρA ◇ A←((ρA),1)ρA
[4]     A←A[O←⍋A;] ◇ X←X[O;]
[5]     I←1,1↓∨/A≠¯1⊖A
[6]     N←I/⍳ρI ◇ N←(1↓N,1+ρI)-N
[7]     R←(+⍀X)[+\N;]◇R←(R-0,[1]¯1 0↓R)
          ÷⍉(⌽ρR)ρN
[8]     R←Dρ(N⌿R)[⍋O;]
      ∇
```

11 Final comments

This paper attempts to provide a unified framework for the teaching and learning of statistical calculations involved in experimental design for univariate or multivariate data. Its philosophy is to avoid at all costs the modern urge to package everything, but instead to make a realistic appraisal of students' programming abilities (whether in APL or not). As such the ideas expressed in the paper are part of a long-term strategy that is emerging from the use of APL in the teaching of Statistics at the University of Wales Swansea.

References

[1] Thomson, N. Understanding ANOVA the APL Way, *APL93 Conference Proceedings, APL Quote Quad*, Vol. 24, No. 1, August 1993.

[2] Thomson, N. *All the ANOVA Commands you will ever need*, APL in Business Workshop, Swansea 1994.

[3] Heiberger, R. M. *Computation for the Analysis of Designed Experiments*, Wiley and Sons, New York, 1989.

[4] McIntyre, D. B. Language as an Intellectual Tool: From Hieroglyphics to APL, *IBM Systems Journal*, Vol. 30, No. 4, 1991.