

EXPLOITING THE INHERENT STRUCTURE IN ROBUST PARAMETER DESIGN EXPERIMENTS

J. Bérubé and V. N. Nair

University of Michigan

Abstract: Robust parameter design methods have been used successfully in industry for some time. Despite this, there has been some skepticism in the statistical literature about the feasibility of conducting industrial experiments to estimate both location and dispersion effects. It has been claimed that a large experimental run size is needed to estimate dispersion effects and that such experiments are not practical in industry where the emphasis is on studying many factors simultaneously using highly fractionated designs. We show in this paper that this misconception arises from the fact that the commonly used methods of analysis ignore the basic structure in parameter design studies and hence are unnecessarily inefficient. We consider different models and methods of analysis and quantify the gains to be made from exploiting the inherent structure in parameter design studies. The consequences of these conclusions for the planning of such studies are also discussed.

Key words and phrases: Design of experiments, dispersion effects, quality improvement, variation reduction.

1. Introduction and Overview

Robust parameter design, proposed by Taguchi (1991), is intended to be a cost-effective approach for improving quality during the design and development of products and processes. The goal of parameter design studies is to choose the settings of the control factors (parameters) so that the performance of a system (product or process) is insensitive to variation in uncontrollable “noise” variables. More specifically, consider the so-called “static” problem, and let $Y = f(x; u)$, where Y denotes the system response, x denotes the setting of system parameters (control factors) and u denotes the uncontrollable or difficult-to-control sources of noise variation. The goal of parameter design is to select the settings of system parameters or control factors x appropriately so that the system response is close to the intended target value and is insensitive to variation in the noise variables.

If the transfer function $f(\cdot)$ is known and easy to evaluate, this can be treated as a numerical optimization problem. In most practical situations, however, $f(\cdot)$ is unknown or hard to evaluate. In this case, statistically designed experiments and data analysis methods have been used to implement parameter design. The commonly used setup is the product array where the control factors are varied

according to a suitably chosen experimental design (control array) and at each setting of the control array, the pre-identified noise variables are systematically varied according to a noise array. The particular data analysis method used will depend on the transfer function $f(\cdot)$. The most commonly considered situation is where the data, possibly after a suitable transformation, follow the location-scale model

$$Y_{ij} = \mu(x_i) + \phi(x_i)\epsilon_{ij}, \quad (1)$$

where Y_{ij} and ϵ_{ij} denote, respectively, the system response and the effect of the noise variables corresponding to the i th row of the control array and j th row of the noise array, and x_i denotes the settings of the control factors at the i th row of the control array, for $i = 1, \dots, n$ and $j = 1, \dots, J$. Here $\mu(x_i)$ is the location parameter and $\phi(x_i)$ is the scale parameter, both of which depend on the control factor settings. For this model, the goal of parameter design experiments can be restated as follows: choose the control factor settings to make the variability (scale parameter) small and the mean response (location parameter) close to target. The popular approach for analyzing data under this model is to first estimate $\mu(x_i)$ by $\bar{Y}_i = \sum_{j=1}^J Y_{ij}/J$, the mean response at the i th control-run and estimate $\phi^2(x_i)$ by

$$S_i^2 = \frac{1}{J-1} \sum_{j=1}^J (Y_{ij} - \bar{Y}_i)^2, \quad (2)$$

the corresponding sample variance. The important location and dispersion effects are then identified by fitting a linear model to the \bar{Y}_i 's and a log-linear model to the S_i^2 's as a function of the control factors and identifying the appropriate control factor settings. Taguchi (1991) actually recommends another performance measure for assessing dispersion that he generically refers to as a signal-to-noise ratio. However, this is essentially equivalent to taking a log-transformation of the original data and analyzing the variance of the transformed data (see, for example, Box (1988); León, Shoemaker and Kacker (1987); and Nair and Pregibon (1986)). As such, this can be viewed as a special case of the above formulation. There are other approaches for data analysis, including the so-called response-model approach, which we will discuss later.

The basic engineering principles and the statistical issues involved in implementing parameter design have generated considerable attention among statisticians and quality professionals in industry over the last decade or so. See, for example, the panel discussion and extensive references in Nair (1992) and subsequent discussions and references in Lunani, Nair and Wasserman (1997) and Miller and Wu (1996). There have been many successful applications of parameter design studies in industry, some of which have been documented in the annual case-studies symposia organized by the American Suppliers Institute.

Despite these, there has been some skepticism in the statistical literature as to the feasibility of conducting experiments for estimating both location and dispersion effects. Multifactor experimental studies have become popular in recent years in industry mainly because practitioners have finally realized that they can study many factors with small run sizes using highly fractioned experiments. It has been argued that parameter design experiments which require noise “replications” to study both location and dispersion effects would require unduly large runs and hence are not practical. For example, Gunter, in his discussion of Box (1988), remarks, “The statistical consequences of Taguchi’s engineering insight is that . . . modeling the response variance – not just the mean – may be vital to achieve high quality and reduced cost.” He goes on to say “Appearances can be deceiving, however. Although such an approach may *appear* to have promise, there are fundamental practical constraints that make the Taguchi strategy difficult or impossible to apply . . . one needs a lot more information to see a change in spread than to see a change in location. The crucial question is, Can we generally afford experimentation that gives us this kind of information. My answer is that I doubt it . . .” He presents some numerical evidence to support his argument that one needs an excessive number of experimental runs to study dispersion effects. Similar comments have also been made by others (see Carroll and Ruppert’s discussion of Box (1988)).

Is such pessimism justified? Are parameter design experiments infeasible in practice? We show in this paper that this misconception arises from the fact that the commonly used methods of analysis do not exploit the inherent structure in parameter design experiments and hence are unnecessarily inefficient. An important element of parameter design experiments is that the major sources of variation (noise variables) are identified up front and they are systematically varied through a noise array. In the product-array setup, the same noise array is repeated across all the control runs. However, many of the commonly discussed methods of analysis in Box (1988), Nair and Pregibon (1986), Taguchi (1991), and others do not take this structure into account.

To be specific, consider the simple case where a single noise has been identified as being important and it is varied at J levels in the parameter design experiment. Then, instead of equation (1), we in fact have the situation

$$Y_{ij} = \mu(x_i) + \phi(x_i)[z_j + \delta_{ij}], \quad (3)$$

where z_j is the j th level of the noise variable that has been identified up front and varied systematically and δ_{ij} represents the remaining sources of variation. This is analogous to the usual blocking setup where the z_j ’s are the J settings of a blocking factor. The critical difference in parameter design is that one is interested in the interactions between the blocking factors (which are the noise

variables in this set-up) and design factors, and in fact it is these interactions that provide us with the opportunity to make the system robust. This is reflected in the above model by the fact that we let $\phi(\cdot)$ depend on the control factors, while traditionally $\phi(\cdot)$ was assumed to be constant.

As the analogy with blocking suggests, the efficiency under this design can be considerably higher than one under complete randomization. In other words, the analysis based on (1) assumes the ϵ_{ij} 's are independent and vary freely from run to run. However, as we can see from (3), only the δ_{ij} 's vary freely while the same z_j 's are repeated across the control runs. Thus, the relevant variation for inference purposes is σ_δ^2 , the variance of the δ_{ij} 's, and not σ_ϵ^2 , the variance of all the noise variables. In particular, if we have captured most of the large sources of variation up front through the z_j 's, then the experiment is likely to be quite a bit more efficient than perceived. In subsequent sections, we will quantify this by considering appropriate methods of analysis for different models and designs. However, the result for the simple case with a single noise variable captures the essence of the conclusion very nicely, so we state it here.

Let σ_P^2 be the (perceived) variance of the estimated dispersion effects calculated under the model in (1) which assumes (incorrectly) that the ϵ_{ij} 's vary freely across control runs. Similarly, let σ_A^2 be the (actual) variance of the estimated dispersion effects under the model in (3) which exploits the structure in parameter design experiments where the same noise array is repeated across control runs. Note that the values of the noise variables z_j 's can be fixed or they can be chosen randomly from the noise distribution. In either case, under the assumption of normally distributed noise variables, we show in Section 2 that

$$\sigma_A^2/\sigma_P^2 \approx (1 - \rho^2), \quad (4)$$

where

$$\rho = \frac{\sigma_z^2}{\sigma_z^2 + \sigma_\delta^2}, \quad (5)$$

the proportion of the total noise variance that can be attributed to the pre-identified source z_j .

We can see from (4) that σ_A^2 , the actual variance of the dispersion effects, can be considerably smaller than σ_P^2 , the perceived variance based on an analysis that assumes that the errors vary independently. Of course, it should be noted that there are some parameter design studies where none of the noise variables are varied systematically. For such situations, the assumptions of independent noise variables is indeed appropriate. However, in most parameter design studies, the same noise array is repeated across the control runs. In such cases, the analyses discussed in Taguchi (1991), Nair and Pregibon (1986), Box (1988) and others that do not take this structure into account will lead to an over-inflation of the

variance of the dispersion effects. Similar conclusions also hold for the location effects as discussed in Section 4.

It should also be clear from this that we do not need as many experimental runs to detect the dispersion effects as noted in Gunter (1988) who did his efficiency calculations under the assumption that the “replications” in the noise array are independent across the control array runs. This result has important implications for the planning of parameter design experiments as well. We can see that the efficiency of parameter design experiments will be increased by making ρ close to one. To do this, we must identify the nature of the noise variation and incorporate as many of the important noise variables as possible into the parameter design experiment. It is very important to note that, from this point of view, we should include all the important sources of variation in the experiment, and not just those that are expected to “interact” with the control factors.

There are various ways one might select and vary the levels of the noise variables. If there are several important noise variables, we could consider them separately or as a compound noise variable. If the distribution of the noise variables are known, the levels of the noise variables can be fixed at certain desired settings or they can be picked at random from the corresponding noise distribution(s). As an example of the latter case, suppose one wants to study manufacturing variation attributable to machine-to-machine variability and operator-to-operator variability. One can then select a random sample of machine/operator combinations and repeat the runs in the control array for each of the machine/operator combinations. There are also many different models one can entertain for estimating the dispersion effects. In subsequent sections, we will consider different models and designs and study the corresponding estimation methods. In each case, we will quantify the gains to be made by appropriately identifying the important noise variables up front during the planning stage and varying them systematically in the parameter design experiment.

We note that related results have also been obtained in Box and Jones (1992) and Steinberg and Bursztyn (1998). Box and Jones (1992) consider different types of split-plot experiments for robust product design and discuss the efficiency of such experiments compared to a completely randomized experiment. Steinberg and Bursztyn (1998) study the power in detecting important effects when the settings of the noise variables are fixed. These technical results differ from ours, but the overall conclusions are qualitatively similar.

2. Single Noise Variable

We first consider the case where there is a single noise variable with J levels. This could arise even in situations with several sources of variation but a single, compound noise variable is used to capture their overall effect.

Suppose the true dispersion effects in model (1) follow, at least approximately, the log-linear model $\log \phi(x_i) = x_i \Phi$, where x_i is the row vector corresponding to the i th row of the control array, $\Phi = (\Phi_1, \dots, \Phi_L)'$ and so for $l \geq 2$, the Φ_l 's correspond to the dispersion effects of the various control factors. We can then fit to the sample variances S_i^2 the log-linear model

$$\log S_i^2 = x_i \Phi + \nu_i \quad (6)$$

and obtain least-squares estimates of the dispersion effects in order to identify the important dispersion effects. To obtain the variance of the estimated dispersion effects, we consider separately the cases where z_j 's, the levels of the noise variable, are chosen randomly or fixed at pre-selected levels.

2.1. Noise settings chosen randomly

Consider first the case where the J settings of the single noise factor are randomly sampled from the noise distribution. The machine/operator set-up discussed earlier is an example of this. The sample variance in (2) can be re-expressed in terms of the z_j 's and δ_{ij} 's as

$$S_i^2 = \frac{\phi^2(x_i)}{J-1} \sum_{j=1}^J (\tilde{z}_j + \tilde{\delta}_{ij})^2,$$

where \tilde{z}_j and $\tilde{\delta}_{ij}$ represent the appropriately centered quantities. We will assume throughout that the δ_{ij} 's are independent and distributed as $N(0, \sigma_\delta^2)$. Further, suppose the z_j 's are also independent $N(0, \sigma_z^2)$ random variables and independent of the δ_{ij} 's. Then, the S_i^2 's have a $\frac{\phi^2(x_i)}{J-1}(\sigma_z^2 + \sigma_\delta^2)\chi_{(J-1)}^2$ distribution with mean and variance given respectively by $E(S_i^2) = \phi^2(x_i)(\sigma_z^2 + \sigma_\delta^2)$ and $\text{Var}(S_i^2) = \frac{2\phi^4(x_i)}{J-1}(\sigma_z^2 + \sigma_\delta^2)^2$. Since the same J settings of the noise factors are repeated across the control runs, the sample variances are not independent from run to run but are correlated with $\text{Covar}(S_i^2, S_{i'}^2) = \frac{2}{J-1}\sigma_z^4\phi^2(x_i)\phi^2(x_{i'})$.

So, if we fit the log-linear model $\log S_i^2 = x_i \Phi + \nu_i$ in (6), the ν_i 's are still marginally distributed as $\log \chi_{(J-1)}^2$ but are now correlated from run to run. We can approximate the variance-covariance structure of the ν_i 's through a first-order approximation and obtain

$$\text{Var}(\log S_i^2) \approx \frac{\text{Var}(S_i^2)}{E(S_i^2)^2} = \frac{2\phi^4(x_i)(J-1)^{-1}(\sigma_z^2 + \sigma_\delta^2)^2}{\phi^4(x_i)(\sigma_z^2 + \sigma_\delta^2)^2} = \frac{2}{J-1}$$

and

$$\begin{aligned} \text{Covar}(\log S_i^2, \log S_{i'}^2) &\approx \frac{\text{Covar}(S_i^2, S_{i'}^2)}{E(S_i^2)E(S_{i'}^2)} = \frac{2(J-1)^{-1}\sigma_z^4\phi^2(x_i)\phi^2(x_{i'})}{\phi^2(x_i)\phi^2(x_{i'})(\sigma_z^2 + \sigma_\delta^2)^2} \\ &= \frac{2}{(J-1)} \left(\frac{\sigma_z^2}{\sigma_z^2 + \sigma_\delta^2} \right)^2 = \frac{2}{(J-1)} \rho^2, \end{aligned}$$

where the covariance is a function of ρ in (5). The approximate variance-covariance matrix of the ν_i 's can now be written as $\mathbf{V} = \frac{2}{(J-1)}[(1-\rho^2)\mathbf{I} + \rho^2\mathbf{J}]$ where \mathbf{I} is the identity matrix and \mathbf{J} is the matrix with all elements equal to 1. Note that the variance-covariance matrix does not depend on the control factor settings.

Consider now the variance of the dispersion effects obtained from least-squares estimation of Φ from the model $\log S^2 = \mathbf{X}\Phi + \nu$, where \mathbf{X} is the control array, *which will be assumed to be orthogonal throughout*. (Since the ν_i 's are correlated, one should actually use generalized least-squares to estimate the dispersion effects. We will come back to this point later in this section.) The approximate variance of the least-squares estimator of the dispersion effects $\hat{\Phi}$ is given by

$$\begin{aligned} \text{Var}(\hat{\Phi}) &\approx (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{V}\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1} \\ &= \frac{2}{n^2(J-1)}\mathbf{X}'[(1-\rho^2)\mathbf{I} + \rho^2\mathbf{J}]\mathbf{X} \\ &= \frac{2(1-\rho^2)}{n(J-1)}\mathbf{I} + \frac{2\rho^2}{(J-1)} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}. \end{aligned}$$

So, for $l \geq 2$,

$$\text{Var}(\hat{\Phi}_l) \approx \frac{2}{n(J-1)}(1-\rho^2). \quad (7)$$

Thus, compared with the usual variance of $\frac{2}{n(J-1)}$ computed assuming the ν_i 's are independent (i.e., the error terms ϵ_{ij} in (1) are assumed to be independent), the estimators are more efficient by a factor of $1/(1-\rho^2)$, thus proving the result noted in Section 1 for the random z_j 's case.

The above calculations were done under the assumption that the various noise factors all followed a normal distribution. Further, the first-order Taylor series approximation is only valid when J is large. We conducted a limited simulation study to assess the sensitivity to these assumptions. Figure 1 shows the result of this study. The solid line in Figure 1 corresponds to the limiting case where J is large and the distributions are all normal so that the ratio of the variances is $(1-\rho^2)$. The four other cases correspond to: (i) all error distributions normal, $J = 10$, (ii) all error distributions normal, $J = 4$, (iii) distribution of z_j 's is t_3 and of δ_{ij} 's is normal and (iv) distribution of z_j 's is t_5 and of δ_{ij} 's is normal. As one can see from Figure 1, the conclusions are still qualitatively the same. In all of these cases, the variances of the estimated dispersion effects taking into account the structure of the parameter design was smaller by a factor close to (if not smaller than) $(1-\rho^2)$.

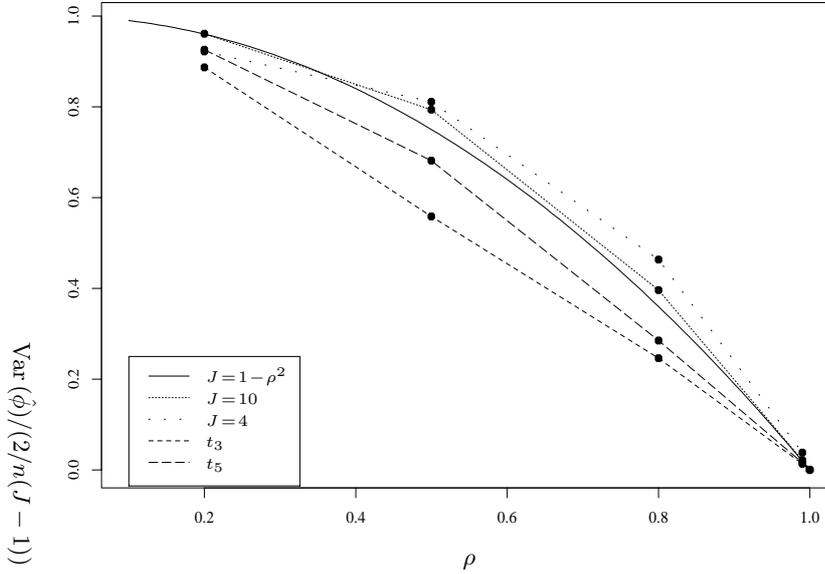


Figure 1. Ratio of the actual variance versus the perceived variance as a function of ρ for several different situations.

As noted earlier, the error terms ν_i 's are correlated, so the use of ordinary least-squares can be questioned. However, when the variance-covariance matrix is of the form $\mathbf{V} = a\mathbf{I} + b\mathbf{J}$ (where a and b are such that \mathbf{V} is non-negative definite) and the design matrix is orthogonal with the first column being a column of one's, it can be readily verified that the generalized least-squares estimators of the Φ_l 's coincide with the ordinary least squares estimators for $l \geq 2$. Thus, in this case, ordinary least squares estimators of the dispersion effects are efficient.

2.2. Noise settings fixed

We now consider the case where the levels of the noise factor are fixed at some prespecified levels. Suppose the noise array consists only of the fixed settings of z_j , $j = 1, \dots, J$, and the values are chosen such that $\sum_{j=1}^J (z_j - \bar{z})^2 / (J - 1) \equiv h^2 \sigma_z^2$ for some fixed h .

Again, the sample variances can be written as

$$S_i^2 = \frac{\phi^2(x_i)}{J-1} \sum_{j=1}^J (\tilde{z}_j + \tilde{\delta}_{ij})^2,$$

but since the z_j 's are fixed, the S_i^2 's now have a non-central χ^2 distribution

$$S_i^2 \sim \frac{\phi^2(x_i) \sigma_\delta^2}{J-1} \chi_{(J-1)}^2(\lambda)$$

with non centrality parameter

$$\lambda = \frac{\sum_{j=1}^J \tilde{z}_j^2}{\sigma_\delta^2} = \frac{(J-1)h^2\sigma_z^2}{\sigma_\delta^2}.$$

Their mean and variance are given respectively by

$$\mathbb{E}(S_i^2) = \phi^2(x_i)\sigma_\delta^2 \left[1 + \frac{h^2\sigma_z^2}{\sigma_\delta^2} \right]$$

and

$$\text{Var}(S_i^2) = \frac{2}{(J-1)}\phi^4(x_i)\sigma_\delta^4 \left[1 + \frac{2h^2\sigma_z^2}{\sigma_\delta^2} \right].$$

Since the noise levels are fixed, the sample variances from run to run are independent.

When we fit the log-linear model $\log S_i^2 = x_i\Phi + \nu_i$ in (6), the ν_i 's are now independent across i and are distributed as the log of a non-central χ^2 . Using a first-order approximation as before, the variance of the ν_i 's is approximated as

$$\begin{aligned} \text{Var}(\nu_i) &\approx \frac{\text{Var}(S_i^2)}{\mathbb{E}(S_i^2)^2} \\ &= \frac{2}{J-1} \left(1 + \frac{2h^2\sigma_z^2}{\sigma_\delta^2} \right) / \left(1 + \frac{h^2\sigma_z^2}{\sigma_\delta^2} \right)^2. \end{aligned}$$

By defining

$$\rho(h) = \frac{h^2\sigma_z^2}{(\sigma_\delta^2 + h^2\sigma_z^2)},$$

the above variance can be written as

$$\text{Var}(\nu_i) \approx \frac{2}{J-1}(1 - \rho^2(h)).$$

This variance does not depend on the setting of the control factors, and we have $\text{Var}(\nu) \approx \frac{2}{J-1}(1 - \rho^2(h))\mathbf{I}$. The variance of the least-squares estimates of the dispersion effects $\hat{\Phi}_l$, for $l \geq 2$, is now approximately

$$\text{Var}(\hat{\Phi}_l) \approx \frac{2}{n(J-1)}(1 - \rho^2(h)). \quad (8)$$

If we choose the levels of the noise factors with a value of $h = 1$, then $\rho(h) = \sigma_z^2/(\sigma_\delta^2 + \sigma_z^2) = \rho$ in (5), and (8) reduces to (7) in Section 2.1. In this case, the variance of the dispersion effects for the fixed noise levels set-up is the same as the one in the random noise levels case, thus proving the result noted in Section 1.

Note in the above that as we increase h , i.e., as we choose the levels of the noise factors further apart, $\rho^2(h)$ increases and so the variance of the estimated dispersion effects decreases. There is of course a physical limit to how extreme the levels of the noise factors can be selected. Also, the model in (6) might no longer be accurate when h is too large. (See Steinberg and Bursztyn (1998) for a similar point.)

2.3. Least-squares analysis

We have restricted attention thus far to an analysis of dispersion effects based on fitting a log-linear model to the sample variances S_i^2 's. When the z_j 's are fixed, an alternative approach is to use least-squares directly and fit the model in (3). This has been referred to as the response-model approach in Shoemaker, Tsui and Wu (1991), although they consider a different model from (3). There are in fact a number of variations to the model in (3) that one could consider, and we will return to this issue more fully in a later section. For now, in order to keep the comparison with the previous analysis meaningful, we restrict attention to the model in (3) with a log-linear structure for $\phi(\cdot)$. Further, instead of simultaneous estimation of location and dispersion effects (as done in Shoemaker et al. (1991) and elsewhere) we will decouple the problem of estimating location and dispersion effects. Specifically, the dispersion effects are estimated in two stages, first estimating $\phi(x_i)$'s in (3) using least-squares and then fitting a log-linear model to the $\hat{\phi}(x_i)$'s. (Note that there is a technical issue here if $\hat{\phi}(x_i)$ is negative and an appropriate modification such as $\log(\hat{\phi}^+(x_i) + \delta)$ should be used to handle such cases). We assume as in Section 2.2 that the z_j 's have been chosen such that $\sum_{j=1}^J (z_j - \bar{z})^2 / (J - 1) \equiv h^2 \sigma_z^2$.

The ordinary least-squares estimator of $\phi(x_i)$, i, \dots, n , is given by

$$\hat{\phi}(x_i) = \frac{\sum_{j=1}^J \tilde{z}_j \tilde{y}_{ij}}{\sum_{j=1}^J \tilde{z}_j^2}.$$

The variance of this estimator is then $\text{Var}(\hat{\phi}(x_i)) = \frac{\phi^2(x_i) \sigma_\delta^2}{(J-1)h^2 \sigma_z^2}$. Now, we fit the log-linear model $\log \hat{\phi}(x_i) = x_i \Phi + \nu_i$ in order to estimate the dispersion effects associated with the control factors. Using a first-order approximation as we did before, we can get an approximate expression for the variance of the error term in the log-linear model,

$$\text{Var}(\log \hat{\phi}(x_i)) \approx \frac{\phi^2(x_i) \sigma_\delta^2 / (J-1)h^2 \sigma_z^2}{\phi^2(x_i)} = \frac{\sigma_\delta^2}{(J-1)h^2 \sigma_z^2}.$$

Note that with this formulation, the variance expression is again independent of the settings of the control factors.

The approximate variance of the least-squares estimates of the dispersion effects $\hat{\Phi}_l$ is then given by

$$\text{Var}(\hat{\Phi}_l) \approx \frac{\sigma_\delta^2}{n(J-1)h^2\sigma_z^2},$$

which can be rewritten in terms of ρ in (5) as

$$\text{Var}(\hat{\Phi}_l) \approx \frac{1}{n(J-1)h^2} \frac{(1-\rho)}{\rho}.$$

Although the relationship is different, the overall conclusion is the same in this case as well. The larger ρ is, the higher the efficiency in estimating the dispersion effects and in detecting factors with important effects. It is interesting to see from Figure 2 that this least-squares approach (with $h = 1$) does not dominate the analysis based on S_i^2 in terms of efficiency.

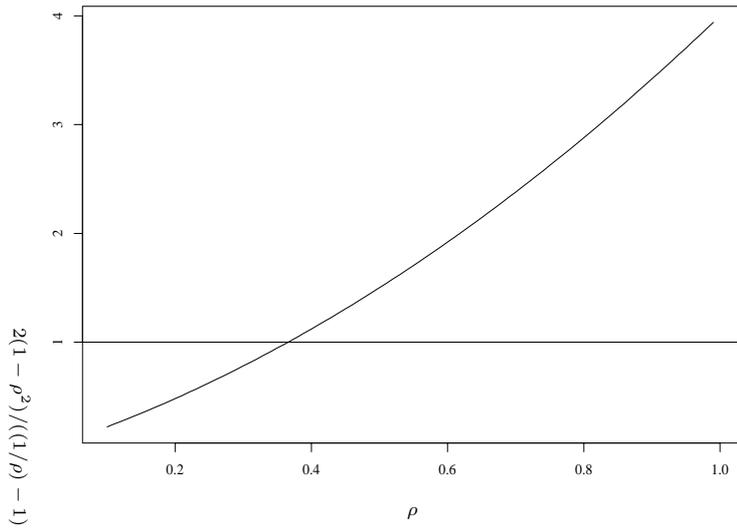


Figure 2. Relative efficiency of regression approach with $h = 1$ with respect to modeling $\log S_i^2$.

3. Multiple Noise Variables

In practice, there are several noise variables that are varied systematically in the noise array. There are many possible models that one can entertain in the situation with multiple noise variables. The model below is reasonably general in that it includes several commonly considered ones as special cases.

Suppose that the data, possibly after a suitable transformation, follow the model

$$Y_{ij} = \mu(x_i) + \phi_1(x_i)z_{1j} + \cdots + \phi_K(x_i)z_{Kj} + \theta(x_i)\delta_{ij}. \quad (9)$$

Here z_{kj} denotes the setting of the k th noise factor for the j th row of the noise array, $j = 1, \dots, J$ and $k = 1, \dots, K$. Further, $\phi_k(x_i)$ measures the dispersion effects that corresponds to the k th noise factor and the $\theta(x_i)$ measures the dispersion effects that correspond to unidentified/uncontrolled noise variables δ_{ij} 's.

3.1. A special case

First, we discuss a special case of this model where the dispersion effects associated with the various noise factors are the same, i.e., $\phi_1(x_i) = \cdots = \phi_K(x_i) = \theta(x_i)$ and equal to $\phi(x_i)$. In this case, (9) reduces to

$$Y_{ij} = \mu(x_i) + \phi(x_i) \left[\sum_{k=1}^K z_{kj} + \delta_{ij} \right], \quad (10)$$

which is essentially (3). It is important to note that Taguchi's SN-ratio analysis in the presence of multiple noise variables can be justified only under this restricted model. We elaborate on this in the next subsection.

Under this special model, the efficiency results for estimating dispersion effects with fixed settings of noise factors can be shown to be the same as those in Section 2.2. The situation with randomly chosen noise levels is more complex as it depends on the actual sampling mechanism for the noise variables. We consider a specific set-up here. Suppose we select two levels randomly for each of the K noise factors. Recall the discussion with machines and operators as a concrete example of this set-up. We use an orthogonal (fractional) factorial design for the noise array with run size J . This will lead to each level of a noise factor being repeated $R = J/2$ times. For example, if we have 3 factors each at two levels and we use a 2^{3-1} design, then we have a total of $J = 4$ runs, and the two levels of a factor each appear twice.

The sample variances can now be expressed as

$$\begin{aligned} S_i^2 &= \frac{\phi^2(x_i)}{J-1} \sum_{j=1}^J \left[\sum_{k=1}^K \tilde{z}_{kj} + \tilde{\delta}_{ij} \right]^2 \\ &= \frac{\phi^2(x_i)}{J-1} \left[R \sum_{j=1}^2 \sum_{k=1}^K \tilde{z}_{kj}^2 + 2 \sum_{j=1}^J \sum_{k=1}^K \tilde{z}_{kj} \tilde{\delta}_{ij} + \sum_{j=1}^J \tilde{\delta}_{ij}^2 \right], \end{aligned}$$

where the \tilde{z}_{kj} 's and the $\tilde{\delta}_{ij}$'s are the centered quantities. The second equality above follows from the fact that, since each factor has only two levels, the centered

z_{kj} 's take on values $\pm a_k$ for some constant a_k , $k = 1, \dots, K$. This together with the orthogonality of the design matrix ensures that the cross-product terms involving the \tilde{z}_{kj} 's are zero.

The S_i^2 's can be seen to be a quadratic form in normal variables, hence its distribution can be obtained as a linear combination of chi-squared random variables. For our purposes, we are only interested in the first two moments. To get this, note that the first term in the square brackets is distributed as a sum of K independent chi-squares, $R\sigma_k^2\chi_1^2$, and the last term is distributed as $\sigma_\delta^2\chi_{(J-1)}^2$. The middle term is not a χ^2 ; it has a mean of 0 and variance $4R\sigma_\delta^2\sum_{k=1}^K\sigma_k^2$. Moreover, under the normality assumption and the special structure here, it is seen to be uncorrelated with the first and third terms. Thus, we obtain the mean and variance of S_i^2 as

$$\begin{aligned} \mathbb{E}(S_i^2) &= \frac{\phi^2(x_i)}{J-1} \left[R \sum_{k=1}^K \sigma_k^2 + (J-1)\sigma_\delta^2 \right] \\ \text{Var}(S_i^2) &= \frac{\phi^4(x_i)}{(J-1)^2} \left[2R^2 \sum_{k=1}^K \sigma_k^4 + 4R\sigma_\delta^2 \sum_{k=1}^K \sigma_k^2 + 2(J-1)\sigma_\delta^4 \right]. \end{aligned}$$

Since the same levels of the noise factors are repeated for each control run, the sample variances are correlated from run to run with covariance given by

$$\text{Covar}(S_i^2, S_{i'}^2) = \frac{2}{(J-1)^2} R^2 \phi^2(x_i) \phi^2(x_{i'}) \sum_{k=1}^K \sigma_k^4.$$

When we fit the log-linear model $\log S_i^2 = x_i\Phi + \nu_i$ in (6), the ν_i 's are therefore correlated from run to run. Again, using a first-order approximation, the variance and covariance of the error term ν_i can be approximated respectively as

$$\text{Var}(\log S_i^2) \approx \frac{2R^2 \sum_{k=1}^K \sigma_k^4 + 4R\sigma_\delta^2 \sum_{k=1}^K \sigma_k^2 + 2(J-1)\sigma_\delta^4}{\left[R \sum_{k=1}^K \sigma_k^2 + (J-1)\sigma_\delta^2 \right]^2}$$

and

$$\text{Covar}(\log S_i^2, \log S_{i'}^2) \approx \frac{2R^2 \sum_{k=1}^K \sigma_k^4}{\left[R \sum_{k=1}^K \sigma_k^2 + (J-1)\sigma_\delta^2 \right]^2}.$$

Again, these expressions do not depend on the settings of the control factors. If we write $\sigma_z^2 = \sum_{k=1}^K \sigma_k^2$, the approximate variance-covariance matrix of the error term in the log-linear model in (6) for this situation can be written as

$$\frac{1}{\left[R\sigma_z^2 + (J-1)\sigma_\delta^2 \right]^2} \left[(4R\sigma_\delta^2\sigma_z^2 + 2(J-1)\sigma_\delta^4)\mathbf{I} + 2R^2 \left(\sum_{k=1}^K \sigma_k^4 \right) \mathbf{J} \right].$$

Consider now least-squares estimates of the dispersion effects in $\log S^2 = \mathbf{X}\Phi + \nu$. Here again, the least-squares estimators of the dispersion effects are equivalent to the generalized least-squares estimators due to the structure of the variance-covariance matrix. Since $\mathbf{X}'\mathbf{J}\mathbf{X}$ is 0 except for its first element, the approximate variance of the least-squares estimator of the dispersion effects $\hat{\Phi}_l$, for $l \geq 2$ is then

$$\text{Var}(\hat{\Phi}_l) = \frac{4R\sigma_\delta^2\sigma_z^2 + 2(J-1)\sigma_\delta^4}{n[R\sigma_z^2 + (J-1)\sigma_\delta^2]^2} = \frac{2}{n(J-1)}(1 - \rho_R^2),$$

where $\rho_R = \frac{R\sigma_z^2}{(J-1)} / (\sigma_\delta^2 + \frac{R\sigma_z^2}{(J-1)})$.

It is interesting to compare the current design where each of the K noise variables are varied separately, each at 2 levels, with a compound noise approach for studying the K noise variables. We do this under the assumption that σ_z^2 , the variance of the compound noise variable, equals $\sum_{k=1}^K \sigma_k^2$. In the compound noise approach, which corresponds to the situation in Section 2.1, the approximate variance of the dispersion effects is $\frac{2}{n(J-1)}(1 - \rho^2)$ where $\rho = \frac{\sigma_z^2}{\sigma_\delta^2 + \sigma_z^2}$. Since $\rho \leq \rho_R$ with equality when $R = 1$, the compound noise approach is more efficient when an S^2 analysis is used. In the compound noise approach, there are J true replications from which to estimate each sample variance S_i^2 . In the separate noise approach, there are also J runs to the noise array, but only 2 replications for the error due to controlled noises, which explains why, in this case, the estimates of total variation in the noise arrays have a greater variance.

3.2. Analysis based on sample variances

We now return to the more general model in (9). As noted earlier, the analysis based on sample variances (or the related signal-to-noise ratios) is not really valid when the dispersion effects associated with the different noise factors are all different as in (9). This can be seen from the fact that, under model (9), the expected value of S_i^2 in (2) is obtained as

$$E(S_i^2) = \phi_1^2(x_i)\sigma_1^2 + \cdots + \phi_K^2(x_i)\sigma_K^2 + \theta^2(x_i)\sigma_\delta^2.$$

Thus, the analysis based on S_i^2 's tries to combine all of the dispersion effects associated with the various noise factors into a single measure and tries to estimate an "overall" set of dispersion effects. If the dispersion effects are all quite different, this analysis can lead to misleading conclusions about the importance and magnitude of the various dispersion effects. This is especially true when there are many dispersion effects and the control array is highly fractionated. In the presence of effect sparsity, however, it is possible that the active effects are

correctly identified, but the magnitude of the effects will still be not estimated correctly.

Consider now the variances of the estimated dispersion effects from this analysis. For simplicity of exposition, consider first the case where $K = 1$ but, unlike the set-up in Section 2, now ϕ is not equal to θ . Routine calculations show that the sample variance for each run of the control array is distributed as

$$S_i^2 \sim [\phi^2(x_i)\sigma_z^2 + \theta^2(x_i)\sigma_\delta^2] \chi_{(J-1)}^2.$$

To identify important dispersion effects, suppose we fit the log-linear model $\log S_i^2 = x_i\Phi + \nu_i$. The error term ν_i has the variance-covariance matrix

$$\text{Var}(\nu) \approx \frac{2}{J-1} \begin{pmatrix} 1 & \rho_1\rho_2 & \cdots & \rho_1\rho_n \\ \rho_1\rho_2 & 1 & \cdots & \rho_2\rho_n \\ \vdots & \vdots & \ddots & \vdots \\ \rho_1\rho_n & \rho_2\rho_n & \cdots & 1 \end{pmatrix}, \quad (11)$$

where

$$\rho_i = \phi^2(x_i)\sigma_z^2 / [\phi^2(x_i)\sigma_z^2 + \theta^2(x_i)\sigma_\delta^2]. \quad (12)$$

We can re-express (11) as

$$\text{Var}(\nu) \approx \frac{2}{J-1} [\mathbf{D}(1 - \rho^2) + \rho\rho^T], \quad (13)$$

where ρ is the vector whose i th element is given by ρ_i in (12) and $\mathbf{D}(1 - \rho^2)$ is the diagonal matrix with diagonal elements given by $(1 - \rho_i^2)$, $i = 1, \dots, n$. The variance of the estimated dispersion effects using least-squares can be easily obtained from this. However, generalized least-squares would be more appropriate here.

In either case, it can be seen from (11) that, as $\rho_i \rightarrow 1$ for all i , the variance of the estimated dispersion effects $\rightarrow 0$. As the proportion of total variation that is due to uncontrolled/unknown noise increases, $\rho_i \rightarrow 0$ and the variance of the estimated dispersion effects $\rightarrow 2/n(J-1)$.

We now turn to the case with K noise factors. We consider only the situation where each factor is selected at two-levels and discuss separately the two situations with:

1. two fixed settings $(\pm (\frac{J-1}{J})^{1/2} h\sigma_k)$ for each of z_1, \dots, z_K ,
2. two randomly selected settings for each of z_1, \dots, z_K .

In both cases, each level is repeated R times in the noise array. In the first case where we have fixed settings, the sample variances

$$S_i^2 = \frac{1}{J-1} \sum_{j=1}^J \left(\sum_{k=1}^K \phi_k(x_i) (\pm (\frac{J-1}{J})^{1/2} h\sigma_k) + \tilde{\delta}_{ij} \right)^2$$

have a non-central chi-square distribution, $\theta^2(x_i)\sigma_\delta^2(J-1)^{-1}\chi_{(J-1)}^2(\lambda)$, with non-centrality parameter given by $\lambda = (J-1)h^2 \sum_{k=1}^K \phi_k^2(x_i)\sigma_k^2/(\theta^2(x_i)\sigma_\delta^2)$. If we now fit the log-linear model in (6), the error term has a diagonal variance-covariance matrix approximately equal to

$$\text{Var}(\nu) \approx \frac{2}{J-1} \mathbf{D}(1 - \rho^2(h)), \quad (14)$$

where

$$\rho_i(h) = \frac{h^2 \sum_{k=1}^K \phi_k^2(x_i)\sigma_k^2}{h^2 \sum_{k=1}^K \phi_k^2(x_i)\sigma_k^2 + \theta^2(x_i)\sigma_\delta^2}.$$

Note that this $\rho_i(h)$ reduces to ρ_i in (12) for $h = 1$ and $K = 1$; note also the relationship between (14) and (13). Again, as $\rho_i(h)$ goes to one, the variance of the estimated dispersion effects decreases, and as $\rho_i(h)$ goes to zero, the variance of the estimated dispersion effects goes to $2/n(J-1)$.

Finally, when two random settings are selected from each noise distribution and varied in a fractional factorial noise array,

$$S_i^2 = \frac{1}{(J-1)} \sum_{j=1}^J \left(\sum_{k=1}^K \phi_k(x_i)\tilde{z}_{kj} + \theta(x_i)\tilde{\delta}_{ij} \right)^2$$

and it now involves summing up over J independent error terms but only 2 independent settings of each noise factor. The variance of S_i^2 is given by

$$2(J-1)^{-2} \left[R^2 \sum_{k=1}^K \phi_k^4(x_i)\sigma_k^4 + 2R\theta^2(x_i)\sigma_\delta^2 \sum_{k=1}^K \phi_k^2(x_i)\sigma_k^2 + (J-1)\theta^4(x_i)\sigma_\delta^4 \right]$$

while the covariance between sample variances for different runs of the control array is

$$\text{Covar}(S_i^2, S_{i'}^2) = \frac{2R^2}{(J-1)^2} \sum_{k=1}^K \phi_k^2(x_i)\phi_k^2(x_{i'})\sigma_k^4.$$

The variance of the estimated dispersion effects can be computed from these expressions using a first-order approximation. However, the expressions do not simplify nicely to allow us to express them in terms of ρ -like expressions as before.

3.3. Least-squares analysis

When the noise settings z_{kj} 's in (9) are all fixed at pre-specified values, then it is natural to absorb the z_{kj} 's into the structural model and estimate the $\phi_k(x_i)$'s directly using least-squares or some other method. This is the response-model approach discussed in Welch et al. (1990), Shoemaker et al. (1991),

and others. However, these authors assumed that $\theta(\cdot)$ in (9) is constant across control runs. This is not an unreasonable assumption if most of the important noise factors have been identified and controlled up front in the parameter design study. A more important difference is that these papers assume the ϕ_k 's follow a linear model in the control factors. The analysis based on log-variances and the related signal-to-noise ratios assumes a log-linear response surface as in (6). These differences should be recognized and taken into account before trying to make comparisons of the results from different analyses.

Let us consider the analysis based on least-squares under the assumption that $\theta(x_i)$ in (9) is constant, which we can take to equal 1. Let the z_{kj} 's be fixed with $\sum_{j=1}^J (z_{kj} - \bar{z}_k)^2 / (J - 1) \equiv h^2 \sigma_k^2$. To decouple the location and dispersion problems, we do a two-stage estimation where we first estimate the $\phi_k(x_i)$, $i = 1, \dots, n$, directly by least-squares for each of the K noise factors. We consider, as in Shoemaker et al. (1991), the situation where the response surfaces for the $\phi_k(x_i)$'s are, at least approximately, linear; i.e., $\phi_k(\mathbf{X}) = \mathbf{X}\Phi_k$, $k = 1, \dots, K$, where \mathbf{X} is the control array, $\Phi_k = (\Phi_{1k}, \dots, \Phi_{Lk})'$ and for $l \geq 2$, the Φ_{lk} 's correspond, for the k th noise factor, to the dispersion effects of the various control factors. One can then fit K separate linear models in the control factors to the estimated $\phi_k(\mathbf{X})$'s to obtain estimates $\hat{\Phi}_{lk}$ of the specific dispersion effects associated with each particular control factor. It is easily seen that the variance of these estimated dispersion effects is then, for $l \geq 2$, $\text{Var}(\hat{\Phi}_{lk}) = \sigma_\delta^2 / [n(J - 1)h^2\sigma_k^2]$, which is the same as the (approximate) variance obtained in Section 2.3. Note, however, that in Section 2.3, we assumed a log-linear response surface for $\phi(x_i)$ and fitted a log-linear model to the least-squares estimates $\hat{\phi}(x_i)$'s. It is interesting that the results for these two different situations turn out to be the same.

The situation becomes more complicated when $\theta(\cdot)$ is not constant. Consider a two-stage analysis as discussed above. If, for each noise factor, $\phi_k(x_i)$, $i = 1, \dots, n$, is estimated by ordinary least-squares, the variance of $\hat{\phi}_k(x_i)$ will now be proportional to $\theta^2(x_i)$, i.e., it will depend on the level of the control factors. So, when we fit a linear model, $\hat{\phi}_k = \mathbf{X}\Phi_k + \nu$, the vector of error, ν , has variance-covariance matrix

$$\text{Var}(\nu) = \frac{\sigma_\delta^2}{(J - 1)h^2\sigma_k^2} \mathbf{D}(\theta^2(x_1), \dots, \theta^2(x_n)).$$

Suppose now one used least-squares to estimate the dispersion effects Φ_k . The variance-covariance matrix of the least-squares estimates is given by $\text{Var}(\hat{\Phi}_k) = \frac{1}{n^2} \mathbf{X}' \text{Var}(\nu) \mathbf{X}$. The variance of the estimated dispersion effects are given by the diagonal elements of this matrix which equal

$$\frac{\sigma_\delta^2}{(J - 1)n^2 h^2 \sigma_k^2} \sum_{i=1}^n \theta^2(x_i).$$

However, the matrix is not diagonal, and the estimates are now correlated. In any case, for this situation, weighted least-squares estimation that takes into account the unequal variances would be more reasonable. An iterative procedure has to be used in this case.

3.4. Variance components analysis

The least-squares analysis discussed earlier is not appropriate when the levels of z_{kj} 's are chosen randomly. In this section, we consider an alternative analysis that can handle both the random and fixed noise settings cases in model (9). To keep the notation simple, we assume without loss of generality that $\sigma_1^2 = \dots = \sigma_k^2 = \sigma_\delta^2 = 1$.

As before, we consider only two level noise factors. Let $R = J/2$ be the number of times each level is repeated in the noise array. All noise factors (identified or not) are assumed to follow a standard normal distribution with variance one, and so the Y_{ij} is distributed as $N(\mu(x_i), \sum_{k=1}^K \phi_k^2(x_i) + \theta^2(x_i))$, $i = 1, \dots, n$. One might consider an analysis based on measures of variability that estimate the contribution of each noise factor to the total variance. As such, the usual components of variance sums of squares can be calculated for each noise factor z_k , $k = 1, \dots, K$, for each row of the control array as

$$S_k^2(x_i) = (\bar{Y}_i(z_{kj} = l_1) - \bar{Y}_i.)^2 + (\bar{Y}_i(z_{kj} = l_2) - \bar{Y}_i.)^2,$$

where \bar{Y}_i represents the average of all J responses for run i of the control array, and $\bar{Y}_i(z_{kj} = l_1)$ represents the average, for row i , of all responses with the k th noise factor set at level 1. For example, if the noise array is a 2^2 factorial design, the above SS for noise 1 would be

$$S_1^2(x_i) = (\bar{Y}_i(z_{1j} = l_1) - \bar{Y}_i(z_{1j} = l_2))^2/2.$$

We consider both the case where the noise variables are pre-selected at fixed values and the case where the levels are randomly selected from the noise distribution and repeated for each control run setting.

3.4.1. Fixed settings

Assume that we have K two-level noise factors, each with levels fixed at $\pm h$. In this case, the K sums of squares corresponding to the K controlled noise factors are

$$\begin{aligned} S_k^2(x_i) &= (\bar{Y}_i(z_{kj} = h) - \bar{Y}_i.)^2 + (\bar{Y}_i(z_{kj} = -h) - \bar{Y}_i.)^2 \\ &= (\phi_k(x_i)h + \bar{\delta}_i(z_{kj} = h) - \bar{\delta}_i.)^2 + (-\phi_k(x_i)h + \bar{\delta}_i(z_{kj} = -h) - \bar{\delta}_i.)^2 \\ &= \frac{2}{J^2} [J\phi_k(x_i)h + \theta(x_i) \left(\sum_{j:z_{kj}=h} \delta_{ij} - \sum_{j:z_{kj}=-h} \delta_{ij} \right)]^2. \end{aligned}$$

Since the noise levels are fixed, the expression in the last parenthesis above is distributed as $N(J\phi_k(x_i)h, J\theta^2(x_i))$ and it follows that the $S_k^2(x_i)$'s are distributed as non-central χ^2 ,

$$S_k^2(x_i) \sim \frac{2}{J^2} J\theta^2(x_i) \chi_1^2(\alpha) = \frac{\theta^2(x_i)}{R} \chi_1^2(\alpha),$$

with non-centrality parameter $\alpha = 2R\phi_k^2(x_i)h^2/\theta^2(x_i)$.

Next, we fit a log-linear model to each $S_k^2(x_i)$, $k = 1, \dots, K$, to identify the dispersion effects associated with the k th noise factor. We fit the log-linear model $\log S_k^2(x_i) = x_i\Phi_k + \nu_i$, where we use a first-order approximation to get the following approximate variance-covariance matrix of the error term,

$$\text{Var}(\nu_i) \approx 2 \left[\frac{\theta^4(x_i) + 4Rh^2\phi_k^2(x_i)\theta^2(x_i)}{(\theta^2(x_i) + 2Rh^2\phi_k^2(x_i))^2} \right] = 2(1 - \rho_{ki}^2(h)),$$

where

$$\rho_{ki}(h) = \frac{2Rh^2\phi_k^2(x_i)}{\theta^2(x_i) + 2Rh^2\phi_k^2(x_i)}. \quad (15)$$

Since the levels of the noise factors are fixed, the sample variances are not correlated from run to run, so the variance-covariance matrix of the ν_i 's is diagonal with the diagonal entries given above. Note again the relationship between these results and those in Section 4.1.

If we use least-squares estimation to obtain the dispersion effects, the approximate variance of the least-squares estimates $\hat{\Phi}_k$ associated with the k th noise factor is then

$$\text{Var}(\hat{\Phi}_k) = \frac{2}{n^2} \mathbf{X}'\mathbf{D}(1 - \rho_{k1}^2(h), \dots, 1 - \rho_{kn}^2(h))\mathbf{X}. \quad (16)$$

Of course, one might consider doing weighted least squares in this situation. Regardless, the same conclusion as before holds about the importance of identifying and controlling the variability associated with the noise variables.

3.4.2. Random levels

Now assume that two settings for each of the K noise factors are chosen randomly from the corresponding noise distribution. The settings are denoted by z_{k1} and z_{k2} and varied according to a (fractional) factorial design. The same settings are repeated across the control runs. Here as above,

$$\begin{aligned} S_k^2(x_i) &= (\bar{Y}_i(z_{kj} = z_{k1}) - \bar{Y}_i.)^2 + (\bar{Y}_i(z_{kj} = z_{k2}) - \bar{Y}_i.)^2 \\ &= \frac{2}{J^2} \left[R\phi_k(x_i)(z_{k1} - z_{k2}) + \theta(x_i) \left(\sum_{j:z_{kj}=z_{k1}} \delta_{ij} - \sum_{j:z_{kj}=z_{k2}} \delta_{ij} \right) \right]^2, \end{aligned}$$

and if the z_k 's and δ_{ij} are all independently normally distributed with variances 1, the expression in the last parenthesis is distributed as $N(0, \frac{1}{2}J^2\phi_k^2(x_i) + J\theta^2(x_i))$. It follows that the K sums of squares are distributed as $(\phi_k^2(x_i) + \frac{\theta^2(x_i)}{R})\chi_1^2$.

With the levels of the noise factors random but repeated for each control run, the ν_i 's in the log-linear model $\log S_k^2(x_i) = x_i\Phi_k + \nu_i$ are correlated. Using a first-order approximation, the variance-covariance matrix is approximated as

$$\text{Var}(\nu) \approx 2 \begin{pmatrix} 1 & \rho_{k1}\rho_{k2} & \cdots & \rho_{k1}\rho_{kn} \\ \rho_{k1}\rho_{k2} & 1 & \cdots & \rho_{k2}\rho_{kn} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k1}\rho_{kn} & \rho_{k2}\rho_{kn} & \cdots & 1 \end{pmatrix},$$

where

$$\rho_{ki} = \frac{R\phi_k^2(x_i)}{R\phi_k^2(x_i) + \theta^2(x_i)}.$$

In this case, the approximate variance-covariance matrix of the estimated dispersion effects associated with the k th noise factor, based on least-squares, is

$$\text{Var}(\hat{\Phi}_k) = \frac{2}{n^2} \mathbf{X}' \left(\mathbf{D}(1 - \rho_{k1}^2, \dots, 1 - \rho_{kn}^2) + \rho_k \rho_k' \right) \mathbf{X}, \quad (17)$$

where $\rho_k = (\rho_{k1}, \dots, \rho_{kn})'$. Note that if the levels of the noise factors are fixed such that $h^2 = 1/2$ then $\rho_{ki}(h)$ in (15) is equal to ρ_{ki} . In this case, the matrix in (17) is equal to the matrix in (16) plus the term $\frac{2}{n^2} \mathbf{X}' \rho_k \rho_k' \mathbf{X}$.

As $\rho_{ki} \rightarrow 1$ (or $\rho_{ki}(h) \rightarrow 1$) for all i , i.e., the proportion of variation controlled through the k th noise factor is large compared to the uncontrolled/unknown variation, the variance of the estimated dispersion effects corresponding to the k th noise factor, $\text{Var}(\hat{\Phi}_{lk})$ decreases. On the other hand, as $\rho_{ki} \rightarrow 0$ (or $\rho_{ki}(h) \rightarrow 0$), $\text{Var}(\hat{\Phi}_{lk}) \rightarrow \frac{2}{n(J-1)}$, which would correspond to the variance of the estimated dispersion effects if the k th noise factor varied independently from run to run.

4. Miscellaneous Remarks

4.1. Estimation of location effects

While the focus of this paper is on estimating dispersion effects, we briefly discuss estimation of location effects in this section. We show that similar conclusions about efficiency also hold for location effects.

Consider the model in (3) with $\mu(x_i) = x_i' \alpha$ where $\alpha = (\alpha_1, \dots, \alpha_L)$ and for $l \geq 2$, the α_l 's denote the location effects. We first consider the use of ordinary least-squares (OLS) estimates to estimate location effects. The properties of weighted least-squares (WLS) estimators, which are more appropriate here because of the unequal variances, is discussed later. The perceived variance of

the estimator when the z_j 's are (wrongly) assumed to vary independently is seen to be $(Jn^2)^{-1} (\sigma_z^2 + \sigma_\delta^2) \mathbf{X}'\mathbf{D}(\phi^2)\mathbf{X}$, where $\phi = (\phi(x_1), \dots, \phi(x_n))'$. Consider now the actual variance. It is clear that if the z_j 's in (3) are fixed during the experiment, the actual variance of the OLS estimator is obtained by replacing the factor $(\sigma_z^2 + \sigma_\delta^2)$ in the above expression by σ_δ^2 . So, the ratio of the actual variance to the perceived variance is $\sigma_A^2/\sigma_P^2 = (1 - \rho)$ where ρ is given by (5). It is interesting to compare this expression with the result for the dispersion effects given in (4).

We briefly turn to WLS estimators. Since the variances are unknown, one would use iteratively reweighted least squares to estimate the location effects. The variances of these estimators can be approximated by the known ϕ_k 's situation. If we do this, the perceived variance of the WLS estimator can be approximated by $J^{-1} (\sigma_z^2 + \sigma_\delta^2) (\mathbf{X}'\mathbf{D}^{-1}(\phi^2)\mathbf{X})^{-1}$. If the z_j 's in (3) are fixed during the experiment, then it is clear that the actual variance of the WLS estimator can be approximated by replacing the factor $(\sigma_z^2 + \sigma_\delta^2)$ in the above expression by σ_δ^2 . Hence, the conclusion is the same as that for the OLS case.

When the z_j 's are random, one can reach the same conclusion as that for the fixed z_j 's in the special case where the ϕ_k 's are constant, i.e., when there are no dispersion effects. In general, however, the results for random z_j 's are complicated, and we will not discuss them in detail here.

4.2. Dynamic characteristics

We have so far considered only static robust design experiments. Robust design studies with the so-called dynamic characteristics are increasingly common in industry. The efficiency results obtained in Sections 2.1 and 2.2 can be shown to hold also in the context of parameter designs with dynamic characteristics. Consider a model as in (3) but where the system's performance (output y) depends on a signal factors (x) and noise Variable (z),

$$Y_{ijk} = \mu(x_i; M_k) + \phi(x_i)[z_j + \delta_{ijk}]. \quad (18)$$

Here, Y_{ijk} denotes the observation corresponding to the i th setting of the control factors, j th setting of the noise factor and k th setting of the signal factor, for $i = 1, \dots, n$, $j = 1, \dots, J$ and $k = 1, \dots, K$. Note that for this model, the signal factor affects the response only through the location parameter.

One approach for estimating the $\phi^2(x_i)$'s in this case might be to consider the signal factor as an extra control factor. The control array can then be thought of as having nK rows since for each setting of the control factors, there are K settings of the signal factor. As in the static case, the noise array of size J is repeated across all the nK control runs. We first obtain nK sample variances and

then for each of the n settings of the control factors, average the corresponding sample variances over the K signal settings,

$$S_i^2 = \frac{1}{K(J-1)} \sum_{k=1}^K \sum_{j=1}^J (Y_{ijk} - \bar{Y}_{ik})^2.$$

This essentially corresponds to the situations analyzed in Sections 2.1 and 2.2 so that the same efficiency results hold true here.

Now, instead of the general model in (18), we consider more specifically a linear signal-response relationship of the form

$$Y_{ijk} = M_k \beta(x_i) + \phi(x_i)[z_j + \delta_{ijk}].$$

Typically, the slope β would first be estimated for each run of the control array by fitting a simple linear regression model with the signal factor as regressor. We then estimate the dispersion effects $\phi^2(x_i)$ by the residual sum of square

$$S_i^2 = (Y_i - M\hat{\beta}(x_i))'(Y_i - M\hat{\beta}(x_i)), \quad (19)$$

where Y_i is a $JK \times 1$ vector of responses corresponding to all combinations of signal and noise levels at the i th control array setting and M is a $JK \times 1$ vector of the K signal levels repeated for each of the J noise levels. In order to identify the distribution of the sample variances, we rewrite (19) as

$$S_i^2 = (Y_i - M\beta(x_i))'(I - H_M)(Y_i - M\beta(x_i)),$$

where $H_M = M(M'M)^{-1}M'$. If the z_j 's are fixed during the experiment, $Y_i - M\beta(x_i)$ has multivariate normal distribution $N_{JK}(\phi(x_i)z, \phi^2(x_i)\sigma_\delta^2\mathbf{I})$ where z is the $JK \times 1$ vector of the J noise settings repeated for each of the K signal levels. It follows that the residual sum of squares S_i^2 has a non-central χ^2 distribution, $\phi(x_i)^2\sigma_\delta^2\chi_{JK-1}^2(\lambda)$ with non-centrality parameter

$$\lambda = \frac{z'(I - H_M)z}{\sigma_\delta^2} = (1/\sigma_\delta^2) \left[K \sum_{j=1}^J z_j^2 - \frac{\sum_{k=1}^K M_k \sum_{j=1}^J z_j}{J \sum_{k=1}^K M_k^2} \right].$$

In the special case where the noise factors are centered and chosen such that $\sum_{j=1}^J (z_j - \bar{z})^2 / (J-1) \equiv h^2\sigma_z^2$ for some fixed h , the non-centrality parameter reduces to $K(J-1)h^2\sigma_z^2/\sigma_\delta^2$. Again, this is similar to the distribution of the sample variances S_i^2 in the static setting analyzed in Section 2.2, and we have essentially equivalent efficiency results.

5. Conclusion

We have considered several different models and methods of estimation for robust parameter design experiments. For all of these situations, we have shown

the efficiency to be gained from carefully planned experiments in which the important sources of noise variation are identified up front and varied systematically. Thus, it is important that the analysis methods explicitly take into account the structure of the noise array. Many of the current analysis methods do not do this and hence are unnecessarily inefficient. Finally, there are many possible methods one can entertain for analyzing data with location and dispersion effects. There have been different approaches that have been advocated in the literature, but the fact that they are all based on different models has not been emphasized enough. These differences should be recognized before one tries to compare the results from different methods of analysis.

Acknowledgements

The authors are grateful to Michael Hamada, David Steinberg, Jeff Wu, and the reviewers for helpful comments. This research was supported by NSF Grants DMS 9404300 and DMI 9501217.

References

- Box, G. E. P. (1988). Signal-to-noise ratios, performance criteria, and transformations (with discussion). *Technometrics* **30**, 1-40.
- Box, G. and Jones, S. (1992). Split-plot designs for robust product experimentation. *J. Appl. Statist.* **19**, 3-26.
- Carroll, R. J. and Ruppert, D. (1988). Discussion of "Signal-to-noise ratios, performance criteria, and transformations". *Technometrics* **30**, 30-31.
- Gunter, B. (1988). Discussion of "Signal-to-noise ratios, performance criteria, and transformations". *Technometrics* **30**, 32-35.
- Kacker, R. N. (1985). Off-line quality control, parameter design, and the Taguchi method. *J. Quality Technology* **17**, 176-188.
- León, R., Shoemaker, A. C. and Kacker, R. N. (1987). Performance measures independent of adjustment: An explanation and extension of Taguchi's signal-to-noise ratios (with discussion). *Technometrics* **29**, 253-285.
- Lunani, M., Nair, V. N. and Wasserman, G. (1997). Graphical methods of data analysis for robust design with dynamic characteristics. *J. Quality Technology*, to appear.
- Miller, A. and Wu, C. F. J. (1996). Parameter design for signal-response systems: A different look at Taguchi's dynamic parameter design. *Statist. Science* **11**, 122-136.
- Nair, V. N. (1992). Taguchi's parameter design: A panel discussion. *Technometrics* **34**, 127-161.
- Nair, V. N. and Pregibon, D. (1986). A data analysis strategy for quality engineering experiments. *AT&T Technical J.* **65**, 73-84.
- Nair, V. N. and Pregibon, D. (1988). Discussion of "Signal-to-noise ratios, performance criteria, and transformations". *Technometrics* **30**, 24-30.
- Phadke, M. S. (1989). *Quality Engineering Using Robust Design*. Prentice-Hall, New Jersey.
- Shoemaker, A. C., Tsui, K. L. and Wu, C. F. J. (1991). Response model analysis for robust design experiments. *Technometrics* **33**, 415-427.
- Steinberg, D. M. and Bursztyn, D. (1998). Noise factors, dispersion effects, and robust design. *Statist. Sinica*, to appear.

- Taguchi, G. (1991). *Taguchi Methods, Vol.-1, Research and Development and Taguchi Methods, Vol.-3, Signal-to-Noise Ratio for Quality Evaluation*. ASI Press, Dearborn, MI.
- Vining, G. G. and Myers, R. H. (1990). Combining Taguchi and response surface philosophies: A dual response approach. *J. Quality Technology* **22**, 38-45.
- Welch, W. J., Yu, T. K., Kang, S. M. and Sacks, J. (1990). Computer experiments for quality control by parameter design. *J. Quality Technology* **22**, 15-22.

Department of Statistics, University of Michigan, Ann Arbor, MI 48109-1027, U. S. A.

E-mail: jber@umich.edu

E-mail: vnn@umich.edu

(Received July 1996; accepted May 1997)