Vijayan Nair

AT&T Bell Laboratories

and

Michael Frenklach

Department of Materials Science and Engineering
Pennsylvania State University[1]

Nair: The paper by Box and Meyer deals with some interesting problems that arise in the off-line quality control methods introduced by Taguchi (see Taguchi and Wu [1][2]). My comments will be restricted to the first part of the paper, viz. estimating dispersion effects.

1) Estimating dispersion effects for quality control

In industrial experiments designed to detect important factors that affect the quality of a manufacturing/production process, the estimation of dispersion effects is as important as the estimation of location effects. In fact in situations where there are readily identifiable signal factors (see [1]), the primary goal is in estimating dispersion effects. The location effects, in this case, play the role of nuisance parameters. León, Shoemaker and Kackar [2] offer an excellent discussion of the statistical formulation of the parameter design problem in industrial experiments.

2) Effect sparsity

The Box-Meyer techniques exploit the notion of effect sparsity to obtain "replicates" in an unreplicated experiment. It is likely that in most cases only a few factors are *highly* significant. However, in many situations, one could also expect many of the other factors included in the experiment to have sizeable effects. This is particularly true when a fair amount of the information about the process is known and is used in the selection of the factors. In such situations, one could not reasonably expect to estimate both location and dispersion effects from an unreplicated experiment.

[1] Michael Frenklach's contribution to the subject stems from work performed in the Department of Chemical Engineering, Louisiana State University.

[2] Figures in brackets indicate literature references.

3) *When to log?*

Box suggests using $\log[S^2(i-)/S^2(i+)]$ as a preliminary estimate of the dispersion effects. An alternative method would be to take log of the squared residuals and do ANOVA with an additive model for the log of the scale parameters. This is the type of analysis usually done in experiments with replications. Some efficiency calculations suggest that the Box-Meyer analysis is more efficient when there are only a few large dispersion effects and less efficient when there are many.

4) Iterating

It is possible that during the first step of the iteration (which does an unweighted analysis) some significant location effects are not detected. So after the dispersion effects are estimated, the location model should be refitted for all the factors.

5) *Transformations*

In replicated experiments, the quasi-likelihood models of Nelder and Pregibon [3] allow one to determine the transformations under which the location effects and the scale effects are approximately additive.

References

[1] Taguchi, G., and Y. Wu, *Introduction to off-line quality control.* Central Japan Quality Control Association (1980).

[2] León, R. V.; A. C. Shoemaker and R. N. Kackar, Performance measures independent of adjustment: An alternative to Taguchi's signal to noise ratio. *AT&T Bell Laboratories Technical Memorandum* (1985).

[3] Nelder, J .A., and D. Pregibon, Quasi-likelihood models and data analysis. Biometrika (to appear).

Frenklach: Professor Box presented a method of analysis of factorial designs for detection of main effects and faulty observations. The approach is analytical and provides numerical measures for what previously has been approached graphically. The availability of the analytical algorithm is important for computerization of the analysis.

The central hypothesis of the method is what the authors call effect sparsity, which states that usually only a small number of input and control process variables would have a significant effect on the process response(s). This situation appears to be true not only in experimental environments but also in computer modeling of various industrial processes and natural phenomena. Mechanistic models usually take the form of differential equations for which no analytical solution is available. The model may contain a large number of (physical) parameters and it is not always obvious from a simple inspection of the computational results what effect each parameter has on a given response or responses. Sensitivity analysis has been used to reveal this information. Among other techniques, the use of screening factorial designs for sensitivity analysis of computer models has been suggested by Box et al. 1978; Frenklach 1984; Frenklach and Bornside, 1984; Miller and Frenklach, 1983; and Morris and Mitchell, 1983 [1–5].

The present experience with chemical kinetic modeling, for example, is as follows. Due to technical difficulties of instrumentation, there are only a few experimental responses available, typically one or two. The cases studied indicate that it is a very small number of chemical reactions, out of hundreds of reactions comprising the model, whose rate coefficient values, within their uncertainty intervals, have significant or "active" effects on the experimentally verifiable model responses. These are exactly the conditions of effect sparsity discussed above. Thus, the method presented by Box is well-suited not only for "real" experimentation, but also for computer modeling.

Computations, however, do not have random errors. Does this fact simplify and economize the analysis? Should special methods and designs be developed or existing ones modified for a most efficient use in screening analysis of computer models?

References

[1] Box, G. E. P.; W. C. Hunter and J. S. Hunter, Statistics for Experimenters. An Introduction to Design, Data Analysis, and Model Building, Wiley, New York, pp. 429–432 (1978).

[2] Frenklach, M., Modeling, Combustion Chemistry (Edited by W. C. Gardiner, Jr.), Springer-Verlag, New York, Chap. 7 (1984).

[3] Frenklach, M., and D. E. Bornside, Shock-initiated ignition in methane-propane mixtures, Combust. Flame **56**, 1–27 (1984).

[4] Miller, D., and M. Frenklach, Sensitivity analysis and parameter estimation in dynamic modeling of chemical kinetics, Int. J. Chem. Kinet. **15**, 677–696 (1983).

[5] Morris, M. D., and T. J. Mitchell, Two-level multifactor designs for detecting the presence of interaction, *Technometrics* **25**, 345–355 (1983).