Letter to the Editor

G. J. McLachlan

A common problem in practice is the analysis of experiments designed to compare treatments. In the situation where an analysis of variance (ANOVA) leads to a significant F-test for the difference between treatment means, there is the problem of isolating those treatments that do not appear to be different. In a recent issue of this journal, Bautista, Smith, and Steiner (1997) considered a new procedure for the grouping of treatments following an analysis of variance, leading to a significant F-test for the differences between treatment means. They subsequently illustrated their procedure in an example involving some data from Steel and Torrie (1980). In particular, they claimed that their "procedure differs from most in that distinct groups are created."

However, in addition to the hierarchical cluster analysis approach of Scott and Knott (1974) referenced in their paper, we wish to point out that various methods have been proposed in the past to partition treatments into distinct groups, including by those by Edwards and Cavalli-Sforza (1965), Jolliffe (1975), Binder (1978, 1981), Cox and Spötvoll (1982), and Caliński and Corsten (1985); see McLachlan and Basford (1988, Chapter 6). Additional references may be found in Carmer and Lin (1983). The possibility of using cluster analysis in place of multiple comparison procedures had been suggested by O'Neill and Wetherill (1971). In the discussion of this paper, Nelder (1971) commented that one of the patterns to look for in the sample means of the treatments was whether "the means divide into two or more groups within which they look like samples from normal distributions." The normal mixture model-based approach to this problem as developed in Aitkin (1980) and Basford and McLachlan (1985) is in keeping with the spirit of this comment.

With this mixture approach, the n treatment means, denoted here by $\bar{x}_1, \ldots, \bar{x}_n$, are assumed to be distributed (independently) about their means $\mu_1, \ldots, \mu_n$ with variances equal to $\pi^2/r_1, \ldots, \pi^2/r_n$, respectively, where r_j denotes the number of observations made on the jth treatment. In addition, s^2 denotes an independent estimate of π^2, distributed as

$$v s^2/\pi^2 \sim \chi^2_v. \tag{1}$$

G. J. McLachlan is with the Department of Mathematics, University of Queensland, St. Lucia, Queensland 4072, Australia.

Journal of Agricultural, Biological, and Environmental Statistics, Volume 6, Number 2, Pages 302–304

Under a normal mixture model with g components, it is assumed further that, conditional on its membership of the ith group, $\bar{x}_j$ is distributed as

$$\bar{x}_j \sim \mathrm{N}(\mu_i, \pi^2/r_j \text{ in } G_i \text{ with probability } \nu_i, \qquad i = 1, \ldots, g. \tag{2}$$

The log likelihood for the vector of unknown parameters formed on the basis of $\bar{x}_1, \ldots, \bar{x}_n$ under the mixture model (2) and also s^2 under (1) is given, up to terms not involving the unknown parameters, by

$$\log L = \sum_{j=1}^{n} \log \sum_{i=1}^{g} \nu_i \varphi\left(\bar{x}_j; \mu_i, \pi^2/r_j\right) - \frac{1}{2} v \log \pi^2 - \frac{1}{2} v s^2/\pi^2, \tag{3}$$

where $\varphi(x; \mu, \pi^2)$ denotes the univariate normal mixture density with mean μ and variance π^2. This mixture model can be fitted to the treatments using the EM algorithm of Dempster, Laird, and Rubin (1977); see also McLachlan and Krishnan (1997). The FORTRAN listing of a program for this purpose is given in the Appendix of McLachlan and Basford (1988). A probabilistic clustering of the treatments is obtained in terms of their fitted posterior probabilities of component membership. An outright clustering into distinct groups is obtained by assigning each treatment mean to the group to which it has the highest posterior probability of belonging.

We applied this approach to the six treatments in the example considered by Bautista et al. (1997). These treatments, designated as 3DOk1, 3DOk5, 3DOk7, COMP, 3DOk4, and 3DOk13, have means equal to 28.82, 23.98, 19.92, 18.70, 14.74, and 13.26, respectively. For $g = 2$, it results in the first four treatments being clustered into one group and the remaining two in another group. For $g = 3$, the first two treatments are grouped together in one cluster, the next three in another, while the last treatment forms a group on its own. These partitions of the treatment means agree with those of Bautista et al. (1997). On the question of whether there should be $g = 2$ or 3 groups, the value of $-2 \log \lambda$ (i.e., twice the increase in the log likelihood for $g = 3$ over $g = 2$) is 4.702. An assessment of the P-value can be obtained by using a resampling approach as in McLachlan (1987). Its application here with $K = 19$ bootstrap replications leads to the null hypothesis of $g = 2$ groups being retained at the 5% level (approximately) since the value of $-2 \log \lambda$ did not exceed its largest bootstrap replication.

[Accepted January 2001.]

REFERENCES

Aitkin, M. A. (1980), "Mixture Applications of the EM Algorithm in GLIM," in *Compstat 1980, Proceedings Computational Statistics,* Vienna: Physica-Verlag, pp. 537–541.

Basford, K. E., and McLachlan, G. J. (1985), "Cluster Analysis in a Randomized Complete Block Design," *Communications in Statistics—Theory and Methods,* 14, 451–463.

Bautista, M. G., Smith, D. W., and Steiner, R. L. (1997), "A Cluster-Based Approach to Means Separation," *Journal of Agricultural, Biological and Environmental Statistics,* 2, 179–197.

Binder, D. A. (1978), "Bayesian Cluster Analysis," *Biometrika,* 65, 31–38.

——— (1981), "Approximations to Bayesian Clustering Rules," *Biometrika,* 68, 275–285.

Caliński, T., and Corsten, L. C. A. (1985), "Clustering Means in ANOVA by Simultaneously Testing," *Biometrics,* 41, 39–48.

Carmer, S. G., and Lin, W. T. (1983), "Type I Error Rates for Divisive Clustering Methods for Grouping Means in Analysis of Variance," *Communications in Statistics—Simulation and Computation,* 12, 451–466.

Cox, D. R., and Spøtvoll, E. (1982), "On Partitioning Means Into Groups," *Scandinavian Journal of Statistics: Theory and Applications,* 9, 147–152.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1997), "Maximum Likelihood From Incomplete Data via the EM Algorithm (with discussion)," *Journal of the Royal Statistical Society, Series B,* 39, 1–38.

Edwards, A. W. F. and Cavalli-Sforza, L. L. (1965), "A Method for Cluster Analysis," *Biometrics,* 21, 362–375.

Jolliffe, I. T. (1975), "Cluster Analysis as a Multiple Comparison Method," in *Applied Statistics,* ed. R. P. Gupta, Amsterdam: North-Holland, pp. 159–168.

McLachlan, G. J. (1987), "On Bootstrapping the Likelihood Ratio Test Statistic for the Number of Components in a Normal Mixture," *Applied Statistics,* 36, 318–324.

McLachlan, G. J., and Basford, K. E. (1988), *Mixture Models: Inference and Applications to Clustering,* New York: Marcel Dekker.

McLachlan, G. J., and Krishnan, T. (1997), *The EM Algorithm and Extensions,* New York: Wiley.

Nelder, J. A. (1971), "Contribution to the Discussion of Paper by R. O'Neill and G. B. Wetherill," *Journal of the Royal Statistical Society, Series B,* 33, 244–246.

Scott, A. J., and Knott, M. (1974), "A Cluster Analysis Method for Grouping Means in the Analysis of Variance," *Biometrics,* 30, 507–512.

Steel, R., and Torrie, J. (1980), *Principles and Procedures of Statistics: A Biometrical Approach* (2nd ed.), San Francisco: McGraw-Hill.