Journal of Research of the National Bureau of Standards
Volume 90, Number 6, November–December 1985

Some New Ideas in the Analysis of Screening Designs

George Box
University of Wisconsin-Madison, Madison, WI 53705
and
R. Daniel Meyer
Lubrizol Corp., Wickliffe, OH 44092

Accepted: July 1, 1985

Consideration of certain aspects of scientific method leads to discussion of recent research on the role of screening designs in the improvement of quality. A projective rationale for the use of these designs in the circumstances of *factor sparsity* is advanced. In this circumstance the possibility of identification of sparse *dispersion* effects as well as sparse *location* effect is considered. A new method for the *analysis of fractional factorial designs* is advanced.

Key words: factor sparsity; fractional factorial designs; screening designs; sparse dispersion; sparse location.

Humans differ from other animals most remarkably in their ability to learn. It is clear that although throughout the history of mankind technological learning has taken place, although until three or four hundred years ago change occurred very slowly. One reason for this was that in order to learn something - for example, how to make fire or champagne - two *rare events* needed to coincide: (a) an informative event had to *occur*, and (b) a person able to draw logical conclusion and to act on them had to be *aware* of that informative event.

Passive surveillance is a way of increasing the probability that the rare informative event will be constructively taken note of and is exemplified by quality charting methods. Thus a Shewhart chart is a means to ensure that possibly informative events are brought to the attention of those who may be able to discover in them an "assignable cause" [1][1] and act appropriately.

About the Author, Paper: George Box, who is with the Research Center for Quality and Productivity Improvement at the University of Wiconsin-Madison, and R. Daniel Meyer are both Statisticians. The work they describe was sponsored by U.S. Army Contract DAAC 29-80-C-0041 and National Science Foundation Grant DMS-8420968.

[1]Figures in brackets indicate literature references.

Active intervention by experimentation aims, in addition, to increase the probability of an informative event *actually occurring*. A designed experiment conducted by a qualified experimenter can dramatically increase the probability of learning because it increases simultaneously the probability of an informative event occurring and also the probability of the event being constructively witnessed. Recently there has been much use of experimental design in Japanese industry particularly by Genichi Taguchi [2] and his followers. In off-line experimentation he has in particular emphasized the use of highly fractionated designs and orthogonal arrays and the minimization of variance.

In the remainder of this paper we briefly outline some recent research on the use of such screening designs.

1. Use of Screening Designs to Identify "Active" Factoring

Table 1 shows in summary a highly fractionated two-level factorial design employed as a screening design in an off-line welding experiment performed by the National Railway Corporation of Japan [2]. In the column to the right of the table is shown the observed tensile strength of the weld, one of several quality characteristics measured.

The design was chosen on the assumption that in addition to main effects only the two-factor interactions AC, AG, AH, and GH were expected to be present. On that supposition, all nine main effects and the four selected two-factor interactions can be separately estimated by appropriate orthogonal contrasts, the two remaining contrasts corresponding to the columns labelled e_1 and e_2 measure only experimental error. Below the table are shown the grand average, the 15 *effect contrasts, and the effects plotted on a dot* diagram. The effects plotted on normal probability paper suggested that, over the ranges studied, only factors B and C affect tensile location by amounts not readily attributed to noise.

If this conjecture is true, then, at least appoximately, the 16 runs could be regarded as four replications of a 2^2 factorial design in factors B and C only. However, when the results are plotted in figure 1 so as to reflect this, inspection suggests the existence of a dramatic effect of a different kind—when factor C is at its plus level the spread of the data appears much larger[3] than when it is at its minus level. Thus, in addition to detecting shift in location due to B and C, the experiment may also have detected what we will call a *dispersion effect* due to C. The example raises the general possibility of analyzing unreplicated designs for dispersion effects as well as for the more usual location effects.

[2]To facilitate later discussion, we have set out the design and labelled the levels somewhat differently from [2].

[3]Data of this kind might be accounted for by the effect of one or more variables other than B that affected tensile strength only at the "plus level" of C (only when the alternative material was used). Analysis of the eight runs made at the plus level of C does not support this possibility, however.

2. Rationales for Using Screening Designs

Before proceeding we need to consider the question, "In what situations are screening designs, such as highly fractionated factorials, useful?"

2.1. Effect Sparsity. A common industrial problem is to find from a rather large number of factors those few that are responsible for *large effects*. The idea is comparable to that which motivates the use in quality control studies of the "Pareto diagram." (See, for example, [3]). The situation is approximated by postulating that only a small proportion of effects will be "*active*" and the rest "*inert*". We call this the postulate of *effect sparsity*. For studying such situations, highly fractionated designs and other orthogonal arrays [2,4,5,6] which can screen moderately large numbers of variables in rather few runs are of great interest. Two main rationalizations have been suggested for the use of these designs; both ideas rely on the postulate of effect sparsity but in somewhat different ways.

2.2. Rationale Based on Prior Selection of Important Interactions. It is argued (see for example [7]) that in some circumstances physical knowledge of the process will make only a few interactions likely and that the remainder may be assumed negligible. For example, in the welding experiment described above there were 36 possible two-factor interactions between the nine factors, but only four were regarded as likely, leaving 32 such interactions assumed

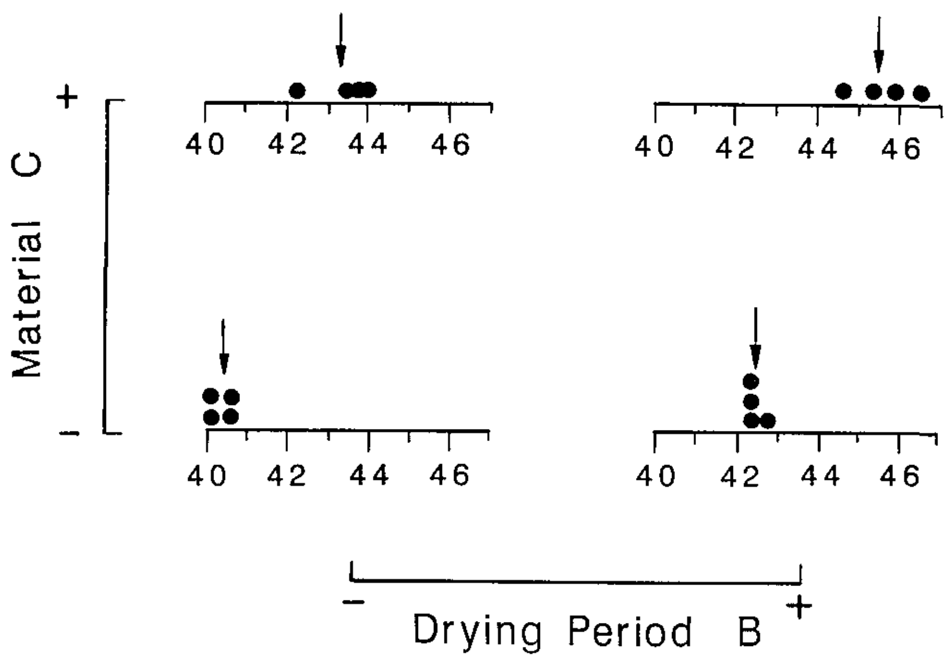


Figure 1—Tensile strength data as four replicates of a 2^2 factorial design in factors B and C only.

negligible. The difficulty with this idea is that in many applications the picking out of a few "likely" interactions is difficult if not impossible. Indeed the investigator might justifiably protest that, in the circumstance where an experiment is needed to determine which *first order* (main) effects are important, it is illogical that he be expected to guess in advance which effects of *second order* (interactions) are important.

2.3. Projective Rationale Factor Sparsity. A somewhat different notion is that of *factor* sparsity. Thus suppose that, of the k factors considered, only a small subset of vaguely known size d, *whose identity is, however, unknown*, will be active in providing main effects and interactions within the subset. Arguing as in [8], a two-level design enabling us to study such a system is a fraction of resolution $R=d+1$ (or in the terminology of [6], an array of strength d) which produces complete factorials (possibly replicated) in every one of the $\binom{k}{d}$ spaces of $d=R-1$ dimensions. For example, we have seen that on the assumption that only factors B and C are important, the welding design could be regarded as four replicates of a 2^2 factorial in just those two factors. But because the design is of resolution $R=3$ the same would have been true for any of the 36 choices of two out of the nine factors tested. Thus the design would be appropriate if it were believed that not more than two of the factors were likely to be "active".

For further illustration we consider again the 16-run orthogonal array of table 1 and, adopting a roman subscript to denote the resolutions of the design, we indicate in table 2 various ways in which that array might be used. It may be shown that

(a) If we associated the 15 contrast columns of the design with 15 factors, we would generate a 2^{15-11}_{III} design providing two-fold replication of 2^2 factorials in every one of the 105 two-dimensional projections.

(b) If we associated only columns 1, 2, 4, 7, 8, 11, 13, and 14 with eight factors we would generate a 2^{8-4}_{IV} design providing two-fold replication of 2^3 factorials in every one of the 56 three-dimensional projections.

(c) If we associated only columns 1, 2, 4, 8, and 15 with five factors we would generate a 2^{5-1}_{V} design providing a 2^4 factorial in every one of the four-dimensional projections.

(d) If we associated only columns 1, 2, 4, and 8 with four factors we would obtain the complete 2^4 design from which this orthogonal array was in fact generated.

Designs (a), (b), and (c) would thus be appropriate for situations where we believed respectively that not more than 2, 3, or 4 factors would be active[4]. Notice that intermediate

[4]The designs give partial coverage for a larger number of factors, for example ([8] (1961)) 56 of the 70 four-dimensonal projections of the 2^{8-4}_{IV} yield a full factorial in four variables.

Table 1. A fractional two-level design used in a welding experiment showing observed tensile strength and effects.

A: Kind of Welding Rods
B: Period of Drying
C: Welded Material
D: Thickness
E: Angle
F: Opening
G: Current
H: Welding Method
J: Preheating

Factor			D	H	e_1	G	F	GH	AC	A	E	AH	e_2	AG	J	B	C	Tensile strength
Column Number		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	kg/mm^2
	1	+	−	−	+	−	+	+	−	−	+	+	−	+	−	−	+	43.7
	2	+	+	−	−	−	−	+	+	−	−	+	+	+	+	−	−	40.2
	3	+	−	+	−	−	+	−	+	−	+	−	+	+	−	+	−	42.4
	4	+	+	+	+	−	−	−	−	−	−	−	−	+	+	+	+	44.7
	5	+	−	−	+	+	−	−	+	−	+	+	−	−	+	+	−	42.4
	6	+	+	−	−	+	+	−	−	−	−	+	+	−	−	+	+	45.9
	7	+	−	+	−	+	−	+	−	−	+	−	+	−	+	−	+	42.2
Run	8	+	+	+	+	+	+	+	+	−	−	−	−	−	−	−	−	40.6
	9	+	−	−	+	−	+	+	−	+	−	−	+	−	+	+	−	42.4
	10	+	+	−	−	−	−	+	+	+	+	−	−	−	−	+	+	45.5
	11	+	−	+	−	−	+	−	+	+	−	+	−	−	+	−	+	43.6
	12	+	+	+	+	−	−	−	−	+	+	+	+	−	−	−	−	40.6
	13	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+	44.0
	14	+	+	−	−	+	+	−	−	+	+	−	−	+	+	−	−	40.2
	15	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	42.5
	16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	46.5
Effect		43.0	.13	−.15	−.30	−.15	.40	−.03	.38	.40	−.05	.43	.13	.13	−.38	2.15	3.10	

Table 2. Some alternative uses of the orthogonal array.

	Columns	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
(a)	2_{III}^{15-11}	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
(b)	2_{IV}^{8-4}	•	•		•			•	•			•		•	•	
(c)	2_{V}^{5-1}	•	•		•				•							•
(d)	2^4	•	•		•				•							

values of k could be accommodated by suitably omitting certain columns. Thus the welding design is a 2_{III}^{9-5} arrangement which can be obtained by omitting six columns from the complete 2_{III}^{15-11}. Notice finally that for intermediate designs we can take advantage of both rationales by arranging, as was done for the welding design, that particular interactions are isolated.

A discussion of the iterative model building process [9] characterized three steps in the iterative data analysis cycle indicated below

→ identification ⟶ fitting ⟶ diagnostic checking ←

Most of the present paper is concerned with model identification - the search for a model worthy to be formally entertained and fitted by an efficient procedure such as maximum likelihood. The situation we now address concerns the analysis of fractional designs such as the welding design in the above context when only a few of the factors are likely to have effects but these may include dispersion effects as well as location effects.

3. Dispersion Effects

We again use the design of table 1 for illustration. There are 16 runs from which 16 quantities—the average and 15 effect contrasts—have been calculated. Now if we were also interested in possible dispersion effects we could also calculate 15 variance ratios. For example, in column 1 we can compute the sample variance s_{1-}^2 for those observations associated with a minus sign and compare it with the sample variance s_{1+}^2 for observations associated with a plus sign to provide the ratio $F_1 = s_{1-}^2 / s_{1+}^2$. If this is done for the welding data we obtain values for $\ln F_i$ given in figure 2(a).[5] It will be recalled that in the earlier analysis a large dispersion effect associated with factor C (column 15) was found, but in figure 2(a) the effect for factor C is not especially extreme, instead the dispersion effect for factor D (column 1) stands out from all the rest. This misleading indication occurs because we have not so far taken account of the aliasing of location and dispersion effects. Since 16 linearly independent location effects have already been calculated for the original data, calculated dispersion effects must be functions

[5] In this figure familiar normal theory significance levels are also shown. Obviously the necessary assumptions are not satisfied in this case, but these percentages provide a rough indication of magnitude.

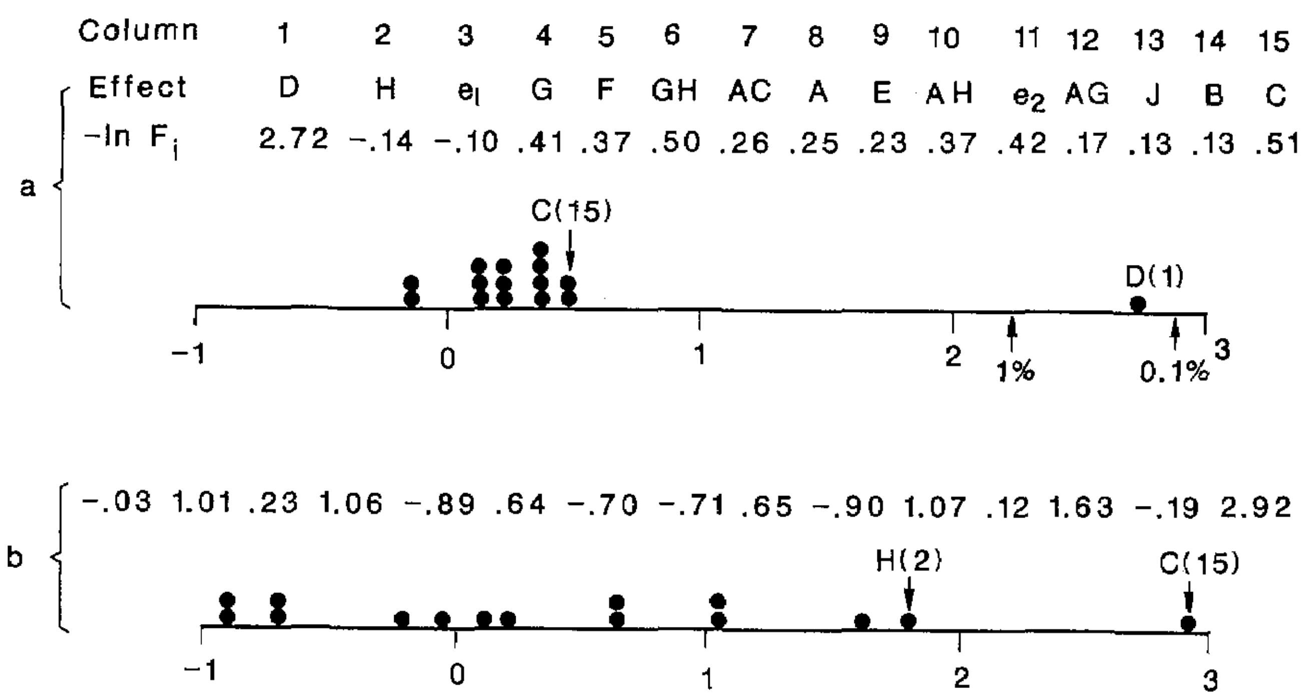


Figure 2—Welding experiment log dispersion effects (a) before, and (b) after elimination of location effects for B and C.

of these. Recently [10] a general theory of location-dispersion aliasing has been obtained for factorials and fractional factorials at two levels. For illustration, in this particular example it turns out that the following identity exists for the dispersion effect F_1, that is the F ratio associated with factor D and hence for column 1 of the design.

$$F_1=\frac{(\hat{2}-\hat{3})^2+(\hat{4}-\hat{5})^2+(\hat{6}-\hat{7})^2+(\hat{8}-\hat{9})^2+(\hat{10}-\hat{11})^2+(\hat{12}-\hat{13})^2+(\hat{14}-\hat{15})^2}{(\hat{2}+\hat{3})^2+(\hat{4}+\hat{5})^2+(\hat{6}+\hat{7})^2+(\hat{8}+\hat{9})^2+(\hat{10}+\hat{11})^2+(\hat{12}+\hat{13})^2+(\hat{14}+\hat{15})^2} \quad (1)$$

Now (see table 1) $\hat{14}=\hat{B}=2.15$ and $\hat{15}=\hat{C}=3.10$ are the two largest location effects, standing out from all the others. The extreme value of F_1 associated with an apparent dispersion effect of factor D(1) is largely accounted for by the squared sum and squared difference of the location effects B and C which appear respectively as the last terms in the denominator and numerator of eq (1). A natural way to proceed is to compute variances from the residuals obtained after eliminating large location effects. After such elimination the alias relations of eq (1) remain the same except that location effects from eliminated variables drop out. That is, zeros are substituted for eliminated variables. Variance analysis for the residuals after eliminating effects of B and C are shown in figure 2(b). The dispersion effect associated with C (factor 15) is now correctly indicated as extreme. It is shown in the paper referenced above how, more generally, under circumstances of effect sparsity a location-dispersion model may be correctly identified when a few effects of both kinds are present.

4. Analysis of Unreplicated Fractional Designs

Another important problem in the analysis of unreplicated fractional designs and other orthogonal arrays concerns the picking out of "active" factors. A serious difficulty is that with unreplicated fractional designs no simple estimate of the experimental error variance against which to judge the effects is available.

In one valuable procedure due to Cuthbert Daniel [11,12] effects are plotted on normal probability paper. For illustration table 3 shows the calculated effects from a 2^{8-4}_{IV} design used in an experiment on injection molding [13, p. 379]. These effects are plotted on normal probability paper in figure 3.

An alternative Bayesian approach [14] is as follows: Let $T_1, T_2, \ldots, T_\nu$ be standardized[6] effects with

$$T_i = e_i \text{ if effect inert}$$

$$T_i = e_i + \tau_i \text{ if effect active}$$

[6]For three-level and mixed two and three level designs for example, this analysis is carried out after the effects are scaled so that they all have equal variances.

Table 3. Calculated effects from a 2^{8-4}_{IV} design showing alias structure assuming three factor and higher order interactions negligible, injection molding experiment.

T_1 =	−0.7 + 1	mold temp.
T_2 =	−0.1 + 2	moisture content
T_3 =	5.5 + 3	holding pressure
T_4 =	−0.3 + 4	cavity thickness
T_5 =	−3.8 + 5	booster pressure
T_6 =	−0.1 + 6	cycle time
T_7 =	0.6 + 7	gate size
T_8 =	1.2 + 8	screw speed

T_9 =	$T_{1.2}$ =	−0.6 + 1.2 + 3.7 + 4.8 + 5.6
T_{10} =	$T_{1.3}$ =	0.9 + 1.3 + 2.7 + 4.6 + 5.8
T_{11} =	$T_{1.4}$ =	−0.4 + 1.4 + 2.8 + 3.6 + 5.7
T_{12} =	$T_{1.5}$ =	4.6 + 1.5 + 2.6 + 3.8 + 4.7
T_{13} =	$T_{1.6}$ =	−0.3 + 1.6 + 2.5 + 3.4 + 7.8
T_{14} =	$T_{1.7}$ =	−0.2 + 1.7 + 2.3 + 6.8 + 4.5
T_{15} =	$T_{1.8}$ =	−0.6 + 1.8 + 2.4 + 3.5 + 6.7

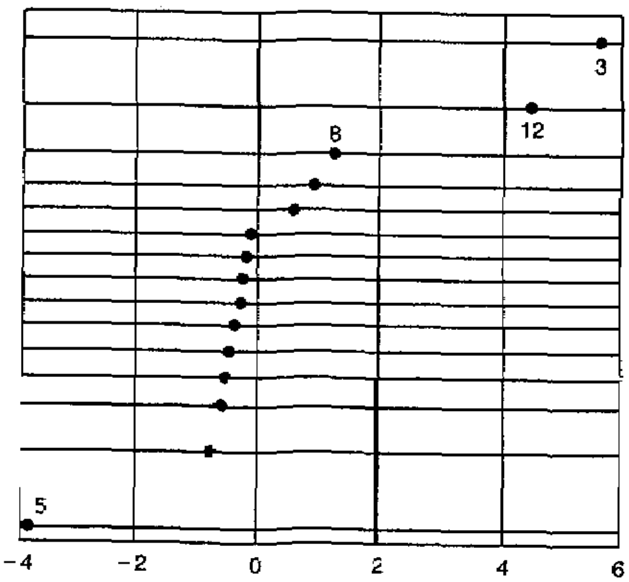
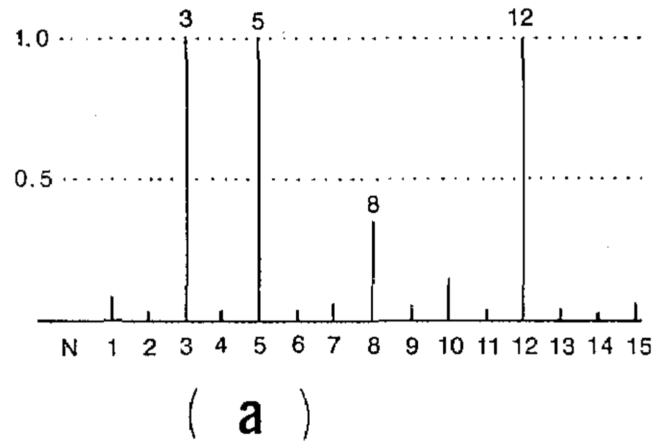


Figure 3—Normal plot of effects. Injection molding experiment.



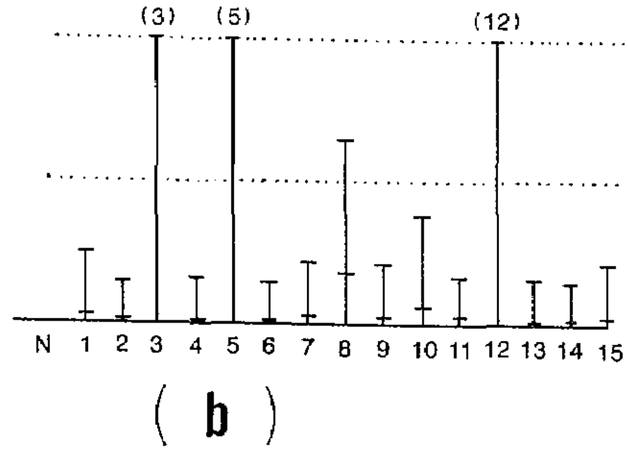


Figure 4—(a) Welding experiment. Posterior probability that factor i is active (a = 0.30, k = 10). (b) Sensitivity analysis for posterior probability (a = .15 – .45, k = 5 – 15).

$$e_i \rightarrow N(0,\sigma^2),\ \tau_i \rightarrow N(0,\sigma_\tau^2)\ k^2 = \frac{\sigma^2 + \sigma_\tau^2}{\sigma^2}\ .$$

Suppose the probability that an effect is active is α.

Let $a_{(r)}$ be the event that a particular set of r of the ν factors are active, and let $\mathbf{T}_{(r)}$ be the vector of estimated effects corresponding to active factors of $a_{(r)}$. Then, [15] with $p(\sigma) \propto 1/\sigma$ the posterior probability that $\mathbf{T}_{(r)}$ are the only active effects is:

$$P[a_{(r)}|\mathbf{T},\alpha,k] \propto \left[\frac{\alpha k^{-1}}{1-\alpha}\right]^r \left\{1-\left(1-\frac{1}{k^2}\right)\frac{S_{(r)}}{S}\right\}^{-\frac{\nu}{2}}\ ,$$

where $S_{(r)} = \mathbf{T}'_{(r)}\mathbf{T}_{(r)}$ and $S = \mathbf{T}'\mathbf{T}$. In particular the marginal probability that an effect i is active give T, α and k is proportional to

$$\sum_{\substack{a_{(r)} \\ i\ active}} \left[\frac{\alpha k^{-1}}{1-\alpha}\right]^r \left\{1-\left(1-\frac{1}{k^2}\right)\frac{S_{(r)}}{S}\right\}^{-\frac{\nu}{2}}\ .$$

A study of the fractional factorials appearing in [7,12,13], suggested that α might range from 0.15-0.45 while k might range from 5 to 15. The posterior probabilities computed with the (roughly average) values, $\alpha=0.30$ and $k=10$ are shown in figure 4(a) in which N denotes the probability (negligible for this example) that there are no active effects. The results from a sensitivity analysis in which α and k were altered to vary over the ranges mentioned above is shown in figure 4(b).

It will be seen that figure 4(a) points to the conclusion that active effects are associated with columns 3, 5 and 12 of the design and that column 8 might possibly also be associated with an active factor. Figure 4(b) suggests that this conclusion is very little affected by widely different choices for α and k. Further research with different choices of prior, with marginization with respect to k, and with different choices of the distribution assumptions is being conducted.

5. Allowance for Faulty Observations

Recent work [16] has shown how a double application of the scale-contamination model (both to the observations themselves as well as to the affects) can make it possible to allow for faulty observations in the analysis of unreplicated factorials or fractional factorials.

References

[1] Shewhart, W.A., Economic Control of Quality of Manufactured Product, D. Van Nostrand Company, Inc. New York, NY (1931).

[2] Taguchi, G., and Y. Wu, Introduction to Off-Line Quality Control, Central Japan Quality Control Association, Nagoya, Japan (1980).

[3] Ishikawa, K., Guide to Quality Control, Asian Productivity Organization, Tokyo (1976).

[4] Finney, D.J., The Fractional Replication of Factorial Arrangements. Annals of Eugenics, 12, 4 291-301 (1945).

[5] Plackett, R.L., and J.P. Burman, Design of Optimal Multifactorial Experiments. Biometrika, 23 305-325 (1946).

[6] Rao, C.R., Factorial Experiments Derivable from Combinatorial Arrangements of Arrays. J. Roy. Statist. Soc., 89 128-140 (1947).

[7] Davies, O.L., The Design and Analysis of Industrial Experiments. London: Oliver and Boyd (1954).

[8] Box, G.E.P., and J.S., Hunter, The 2^{k-p} Fractional Factorial Designs. Technometrics, 3 311-351, 449-458 (1961).