Center for Quality and Productivity Improvement
University of Wisconsin-Madison
610 Walnut Street
Madison, WI 53705

Report No. 3

STUDIES IN QUALITY IMPROVEMENT: ANALYSIS OF UNREPLICATED FACTORIALS ALLOWING FOR POSSIBLY FAULTY OBSERVATIONS*

George E.P. Box and R. Daniel Meyer**

February 1986

---

* This article also appeared as Technical Summary Report #2799, Mathematics Research Center, University of Wisconsin-Madison, 53705.

** Current address: Lubrizol Corporation, 29400 Lakeland Boulevard, Wickliffe, OH 44092.

This research was sponsored by the National Science Foundation Grant No. DMS-8420968, by the United States Army under Contract No. DAAG29-80-C-0041, and by the Vilas Trust of the University of Wisconsin-Madison; and was aided by access to the Statistics Department Research Computer at the University of Wisconsin-Madison.

CENTER FOR QUALITY AND PRODUCTIVITY IMPROVEMENT
UNIVERSITY OF WISCONSIN-MADISON

Report No. 3

STUDIES IN QUALITY IMPROVEMENT; ANALYSIS OF UNREPLICATED FACTORIALS ALLOWING FOR POSSIBLY FAULTY OBSERVATIONS*

George E.P. Box and R. Daniel Meyer**

PRACTICAL SIGNIFICANCE

It is well known that analysis of unreplicated factorial experiments is sensitive to faulty data values. These may be due to oversights in experimental procedure, gross measurement errors, or mistakes in recording the observations. In this article it is shown how the model previously introduced by Box and Meyer (1985) may be extended to detect and allow for such faulty values.

Key words: Unreplicated factorial; normal probability plot; faulty observations; detection, accommodation.

* This article also appeared as Technical Summary Report #2799, Mathematics Research Center, University of Wisconsin-Madison, 53705.
** Current address: Lubrizol Corporation, 29400 Lakeland Boulevard, Wickliffe, OH 44092.

This research was sponsored by the National Science Foundation Grant No. DMS-8420968, by the United States Army under Contract No. DAAG29-80-C-0041, and by the Vilas Trust of the University of Wisconsin-Madison; and was aided by access to the Statistics Department Research Computer at the University of Wisconsin-Madison.

STUDIES IN QUALITY IMPROVEMENT: ANALYSIS OF UNREPLICATED FACTORIALS ALLOWING FOR POSSIBLY FAULTY OBSERVATIONS*

George E.P. Box and R. Daniel Meyer**

1. Introduction

Normal probability plotting of orthogonal contrasts (Daniel (1959)) has become a standard technique for interpreting and criticizing unreplicated factorial and fractional factorial experiments. While the primary objective of the normal plot is to determine which effects are distinguishable from noise, Daniel has pointed out that this is not its only function. Various departures from assumptions may be detected by critical inspection of the normal plot. In particular, the presence of one or more faulty observations is marked by a characteristic pattern among the plotted points.

1.1 An Example

To illustrate, consider the data in Table 1 for a full 2^4 factorial experiment taken from Box and Draper (1986). With the factors denoted by 1, 2, 3 and 4 this shows the design array, the original observations and the estimated effects. A normal plot of the effects is shown in Figure 1. From this plot it will be seen that while the main effects 2 and 3 are largest in absolute magnitude they do not deviate very much from a line drawn through all the remaining points. One would hesitate therefore to conclude on this basis that they were distinguishable from the noise. There is, however, another feature of the plot which bears further consideration. The points falling

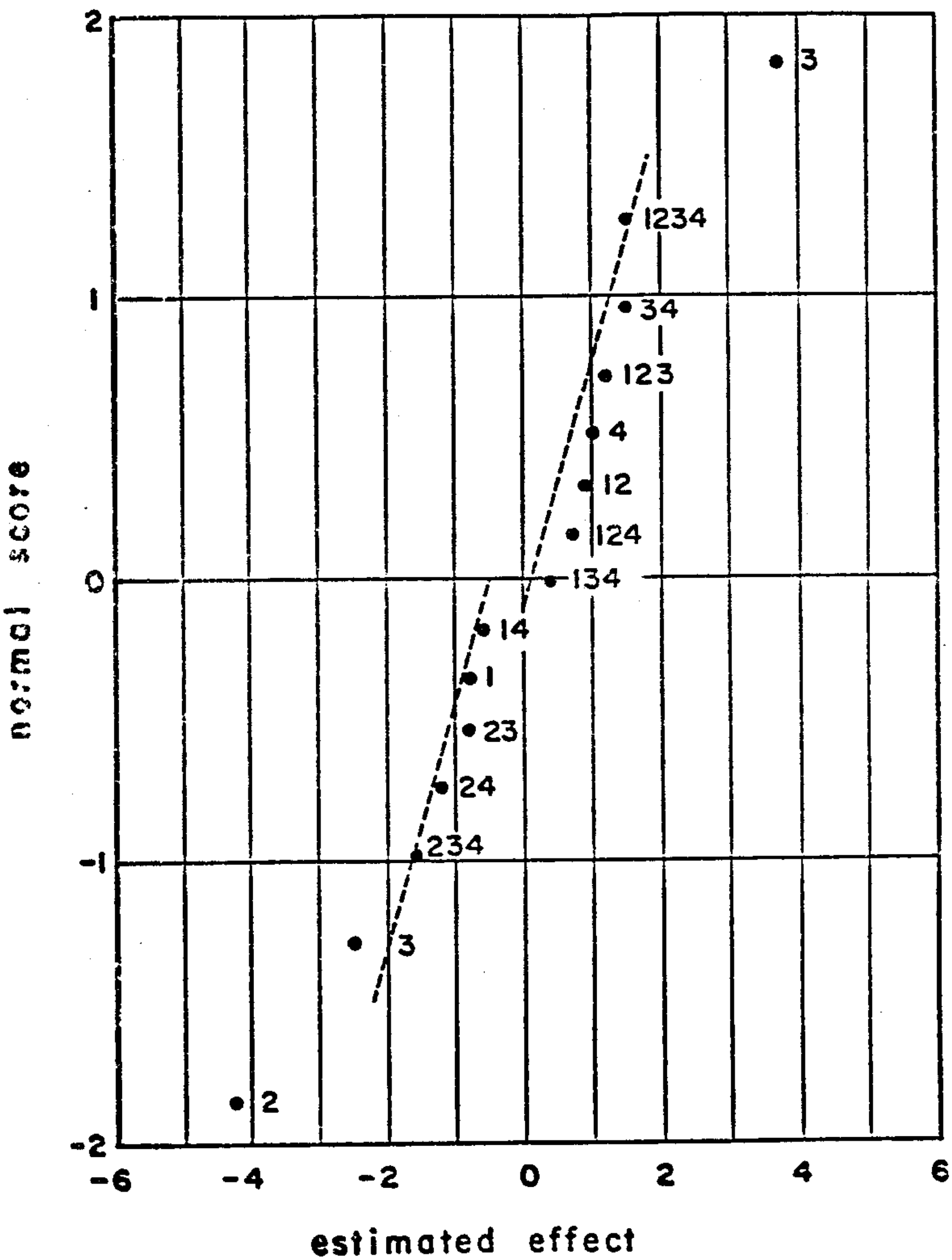
* This article also appeared as Technical Summary Report #2799, Mathematics Research Center, University of Wisconsin--Madison, 53705.

** Current address: Lubrizol Corporation, 29400 Lakeland Boulevard, Wickliffe, OH 44092.

EST.	-0.80	-4.22	3.71	1.01	0.91	-2.49	-0.58	-0.80	-1.18	-1.49	1.20	0.72	0.40	-1.58	1.52	
	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234	
1	-	-	-	-	+	+	+	+	+	+	-	-	-	-	+	47.46
2	+	-	-	-	-	-	-	+	+	+	+	+	+	-	-	49.62
3	-	+	-	-	-	+	+	-	-	+	+	+	-	+	-	43.13
4	+	+	-	-	+	-	-	-	-	+	-	-	+	+	+	46.31
5	-	-	+	-	+	-	+	-	+	-	+	-	+	+	-	51.47
6	+	-	+	-	-	+	-	-	+	-	-	+	-	+	+	48.49
7	-	+	+	-	-	-	+	+	-	-	-	+	+	-	+	49.34
8	+	+	+	-	+	+	-	+	-	-	+	-	-	-	-	46.10
9	-	-	-	+	+	+	-	+	-	-	-	+	+	+	-	46.76
10	+	-	-	+	-	-	+	+	-	-	+	-	-	+	+	48.56
11	-	+	-	+	-	+	-	-	+	-	+	-	+	-	+	44.83
12	+	+	-	+	+	-	+	-	+	-	-	+	-	-	-	44.45
13	-	-	+	+	+	-	-	-	-	+	+	+	-	-	+	59.15
14	+	-	+	+	-	+	+	-	-	+	-	-	+	-	-	51.33
15	-	+	+	+	-	-	-	+	+	+	-	-	-	+	-	47.02
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	47.90

Table 1. The data for the example with the original observations in the right-hand column and the estimated effects along the top.

Figure 1. Normal plot of the estimated effects for the data in Table 1. Points near zero appear to follow two different parallel lines, represented by the dashed lines, rather than one.



near zero appear to follow two different parallel lines rather than one, with negative values on one line and positive values on the other. Daniel points out that such behavior suggests the possibility of a faulty observation. He also suggests (Daniel (1976)) how that observation may be identified using the fact that if a particular observation is biased by, say, a positive amount, those contrasts in which the observation enters positively are shifted to the right, and those contrasts in which the observation enters negatively are shifted to the left. This produces a "gap" such as the one seen in Figure 1. Thus the observation, if it exists, which enters positively in the small positive contrasts and negatively in the small negative contrasts will be under suspicion.

Examination of the design array, Table 1, shows that the row of signs corresponding to $y_{13} = 59.15$ matches the signs of all contrasts save one, suggesting that this thirteenth observation is too large. In practice, such a discovery should lead to reconsideration of the data and in particular of what effects might show up if this observation were appropriately adjusted. In an ongoing investigation it should lead the experimenter to consider any special circumstances which might have surrounded the making of this observation and possibly to a repetition of the observation or of some selected part of the design involving this observation.

2. A More Formal Solution

The remainder of this paper summarizes some recent work on a more formal study which however closely follows the spirit of Daniel's analysis. We emphasize that we regard this work not as replacing this analysis but as perhaps a useful adjunct to it.

The possibility of model inadequancies poses a dilemma for the experimenter rather like that faced by a small country which believes itself in danger of air attack and wonders how it should spend a limited budget on radar apparatus. While some resources should be spent on highly directional radars to monitor with great sensitivity the direction (or directions) regarded as most likely, it might be wise to spend the rest on nondirectional instruments which, while less sensitive, could monitor the whole horizon. Graphical analysis can perform a task like global radar making it possible that the investigator is alerted to contingencies not initially bargained for (see also Box (1980)).

2.1 A model based on the effect sparsity hypothesis

Clearly implied by Daniel's normal plot analysis is a hypothesis of effect sparsity - that most of what is occurring can be accounted for by a few active effects. Suppose X is the $n \times n$ design matrix from which the $n - 1$ usual estimated effects are calculated, and y is the $n \times 1$ vector of observations. If $X_{(a)}$ denotes the columns of X which correspond to active effects $\tau_{(a)}$, then y may be described by the relationship

$$y = X_{(a)}\tau_{(a)} + \varepsilon$$

where ε is the $n \times 1$ vector of normally distributed errors with zero mean and variance σ^2. Let α_1 be the prior probability that an effect is active, and let $a_{(r_1)}$ be the event that a particular set of r_1 of the $n - 1$ effects is active; $X_{(r_1)}$ and $\tau_{(r_1)}$ are the columns of X and the effects corresponding to $a_{(r_1)}$. The prior distribution of each active effect τ, given σ^2, is an independent normal with mean zero and variance $\gamma^2\sigma^2$; the prior distributions of the mean τ_0 and $\log(\sigma)$ are locally uniform (see, e.g., Box and Tiao (1968), (1973)). Thus the posterior probability of the event $a_{(r_1)}$ can be written

$$p(a_{(r_1)}|y) \propto \left(\frac{\alpha_1}{1-\alpha_1}\right)^{r_1} \gamma^{-r_1} \frac{|X'_{(0)}X_{(0)}|^{1/2}}{|\Gamma_{r_1}+X'_{(r_1)}X_{(r_1)}|^{1/2}} \times$$

$$\left[\frac{S(\hat{\tau}_{(r_1)}) + \hat{\tau}'_{(r_1)}\Gamma_{r_1}\hat{\tau}_{(r_1)}}{S(\hat{\tau}_{(0)})}\right]^{-(n-1)/2} ,$$

where

$$\Gamma_{r_1} = \frac{1}{\gamma^2}\begin{bmatrix} 0 & \underset{\sim}{0}' \\ \underset{\sim}{0} & I_{r_1} \end{bmatrix} , \quad I_{r_1} = r_1 \times r_1 \ \text{identity matrix}$$

$$\hat{\tau}_{(r_1)} = \left(\Gamma_{r_1} + X'_{(r_1)}X_{(r_1)}\right)^{-1}X'_{(r_1)}\, y$$

$$S(\hat{\tau}_{(r_1)}) = \left(y-X_{(r_1)}\hat{\tau}_{(r_1)}\right)'\left(y-X_{(r_1)}\hat{\tau}_{(r_1)}\right) .$$

Then, for example, the posterior probability that an effect i is active is

$$p_i = P[\text{effect } i \text{ active} \mid y] = \sum_{(r_1):i \text{ active}} p(a_{(r_1)}|y) .$$

In an earlier paper Box and Meyer (1985) follow such an approach to provide an alternative means of locating active effects and they show that the statistical literature suggests average values for these parameters of $\alpha_1 =$ 0.2,

$\gamma = 2.5$. They furthermore show that the conclusions about which effects are active are usually insensitive to variations over the ranges of values of (α_1,γ) which appear to be actually encountered.

2.2 Faulty observations

To allow for the possibility of faulty observations (Meyer and Box (1985)), we suppose that the errors associated with such values have an inflated error variance $k^2\sigma^2$ $(k>1)$ and occur with some small probability α_2. Thus the error ε is supposed to follow the scale-contaminated normal distribution $(1-\alpha_2)N(0,\sigma^2) + \alpha_2 N(0,k^2\sigma^2)$. Let $a_{(r_1,r_2)}$ be the event that a particular set of r_1 effects are active and a particular set of r_2 observations are faulty $X_{(r_1,r_2)}$ is the matrix of columns and rows of X corresponding to active effects and faulty observations, and $y_{(r_2)}$ the elements of y supposed to be faulty. Then the posterior probability of the event $a_{(r_1,r_2)}$ can be written

$$p(a_{(r_1,r_2)} \mid y) \propto \left(\frac{\alpha_1}{1-\alpha_1}\right)^{r_1} \left(\frac{\alpha_2}{1-\alpha_2}\right)^{r_2} \gamma^{-r_1} k^{-r_2} \times$$

$$\frac{|X'_{(0)}X_{(0)}|^{1/2}}{|\Gamma_{r_1}+X'_{(r_1)}X_{(r_1)}-\varphi X'_{(r_1,r_2)}X_{(r_1,r_2)}|^{1/2}} \left[\frac{S(\hat{\tau}_{(r_1,r_2)})+\hat{\tau}'_{(r_1,r_2)}\Gamma_{r_1}\hat{\tau}_{(r_1,r_2)}}{S(\hat{\tau}_{(0)})}\right]^{-(n-1)/2}$$

where

$$\varphi = 1 - \frac{1}{k^2} ,$$

$$\hat{\tau}_{(r_1,r_2)} = \left(\Gamma_{r_1}+X'_{(r_1)}X_{(r_1)}-\varphi X'_{(r_1,r_2)}X_{(r_1,r_2)}\right)^{-1}\left(X'_{(r_1)}y-\varphi X_{(r_1,r_2)}y_{(r_2)}\right)$$

$$S(\hat{\tau}_{(r_1,r_2)}) = \left(y-X_{(r_1)}\hat{\tau}_{(r_1,r_2)}\right)'\left(y-X_{(r_1)}\hat{\tau}_{(r_1,r_2)}\right) -$$

$$\varphi\left(y_{(r_2)}-X_{(r_1,r_2)}\hat{\tau}_{(r_1,r_2)}\right)'\left(y-X_{(r_1,r_2)}\hat{\tau}_{(r_1,r_2)}\right) .$$

Then, for example, the posterior probability that effect i is active is

$$p_i = \sum_{(r_1,r_2):i \text{ active}} p(a_{(r_1,r_2)}|y)$$

and the posterior probability that observation y_j is faulty is

$$q_j = \sum_{(r_1,r_2):y_j \text{ faulty}} p(a_{(r_1,r_2)}|y) .$$

Computing the $\{p_i\}$ and $\{q_j\}$ over all combinations $a_{(r_1,r_2)}$ will generally not be feasible. Instead we employ the following iterative approximation. We first compute the probabilities $\{p_i\}$ assuming there are no faulty observations. Then temporarily choose the active effects as those with $p_i > P$. The probabilities $\{q_j\}$ are computed with the active effects held fixed by the above choice. The probabilities $\{p_i\}$ are then recomputed, assuming all observations with $q_j > Q$ to have variance $k^2\sigma^2$, and so on. In most cases convergence is achieved in one or two iterations, with $P = Q = 0.5$. Alternatively, P and Q may be chosen after observing the results of the first iteration as a more exploratory approach. As computing power increases, the simultaneous summation over all combinations of active columns and faulty observations will be the most desirable method of computation.

Analysis of data assuming no possibility of faulty observations

Figure 2 shows the posterior probabilities $\{p_i\}$ for the data of table 1 with $\alpha_1 = 0.2$, $\gamma = 2.5$ when we do not allow for faulty observations ($\alpha_2 = 0$). The probabilities suggest, as did the normal plot, that although main effects 2 and 3 are largest in absolute magnitude the evidence for these effects being active is rather slight.

Analysis of data assuming faulty observations possible

Earlier work (Chen and Box (1979)) suggested that this kind of analysis for faulty values is chiefly affected by the parameter $G = \alpha_2 k^{-1}/(1-\alpha_2)$ and that it is fairly insensitive to change. Relying on this work we employ the values $\alpha_2 = 0.05$ and $k = 5$. In practice the analyst may use the computer to experiment somewhat with other values and thus to check on the stability of the conclusions.

In Figure 3, using $\alpha_2 = 0.05$ and $k = 5$, the posterior probabilities $\{p_i\}$ and $\{q_j\}$ are plotted. The value of q_{13} is very close to one, suggesting strongly that observation y_{13} is faulty. The affect on the probabilities $\{p_i\}$ of the automatic downweighting of y_{13} achieved by this

Figure 2. Posterior probabilities $\{p_i\}$ that each effect is active, assuming no faulty observations, with $\alpha_1 = 0.2$, γ -2.5.

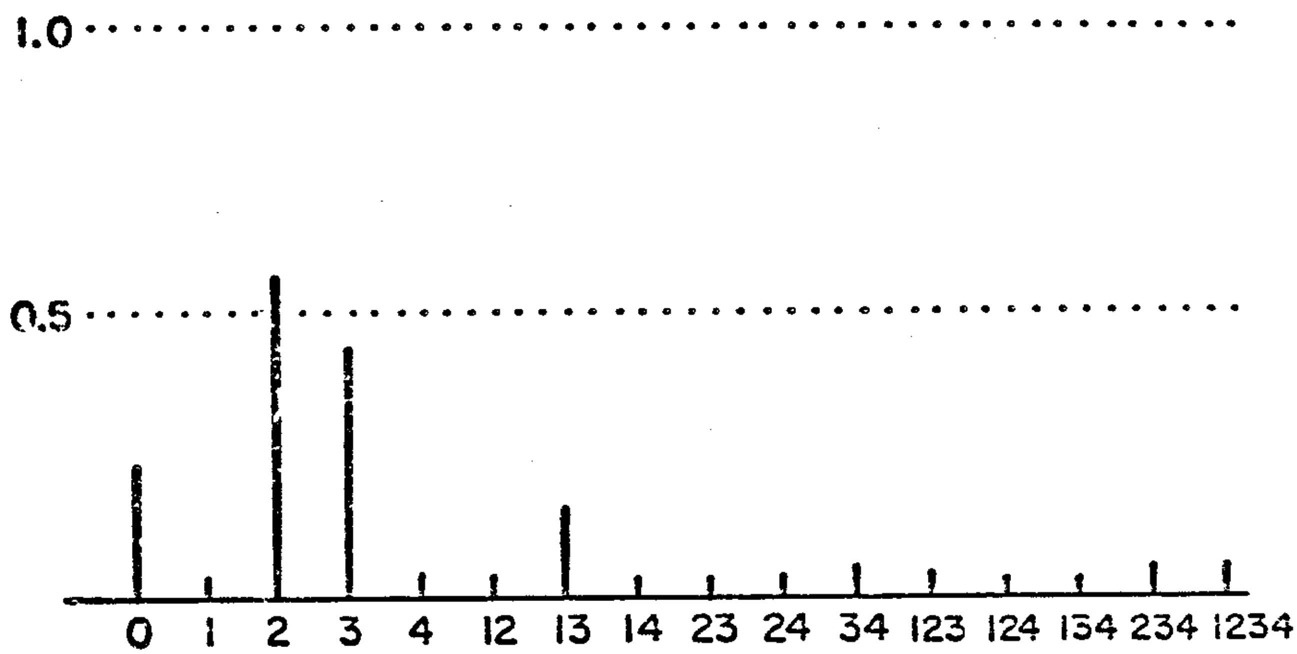
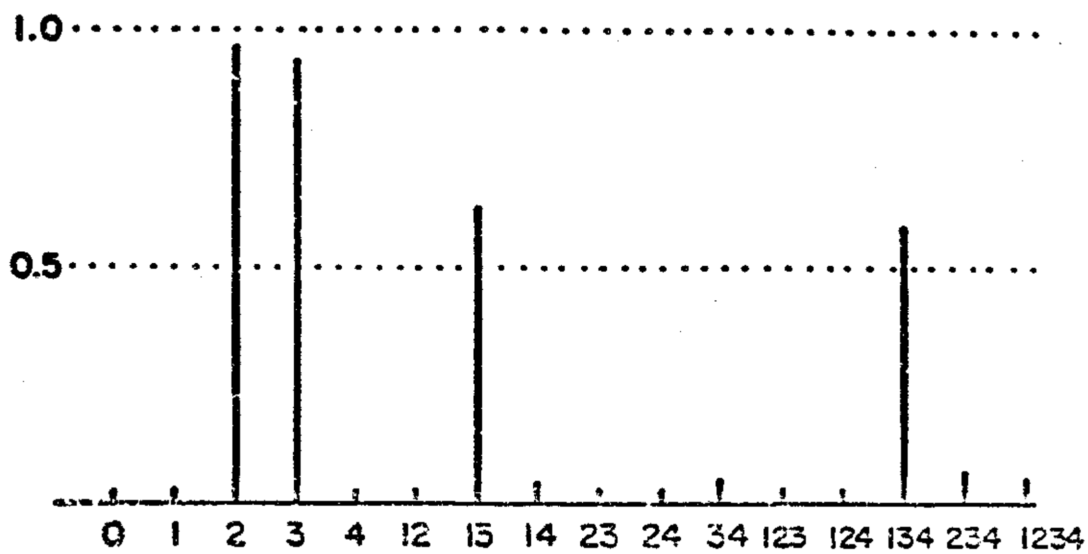
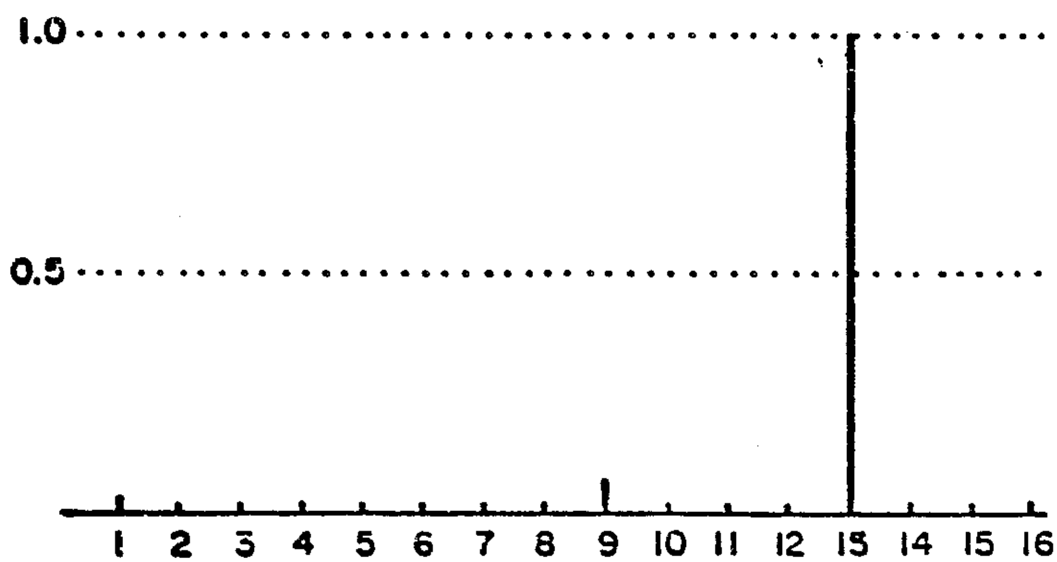


Figure 3. a) Posterior probabilities $\{p_i\}$ that each effect is active, assuming $y_{13} = 59.15$ has variance $k^2\sigma^2$, with $\alpha_1 = 0.2$, $\gamma = 2.5$, $k = 5$.

b) Posterior probabilities $\{q_j\}$ that each observation is faulty, assuming main effects 2 and 3 and interactions 13 and 134 are active, with $\gamma = 2.5$, $\alpha_2 = 0.05$, $k = 5$.





analysis is to make the posterior probabilities for main effects 2 and 3 much closer to one, and the probabilities for interactions 13 and 134 also much larger. The conclusions are similar to those suggested by a Daniel plot with data in which y_{13} has been suitably adjusted. Revised values of the estimated effects are given in Table 2.

Table 2. Posterior probabilities $\{p_i\}$ and Bayesian estimates of effects with $\alpha_1 = 0.2$, $\gamma = 2.5$, $\alpha_2 = 0.05$, $k = 5$.

	assuming no outliers		allowing for outliers	
column	estimated effect	post. prob.	estimated effect	post. prob.
1	-0.79	.029	0.25	.029
2	-4.18	.557	-3.17	.960
3	3.67	.432	2.68	.931
4	1.00	.032	-0.02	.026
5	0.90	.031	-0.13	.026
6	-2.47	.151	-1.49	.628
7	-0.57	.027	0.48	.043
8	-0.79	.029	0.25	.029
9	-1.17	.036	-0.17	.028
10	1.48	.046	0.52	.051
11	1.19	.036	0.19	.028
12	0.71	.028	-0.33	.032
13	0.40	.025	1.42	.587
14	-1.56	.051	-0.62	.069
15	1.50	.048	0.55	.056

Conclusion

We feel that with the increase in computational power now becoming available analysis of the kind we suggest here is a practical possibility. Furthermore experimentation with the parameters α_1, α_2, γ and k can indicate to what extent the conclusions are insensitive to reasonable changes in the probability model. Experience may show that such analysis can usefully augment the highly successful graphical methods of Daniel.

References

Daniel, C. (1959). Use of half-normal plots in interpreting factorial two-level experiments. Technometrics 1, 4, 311-341.

Box, G. E. P. and N. R. Draper (1986). Response Surfaces and Empirical Model Building. John Wiley and Sons, New York.

Daniel, C. (1976). Applications of Statistics to Industrial Experimentation. John Wiley and Sons, New York.

Box, G. E. P. (1980). Sampling and Bayes' inference in scientific modelling and robustness. J. Roy. Statis. Soc. A 143, 4, 383-430.

Box, G. E. P. and G. C. Tiao (1968). A Bayesian approach to some outlier problems. Biometrika 55, 1, 119-129.

Box, G. E. P. and G. C. Tiao (1973). Bayesian Inference in Statistical Analysis. Addison-Wesley, Reading, Massachusetts.

Box, G. E. P. and R. D. Meyer (1985). Studies in quality improvement II: an analysis for unreplicated fractional factorials.

Meyer, R. D. and G. E. P. Box (1985). Detection and accomodation of bad values in fractional factorials. University of Wisconsin-Madison, Mathematics Research Center Technical Summary Report.

Chen, G. and G. E. P. Box (1979). Further study of robustification via a Bayesian approach. University of Wisconsin-Madison, Mathematics Research Center Technical Summary Report #1998.

Acknowledgement

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041, by the National Science Foundation Grant No. DMS-8420968, and by the Vilas Trust of the University of Wisconsin--Madison; and aided by access to the Statistics Department Research Computer at the University of Wisconsin--Madison.