OPTIMAL STRATEGIES IN FACTORIAL EXPERIMENTS[1]

By S. Ehrenfeld and S. Zacks[2]

New York University, Columbia University, and Technion, Israel Institute of Technology

1. Introduction and summary. The present study is an extension of the previous work of the authors [3], [5], in which two randomization procedures in fractional factorial experiments were investigated. The general problem is to choose, in an optimal manner, a fractional replication of a full factorial system and an estimator, for the purpose of making inferences concerning a subset of pre-assigned parameters. We consider factorial systems of order 2^m, which consist of m factors each at 2 levels. The results can be generalized for factorial systems of order p^m, where $p > 2$. The factorial model adopted for the present study is the same as that adopted in the previous work, and the reader is referred to the preceeding papers [3], [5] for details and properties. The first two sections of [5] are essential for the present paper.

The statistical properties of the conditional least-squares estimators (c.l.s.e.) were studied in [5] with respect to two specified randomization procedures (R.P.I. and R.P.II.), which are particular types of random allocation designs. It was shown that the c.l.s.e.'s constitute a complete class of linear unbiased estimators. The c.l.s.e. can be characterized as least-square estimators *adjusted* for the block of treatment combinations chosen and the information available on the nuisance parameters. In the present study we extend the investigation into the comparison of different randomization procedures. We consider a general class of procedures, characterized as follows: By some confounding method we construct $M = 2^{m-s}$ blocks, each one containing $S = 2^s$ treatment combinations. We choose one of the blocks with an arbitrary probability vector, ξ, and observe the associated random variables. R.P.I. is the special case where every block has the same probability of being chosen. A fixed fractional replication procedure is the special case where one of the blocks is chosen with probability one. Each randomization procedure is represented (uniquely) by a probability vector, ξ, and each c.l.s.e. is represented by a vector γ in $(2^m - 2^s)$-dimensional Euclidean space. A strategy of the Statistician is thus represented by a pair of vectors (ξ, γ). Minimax strategies are studied in the present paper for two states of information concerning the nuisance parameters: (i) all the nuisance parameters are bounded; (ii) all the nuisance parameters are bounded and all their signs are known. As proven in the present paper, the minimax strategy corresponding to Case (i) consists of R.P.I. with an unadjusted c.l.s.e. and is thus independent of the actual bounds

Received March 7, 1962; revised March 7, 1963.

[1] This paper was prepared with the support of the Air Force Office of Scientific Research, under contract AF-AFOR-62-153 with New York University.

[2] Now at Stanford University.

of the nuisance parameters. On the other hand, the minimax strategy for Case (ii) consists of some fixed fractional replication with an adjusted c.l.s.e. which depends on the actual bounds of the nuisance parameters. These minimax theorems are proven with respect to a mean-square-error loss function, defined as the trace of the mean-square-error matrix. A closeness loss function is considered too. The closeness of an estimator is defined as the probability that the values of the estimator will lie in a prescribed neighborhood of the true values of the parameters. In Section 2 the mean-square-error matrix and the closeness of a c.l.s.e., under an arbitrary randomization procedure, ξ, are derived. The mean-square-error and the closeness loss functions are defined in Section 3. Formulae for mean-square-error Bayes strategies, against any given a priori distribution of the nuisance parameters, are then derived. The closeness risk function is approximated by a similar function which has the same Bayes strategies as the mean-square-error loss function. It is shown that when the size of the experiment, $S = 2^s$, grows the closeness risk function and the approximative function converge. The minimax theorems are stated and proved in Section 4 only with respect to the mean-square-error loss function. It can be shown that the minimax closeness strategies, for the states of information studied, are the same as those for the mean-square-error loss function. This can be concluded also from the results of Section 3. In Section 5 we present a numerical example to illustrate the results and the computations involved.

2. Mean-square-error and closeness of conditional least-squares estimators, under randomized fractional replication procedures. Consider a *conditional least-squares estimator* (c.l.s.e.) of a subset of $S = 2^s$ pre-assigned parameters of a $N = 2^m$ factorial system ($s < m$) with a *randomized fractional replication procedure* (R.F.R.P.). As in the preceeding paper [5] assume, without loss of generality, that the subvector, α, of $S = 2^s$ pre-assigned parameter consists of the first parameters, namely $\beta_0, \cdots, \beta_{S-1}$ of the linear factorial model. Furthermore, assume that the $m - s$ defining parameters are the main-effects $\beta_S, \beta_{2S}, \cdots, \beta_{N-S}$. Accordingly, the treatment combinations of the full factorial system are classified into 2^{m-s} blocks $\{X_v : v = 0, \cdots, M - 1\}$, where $X_v = \{x_t : t = i + vS; i = 0, \cdots, S - 1\}$. Denote by $Y(X_v)$ the vector of random variables (observation vector) associated with block

$$X_v \ (v = 0, \cdots, M - 1).$$

Then, the factorial model for $Y(X_v)$ is

$$Y(X_v) = (C^{(S)})\alpha + (H_v)\beta + \epsilon \qquad \text{for all } v = 0, \cdots, M - 1 \tag{2.1}$$

$(C^{(S)})$ is the matrix of coefficients corresponding to a S factorial system, having the properties: $(C^{(S)})'(C^{(S)}) = SI^{(S)}$, and $1^{(S)\prime}(C^{(S)}) = (S, 0, \cdots, 0)$. β designates the vector of $N - S$ nuisance parameters ($N = 2^m$), i.e., $\beta' = (\beta_S, \cdots, \beta_{N-1})$. (H_v) is a rectangular matrix of order $S \times (N - S)$, given for each $v = 0, \cdots, M - 1$ by:

$$(H_v) = (c_{v1}^{(M)}, \cdots, c_{v(M-1)}^{(M)}) \otimes (C^{(S)}) \tag{2.2}$$

and where $(C^{(M)}) = \|c_{vu}^{(M)}\|; v, u = 0, \cdots, M - 1$. ϵ is a random vector, of order $S \times 1$, independent of X_v $(v = 0, \cdots, M - 1)$, with $E\{\epsilon\} = 0$ and $E\{\epsilon\epsilon'\} = \sigma^2 I^{(S)}$.

A R.F.R.P. is a procedure according to which a block (or generally n blocks, $1 \leqq n < M$) of treatment combinations, X_v, is chosen with a specified probability ξ_v. Thus, a R.F.R.P. is represented by a probability vector $\xi = (\xi_0, \cdots, \xi_{M-1})'$ of order M. As proven in the preceding paper, if all the blocks $X_v (v = 0, \cdots, M - 1)$ are chosen with equal probabilities, then the class of all c.l.s.e. constitutes a complete class of unbiased linear estimators of the preassigned parameters α. However, if $\xi \neq (1/M)1^{(M)}$ there is no unbiased linear estimator of α. Thus, we measure the precision of a c.l.s.e., under an arbitrary R.F.R.P. ξ, by its mean-square-error matrix, and by its closeness (in probability) to the true point α. We derive first the expression for the mean-square-error matrix.

A c.l.s.e. of α, designated by $\hat{\alpha}(\gamma)$, is given by the formula:

$$\hat{\alpha}(\gamma) = (1/S)(C^{(S)})'[Y(X_v) - (H_v)\gamma] \tag{2.3}$$

where γ is any fixed point in an $(N - S)$-dimensional Euclidean space. Substituting (2.1) for $Y(X_v)$ in (2.3) we get:

$$\hat{\alpha}(\gamma) = \alpha + (1/S)(C^{(S)})'(H_v)(\beta - \gamma) + (1/S)(C^{(S)})'\epsilon. \tag{2.4}$$

Let $D_\xi(\hat{\alpha}(\gamma))$ denote the mean-square-error matrix of $\hat{\alpha}(\gamma)$, under R.F.R.P. ξ defined to be

$$D_\xi(\hat{\alpha}(\gamma)) = E_\xi\{(\hat{\alpha}(\gamma) - \alpha)(\hat{\alpha}(\gamma) - \alpha)'\}. \tag{2.5}$$

According to (2.4) and the assumptions about ϵ, we obtain:

$$\begin{aligned} D_\xi(\hat{\alpha}(\gamma)) &= (\sigma^2/S)I^{(S)} \\ &\quad + (1/S^2)(C^{(S)})'E_\xi\{(H_v)(\beta - \gamma)\cdot(\beta - \gamma)'(H_v)'\}(C^{(S)}). \end{aligned} \tag{2.6}$$

By (2.2) we can write, for each $v = 0, \cdots, M - 1$,

$$(1/S)(C^{(S)})'(H_v)(\beta - \gamma) = \sum_{u=1}^{M-1} c_{vu}^{(M)}(\beta - \gamma)_u \tag{2.7}$$

where $(\beta - \gamma)' = ((\beta - \gamma)_0', (\beta - \gamma)_1', \cdots, (\beta - \gamma)_{M-1}')$ i.e.,

$$(\beta - \gamma)_u (u = 1, \cdots, M - 1)$$

are subvectors of order $S \times 1$. Substituting (2.7) in (2.6) we attain:

$$D_\xi(\hat{\alpha}(\gamma)) = (\sigma^2/S)\, I^{(S)} + \sum_{u_1=1}^{M-1} \sum_{u_2=1}^{M-1} \left(\sum_{v=0}^{M-1} \xi_v c_{vu_1}^{(M)} c_{vu_2}^{(M)} \right) \cdot (\beta - \gamma)_{u_1} (\beta - \gamma)_{u_2}'. \tag{2.8}$$

In particular, for $\xi^* = (1/M)1^{(M)}$ we have

$$D_{\xi^*}(\hat{\alpha}(\gamma)) = (\sigma^2/S)I^{(S)} + \sum_{u=1}^{M-1} (\beta - \gamma)_u(\beta - \gamma)_u' . \tag{2.9}$$

This results from the orthogonality of the row and column vectors of $(C^{(M)})$. The closeness of a c.l.s.e., $\hat{\alpha}(\gamma)$, to α is the probability, under R.F.R.P. ξ, that $\hat{\alpha}(\gamma)$ lies within a prescribed neighborhood of α. Let N_ρ denote a spherical neighborhood about α, with radius $\rho \geqq 0$, i.e., $N_\rho = \{\hat{\alpha}: |\hat{\alpha} - \alpha| \leqq \rho\}$. Denote by $\psi_\xi(\rho; \alpha, \beta, \gamma)$ the closeness of $\hat{\alpha}(\gamma)$ to α under ξ. Then,

$$\psi_\xi(\rho; \alpha, \beta, \gamma) = P_\xi\{|\hat{\alpha}(\gamma) - \alpha| \leqq \rho\}. \tag{2.10}$$

To evaluate (2.10) we have to specify besides ξ also the distribution of ϵ. Thus, as commonly practiced, assume that ϵ has a multivariate normal distribution, with mean zero and dispersion matrix $\sigma^2 I^{(S)}$; i.e., $\mathcal{L}(\epsilon) = \mathfrak{N}(0, \sigma^2 I^{(S)})$. Furthermore, from (2.4) we obtain that

$$\begin{aligned} |\hat{\alpha}(\gamma) - \alpha|^2 = (1/S)(\beta - \gamma)'(H_v)'(H_v)(\beta - \gamma) \\ + (2/S)(\beta - \gamma)'(H_v)'\epsilon + (1/S)\epsilon'\epsilon. \end{aligned} \tag{2.11}$$

Hence, the conditional distribution of $|\hat{\alpha}(\gamma) - \alpha|^2$ given X_v, is like that of $(\sigma^2/S)\chi^2[S; \lambda_v(\beta, \gamma)]$; where $\chi^2[f; \lambda]$ denotes the non-central chi-square statistic, with f degrees of freedom and a parameter of non-centrality λ. From (2.11) we obtain that, for each $v = 0, \cdots, M - 1$,

$$\lambda_v(\beta, \gamma) = \frac{S}{2\sigma^2} \sum_{i=0}^{S-1} \left[\sum_{u=1}^{M-1} c_{vu}^{(M)} (\beta_{i+uS} - \gamma_{i+uS}) \right]^2 . \tag{2.12}$$

Denote by $F(r \mid f; \lambda)$ the c.d.f. of $\chi^2[f; \lambda]$ at the point r. $F(r \mid f; \lambda)$ can be represented as a mixture of central chi-square c.d.f.'s, $F(r \mid f_m; 0)$, by the formula:

$$F(r \mid f; \lambda) = e^{-\lambda} \sum_{m=0}^{\infty} (\lambda^m/m!) F(r \mid f + 2m; 0) \tag{2.13}$$

where, for every $m = 0, 1, \cdots$

$$\begin{aligned} F(r \mid f + 2m; 0) &= 0, & \text{if } r < 0 \\ &= \frac{1}{2^{m+f/2}\,\Gamma(m + f/2)} \int_0^r e^{-t/2} t^{m+f/2-1}\, dt, & \text{if } r \geqq 0 \end{aligned} \tag{2.14}$$

(see Kempthorne [4], pp. 221).

It follows that the closeness of a c.l.s.e., $\hat{\alpha}(\gamma)$, is independent of α, and is given by

$$\psi_\xi(\rho; \beta, \gamma) = E_\xi\{F(r \mid S; \lambda_v(\beta, \gamma))\} = \sum_{v=0}^{M-1} \xi_v F(r \mid S; \lambda_v(\beta, \gamma)) \tag{2.15}$$

where $r = \rho^2 S/\sigma^2$. In particular, $\psi_\xi(\rho; \beta, \beta) = F(r \mid S; 0) > \psi_\xi(\rho; \beta, \gamma)$ for all $\beta \neq \gamma$, since $F(r \mid f; \lambda)$ is a monotonically decreasing function of λ, at every r and f.

3. The decision framework and Bayes strategies. The decision problem for the Statistician is to choose a R.F.R.P. and a c.l.s.e. in some optimal manner. The

Statistician's strategies for this decision problem are thus represented by pairs of vectors (ξ, γ). ξ is a $M \times 1$ probability vector, γ is any vector in $E^{(N-S)}$. We consider two kinds of loss functions. One based on the mean-square-error matrix; the other based on the closeness of $\hat{\alpha}(\gamma)$ under ξ. Accordingly, let

$$L_1(\xi, \gamma; \beta) = \operatorname{tr}\left\{D_\xi(\hat{\alpha}(\gamma)) - \frac{\sigma^2}{S} I^{(S)}\right\} = \sum_{v=0}^{M-1} \xi_v \sum_{i=0}^{S-1} \left[\sum_{u=1}^{M-1} c_{vu}^{(M)}(\beta_{i+uS} - \gamma_{i+uS})\right]^2 \tag{3.1}$$

be the *mean-square-error loss function*; and let

$$L_2(\xi, \gamma; \beta) = F(r \mid S; 0) - \sum_{v=0}^{M-1} \xi_v F(r \mid S; \lambda_v(\beta, \gamma)) \tag{3.2}$$

be the *closeness loss function*. Both $L_1(\xi, \beta; \beta) = 0$ and $L_2(\xi, \beta; \beta) = 0$ for every ξ. Indeed, when β is known the present decision problem has a trivial solution.

Let $\eta(\beta)$ be an a priori distribution (c.d.f.) of β in $E^{(N-S)}$. Define the corresponding risk functions,

$$R_1(\xi, \gamma; \eta) = \sum_{v=0}^{M-1} \xi_v \int \sum_{i=0}^{S-1} \left[\sum_{u=1}^{M-1} c_{vu}^{(M)} (\beta_{i+uS} - \gamma_{i+uS})\right]^2 d\eta(\beta) \tag{3.3}$$

and

$$R_2(\xi, \gamma; \eta) = F(r \mid S; 0) - \sum_{v=0}^{M-1} \xi_v \int F(r \mid S; \lambda_v(\beta, \gamma))\, d\eta(\beta). \tag{3.4}$$

Assume that $E_\eta\{\lambda_v^2(\beta, \gamma)\} \leqq K < \infty$ for all $v = 0, \cdots, M-1$; i.e., the risk function $R_1(\xi, \gamma; \eta)$ is bounded. For S sufficiently large, i.e., $S \geqq S_0(\epsilon, r, \Lambda_v(\eta, \gamma))$, we can approximate $R_2(\xi, \gamma; \eta)$ by

$$R_2^*(\xi, \gamma; \eta) = F(r \mid S; 0) - \sum_{v=0}^{M-1} \xi_v F(r \mid S; \Lambda_v(\eta, \gamma)) \tag{3.5}$$

so that $|R_2(\xi, \gamma; \eta) - R_2^*(\xi, \gamma; \eta)| < \epsilon$, where

$$\Lambda_v(\eta, \gamma) = \int \lambda_v(\beta, \gamma)\, d\eta(\beta), \qquad \text{for } v = 0, \cdots, M-1. \tag{3.6}$$

Indeed, since $(d/d\lambda)F(r \mid S; \lambda) = F(r \mid S+2; \lambda) - F(r \mid S; \lambda) < 0$ for all λ we obtain, by expanding $F(r \mid S; \lambda_v(\beta, \gamma))$ about $\Lambda_v(\eta, \gamma)$, the inequality

$$\int F(r \mid S; \lambda_v(\beta, \gamma))\, d\eta(\beta) < F(r \mid S; \Lambda_v(\eta, \gamma)) \tag{3.7}$$

for each $v = 0, \cdots, M-1$. With the assumption that $E_\eta\{\lambda_v^2(\beta, \gamma)\}$, $v = 0, \cdots, M-1$, are uniformly bounded we obtain:

$$\int F(r \mid S; \lambda_v(\beta, \gamma))\, d\eta(\beta) = F(r \mid S; \Lambda_v(\eta, \gamma)) + o(F''(r \mid S; \Lambda_v(\eta, \gamma))), \tag{3.8}$$

as

$$F''(r \mid S; \Lambda_v(\eta, \lambda)) \to 0;$$

where $F''(r \mid S; \Lambda_v(\eta,\gamma)) = (d^2/d\lambda^2)F(r \mid S; \lambda) |_{\lambda=\Lambda_v(\eta,\gamma)}$. However, $F''(r \mid S; \lambda) = F(r \mid S + 4; \lambda) - 2F(r \mid S + 2, \lambda) + F(r \mid S; \lambda) \to 0$ as $S \to \infty$. Since $S = 2^s$ grows exponentially with s, we expect the approximation to be adequate even for experiments of moderate size.

THEOREM 3.1. *Let $\eta(\beta)$ be an a priori distribution of β, then the Bayes strategies $(\xi^0(\eta), \gamma^0(\eta))$ against $\eta(\beta)$, with respect to $R_1(\xi, \gamma; \eta)$ and $R_2^*(\xi, \gamma; \eta)$ are the same.*

PROOF. Let $\xi^{(v)}$, $v = 0, \cdots, M - 1$, denote the probability vector whose vth component is 1 and all the other components are 0. Thus,

$$(3.9) \quad \begin{aligned} R_1(\xi^{(v)}, \gamma; \eta) &= \sum_{i=0}^{S-1} \int \left[\sum_{u=1}^{M-1} c_{vu}^{(M)} (\beta_{i+uS} - \gamma_{i+uS}) \right]^2 d\eta(\beta) \\ &= \frac{2\sigma^2}{S} \int \lambda_v(\beta, \gamma)\, d\eta(\beta) \\ &= \frac{2\sigma^2}{S} \Lambda_v(\eta, \gamma), \qquad \text{for } v = 0, \cdots, M - 1. \end{aligned}$$

$R_1(\xi^{(v)}, \gamma; \eta)$, for all $v = 0, \cdots, M - 1$, are strictly convex functions of γ, attaining their minimum at $\gamma^0(\eta)$, whose components are given by

$$(3.10) \quad \gamma_{i+uS}^0(\eta) = \int \beta_{i+uS}\, d\eta(\beta), \quad \text{for all } i = 0, \cdots, S - 1;$$

and all $u = 1, \cdots, M - 1$. Let

$$(3.11) \quad R_1^0(\eta) = \min_{v=0,\cdots,M-1} R_1(\xi^{(v)}, \gamma^0(\eta); \eta)$$

and let v_0 be any integer from 0 to $M - 1$ for which $R_1(\xi^{(v_0)}, \gamma^0(\eta); \eta) = R_1^0(\eta)$. Then $(\xi^{(v_0)}, \gamma^0(\eta))$ is a Bayes strategy against $\eta(\beta)$, relative to $R_1(\xi, \gamma; \eta)$. Finally, according to (3.9), $\gamma^0(\eta)$ minimizes $\Lambda_v(\eta, \gamma)$ for all $v = 0, \cdots, M - 1$. Hence $R_2^*(\xi^{(v)}, \gamma^0(\eta); \eta) = \min_\gamma R_2^*(\xi^{(v)}, \gamma; \eta)$, for all $v = 0, \cdots, M - 1$, because $F(r \mid S; \Lambda_v(\eta, \gamma)$ is a monotonically decreasing function of $\Lambda_v(\eta, \gamma)$. Finally, if $\Lambda_{v_1}(\eta, \gamma) < \Lambda_{v_2}(\eta, \gamma)$ then $F(r \mid S; \Lambda_{v_1}(\eta, \gamma)) > F(r \mid S; \Lambda_{v_2}(\eta, \gamma))$ or, equivalently, $R_2^*(\xi^{(v_1)}, \gamma^0(\eta); \eta) < R_2^*(\xi^{(v_2)}, \gamma^0(\eta); \eta)$. Thus, if $(\xi^{(v_0)}, \gamma^0(\eta))$ is a Bayes strategy against $\eta(\beta)$, relative to $R_1(\xi, \gamma; \eta)$, then it is also a Bayes strategy against $\eta(\beta)$, relative to $R_2^*(\xi, \gamma; \eta)$.

In the following section we prove two minimax theorems, with respect to the mean-square-error loss function $L_1(\xi, \gamma; \eta)$. As a result of Theorem 3.1 and the preceding discussion on the approximation of $R_2^*(\xi, \gamma; \eta)$ to the closeness loss-function we expect the same minimax strategies to hold also for the closeness loss function. Indeed, one can prove it directly by studying the closeness loss function $L_2(\xi, \gamma; \beta)$. To save space, we shall present the minimax theorems only for the mean-square-error loss function.

4. Minimax mean-square-error strategies. Let $(E^{(N-S)}, \mathcal{B}^{(N-S)})$ be a probability space, where $\mathcal{B}^{(N-S)}$ is the Borel field of sets over $E^{(N-S)}$. Let $B \varepsilon \mathcal{B}^{(N-S)}$ be a set in $\mathcal{B}^{(N-S)}$ and $\mathcal{H}(B)$ the class of all probability measures over $(E^{(N-S)}, \mathcal{B}^{(N-S)})$, which assign probability zero to points outside B. In the present section we study the minimax strategies relative to the case where: (i) B is an $(N - S)$-dimensional hypersphere, i.e., $B = \{\beta: |\beta - \beta^0| \leqq R\}$ for some β^0 and R; and (ii) B is the positive orthant of an N-dimensional hypersphere, i.e., $B = \{\beta: |\beta - \beta^0| \leqq R, \text{ and } (\beta_t - \beta_t^0) \geqq 0 \text{ for all } t = 0, \cdots, N - S - 1\}$. The result is later generalized to an arbitrary orthant of an N-dimensional hypersphere.

THEOREM 4.1. *If B is an $(N - S)$-dimensional hypersphere, centered at β^0, with radius R then, a minimax and admissible strategy is (ξ^*, β^0), where $\xi^* = (1/M)1^{(M)}$; and the minimax risk is R^2.*

PROOF. Without loss of generality, let $\beta^0 = 0$. First we notice that for every $\eta \varepsilon \mathcal{H}(B)$, where $B = \{\beta: |\beta| \leqq R\}$, the following holds

$$R_1(\xi^*, 0; \eta) = \int_B |\beta|^2 d\eta(\beta) \leqq R^2. \tag{4.1}$$

Accordingly, every minimax strategy of nature is some mixture of the boundary points of B. In other words, it suffices to consider the subclass of strategies $\mathcal{H}(B^*)$, where $B^* = \{\beta: |\beta| = R\}$. Thus, $R_1(\xi^*, 0; \eta) = R^2$ for all $\eta \varepsilon \mathcal{H}(B^*)$. It remains, hence, to show that $(\xi^*, 0)$ is a Bayes strategy against some $\eta \varepsilon \mathcal{H}(B^*)$. Let $\eta^0 \varepsilon \mathcal{H}(B^*)$ be a d.f. concentrating on two points $(\beta^*, -\beta^*)$, i.e.,

$$\begin{aligned} \eta^0(\beta) &= \tfrac{1}{2}, & &\text{if } \beta = \pm\beta^*, \beta^* \varepsilon B^* \\ &= 0, & &\text{otherwise.} \end{aligned} \tag{4.2}$$

The Bayes strategy against $\eta^0(\beta)$, where β^* is any point on the boundary of B, is $\gamma^0(\eta^0)$ whose components are

$$\gamma_t^0(\eta^0) = 0 \quad \text{for all} \quad t = 0, \cdots, N - S - 1. \tag{4.3}$$

Thus $R_1(\xi, \gamma; \eta^0) \geqq R_1(\xi, 0; \eta^0)$ for all ξ and γ. It remains to show that $R_1(\xi, 0; \eta^0) \geqq R_1(\xi^*, 0; \eta^0)$ for all probability vectors ξ. Indeed,

$$R_1(\xi, 0; \eta^0) = \sum_{v=0}^{M-1} \xi_v \sum_{i=0}^{S-1} \left(\sum_{u=1}^{M-1} c_{vu}^{(M)} \beta_{i+uS}^* \right)^2. \tag{4.4}$$

If $\xi \neq \xi^*$ then $\xi_{v_0} = \max_{v=0,\cdots,M-1} \{\xi_v\} > 1/M$. Since β^* can be any point on the boundary of B, let

$$\beta_{i+uS}^* = c_{v_0 u}^{(M)}[R/(N - S)^{\frac{1}{2}}], \tag{4.5}$$

for all $i = 0, \cdots, S - 1$; and all $u = 1, \cdots, M - 1$. Then, substituting (4.5) in (4.4) we get

$$\begin{aligned} R_1(\xi, 0; \eta^0) &= \sum_{v=0}^{M-1} \xi_v \sum_{i=0}^{S-1} \left(\sum_{u=1}^{M-1} c_{vu}^{(M)} c_{v_0 u}^{(M)} \frac{R}{(N - S)^{\frac{1}{2}}} \right)^2 \\ &= R^2 \xi_{v_0} \frac{S(M - 1)^2}{(M - 1)S} + \frac{R^2}{(M - 1)} \sum_{v \neq v_0} \xi_v \left(\sum_{u=1}^{M-1} c_{vu}^{(M)} c_{v_0 u}^{(M)} \right)^2. \end{aligned} \tag{4.6}$$

But if $v \neq v_0$ then $\sum_{u=1}^{M-1} c_{vu}^{(M)} c_{v_0 u}^{(M)} = -1$. Hence,

$$(4.7) \quad R_1(\xi, 0; \eta^0) = R^2\{\xi_{v_0}(M-1) + [1/(M-1)](1-\xi_{v_0})\} \geqq R^2.$$

Equality holds if, and only if, $\xi_{v_0} = 1/M$. Thus, $(\xi^*, 0)$ is the unique minimax strategy. Therefore it is an admissible one. To prove that the minimax risk is R^2, substitute ξ^* in (4.7).

We turn now to the case where, in addition to the information that β is bounded, all the signs of the components of β are known. This case is represented by the set

$$(4.8) \quad B = \{\beta: |\beta| \leqq R \text{ and } \operatorname{sgn}(\beta_t) = (-1)^{i_t} \text{ for all } t = 0, \cdots, N-S-1; \quad \text{where } i_t = 0, 1\}.$$

There is no loss in generality to assume that all the components of β are positive. Indeed, by applying the orthogonal transformation,

$$(4.9) \quad \beta^* = \begin{bmatrix} (-1)^{i_0} & & & 0 \\ & \cdot & & \\ & & \cdot & \\ 0 & & & (-1)^{i_{N-S-1}} \end{bmatrix} \beta$$

the set (4.8) is rotated into the positive orthant. Furthermore, it is sufficient to prove the result for the positive simplex generated by the points $\beta^{(k)}$, $k = 0, \cdots, N-S$; where $\beta^{(0)} = 0$ and $\beta^{(k)}$ for $k = 1, \cdots, N-S$ are the orthogonal unit vectors, whose components are given by:

$$(4.10) \quad \begin{aligned} \beta_t^{(k)} &= 1, && \text{if } t = k-1 \\ &= 0, && \text{otherwise} \end{aligned}$$

for all $t = 0, \cdots, N-S-1$.

THEOREM 4.2. *If B is the simplex*

$$(4.11) \quad B = \{\beta: \beta_t \geqq 0 \text{ for all } t = 0, \cdots, N-S-1; \quad \text{and } \sum_{t=0}^{N-S-1} \beta_t \leqq R\}$$

then $(\xi^{(M-1)}, [R/(N-S)^{\frac{1}{2}}]1^{(N-S)})$ is a minimax and admissible strategy, with minimax risk $R^2(1-1/S)$.

PROOF. Without loss of generality, assume $R = 1$. Any good strategy of Nature consists of some mixture of the $N - S + 1$ extreme points

$$\beta^{(k)} \quad (k = 0, \cdots, N-S).$$

Let $\mathcal{H}^*(B)$ be the set of all $(N-S+1) \times 1$ probability vectors, $\eta' = (\eta_0, \cdots, \eta_{N-S})$, representing mixtures of the points $\beta^{(k)}$. Furthermore, let $\mathcal{H}_0^*(B) \subset \mathcal{H}^*(B)$ be the subset of all $(N-S+1) \times 1$ probability vectors whose first component $\eta_0 = 0$. For every $\eta \varepsilon \mathcal{H}^*(B)$ we have:

$$R_1(\xi, \gamma; \eta) = \sum_{v=0}^{M-1} \xi_v \sum_{k=0}^{N-S} \eta_k L_1(\xi^{(v)}, \gamma; \beta^{(k)}) \tag{4.12}$$

where, for all $v = 0, \cdots, M - 1$,

$$\begin{aligned} L_1(\xi^{(v)}, \gamma; \beta^{(k)}) &= \sum_{i=0}^{S-1} \left(\sum_{u=1}^{M-1} c_{vu}^{(M)} \gamma_{i+uS} \right)^2, & \text{if } k = 0 \\ &= 1 + \sum_{i=0}^{S-1} \left(\sum_{u=1}^{M-1} c_{vu}^{(M)} \gamma_{i+uS} \right)^2 \\ &\quad - 2c_{vu_k}^{(M)} \left(\sum_{u=1}^{M-1} c_{vu}^{(M)} \gamma_{i_k+uS} \right), & \text{if } k \geqq 1 \end{aligned} \tag{4.3}$$

with $i_k \equiv k - 1 \pmod{S}$ and $u_k = 1 + [(k-1)/S]$ for all $k = 1, \cdots, N - S$. In particular, for $\gamma^0 = [1/(N - S)]1^{(N-S)}$ we have, for all $v = 0, \cdots, M - 2$,

$$\begin{aligned} L_1(\xi^{(v)}, \gamma^0; \beta^{(\kappa)}) &= 1/S(M-1)^2, & \text{if } k = 0 \\ &= [(S-1)/S][1 + 1/(M-1)]^2 & \text{if } k \geqq 1 \end{aligned} \tag{4.14}$$

and for $v = M - 1$ we have,

$$\begin{aligned} L_1(\xi^{(M-1)}, \gamma^0; \beta^{(k)}) &= 1/S, & \text{if } k = 0 \\ &= (S-1)/S, & \text{if } k \geqq 1. \end{aligned} \tag{4.15}$$

It follows that, for every ξ and η

$$\begin{aligned} &R_1(\xi, \gamma^0; \eta) \\ &\quad = (1 - \eta_0)\{(1 - \xi_{M-1})(1 - 1/S)[1 + 1/(M-1)]^2 \\ &\quad + \xi_{M-1}(1 - 1/S)\} + \eta_0\{(1 - \xi_{M-1})[1/S(M-1)^2] + \xi_{M-1}(1/S)\}. \end{aligned} \tag{4.16}$$

Thus, a good strategy of Nature, against any strategy (ξ, γ^0) of the Statistician, is in $\mathcal{H}_0^*(B)$. Moreover, $R_1(\xi, \gamma^0; \eta)$ is independent of the η's in $\mathcal{H}_0^*(B)$. It remains to show that $(\xi^{(M-1)}, \gamma^0)$ is Bayes against some η in $\mathcal{H}_0^*(B)$. Let $\eta^0 = (0, 1/(N - S), \cdots, 1/(N - S))$. According to (3.10) γ^0 is Bayes against η^0, for every $\xi^{(v)}$, and thus for every ξ. Finally, according to (4.16), $\inf_\xi R_1(\xi, \gamma^0; \eta^0) = (1 - 1/S)$ is attained by $\xi^{(M-1)} = (0, 0, \cdots, 0, 1)'$. Both $\xi^{(M-1)}$ and γ^0 are unique minimization points of $R_1(\xi, \gamma; \eta^0)$. Thus $(\xi^{(M-1)}, \gamma^0)$ is an admissible strategy.

To conclude the present case we mention that the minimax strategy for an arbitrary set (4.8) is represented by $(\xi^{(v_0)}, \gamma^0)$; where $\gamma^0 = [R/(N - S)]1^{(N-S)}$ as before, and $\xi^{(v_0)}$ is found as follows. We determine for every $v = 0, \cdots, M - 1$ the Bayes risks

$$R_1(\xi^{(v)}, \gamma^0; \eta^0) = R^2 \left[1 - \frac{1}{(M-1)^2 S^2} \sum_{i=0}^{S-1} \left[\sum_{u=1}^{M-1} c_{vu}^{(M)} \operatorname{sgn}(\beta_{i+uS}) \right]^2 \right]. \tag{4.17}$$

$\xi^{(v_0)}$ is the vector which minimizes (4.17) overall $v = 0, \cdots, M - 1$. As in (4.16) it can be shown that $\xi^{(v_0)}$ is unique.

5. Numerical example. The following numerical example illustrates the results of the present study, and the actual computations involved. We shall consider the problem of estimating 4 parameters of a 2^4 factorial system. For the purpose of actual comparisons we assume that $\sigma^2 = 0$, i.e., the experimental results given are the expected treatment yields. We sometimes use capital letters A, B, C and D and their combinations to present main-effects and interactions, and small letters $a, b, c, d, ab, \cdots$ to denote treatment combinations. The following data, taken from O. L. Davies [2] p. 275, represent a 2^4 factorial system comprised of the following factors:

A: acid strength—levels: 87%, 93%.
B: reaction time—levels: 15 min., 30 min.
C: amount of acid—levels: 35 ml., 45 ml.
D: reaction temperature—levels: 60°C, 70°C.

The response of the system is the amount of Isatin Derivative (g. per 10 g. of base materials). This variable is denoted by $Y(x_v)$. The response values in the given experiment are represented in Table 1.

Let $\{M, A, B, C\}$ be the subset of 4 pre-assigned parameters. Let $\{M, ABC, D, ABCD\}$ be the specified subgroup of defining parameters. The 4 blocks of treatment combinations, corresponding to this subgroup of defining parameters, are presented in Table 2 with the associated expected treatment yields.

The estimates of the pre-assigned parameters, using the conditional least-

TABLE 1
Yield of Isatin Derivative in a 2^4 factorial experiment

Acid Strength A	Reaction Time B	Temperature D: 60°C, Amount acid C: 35 ml.	60°C, 45 ml.	70°C, Amount acid C: 35 ml.	70°C, 45 ml.
87%	15 min.	6.08 (1)	6.31 (*c*)	6.79 (*d*)	6.77 (*cd*)
	30 min.	6.53 (*b*)	6.12 (*bc*)	6.73 (*bd*)	6.49 (*bcd*)
93%	15 min.	6.04 (*a*)	6.09 (*ac*)	6.68 (*ad*)	6.38 (*acd*)
	30 min.	6.43 (*ab*)	6.36 (*abc*)	6.08 (*abd*)	6.23 (*abcd*)

TABLE 2
Possible vectors of observations, $y(X_v)(v = 0, 1, 2, 3)$ in a $\frac{1}{4}$ replicate of a 2^4 factorial system

$y(X_0)$	$y(X_1)$	$y(X_2)$	$y(X_3)$
6.08 (1)	6.04 (*a*)	6.79 (*d*)	6.68 (*ad*)
6.43 (*ab*)	6.53 (*b*)	6.08 (*abd*)	6.73 (*bd*)
6.09 (*ac*)	6.31 (*c*)	6.38 (*acd*)	6.77 (*cd*)
6.12 (*bc*)	6.36 (*abc*)	6.49 (*bcd*)	6.23 (*abcd*)

squares estimator $\hat{\alpha}(0)$ (for details see Zacks [5], Section 5) are given, for each of the four possible blocks X_v ($v = 0, 1, 2, 3$) in Table 3.

The expectations and variances of the estimates in Table 3 are the arithmetic means of the possible estimates, and the variances about these means. According to the assumption that $\sigma^2 = 0$, these are the expected values and variances of the components of the c.l.s.e. $\hat{\alpha}(0)$, under ξ^* (each block chosen at random, with equal probabilities). The covariances of these estimates; under ξ^*; are: cov $(\hat{M}, \hat{A}) = -0.0129$; cov $(\hat{M}, \hat{B}) = -0.0319$; cov $(\hat{M}, \hat{C}) = -0.0023$; cov $(\hat{A}, \hat{B}) = 0.0094$; cov $(\hat{A}, \hat{C}) = -0.0021$; cov $(\hat{B}, \hat{C}) = 0.0018$.

According to Theorem 4.1 these estimates under ξ^* constitute a minimax strategy when the 12 nuisance parameters may assume arbitrary bounded values. We turn now to the case where the signs of the nuisance parameters are known. This state of information corresponds to an orthant of a 12-dimensional hypersphere, centered at the origin. Assume that all the nuisance parameters are

TABLE 3

The possible estimates of M, A, B, C; their expectations and variances under ξ^* (R.P.I.)

Pre-assigned Parameters	Possible Blocks				Expectations	Variances
	X_0	X_1	X_2	X_3		
M	6.180	6.311	6.435	6.602	6.382	0.0244
A	0.080	−0.110	−0.205	−0.147	−0.095	0.0114
B	0.095	0.135	−0.152	−0.122	−0.011	0.0160
C	−0.075	0.025	0.000	−0.102	−0.038	0.0027

TABLE 4

Minimax adjustments of $\hat{\alpha}(0)$

Pre-assigned Parameters	Possible Blocks			
	X_0	X_1	X_2	X_3
M	0.0114	0.0114	0.0114	−0.0342
A	−0.0342	0.0114	0.0114	0.0114
B	0.0114	−0.0342	0.0114	0.0114
C	−0.0114	−0.0114	−0.0114	0.0342

TABLE 5

Minimax estimates of M, A, B, C

Pre-assigned Parameters	Possible Blocks				Expectation	Variance
	X_0	X_1	X_2	X_3		
M	6.1914	6.3224	6.4464	6.5678	6.3820	0.0196
A	0.0458	−0.0986	−0.1936	−0.1356	−0.0950	0.0079
B	0.1064	0.1008	−0.1406	−0.1076	−0.0110	0.0016
C	−0.0864	−0.0136	−0.0114	−0.0678	−0.0380	0.0017

bounded by $R = 0.1370$, and the signs of the nuisance parameters are:

$$\begin{array}{lll} \text{sgn}\,(ABC) = +1, & \text{sgn}\,(D) = +1, & \text{sgn}\,(ABCD) = +1 \\ \text{sgn}\,(BC) = -1, & \text{sgn}\,(AD) = -1, & \text{sgn}\,(BCD) = +1 \\ \text{sgn}\,(AC) = +1, & \text{sgn}\,(BD) = -1, & \text{sgn}\,(ACD) = -1 \\ \text{sgn}\,(AB) = -1, & \text{sgn}\,(CD) = -1, & \text{sgn}\,(ABD) = -1 \end{array}$$

According to (2.3) and Theorem 4.2, the minimax adjustments of the estimates in Table 3 are given by the components of $-\frac{1}{4}(c_{v1}^{(4)}I^{(4)}, c_{v2}^{(4)}I^{(4)}, c_{v3}^{(4)}I^{(4)})'\gamma^0$, where $\gamma_t^0 = (0.1370/12)\,\text{sgn}\,(\beta_t)$ for all $t = 0, \cdots, 11$; and $C_{vu}^{(4)}$ $(v = 0, 1, 2, 3; u = 1, 2, 3)$ are the elements of $(C^{(4)})$. These minimax adjustments, corresponding to different possible X_v $(v = 0, 1, 2, 3)$ are given in Table 4.

The sum of the minimax adjustment values, over the possible samples is always zero. This is due to the property of the (H_v) matrices under ξ^* (R.P.I.). Adding the minimax adjustment values of Table 4 to the estimates in Table 3, we obtain the minimax estimates of the pre-assigned parameters in Table 5.

Comparing the variances of the unadjusted estimators $\hat{\alpha}(0)$ given in Table 3, to those of the adjusted estimators, presented in Table 5, we see that the total reduction in the *actual* variances due to available information about the signs of the nuisance parameters, and using the appropriate minimax estimators is from 0.0545 to 0.0308, i.e., 43.5%. In reality we cannot compare the actual total reduction in the variance, since we do not know all the possible estimates. We can compare, however, the minimax risks corresponding to different states of information on the nuisance parameters. In the present example of 12 nuisance parameters the reduction in the minimax risk is of 8.3%. The use of minimax adjusted estimators protects the statistician against the least favorable a priori distribution of the β's. The actual gain would be generally higher than that expected by comparing the minimax risks.

REFERENCES

[1] Blackwell, D., and Girshick, M. A. (1954). *Theory of Games and Statistical Decisions*. Wiley, New York.

[2] Davies, O. L., ed. (1954). *Design and Analysis of Industrial Experiments*. Oliver and Boyd, London.

[3] Ehrenfeld, S., and Zacks, S. (1961). Randomization and factorial experiments. *Ann. Math. Statist.* **32** 270–297.

[4] Kempthorne, O. (1952). *The Design and Analysis of Experiments*. Wiley, New York.

[5] Zacks, S. (1963). On a complete class of linear unbiased estimators for randomized factorial experiments. *Ann. Math. Statist.* **34** 769–779.