REGULAR ARTICLE

A SAS macro for generating letter displays of pairwise mean comparisons

Hans-Peter Piepho

Bioinformatics Unit, Institute of Crop Science, University of Hohenheim, Fruwirthstrasse 23,
70599 Stuttgart, Germany.
E-mail: piepho@uni-hohenheim.de

ABSTRACT

Data from many experiments are routinely subjected to analysis of variance, followed by a multiple comparison of means. A popular way to present the result of mean comparisons is by attaching super-scripted letters to the means, with a common letter on two means indicating that they are not significantly different. When the standard error of a difference is not constant, the algorithm traditionally used for generating such displays may fail to represent all significant differences. This paper reports on a so-called insert-and-absorb algorithm that is guaranteed to always truthfully represent all significant differences. This algorithm was implemented in a SAS macro %MULT. Its usage is illustrated using three examples.

Key Words: *multiple comparison; insert-and-absorb algorithm; lines display; letter display; compact letter display; mixed model; generalized linear mixed model; standard error of a difference; analysis of variance; %MULT macro.*

INTRODUCTION

Agricultural experiments are often analysed by linear model procedures, including a multiple comparison of treatment means. Results may be reported as lines or letter displays. Such displays are straightforward to generate when the standard error of a difference (SED) among treatment means is constant. When the SED is not constant, however, a traditional lines display is no longer guaranteed to correctly represent all significant differences (Piepho, 2000). The SAS procedures GLM and GLIMMIX (Version 9.3) provide a LINES option to the LSMEANS and SLICE statements that generates a lines display using the same algorithm that is commonly used when the SED is constant (Steel and Torrie, 1980). This method may require suppressing some significant differences in order to derive lines connecting means that are judged not significantly different (Piepho, 2000).

Piepho (2004) proposed an 'insert-and-absorb' algorithm that can generate a letter display to represent all significant differences even in cases where the ordinary lines display fails due to non-constant SED. The algorithm operates on the set of p-values for all pairwise comparisons. In comparison to the traditional lines display, the algorithm often generates 'broken lines' in order to represent all significant differences. The purpose of this paper is to illustrate the usage of the SAS macro %MULT, which implements the algorithm. The paper starts with a brief description of the macro and then illustrates its use with a single-factor experiment and two factorial experiments, one laid out according to a split-plot design and the other a trial involving repeated measures.

THE %MULT MACRO

The main idea of the insert-and-absorb algorithm is to start with a false letter display that attaches the same letter to all treatments and then to successively correct this display by inserting significant differences. This is illustrated in Table 1 for a small example of four treatments, where treatments 1 and 3 are significantly different. The start configuration attaches an 'a' to all treatments, then duplicates this column of letters into a column of b's, and finally inserts the significant difference by dropping an 'a' from treatment 1 and a 'b' from treatment 3. Further algorithmic details can be found in Piepho (2004).

Table 1. Illustration of the 'insert'-step of the insert-and-absorb algorithm (Piepho, 2004). Treatments 1 and 3 are significantly different.

Start configuration		Duplication		Insertion	
Treatment	Letters	Treatment	Letters	Treatment	Letters
1	a	1	a b	1	b
2	a	2	a b	2	a b
3	a	3	a b	3	a
4	a	4	a b	4	a b

The macro %MULT is available at https://www.uni-hohenheim.de/bioinformatik/beratung/toolsmacros/sasmacros/mult.sas. To use it, the macro must be made available either by loading it into a program editor window and then submitting the code, or by using the %INCLUDE statement. The macro requires the SAS/IML module and can process output on least squares means and differences from the MIXED, GLIMMIX, and GENMOD procedures as generated via ODS (table names 'lsmeans' and 'diffs'). It can handle adjusted p-values generated by these procedures or by post-processing, e.g., using the MULTTEST procedure. The macro allows up to three by-variables for factorial experiments, so it can analyse experiments with up to four treatment factors. In addition to the letter display, the macro also computes the average, minimum and maximum values of LSD and SED. It can process least squares means for one effect only. If least squares means are needed for several effects, the linear model procedure must be run several times, each time using only one LSMEANS statement with only one effect.

The macro has the following options:

trt=   Specifies the treatment factor for which means are to be compared. Generally, also with factorial experiments, you can only use one factor here.

by=   You can define up to three factors for slicing the mean comparisons.
by2=   For example, if you computed A*B means, you can use the specification
by3=   "trt=A, by=B" to compare A*B means by levels of B. But you can't use "trt=A*B".

alpha=   Specifies the type I error rate (default is alpha=0.05)

p=   Specifies the variable containing the p-values; default: p = probt, the p-value of pairwise t-tests as generated by the LSMEANS statement.

descending   =0 smallest mean will get the letter 'a', etc.
             =1 largest mean will get the letter 'a', etc. (default)

Several examples illustrating the use of the macro can be found under https://www.uni-hohenheim.de/bioinformatik/beratung/toolsmacros/sasmacros/mult_examples.sas.

A SINGLE-FACTOR EXPERIMENT

A randomized complete block experiment was performed with 11 varieties of lima beans, which were compared for ascorbic acid content (y) (Steel and Torrie, 1980, p. 411). The percentage of dry matter of freshly harvested beans (x) was assessed as a covariate to account for differences in maturity at harvest. The data are reproduced in Table 2.

Table 2. Lima bean data of Steel and Torrie (1980, p.411), where y = ascorbic acid content, x = percentage dry matter.

Variety	Block 1		Block 2		Block 3		Block 4		Block 5	
	x	y	x	y	x	y	x	y	x	y
1	34.0	93.0	33.4	94.8	34.7	91.7	38.9	80.8	36.1	80.2
2	39.6	47.3	39.8	51.5	51.2	33.3	52.0	27.2	56.2	20.6
3	31.7	81.4	30.1	109.0	33.8	71.6	39.6	57.5	47.8	30.1
4	37.7	66.9	38.2	74.1	40.3	64.7	39.4	69.3	41.3	63.2
5	24.9	119.5	24.0	128.5	24.9	125.6	23.5	129.0	25.1	126.2
6	30.3	106.6	29.1	111.4	31.7	99.0	28.3	126.1	34.2	95.6
7	32.7	106.1	33.8	107.2	34.8	97.5	35.4	86.0	37.8	88.8
8	34.5	61.5	31.5	83.4	31.1	93.9	36.1	69.0	38.5	46.9
9	31.4	80.5	30.5	106.5	34.6	76.7	30.9	91.8	36.8	68.2
10	21.2	149.2	25.3	151.6	23.5	170.1	24.8	155.2	24.6	146.1
11	30.8	78.7	26.4	116.9	33.2	71.8	33.5	70.3	43.8	40.9

The data were subjected to analysis of covariance using ascorbic acid content as the response and percentage dry matter as the covariate. Due to the covariate adjustment, pairwise differences of adjusted means do not have a common SED. Analysis of covariance can be performed using the GLIMMIX procedure as shown in Box 1 The note at the bottom of the output indicates that variety comparisons 5 vs. 9 and 1 vs. 9 are not represented by the lines display but need to be suppressed in order to generate a lines display.

Box 1.GLIMMIX statements and output for analysis of Lima bean data (Steel and Torrie, 1980, p.411) with mean comparisons using the LINES option.

```
proc glimmix data=lima;
  class block variety;
  model y=x block variety;
  lsmeans variety/pdiff lines;
run;

       Conservative T Grouping for variety Least Squares Means (Alpha=0.05)

          LS-means with the same letter are not significantly different.

                    variety    Estimate

                       10       122.78        A

                        7        99.9787       B
                                               B
                        6        97.5068       B
                                               B
                        5        95.9831    C  B
                                           C  B
                        1        92.5873    C  B
                                           C
                        4        84.5301    C  D
                                           C  D
                        9        81.1467    C  D  E
                                           C  D  E
                        2        79.1164    C  D  E
                                              D  E
                        3        78.1031       D  E
                                               D  E
                       11        74.3191       D  E
                                                  E
                        8        72.0447          E

 The LINES display does not reflect all significant comparisons. The
 following additional pairs are significantly different: (5,9), (1,9).
```

Box 2. GLIMMIX statements and output for multiple comparisons using the %MULT macro.

```
%include '<=insert drive here=>:\<=insert path here=>\mult.sas';
proc glimmix data=lima;
  ods output lsmeans=lsmeans diffs=diffs;
  class rep variety;
  model y=x rep variety;
  lsmeans variety/pdiff;
  %mult(trt=variety);
run;

                Letter display

  trt        label    lsmean letters

  variety       1 92.587327        b c
                2 79.116425          c   e
                3 78.103108           d e
                4 84.530117          c d
                5 95.983054        b c
                6 97.506844        b
                7 99.978678        b
                8 72.044748            e
                9 81.146722         d e
               10 122.78384     a
               11 74.319134         d e

Standard errors of a difference (SED) and Least significant differences (LSD)

                   The MEANS Procedure

         Variable        Minimum          Mean        Maximum

         SED           4.7581045      5.8194515     10.3449555
         LSD           9.6241747     11.7709518     20.9246476
```

The SAS code for using the %MULT macro and the associated output are shown in Box 2.The letter display represents all significant differences at $\alpha = 5\%$, and it is the same as that obtained from GLIMMIX using the LINES option except that the letter "c" on variety 9 and the letter "d" on variety 2 are dropped compared to the output in Box 1. Note that by dropping these letters, we 'insert' the suppressed significances of the comparisons1 vs. 9 and 5 vs. 9 into the imperfect lines display in Box 1. In addition, the %MULT macro also shows the average LSD and the average SED. The minimum and maximum for both statistics are quite different, so in this case one may not want to report the average SED or LSD, unless with a clear qualification that these statistics are given just for descriptive purposes. In fact, in case of heterogeneity it may be prudent to also report minimum and maximum of these statistics for clarity.

A SPLIT-PLOT EXPERIMENT

The letter display can also be used for factorial experiments and with generalized and/or mixed linear models (Piepho, 2004). To illustrate the use of the %MULT macro and compare it to the SLICE statement with LINES option of the GLIMMIX procedure, I will use a split-plot experiment with oats (Steel and Torrie, 1980, p.384). The main plot factor was seed lot and the sub plot factor was seed treatment. Main plots were randomized in complete blocks. The response was yield in bushels per acre. A mixed linear model with two random error terms (main plot and sub plot error) needs to be used for analysis of this trial. The interactions are significant in this example, so it is useful to compare seed lot means separately for each level of seed treatment and vice versa. This is easily done using the new

SLICE statement in conjunction with the LINES option (SAS Version 9.3) as shown in Box 3. The same analysis is obtained by the %MULT macro using the statements shown in Box 4.

Box 3.GLIMMIX statements to analyse seed lot × seed treatment means of oats experiment using the LINES option. Part of output: Comparison of seed lots for seed treatment 'Panogen'.

```
proc glimmix data=oats;
  class block seed_treatment seed_lot;
  model yield=block seed_treatment|seed_lot;
  random block*seed_lot;
  slice seed_lot*seed_treatment/sliceby=seed_treatment lines;
  slice seed_lot*seed_Treatment/sliceby=seed_lot lines;
run;

T Grouping for seed_treatm*seed_lot Least Squares Means Slice (Alpha=0.05)

      LS-means with the same letter are not significantly different.

        Slice                    seed_lot    Estimate

        seed_treatment Panogen   Branch       57.6750        A
        seed_treatment Panogen                               A
        seed_treatment Panogen   Clinton      55.8750        A
        seed_treatment Panogen                               A
        seed_treatment Panogen   Vicaland2    53.1000    B   A
        seed_treatment Panogen                          B
        seed_treatment Panogen   Vicaland1    45.8500    B
```

Box 4. GLIMMIX statements to analyse seed lot × seed treatment means of oats experiment using the %MULT macro. Part of output: Comparison of seed lots for seed treatment 'Panogen'.

```
proc glimmix data=oats;
  ods output lsmeans=lsmeans diffs=diffs;
  class block seed_treatment seed_lot;
  model yield=block seed_treatment|seed_lot;
  random block*seed_lot;
  lsmeans seed_treatment*seed_lot/pdiff;
  %mult(trt=seed_lot, by=seed_treatment);
  %mult(trt=seed_treatment, by=seed_lot);
run;

                      Letter display

  trt       by               level    label      lsmean letters

  seed_lot seed_treatment "Panogen" Branch      57.675        a
                                    Clinton     55.875        a
                                    Vicaland1    45.85          b
                                    Vicaland2     53.1        a b

Standard errors of a difference (SED) and Least significant differences (LSD)

                      The MEANS Procedure

        Variable        Minimum            Mean            Maximum

        SED            4.0254335        4.0254335        4.0254335
        LSD            8.1639575        8.1639575        8.1639575
```

The output for one of the comparisons (seed lots at seed treatment 'Panogen') using both approaches is shown in Boxes 3 and 4. Results are identical by both procedures, and no significances need to be suppressed with the SLICE statement because the data are balanced and the random part of the model has a simple variance components form. An advantage of the %MULT macro in this case is that it reports an average LSD and SED. These two statistics are constant across all comparisons because the data are balanced and so they can be reported along with the means and letter display.

A REPEATED-MEASURES EXPERIMENT

A completely randomized experiment with four plant protection treatments was conducted to assess the disease progress of lettuce drop (trial LD8 in Simko and Piepho, 2012). The percentage of diseased leaf area was assessed on the same plots on eight consecutive dates. The repeated measures per plot were analysed using an unstructured variance-covariance model for plot error. This model allows for two important properties of the data: (i) serial correlation among repeated measures in the same plot and (ii) heterogeneity of variance between different time points due to the fact that spread of the disease progresses over time. With percentage data, the homogeneity of variance and normality assumptions need to be critically checked. While the fitted model accounts for heterogeneity of variance between time points, it assumes constancy of variance within time points. In this case, the residual plots for the fitted mixed model showed no evidence of gross departures from assumptions, so the analysis seems acceptable. The statements for analysing this experiment are given in Boxes 5 and 6. The treatment × time interaction is significant, so time means are compared separately for each treatment. The mean comparison for treatment T-02 by both approaches is shown in Boxes 5 and 6. The LINES option to the SLICE statement suppresses the significant difference between time points 6 and 8 (Box 5). This significance is correctly represented by the letter display obtained with the %MULT macro (Box 6). The LSDs and SEDs are very heterogeneous because of the unstructured variance-covariance model. Thus, their average values would not be reported in this case.

Box 5. GLIMMIX statements to analyse treatment × time means of lettuce experiment using the LINES option. Part of output: Comparison of time points for treatment 'T-02'.

```
proc glimmix data=LD8;
  class treatment time plot;
  model response=treatment|time/ddfm=KR;
  random time/subject=plot type=unr residual;
  slice treatment*time/sliceby=treatment lines;
run;

Conservative T Grouping for treatment*time Least Squares Means Slice
(Alpha=0.05)

          LS-means with the same letter are not significantly different.

                  Slice              time    Estimate

                  treatment T-02     8       58.9286    A
                  treatment T-02                        A
                  treatment T-02     7       57.1429    A
                  treatment T-02                        A
                  treatment T-02     6       55.3571    A
                  treatment T-02                        A
                  treatment T-02     5       54.4643    A
                  treatment T-02                        A
                  treatment T-02     4       53.5714    A
                  treatment T-02
                  treatment T-02     3       37.5000    B
                  treatment T-02
                  treatment T-02     2        6.2500    C
                  treatment T-02                        C
                  treatment T-02     1        2.6786    C

The LINES display does not reflect all significant comparisons. The following
additional pairs are significantly different: (T-02 8,T-02 6).
```

Box 6. GLIMMIX statements to analyse treatment × time means of lettuce experiment using the %MULT macro. Part of output: Comparison of time points for treatment 'T-02'.

```
proc glimmix data=LD8;
  ods output lsmeans=lsmeans diffs=diffs;
  class treatment time plot;
  model response=treatment|time/ddfm=KR;
  random time/subject=plot type=unr residual;
  lsmeans treatment*time/pdiff;
  %mult(trt=time, by=treatment);
run;


                    Letter display


trt   by          level     label    lsmean letters

time treatment "T-02"          1 2.6785714              d
                               2      6.25              d
                               3      37.5             c
                               4 53.571429         a b
                               5 54.464286         a b
                               6 55.357143           b
                               7 57.142857         a b
                               8 58.928571         a


Standard errors of a difference (SED) and Least significant differences (LSD)


                     The MEANS Procedure


        Variable        Minimum          Mean        Maximum
        ------------------------------------------------------
        SED           1.0777160     4.4535985      7.5507163
        LSD           2.2242966     9.1917756     15.5839125
        ------------------------------------------------------
```

DISCUSSION

The likelihood of lost significances increases with the degree of imbalance and with the number of treatments being compared. Also, the problem is more likely to occur with complex variance-covariance structures, e.g., when analysing repeated measures, than with simple variance component models. If the problem does not occur, use of the new SLICE statement for comparison of simple means in factorial experiments and of the LSMEANS statement for comparison of marginal means in factorial and single-factor experiments using the LINES option may be preferred for simplicity, unless an average LSD or SED is needed, in which case one may want to use the %MULT macro. When the LINES option needs to suppress significances, use of the %MULT macro is preferable, because this is guaranteed to always truthfully represent all significant differences.

Saxton (1998) provided a SAS macro %PDMIX800 for generating letter displays. His algorithm starts from a display that assigns a different letter to each pair of non-significantly different treatments and then successively merges groups of treatments that are not significantly different. The insert-and-absorb algorithm starts from the opposite end: It first declares all means to be non-significantly different, i.e., all means get an 'a'. It then inserts significant differences of treatments X and Y by duplicating columns of letters connecting X and X and dropping one letter from each column, i.e. the letter for X in the one and the letter for Y in the other. Finally, the display is swept for redundant letters. When developing this

algorithm, I had also experimented with an algorithm similar to that implemented by Saxton (1998) (alas, I was unaware of his fine macro at the time), but eventually abandoned it because I found the insert-and-absorb algorithm to be somewhat more efficient in terms of computing time. Note that when there are no differences among t means, the initial configuration of Saxton's algorithm requires $t(t-1)/2$ letters, and the computational effort to reduce this display can become non-negligible when t is large.

Box 7 shows the output of Saxton's macro when applied to the lima bean example. The result is the same as that from the 'insert-and-absorb' algorithm (Box 2), except for the additional 'D' on variety 2. That letter on variety 2 is redundant and can be dropped without losing any significant difference. This is easily verified by checking the comparisons of variety 2 with varieties 3, 4, 9 and 11, the other varieties carrying a 'D' in Box 7. Comparisons 2 vs. 3, 2 vs. 9, and 2 vs. 11 are declared non-significant by the letter 'E', while comparison 2 vs. 4 is declared non-significant by letter 'C', so the letter 'D' on variety 2 can be dropped without losing a significance. It would probably be easy to extend the macro by a sweeping step that drops redundant letters. In this example, sweeping would result in the same display as that obtained by %MULT.

Box 7. Output of macro %PDMIX800 (Saxton, 1998) when applied to the lima bean example.

```
-------------- Effect=variety   Method=LSD(P<.05)   Set=1 ----------------

                                        Standard     Letter
        Obs     variety    Estimate     Error        Group

         1         1        92.5873      3.4064        BC
         2         2        79.1164      6.2754        CDE
         3         3        78.1031      3.5086        DE
         4         4        84.5301      3.9503        CD
         5         5        95.9831      4.9680        BC
         6         6        97.5068      3.5889        B
         7         7        99.9787      3.3798        B
         8         8        72.0447      3.3643        E
         9         9        81.1467      3.3904        DE
        10        10       122.78        5.1403        A
        11        11        74.3191      3.3659        DE
```

The insert-and-absorb algorithm is also implemented in the R package **multcomp** available at http://CRAN.R-project.org. For an illustration see Bretz et al. (2010).

REFERENCES

Bretz, F., Hothorn, T., Westfall, P. (2010). *Multiple comparisons using R*. CRC Press, Boca Raton.

Piepho, H.P. (2000). Multiple treatment comparisons in linear models when the standard error of a difference is not constant. *Biometrical Journal* 42, 823–835.

Piepho, H.P. (2004). An algorithm for a letter-based representation of all-pairwise comparisons. *Journal of Computational and Graphical Statistics* 13, 456–466.

Saxton, A.M. (1998). A macro for converting mean separation output to letter groupings in Proc Mixed. *Proc. 23rd SAS Users Group Intl., SAS Institute, Cary, NC*, pp 1243-1246.

Simko, I., Piepho, H.P. (2012). The area under the disease progress stairs (AUDPS): calculation, advantages, and application. *Phytopathology* 102, 381-389.

Steel, R.G.D., Torrie, J.H. (1980). *Principles and procedures of statistics. A biometrical approach.* 2nd edition. McGraw-Hill, New York.