# MULTIPARAMETER PROBLEMS FROM A BAYESIAN POINT OF VIEW[1]

By G. E. P. Box and George C. Tiao

*University of Wisconsin*

**1. Introduction.** It has long been known that when attention is focused on a single parameter or comparison of parameters such as the difference in means or the ratio of variances from the normal samples, then using "non-informative" prior distributions Bayesian results exactly paralleling classical procedures and involving the standard $t$ and $F$ distributions can be obtained. When there are many parameters certain "portmanteau" multi-comparison procedures can be derived on the sampling theory approach and they are frequently of great value when used to supplement individual comparisons. These include the $\chi^2$-goodness of fit test, the analysis of variance test and Bartlett's test to compare variances (when it is suitably modified so as to be robust to non-normality). This paper shows that these portmanteau procedures do have a simple and natural Bayesian interpretation. Furthermore, using Bayesian methods, it is possible to cover important cases not amenable to treatment by classical techniques.

Suppose in a particular investigation we have computed an appropriate posterior distribution $p(\boldsymbol{\theta} \mid \mathbf{y})$ where $\boldsymbol{\theta}$ is a $k$-dimensional vector of parameters of interest and $\mathbf{y}$ is an $n$-dimensional vector of observations. Then from the Bayesian point of view, all inferential problems concerning $\boldsymbol{\theta}$ may be answered in terms of $p(\boldsymbol{\theta} \mid \mathbf{y})$. In practice, inference involves a communication with the mind, and usually it is difficult to comprehend a function in $k$-dimensions. Fortunately, there are often specific individual features of $p(\boldsymbol{\theta} \mid \mathbf{y})$ of interest which can be appreciated by one, two or three dimensional thought. For example, marginal distributions may be of interest. Or we may inspect conditional distributions of a small subset of the parameters for specific values of the other parameters. With high speed electronic computers available, print-outs of two dimensional sections of such distributions can be readily obtained. The value of such appraisals of the estimation situation is very great, as has been repeatedly pointed out by Barnard in connection with the likelihood principle.

Another feature of $p(\boldsymbol{\theta} \mid \mathbf{y})$ which is of value is a posterior probability region for the parameters. Often the region over which the posterior density is non-zero extends over infinite ranges in the parameter space. Nevertheless over a substantial part of the parameter space the density may be small or negligible. It is therefore possible to delineate a comparatively small region which contains most (say 95%) of the probability mass. Obviously there is an infinite number of ways such a region can be chosen. We must therefore decide what properties we would like the region to have. Either of the following two principles seems to be

Received 8 February 1965; revised 12 April 1965.

[1] This research was supported by the Office of Naval Research under Contract Nonr-1202(17), Project Nro 42-222.



Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org