The Annals of Statistics
1981, Vol. 9, No. 6, 1168–1177

ON THE ROBUSTNESS AND EFFICIENCY OF SOME RANDOMIZED DESIGNS[1]

By Chien-Fu Wu

University of Wisconsin-Madison

A concept of model-robustness is defined in terms of the performance of the design in the presence of model violations. The robustness problem is discussed for several randomization procedures commonly used in experimental design situations. Among them, the balanced completely randomized design, the randomized complete block design and the randomized Latin square design are shown to be model-robust in their own settings. To compare different randomization procedures, we define a concept of efficiency which depends on the particular "pattern" of model violations. This concept, when applied to different designs, gives results which are consistent with the intuitive grounds on which the designs are suggested.

1. Introduction. Experimental randomization is one of the greatest contributions of R. A. Fisher to science and statistics. Among the most popular of the arguments favoring the use of randomization are the following: It provides a solid basis for statistical inference; it ensures impartiality; it is a source of robustness against model inadequacies. The first argument has been discussed very extensively in the literature (Cox, 1958; Harville, 1975; Kempthorne, 1955, 1975 and references therein). The main contention is that the necessary assumptions for the randomization models are much less restrictive than for the ordinary normal-theory models. The second argument contends that the use of randomization ensures that the choice of design is not affected by any bias or preconceived notion on the part of the experimenter (Cox, 1958; Harville, 1975). Both arguments seem to be well defined and accepted. The third argument on the model robustness aspect of randomization is also well accepted, but there has not been given a formal definition and rigorous justification. It is the purpose of the present paper to take up this task. Our approach is motivated by and related to works on robust estimation and design (Bickel and Herzberg, 1979; Box and Draper, 1959; Huber, 1972, 1975). Since a very different and new viewpoint is taken in this paper, our criterion for comparing designs (Section 3) is different from the traditional ones, e.g. Yates's criterion for an unbiased error, "old" (in contrast to our "new") measure for the efficiency of a design (Kempthorne, 1955, 1975). A Bayesian justification of randomization is given in Rubin (1978).

In this paper we restrict our attention to comparative experiments, in which the experimental material is divided into N experimental units, to each of which any one of the T treatments can be applied. An experimental design is an assignment of treatments to units. If the experimenter's model assumption is exact and correct, he will take the optimal design approach (Kiefer, 1959; Fedorov, 1969) and choose a systematic design. This is certainly not a good situation for justifying the use of randomization. In fact, the experimenter's information about the model is never perfect. When a model is proposed, there is always the possibility that the "true" model deviates from the assumed model. Let

Received August, 1977; revised April, 1980.

[1] Sponsored by the United States Army under Contract No. DAAG29-80-C-0041 and NSF Grant No. MCS-79-01846.

AMS 1970 *Subject Classifications.* Primary 62K99, secondary 62A05, 62C20.

Key words and phrases. Model robustness, minimaxity, invariance, efficiency, systematic design, balanced completely randomized design, randomized complete block design, Latin square design.

G be the collection of all the possible "true" models. The concept of model-robustness with respect to G is defined in terms of minimizing the maximum possible mean squared error (m.s.e.) of the corresponding best linear unbiased estimator (B.L.U.E.) over G. In Section 2, some randomized designs, including the balanced completely randomized design, the randomized complete block design and the randomized Latin square design, are shown to be model-robust with respect to any G which possesses an invariance property. A similar result for simple random sampling was obtained by Blackwell and Girshick (1954). This optimality property also holds for other types of problems, including violations of the homoscedasticity assumption of the error terms and the estimation of variance under either type of violation. As a by-product, several randomization procedures for the Latin square design are reassessed in this framework. The standard practice of choosing a Latin square arbitrarily and then randomly permuting its rows, columns and treatments is shown to possess some desirable properties.

In Section 3, the efficiency comparisons of several procedures (systematic, partially randomized, completely randomized) are made in terms of the maximum bias square of the corresponding B.L.U.E. over a particular choice (3.4) of G. The efficiency of the systematic arrangements relative to the completely randomized arrangements is inversely proportional to the number of replications of each treatment. A similar result holds for the comparison of randomized and systematic Latin square designs. The randomized block design is shown to be very efficient when the block size is moderate or large. Conditions for these designs to be superior to the completely randomized design (CRD) are also given in terms of the special patterns of G. These patterns are quite consistent with the intuitive grounds on which these designs are suggested. For example, the randomized block design is better than the CRD if block effects account for most of the unknown effects in G. The efficiency concept introduced here thus makes it possible to compare different randomization procedures quantitatively.

2. Model-robustness of some randomized designs. T different treatments are compared on N experimental units with n_t units assigned to the tth treatment, $\sum_{t=1}^{T} n_t = N$. Let y_{ut} be the yield (or response) of the uth unit when treatment t is applied. In this paper we only consider the following additivity assumption (Kempthorne, 1955):

$$y_{ut} = \alpha_t + g_u + \epsilon_{ut}, \qquad u = 1, \cdots, N \tag{2.1}$$

where α_t is the tth treatment effect, g_u is the uth unit effect, ϵ_{ut} is the random error with zero mean and equal and uncorrelated variances σ^2. No interaction between treatment and unit is assumed. In standard randomization models (Scheffé, 1958, Chapter 9), g_u and ϵ_{ut} are called respectively *unit error* and *technical error*. Similar remarks apply to models (2.7), (2.10), (2.11). The unit effect g_u can be related to any particular feature of the uth unit, e.g., fertility gradient, initial weight, income level. Our main concern is in comparing the α_t. If we can apply both treatments s and t to each experimental unit and assume that the outcomes on u do not interact with each other (e.g. there is no carry-over effect), then $\alpha_s - \alpha_t$ can be estimated by $y_{us} - y_{ut}$ or its average over u. But in practice this is rarely the case. In this paper we assume that each unit can receive only one treatment. Therefore, the accuracy of estimating α_t is complicated by the uneasy presence of g_u.

If a component of g_u like a block effect or covariate effect is envisaged, we will include these effects in a more complicated model than (2.1), which will be discussed later. For the moment we assume that g can be any element from a set G. The special case $g \equiv c$, a constant, corresponds to the usual model assumption underlying the least squares theory. We call G a *neighborhood of model violations*. In this paper we assume that *G is invariant under a certain transformation group H*, i.e.,

$$g \in G \Rightarrow \pi g \in G \quad \text{for all} \quad \pi \in H, \tag{2.2}$$

where $\pi g = \{g_{\pi^{-1}(u)}\}_u$, π^{-1} is the inverse of π in the group H. The element $g \equiv c$ should also be in G. To avoid triviality, G is assumed to be bounded. *The invariance assumption reflects the vagueness of the experimenter's knowledge about* g_u. In this paper we only consider permutation groups. For response surface design, rotation groups may be considered.

Let $I_t = \{u : u \text{ is assigned to treatment } t\}$ and call $I = \{I_t\}_{t=1}^{T}$ a pattern. In design terminology I corresponds to a non-randomized design with treatment group sizes $|I_t| = n_t$, $t = 1, \cdots, T$. Let $\mathscr{I}$ be the collection of such I's. A randomized design η is defined as a probability measure over $\mathscr{I}$, i.e., $\{\eta(I)\}_{I\in\mathscr{I}}$ with $\eta(I) \geq 0$ and $\sum_{I\in\mathscr{I}} \eta(I) = 1$.

The vagueness of our knowledge about model violations, as indicated by the invariance assumption (2.2), suggests that we use the best linear unbiased estimator when $g \equiv c$ in comparing designs. (Such an approach was taken by Bickel and Herzberg (1979) for a robust design problem. See also Box and Draper (1959), Huber (1975). If one has a definite idea about model violations, say, block effect or time trend, this should be included as parameters in the model and analysis should be made accordingly.) Therefore

$$\hat{\alpha}_s - \hat{\alpha}_t = y_{\cdot s} - y_{\cdot t}, \quad \text{where} \quad y_{\cdot s} = n_s^{-1} \textstyle\sum_{u\in I_s} y_{us}.$$

For any systematic assignment $I = \{I_t\}_{t=1}^{T}$ with $|I_t| = n_t$, we have, under (2.1),

$$\begin{aligned} a(I, g) &= \textstyle\sum_{s<t} E(\hat{\alpha}_s - \hat{\alpha}_t - \alpha_s + \alpha_t)^2 \\ &= T \textstyle\sum_{t=1}^{T} (g_{\cdot t} - g_{\cdot\cdot})^2 + \sigma^2(T-1) \textstyle\sum_{t=1}^{T} n_t^{-1}, \end{aligned} \tag{2.3}$$

where

$$g_{\cdot s} = n_s^{-1} \textstyle\sum_{u\in I_s} g_u, \; g_{\cdot\cdot} = T^{-1} \textstyle\sum_{t=1}^{T} g_{\cdot t}.$$

For a randomized design η, the expected mean squared error is then

$$r(\eta, g) = \textstyle\sum_{I\in\mathscr{I}} \eta(I) a(I, g).$$

This becomes a game with "risk" $r(\eta, g)$ between the experimenter and nature. The experimenter can choose any design measure η over $\mathscr{I}$ and nature can choose any g from G, corresponding to an unknown true model. The robust design problem can now be formulated in a decision-theoretic framework.

A design η^* *is called minimax, or model-robust, with respect to* G *if it achieves*

$$\min_\eta \max_{g\in G} r(\eta, g).$$

For any permutation π, define $\pi I_t = \{\pi(u) : u \in I_t\}$, $\pi I = \{\pi I_t\}_{t=1}^{T}$ and $\eta_\pi(I) = \eta(\pi I)$. From any design measure η, we can construct a new design measure

$$\tilde{\eta} = \frac{1}{N!} \textstyle\sum_{\pi\in P} \eta_\pi, \tag{2.4}$$

where P is the permutation group on N units. Denote a systematic design I by δ_I, i.e. $\delta_I(I) = 1$ and $\delta_I(K) = 0$ for other K in $\mathscr{I}$. Then $\tilde{\delta}_I$ is the design which assigns treatment t to a set of $n_t (= |I_t|)$ units completely randomly, $t = 1, \cdots, T$.

PROPOSITION 1. *Suppose* G *satisfies the invariance property* (2.2) *for all permutations of* $\{1, \cdots, N\}$. *Then for any* η,

$$\max_{g\in G} r(\tilde{\eta}, g) \leq \max_{g\in G} r(\eta, g).$$

PROOF. From (2.2) and $a(\pi I, \pi g) = a(I, g)$,

$$\begin{aligned} \max_{g\in G} r(\tilde{\eta}, g) &\leq \frac{1}{N!} \textstyle\sum_{\pi\in P} \max_{g\in G} \{\textstyle\sum_{I\in\mathscr{I}} \eta(\pi I) a(I, g)\} \\ &= \frac{1}{N!} \textstyle\sum_{\pi\in P} \max_{g\in G} \{\textstyle\sum_{J\in\mathscr{I}} \eta(J) a(J, \pi g)\} = \max_{g\in G} r(\eta, g). \qquad \square \end{aligned}$$

To assign N units to T treatment groups, it is most balanced to choose treatment group sizes n_t as equally as possible, i.e. $| n_t - (N/T) | < 1$ for all t. Such a choice minimizes $\sum_{t=1}^{T} n_t^2$ and $\sum_{t=1}^{T} n_t^{-1}$ subject to $\sum_{t=1}^{T} n_t = N$. A *balanced completely randomized design* is defined to be the design which assigns treatment t to a set of n_t units at random, where $\{n_t\}_{t=1}^{T}$ are the most balanced group sizes as defined above.

THEOREM 1. *The balanced completely randomized design is minimax with respect to any G, which satisfies the invariance property* (2.2) *for all permutations of* $\{1, \cdots, N\}$.

PROOF. Any η and $\tilde{\eta}$ can be expressed as $\sum_{I\in\mathscr{I}} \eta(I)\delta_I$ and $\sum_{I\in\mathscr{I}} \eta(I)\tilde{\delta}_I$ respectively. Since $r(\tilde{\eta}, g)$ is linear in $\tilde{\eta}$, $r(\tilde{\eta}, g) = \sum_{I\in\mathscr{I}} \eta(I) r(\tilde{\delta}_I, g)$. It is proved in the appendix that

$$r(\tilde{\delta}_I, g) = \left[\frac{T-1}{N-1} \sum_{u=1}^{N} (g_u - \bar{g})^2 + \sigma^2(T-1)\right] \sum_{t=1}^{T} \frac{1}{n_t}, \tag{2.5}$$

where $\bar{g} = N^{-1} \sum_{u=1}^{N} g_u$ and $|I_t| = n_t$. Therefore $\max_{g\in G} r(\tilde{\delta}_I, g)$ is attained by any g^* maximizing $\sum_{u=1}^{N} (g_u - \bar{g})^2$ irrespective of the choice of I, and is minimized by choosing $\{n_t\}_{t=1}^{T}$ to minimize $\sum_{t=1}^{T} n_t^{-1}$. These and Proposition 1 imply the result. □

Theorem 1 justifies *randomization* as well as *balance* from the model-robustness viewpoint. However, by using the squared error loss function in (2.3), we assume implicitly that all the treatments are of equal interest. Otherwise, the results in Theorems 1, 2 and 3 may not be true. Such a situation needs further investigation.

Besides estimating the contrasts, we may also be interested in estimating the error variance σ^2. Due to the invariance assumption (2.2), we may assume $g \equiv c$ in estimating σ^2 in (2.1), i.e.

$$\hat{\sigma}^2 = (N-T)^{-1} \sum_{t=1}^{T} \sum_{u\in I_t} (y_{ut} - y_{\cdot t})^2.$$

The true bias under (2.1) is

$$b(I, g) = E(\hat{\sigma}^2) - \sigma^2 = (N-T)^{-1} \sum_{t=1}^{T} \sum_{u\in I_t} (g_u - g_{\cdot t})^2,$$

where $g_{\cdot t} = n_t^{-1} \sum_{u\in I_t} g_u$. *A design* η^* *is called minimax with respect to G for estimating* σ^2 *if it achieves* $\min_\eta \max_{g\in G} r_1(\eta, g)$, where $r_1(\eta, g) = \sum_{I\in\mathscr{I}} \eta(I) b(I, g)$.

THEOREM 2. *Under the condition on G in Theorem* 1, *any design* $\tilde{\eta}$ *as defined in* (2.4) *is minimax with respect to G for estimating* σ^2.

PROOF. From $b(\pi I, \pi g) = b(I, g)$ and (2.2), it can be shown by following the proof of Proposition 1 that $\max_{g\in G} r_1(\tilde{\eta}, g) \le \max_{g\in G} r_1(\eta, g)$. Again from the linearity of $r_1(\tilde{\eta}, g)$ in $\tilde{\eta}$, $r_1(\tilde{\eta}, g) = \sum_{I\in\mathscr{I}} \eta(I) r_1(\tilde{\delta}_I, g)$. It is proved in the appendix that

$$r_1(\tilde{\delta}_I, g) = \frac{1}{N-1} \sum_{u=1}^{N} (g_u - \bar{g})^2, \bar{g} = \frac{1}{N} \sum_{u=1}^{N} g_u, \tag{2.6}$$

independent of the choice of I. Therefore $\max_{g\in G} r_1(\tilde{\eta}, g)$ is a constant and the theorem follows. □

Since $\tilde{\eta}$ can involve different sets of $\{n_t\}_{t=1}^{T}$, in contrast to Theorems 1 and 2, the sizes of the treatment groups are irrelevant to the study of model-robustness for estimating σ^2.

We may also consider the possibility that the assumptions on errors ϵ_{ut} in (2.1) are violated. Suppose the true model is

$$y_{ut} = \mu + \alpha_t + \epsilon_{ut}, \quad u = 1, \cdots, N, \tag{2.7}$$

where ϵ_{ut} has mean zero and variance-covariance matrix V, V can be any element from a set $\mathscr{V}$, which contains $\sigma^2 I_{N\times N}$. For any permutation π, define $(\pi V)_{ij} = (V)_{\pi^{-1}i, \pi^{-1}j}$ for all

i, j. The set $\mathscr{V}$ has the invariance property (2.2) for any permutation π. Due to this invariance assumption on $\mathscr{V}$, we estimate α_t by assuming $V = \sigma^2 I_{N\times N}$. Therefore

$$\hat{\alpha} = \{\hat{\alpha}_t\}_{t=1}^T = \{y_{\cdot t} - y_{\cdot\cdot}\}_{t=1}^T = (I_{T\times T} - T^{-1}J_{T\times T})P\mathbf{y},$$

where

$$\mathbf{y} = (y_1, \cdots, y_N)', \qquad y_{\cdot t} = n_t^{-1}\sum_{u\in I_t} y_{ut}, \qquad y_{\cdot\cdot} = T^{-1}\sum_{t=1}^T y_{\cdot t},$$

$I_{T\times T}$ is the $T \times T$ identity matrix and $J_{T\times T}$ is the $T \times T$ matrix of ones, $P = [a_{tj}]_{T\times N}$ with $a_{tj} = n_t^{-1}$ if $j \in I_t$ and 0 otherwise. For any systematic assignment $I = \{I_t\}_{t=1}^T$, the squared error loss under (2.7) is

$$c(I, V) = \sum_{s<t} E(\hat{\alpha}_s - \hat{\alpha}_t - \alpha_s + \alpha_t)^2 = T \operatorname{tr} \operatorname{Var}\hat{\alpha} = T \operatorname{tr}\{VP'(I_{T\times T} - T^{-1}J_{T\times T})P\}, \tag{2.8}$$

where $c(I, V)$ depends on I through P and tr A is the trace of matrix A. *A design* η^* *is called minimax with respect to* $\mathscr{V}$ *if it achieves*

$$\min_\eta \max_{V\in\mathscr{V}} r_2(\eta, V), \quad \text{where} \quad r_2(\eta, V) = \sum_{I\in\mathscr{I}} \eta(I)c(I, V).$$

THEOREM 3. *Suppose* $\mathscr{V}$ *satisfies the invariance property* (2.2), *i.e.* $V \in \mathscr{V} \Rightarrow \pi V \in \mathscr{V}$ *for all permutations* π, *then the balanced completely randomized design is minimax with respect to* $\mathscr{V}$.

PROOF. From (2.2) and $c(\pi I, \pi V) = c(I, V)$, it can be shown by following the proof of Proposition 1 that $\max_{V\in\mathscr{V}} r_2(\tilde{\eta}, V) \le \max_{V\in\mathscr{V}} r_2(\eta, V)$. From the linearity of $r_2(\tilde{\eta}, V)$ in $\tilde{\eta}$, $r_2(\tilde{\eta}, V) = \sum_{I\in\mathscr{I}} \eta(I) r_2(\tilde{\delta}_I, V)$. It is proved in the Appendix that

$$r_2(\tilde{\delta}_I, V) = \frac{T-1}{N-1}\operatorname{tr}\left\{V\left(I_{N\times N} - \frac{1}{N}J_{N\times N}\right)\right\}\sum_{t=1}^T \frac{1}{n_t}, \tag{2.9}$$

where $|I_t| = n_t$. The maximum of (2.9) is attained by a V^* irrespective of the choice of I and is minimized by choosing $\{n_t\}_{t=1}^T$ to minimize $\sum_{t=1}^T n_t^{-1}$. The minimax design is any $\tilde{\delta}_I$ with the most balanced $I = \{I_t\}_{t=1}^T$, i.e. the balanced completely randomized design. □

If the experimental units can be divided into blocks of units such that the units within blocks are more homogeneous than the units between blocks, then model (2.1) can be refined as

$$y_{ut} = \alpha_t + \beta_i + g_u + \epsilon_{ut}, \tag{2.10}$$

where α_t, g_u and ϵ_{ut} are defined as in (2.1) and β_i is the ith block effect. Since the block effects, which are part of the "old" unit effects in (2.1), have been explicitly included in (2.2), our knowledge of the "new" unit effects in (2.2) is vague. We therefore assume that $\{g_u\}$ can be any element in an invariant set G to be defined below. For simplicity, we assume that there are b blocks of T units and in each block T different treatments are assigned to the T units. Let each such systematic complete block design be denoted by I. From the vagueness of our knowledge on g, we will use the best linear unbiased estimator when $g \equiv c$ in comparing designs. Since the designs under consideration are equi-replicated within each block, we have, for each I, $\hat{\alpha}_s - \hat{\alpha}_t = y_{\cdot s} - y_{\cdot t}$ where $y_{\cdot s} = T^{-1}\sum_{u\in I_s} y_{us}$ and I_s is the collection of units receiving treatment s under I. The definitions of $a(I, g)$, $\mathscr{I}$, η, $r(\eta, g)$ and minimaxity with respect to G are analogous to those for model (2.1). Suppose G satisfies the condition that the g_u's in each block are invariant with respect to the group of permutations of $\{1, \cdots, T\}$, then the randomized complete block design (design with uniform measure on $\mathscr{I}$) is minimax with respect to G. The proof is analogous to that of Proposition 1.

If the class of competing designs is enlarged to include connected designs with unequal replications within blocks, the form of the best linear unbiased estimator under $g \equiv c$ is too complicated to render a simple proof like the one for Proposition 1. But the minimaxity of the randomized complete block design over the enlarged class may still be true.

The following model for Latin square designs is proposed in the same spirit as models (2.1) and (2.10):

$$y_{ut} = \alpha_t + \beta_i + r_j + g_u + \epsilon_{ut}, \tag{2.11}$$

where α_t, β_i, r_j are the t-th treatment effect, the ith row effect and the jth column effect ($i, j, t = 1, \cdots, T$), g_u is the uth unit effect, $u = (i, j)$; $g = \{g_u\}_u$ can be any element in a set G. We also make the usual assumptions on $\{\epsilon_{ut}\}_u$.

Since there does not exist a simple transformation which maps a transformation set of Latin squares (L.S.) into another set, the invariance technique used before will be applied to each transformation set. Let $\mathscr{I} = \cup_{i=1}^k \mathscr{I}_i$ be the totality of transformation sets of Latin squares. Formally we define, for $i = 1, \cdots, k$,

$$\mathscr{I}_i = \{\pi S_i : \pi \in Q\},$$

where S_i is a generating L.S. for $\mathscr{I}_i$ and Q is the group of permutations of rows, columns and treatments. A randomized Latin square design is defined as a probability measure η over $\mathscr{I}$. Note that here we do not identify the identical L.S.'s among the $(T!)^3$ squares, i.e. every square in $\mathscr{I}_i$ is considered "different".

Motivation of the following definitions were given before for models (2.1) and (2.10). For any systematic Latin square design I, we use the estimator $\hat{\alpha}_s - \hat{\alpha}_t = y_{\cdot s} - y_{\cdot t}$ where $y_{\cdot s}$ is the average of observations receiving treatment s. The definitions of $a(I, g)$, $r(\eta, g)$ and minimaxity with respect to G are analogous to those for model (2.1). For example,

$$a(I, g) = T \textstyle\sum_{t=1}^T (g_{\cdot t} - \bar{g})^2 + \sigma^2(T-1) \quad \text{with} \quad \bar{g} = T^{-2} \textstyle\sum_{u=1}^{T^2} g_u$$

and

$$r(\eta, g) = T^{-1}\{\textstyle\sum_{u=1}^{T^2} (g_u - \bar{g})^2 + \textstyle\sum_{u \neq u'}^{T^2} (g_u - \bar{g})(g_{u'} - \bar{g})\pi_{uu'}\} + \sigma^2(T-1) \tag{2.12}$$

with $\pi_{uu'} = \text{Prob}(u \text{ and } u' \text{ receive the same treatment})$.

Write $\eta(I) = \eta(I \mid \mathscr{I}_i)c_i$ with c_i equal to the probability of choosing the transformation set $\mathscr{I}_i$ containing I and $\eta(I \mid \mathscr{I}_i)$ equal to the conditional probability of choosing I within $\mathscr{I}_i$. To simplify the notation, we write $\eta(I \mid i)$ for $\eta(I \mid \mathscr{I}_i)$.

THEOREM 4. *Suppose G satisfies the invariance property* (2.2) *for all $\pi \in Q$. Then any design with equal $\pi_{uu'}$ for all $u \neq u'$ is minimax with respect to G.*

PROOF. For each η, define $\tilde{\eta}(I \mid i) = (T!)^{-3} \sum_{\tau \in Q} \eta(\tau I \mid i)$ and $\tilde{\eta}(I) = \tilde{\eta}(I \mid i)c_i$ for all $I \in \mathscr{I}_i$ and $i = 1, \cdots, k$. Then, using (2.2) together with the fact that $a(\tau I, \tau g) = a(I, g)$, we have

$$\begin{aligned} \max_{g \in G} r(\tilde{\eta}, g) &= \max_{g \in G} \textstyle\sum_{i=1}^k \sum_{\pi \in Q} c_i \tilde{\eta}(\pi S_i \mid i) a(\pi S_i, g) \\ &\leq (T!)^{-3} \textstyle\sum_{\tau \in Q} \max_{g \in G}\{\sum_{\tau \in Q} \sum_{i=1}^k c_i \eta(\tau\pi S_i \mid i) a(\pi S_i, g)\} \\ &= (T!)^{-3} \textstyle\sum_{\tau \in Q} \max_{g \in G}\{\sum_{\lambda \in Q} \sum_{i=1}^k c_i \eta(\lambda S_i \mid i) a(\lambda S_i, g)\} \\ &= \max_{g \in G}\{\textstyle\sum_{i=1}^k c_i \sum_{I \in \mathscr{I}_i} \eta(I \mid i) a(I, g)\} = \max_{g \in G} r(\eta, g). \end{aligned} \tag{2.13}$$

Since $\tilde{\eta}(I \mid i) = (\#\mathscr{I}_i)^{-1}$ for all $I \in \mathscr{I}_i$, $i = 1, \cdots, k$, (2.13) reduces the candidates for minimax designs to the subclass of designs $\tilde{\eta}$ with $\tilde{\eta}(\cdot \mid \mathscr{I}_i)$ the uniform measure on $\mathscr{I}_i$ for each i. From (2.12), $\max_{g \in G} r(\eta, g)$ depends only on $\{\pi_{uu'}\}_{u \neq u'}$. The proof is completed by noting that any $\tilde{\eta}$ with $\tilde{\eta}(\cdot \mid \mathscr{I}_i)$ the uniform measure on $\mathscr{I}_i$, $i = 1, \cdots, k$, give equal $\pi_{uu'}$ for all $u \neq u'$. □

In particular, complete randomization within a transformation set is minimax. This will greatly simplify the randomization procedure for Latin squares in the Fisher-Yates Tables (1953). Instead of randomizing first over the class of transformation sets and then within a particular set (the Fisher-Yates "recipe"), it is sufficient to consider any procedure η

with equal π_{uv}. The simplest one is to choose any Latin square and then randomly permute its rows or columns. This is especially convenient for higher order Latin squares where the class of transformation sets is not available. It is interesting to note that both the equal $\pi_{uu'}$ procedure and the Fisher-Yates procedure were mentioned in Fisher's definition of Latin Square in his 1926 paper. However, no justification for the equal $\pi_{uu'}$ procedure was given there:

> Consequently, the term Latin Square should only be applied to a process of randomization by which one is selected at random out of the total number of Latin Squares possible; or, at least, to specify the agricultural requirement more strictly, out of a number of Latin Squares in the aggregate, of which every pair of plots, not in the same row or column, belongs equally frequently to the same treatment (Fisher, 1926).

The same invariance technique can be used to show that some other randomized designs are minimax over a specified class of competing designs with respect to an invariant G. For example, in the BIBD case, the standard method of randomizing the blocks, the units within each block and the treatment numbers is minimax with respect to such a G; in the first order multi-factor design, the device of "angular randomization" (Box, 1952) makes the design minimax with respect to the set of second order models.

3. Comparisons of some randomized designs in terms of maximum bias squares. In this section, we attempt to evaluate more precisely the gains and losses in using various randomization procedures. The efficiency comparisons will be made for model (2.1) and $n_t = n$, under which

$$r(\eta, g) = Tn^{-2}s(\eta, g) + \sigma^2 T(T-1)n^{-1} \tag{3.1}$$

and

$$s(\eta, g) = \textstyle\sum_{u=1}^{N} (g_u - \bar{g})^2 + \sum_{u\neq v} (g_u - \bar{g})(g_v - \bar{g})\pi_{uv}, \tag{3.2}$$

where $\bar{g} = N^{-1} \sum_{u=1}^{N} g_u$ and π_{uv} is the probability that units u and v receive the same treatment under η. Since the variance of technical error in (3.1) is independent of η, comparisons of designs will be based solely on the bias squares. (The randomization models considered by Kempthorne and his coauthors do not include the technical error term, i.e. $\sigma^2 = 0$. Our efficiency criterion (3.3) is closer to the traditional one in such models.) More specifically, for designs η_1, η_2, define the relative efficiency of η_1 to η_2 to be

$$\max_{g\in G} s(\eta_2, g)/\max_{g\in G} s(\eta_1, g) \tag{3.3}$$

for model (2.1), where

$$G = \{g : | g_u - \bar{g} | \leq c, \qquad u = 1, \cdots, N\}. \tag{3.4}$$

Comparison of designs in terms of bias squares is no stranger in experimental design theory. Box and Draper (1959) in their pioneering paper on robust designs used this kind of criterion in measuring the model-robustness of various designs (see also Karson et al, 1969).

Since $[\pi_{uv}]_{u,v}$ is a nonnegative definite matrix, $s(\eta, g)$ in (3.2) is a convex function in g for any η and its maximum over G is attained at one of the extreme points of G. The set $\mathscr{E}$ of extreme points of G is

$$\{c(1, \cdots, 1, -1, \cdots, -1) \text{ and its permutations}\} \text{ for } N \text{ even}, \tag{3.5}$$

$$\{c(0, 1, \cdots, 1, -1, \cdots, -1) \text{ and its permutations}\} \text{ for } N \text{ odd}, \tag{3.6}$$

where $N/2$ of the components are 1 in (3.5) and $(N-1)/2$ of the components are 1 in (3.6). Since the relative efficiency (3.3) is independent of c, the radius of G, we will assume $c = 1$ in the following efficiency comparisons. The computations are based on (3.2), (3.5) and (3.6). For simplicity details are omitted (see C. F. Wu (1977)).

Example 1. Completely Randomized Design (CRD).

$$\max_{g\in G} s(\eta, g) = \begin{cases} \dfrac{N-n}{N-1}N & \text{for } N \text{ even,} \\ N-n & \text{for } N \text{ odd.} \end{cases}$$

Note that the maximizing g can be any point from the set $\mathscr{E}$.

This calculation amounts to saying that all procedures with equal π_{uv}'s are minimax with respect to G (3.4). In particular, the CRD is minimax.

Example 2. Systematic Design.

The design measure for the systematic design is a point mass $\delta_{\tilde{I}}$ with $\delta_{\tilde{I}}(\tilde{I}) = 1$, $\tilde{I}$ is a pattern. We have

$$\max_{g\in G} s(\delta_{\tilde{I}}, g) = Nn \text{ for } T \text{ even, } Nn - n^2 \text{ for } T \text{ odd.}$$

The relative efficiency with respect to the minimax value is $(N-n)/\{n(N-1)\} < 1/n$ for T even, $T/(nT-1) \simeq 1/n$ for T odd and n even, $(N-n)/(T-1)n^2 = 1/n$ for T, n odd. Therefore the loss of efficiency for the systematic design is proportional to the number of replications of each treatment. For moderate or large n, unless a specific pattern of the model violations is known, it is not advisable to use a systematic design.

Example 3. Randomized Block Design (RBD).

Divide the N units into ℓ blocks $\{I^{(i)}\}_{i=1}^{\ell}$ and assign T treatments each with n/ℓ replications (n being divisible by ℓ) completely randomly to each of the ℓ blocks. Let m_i (or n_i) be the number of u from $I^{(i)}$ with $g_u - \bar{g}$ equal to 1 (or -1). Then $s(\eta, g)$ is

$$\frac{N-n}{N-\ell}\sum_{i=1}^{\ell}(m_i + n_i) - \frac{\ell(N-n)}{N(N-\ell)}\sum_{i=1}^{\ell}(m_i - n_i)^2.$$

This is maximized by taking $|m_i - n_i|$ as small as possible and $m_i + n_i$ as large as possible. We obtain

$$\begin{aligned} \max_{g\in G} s(\eta, g) &= N\frac{N-n}{N-\ell} \quad \text{for } \frac{N}{\ell} \text{ even,} \\ &= (N-n)(N+\ell)/N \quad \text{for } \frac{N}{\ell} \text{ odd and } \ell \text{ even,} \\ &= (N-n)(N+\ell-1)/N \quad \text{for } \frac{N}{\ell} \text{ odd and } \ell \text{ odd.} \end{aligned}$$

The relative efficiency to a CRD is $(N-\ell)/(N-1)$ for N/ℓ even, $N^2/((N+\ell)(N-1))$ for N/ℓ odd and ℓ even, $N/(N+\ell-1)$ for N/ℓ odd and ℓ odd. In particular, $\ell = 1$ is equivalent to a CRD and the relative efficiency is 1. When ℓ/N is close to 0, the relative efficiency is close to 1. For moderate ℓ/N, the loss of efficiency for the randomized block design is still very small.

If G is not of the form (3.4), the relative efficiency of an RBD to a CRD can be greater than 1. For example, if an extreme point g is chosen such that

$$\sum_{i=1}^{\ell}(m_i - n_i)^2 > N, \tag{3.7}$$

then

$$s(\eta_{\mathrm{RBD}}, g) < \frac{N-n}{N-\ell}N - \frac{\ell(N-n)}{N(N-\ell)}N = N-n \le s(\eta_{\mathrm{CRD}}, g).$$

Let $\tilde{G}$ be the smallest convex set spanned by some (or all) extreme points g satisfying (3.7); $\tilde{G}$ is not invariant under the permutation group of $\{1, \cdots, N\}$. Then

$$\max_{g\in\tilde{G}} s(\eta_{\mathrm{RBD}}, g) < \max_{g\in\tilde{G}} s(\eta_{\mathrm{CRD}}, g).$$

Condition (3.7) can be best explained by the effect of blocking. In order to satisfy (3.7), $|m_i - n_i|$ should be large. The extreme case $|m_i - n_i| = N/\ell$ for all i gives the same g_u value for all u in the same block. The common g_u value for the ith block is the ith block effect. For the general situation (3.7), "block effect" still accounts for most of the effects in $\tilde{G}$ and thus makes RBD the more efficient design.

The efficiency comparisons of the Latin square design will be made under model (2.11) and

$$G = \{g: | g_u - \bar{g} | \le 1,\ u = (i, j),\ 1 \le i, j \le T\}.$$

Example 4. Randomized Latin Square Design.

For the randomized Latin square design, $\pi_{uv} = (T-1)^{-1}$ for u, v not in the same row and column and $\max_{g\in G} s(\eta, g) = T^3/(T-1)$ for T even, $(T^3 - 3T + 4)/(T-1)$ for T odd.

Example 5. Systematic Latin Square Design.

For the systematic Latin square design, $\max_{g\in G} s(\eta, g) = T^3$ for T even, $(T-1)T^2$ for T odd. The relative efficiency to a randomized Latin square design is

$$\frac{1}{T-1} \quad \text{for} \quad T \text{ even}, \qquad \frac{T^3 - 3T + 4}{T^3 - T^2}\,\frac{1}{T-1} \quad \text{for} \quad T \text{ odd}.$$

It is also worth noting that the permutations of treatments only gives the same π_{uv}, and thus the same relative efficiency, as the systematic design and hence is not advisable.

Acknowledgments. I thank D. R. Cox, A. Herzberg, J. Kiefer, and a referee for comments. The improvement of results in Section 2 is due to a question raised by the referee. This paper is dedicated to the memory of my teacher and friend, Professor Jack Kiefer.

APPENDIX

For any systematic assignment pattern $I = \{I_t\}_{t=1}^{T}$ with $|I_t| = n_t$, define $g_{\cdot t}(I) = n_t^{-1} \sum_{u\in I_t} g_u$. It can be verified that

$$\frac{1}{N!} \sum_{\pi\in P} g_{\cdot t}^2(\pi I) = \frac{1}{Nn_t} \sum_{u=1}^{N} g_u^2 + \frac{1}{N(N-1)}\left(1 - \frac{1}{n_t}\right) \sum_{u\ne v}^{N} g_u g_v, \tag{4.1}$$

$$\frac{1}{N!} \sum_{\pi\in P} g_{\cdot s}(\pi I) g_{\cdot t}(\pi I) = \frac{1}{N(N-1)} \sum_{u\ne v}^{N} g_u g_v, \tag{4.2}$$

where P is the group of permutations of $\{1, \cdots, N\}$.

Proof of (2.5). This follows from

$$\sum_{t=1}^{T} (g_{\cdot t} - g_{\cdot\cdot})^2 = (1 - T^{-1}) \sum_{t=1}^{T} g_{\cdot t}^2 - T^{-1} \sum_{s\ne t}^{T} g_{\cdot s} g_{\cdot t},$$

(4.1) and (4.2).

Proof of (2.6). This follows from

$$\sum_{t=1}^{T} \sum_{u\in I_t} (g_u - g_{\cdot t})^2 = \sum_{u=1}^{N} g_u^2 - \sum_{t=1}^{T} n_t g_{\cdot t}^2$$

and (4.1).

Proof of (2.9). Without loss of generality, we can assume $I_t = \{\sum_{i=1}^{t-1} n_i + j\}_{j=1}^{n_t}$, $t = 1, \cdots, T$. Then the matrix $P'(I_{T\times T} - T^{-1}J_{T\times T})P$ in (2.8) is equal to $\sum_{i,j=1}^{T} b_{ij} M_{n_i\times n_j}$, where $b_{ii} = (1 - T^{-1})n_i^{-2}$, $b_{ij} = -T^{-1}n_i^{-1}n_j^{-1}$ for $i \ne j$, and $M_{n_i\times n_j} = [m_{k\ell}]_{1\le k,\ell\le N}$ with $m_{k\ell} = 1$ if $k \in I_i$ and $\ell \in I_j$, and 0 otherwise. For any matrix $A = [a_{ij}]$, define $\pi A = [a_{\pi^{-1}i,\pi^{-1}j}]$. Then it can be verified that

(4.3) $$\frac{1}{N!}\sum_{\pi\in P}\pi M_{n_i\times n_i}=\frac{n_i(N-n_i)}{N(N-1)}I_{N\times N}+\frac{n_i(n_i-1)}{N(N-1)}J_{N\times N},$$

(4.4) $$\frac{1}{N!}\sum_{\pi\in P}\pi M_{n_i\times n_j}=\frac{n_in_j}{N(N-1)}(J_{N\times N}-I_{N\times N})\quad\text{for}\quad i\neq j.$$

Now (2.9) follows easily from (4.3) and (4.4).

REFERENCES

BICKEL, P. J. and HERZBERG, A. M. (1979). Robustness of design against autocorrelation in time I: asymptotic theory, optimality for location and linear regression. *Ann. Statist.* **7** 77–95.

BLACKWELL, D. and GIRSHICK, M. A. (1954). *Theory of Games and Statistical Decisions.* Wiley, New York.

BOX, G. E. P. (1952). Multi-factor designs of first order. *Biometrika* **39** 49–57.

BOX, G. E. P. and DRAPER, N. R. (1959). A basis for the selection of a response surface design. *J. Amer. Statist. Assoc.* **54** 622–654.

COX, D. R. (1958). *Planning of Experiments.* Wiley, New York.

FEDOROV, V. V. (1969, Transl. 1972). *Theory of Optimal Experiments.* Academic, New York.

FISHER, R. A. (1926). The arrangement of field experiments. *J. Min. of Agric.* **33** 503–513.

FISHER, R. A. and YATES, F. (1953). *Statistical Tables for Biological, Agricultural and Medical Research.* (Fourth Ed.) Oliver and Boyd, Edinburg.

HARVILLE, D. A. (1975). Experimental randomization. Who needs it? *Amer. Statist.* **29** 27–31.

HUBER, P. J. (1972). Robust Statistics: A review. *Ann. Math. Statist.* **43** 1041–1067.

HUBER, P. J. (1975). Robustness and designs. In *A Survey of Statistical Designs and Linear Models* (J. N. Srivastava, ed.) 287–301. North Holland, Amsterdam.

KEMPTHORNE, O. (1952). *The Design and Analysis of Experiments.* Wiley, New York.

KEMPTHORNE, O. (1955). The randomization theory of experimental inference. *J. Amer. Statist. Assoc.* **50** 946–967.

KEMPTHORNE, O. (1975). Inference from experiments and randomization. In *A Survey of Statistical Designs and Linear Models* (J. N. Srivastava, ed.) 303–331. North Holland, Amsterdam.

KARSON, N. J., MANSON, A. R. and HADER, R. J. (1969). Minimum bias estimation and experimental designs for response surface. *Technometrics* **11** 461–475.

KIEFER, J. (1959). Optimal experimental designs. *J. Roy. Statist. Soc. Ser. B* **21** 272–319.

RUBIN, D. B. (1978). Bayesian inference for causal effects: the role of randomization. *Ann. Statist.* **6** 34–58.

SCHEFFÉ, H. (1958). *The Analysis of Variance.* Wiley, New York.

WU, C. F. (1977). Robustness and efficiency problems of some randomization procedures in experimental designs. Technical Report No. 1797, Mathematics Research Center, University of Wisconsin, Madison.

1210 W. Dayton St.
DEPARTMENT OF STATISTICS
MADISON, WISCONSIN 53706