The Annals of Statistics
1980, Vol. 8, No. 5, 1142–1155

DEFICIENCIES BETWEEN LINEAR NORMAL EXPERIMENTS[1]

BY ANDERS RYGH SWENSEN

University of Oslo and University of California, Berkeley

Let $X_1, \cdots, X_n$ be independent and normally distributed variables, such that $0 < \operatorname{Var} X_i = \sigma^2$, $i = 1, \cdots, n$ and $E(X_1, \cdots, X_n)' = A'\beta$ where A is a $k \times n$ matrix with known coefficients and $\beta = (\beta_1, \cdots, \beta_k)'$ is an unknown vector. σ may be known or unknown. Denote the experiment obtained by observing $X_1, \cdots, X_n$ by $\mathcal{E}_A$. Let A and B be matrices of dimension $n_A \times k$ and $n_B \times k$.

The deficiency $\delta(\mathcal{E}_A, \mathcal{E}_B)$ is computed when σ is known and for some cases, including the case $BB' - AA'$ positive semidefinite and AA' nonsingular, also when σ is unknown. The technique used consists of reducing to testing a composite hypotheses and finding a least favorable distribution.

1. Introduction. An experiment $\mathcal{E}$ is a pair $((\mathcal{X}, \mathcal{A}), (P_\theta : \theta \in \Theta))$ where $(\mathcal{X}, \mathcal{A})$ is a measurable space and $(P_\theta : \theta \in \Theta)$ is a family of probability measures on $(\mathcal{X}, \mathcal{A})$. Between experiments indexed by the same parameter set, Le Cam [4] defined the deficiency in terms of risk functions obtainable in the two experiments.

This naturally leads to considerations where deficiencies are used as measures of information. For countable parameter sets there is a recent paper by Torgersen [11] where he compares an experiment with the totally informative experiment (i.e., all the measures P_θ are mutually singular) and the least informative experiment (i.e., the P_θ's are identical) and studies the behaviour of the deficiencies under replications of the experiment. The finite state Markov chain has been treated by Lindqvist in [6] and [7] where the deficiency is used to measure the loss of memory of the initial state X_0 in the tail $(X_n, X_{n+1}, \cdots)$ of the Markov chain $(X_0, X_1, \cdots)$. In [8] Lindqvist shows a relation between Dobrushin's ergodic coefficient and the deficiency of a Markov chain with arbitrary state space with respect to the least informative experiment.

In this paper we will treat a case where the parameter set is uncountable, namely the linear normal experiments and hopefully the results might be of interest where several experiments of this kind are involved as in design of experiments and in choosing regression coefficients for additional observations. A short discussion and some examples are given in Section 5. In Sections 3 and 4 the deficiencies are evaluated for known and unknown variance respectively.

Among the several equivalent formulations of deficiencies given in [4], of which one is referred to above, we will use the following. Let $\mathcal{E} = ((\mathcal{X}, \mathcal{A}), (P_\theta : \theta \in \Theta))$ and $\mathcal{F} = ((\mathcal{Y}, \mathcal{B}), (Q_\theta : \theta \in \Theta))$ be two experiments where $(P_\theta : \theta \in \Theta)$ is assumed to be a dominated family of probability measures, $\mathcal{Y}$ a Borel set of a

Received April 1977; revised February 1979.

[1]This paper was prepared with the partial support from Norges almenvitenskapelige forskningsråd.

AMS 1970 *subject classifications.* Primary 62B15; secondary 62K99.

Key words and phrases. Deficiencies, invariant kernels, normal models, additional observations.

complete separable metric space and $\mathcal{B}$ the Borel sets in $\mathcal{Y}$. Then the deficiency of $\mathcal{E}$ with respect to $\mathcal{F}$ is defined as

$$\delta(\mathcal{E}_A, \mathcal{E}_B) = \inf \sup_\theta \|P_\theta M - Q_\theta\|$$

where the infimum is taken over all Markov kernels $\mathfrak{M}$ from $(\mathcal{X}, \mathcal{A})$ to $(\mathcal{Y}, \mathcal{B})$, and $\| \ \|$ is the total variation norm.

By letting $\Delta(\mathcal{E}, \mathcal{F}) = \max(\delta(\mathcal{E}, \mathcal{F}), \delta(\mathcal{F}, \mathcal{E}))$ we get a pseudometric between experiments.

To be able to handle the complicated expression for the deficiency, we shall use some symmetry properties present in linear models and appeal to results from [4] and [10] on invariance. These results may be summarized as follows.

Let G be a group of transformations acting on Θ, $\mathcal{X}$, $\mathcal{Y}$ such that $x \to g(x), y \to g(y)$ where $g \in G$, are measurable mappings, and the families $(P_\theta : \theta \in \Theta)$ and $(Q_\theta: \theta \in \Theta)$ are invariant under G. Let $\mathfrak{M}_G$ be the set of invariant Markov kernels, i.e.,

$$\mathfrak{M}_G = \{M \in \mathfrak{M} : M(g(B)|g(x)) = M(B|x) \text{ if } g \in G, B \in \mathcal{B}, x \notin N \text{ where } gN = N, g \in G \text{ and } P_\theta(N) = 0, \theta \in \Theta\}.$$

If the following conditions are satisfied, the infimum in the definition of deficiency may be taken over $\mathfrak{M}_G$.

(i) There is a σ-field $\mathcal{G}$ in G such that the mappings $(x, g) \to g(x)$, $(y, g) \to g(y)$ are $\mathcal{A} \times \mathcal{G}$ and $\mathcal{B} \times \mathcal{G}$ measurable, respectively.

(ii) There is a σ-finite measure τ on $(G, \mathcal{G})$ such that $\tau(B) = 0$ implies $\tau(Bg) = 0$ when $B \in \mathcal{B}$ and $g \in G$.

(iii) The group has an invariant mean, which is the case if it is solvable. If in addition there is at least one invariant Markov kernel for which the exceptional set $N = \varnothing$, we may also restrict attention to invariant kernels having this additional property.

For a particular class of experiments, the translation experiments, the conditions take a particularly simple form. Let $\Theta = \mathcal{X}$ be a locally compact topological group which is Hausdorff, $\mathcal{A}$ the Borel sets and

$$P_\theta(A) = P(A\theta^{-1})$$

where P is a probability measure. Clearly the group G described above corresponds to translations $x \to x\theta$ where $\theta \in \Theta$. Now (i) is automatically satisfied, and by taking τ as a Haar measure, (ii) is also satisfied. Hence the only condition which needs to be verified is (iii). We will denote an experiment of the form above by $\mathcal{E}_P$.

In [10] Torgersen showed that for two translation experiments on $(\mathcal{X}, \mathcal{A})$ every invariant Markov kernel with ϕ as exceptional set may be written $M(B|x) = \mu(Bx^{-1})$ where μ is a probability measure on $(\mathcal{X}, \mathcal{A})$. Hence if $\mathcal{E}_P$ is dominated,

$$\delta(\mathcal{E}_p, \mathcal{E}_Q) = \inf_\mu \|\mu * P - Q\|$$

where $*$ denotes convolution, i.e., $\mu * P(A) = \mu \times P\{(x_1, x_2) : x_1 x_2 \in A\}$.

Furthermore, if μ_0 is a least favorable distribution at all levels $\alpha \in [0, 1]$ for testing

$$H: P_\theta'', \theta \in \Theta \qquad \text{against} \quad K: Q$$

where $P_\theta''(A) = P(\theta^{-1}A)$ for $\theta \in \Theta, A \in \mathcal{C}$, then

$$\delta(\mathfrak{S}_P, \mathfrak{S}_Q) = \|\mu_0 * P - Q\|.$$

This result will enable us to compute the deficiency between certain linear normal experiments.

We can describe them in the following way. Let A be a known $k \times n_A$ matrix, and $\mathfrak{S}_A$ the experiment given by the independent normally distributed random variables $X_1, \cdots, X_{n_A}$ with $\operatorname{Var} X_i = \sigma^2$, $i = 1, \cdots, n_A$ and $E(X_1, \cdots, X_{n_A})' = A'\beta$ where $\beta = (\beta_1, \cdots, \beta_k)' \in \mathbb{R}^k$. $'$ denotes transpose. Thus to different design matrices A and B there correspond experiments $\mathfrak{S}_A$ and $\mathfrak{S}_B$ respectively.

To avoid trivial cases we will assume that $n_A, n_B \geqslant k \geqslant 1$.

We will treat both the case where σ is known and the case where it is unknown. In the former the parameter set is $]-\infty, \infty[^k = \mathbb{R}^k$ and in the latter $]-\infty, \infty[^k \times]0, \infty[= \mathbb{R}^k \times \mathbb{R}^+$.

2. A special case and reduction to a canonical form.

PROPOSITION 2.1. *Let A, B be design matrices corresponding to experiments $\mathfrak{S}_A$ and $\mathfrak{S}_B$ respectively. If row $[B'] \not\subset$ row $[A']$,*

$$\delta(\mathfrak{S}_A, \mathfrak{S}_B) = 2$$

whether σ is known or not.

PROOF. Suppose σ is known and let P_β, Q_β be the probability measures corresponding to the parameter value β in $\mathfrak{S}_A$ and $\mathfrak{S}_B$ respectively.

By assumption there is a β_0 such that $A'\beta_0 = 0$ and $B'\beta_0 \neq 0$. Then

$$\begin{aligned}\delta(\mathfrak{S}_A, \mathfrak{S}_B) &= \inf_M \sup_\beta \|P_\beta M - Q\| \\ &\geqslant \inf_M \sup_{t \in \mathbf{R}} \|P_{t\beta_0} M - Q_{t\beta_0}\| \\ &= \inf_M \sup_t \|P_0 M - Q_{t\beta_0}\|.\end{aligned}$$

But $\|P_0 M - Q_{t\beta_0}\| \to 2$ as $t \to \infty$ for all Markov kernels. Hence $\delta(\mathfrak{S}_A, \mathfrak{S}_B) = 2$.

Consider now the case where σ is unknown. By fixing this parameter we obtain experiments for which $\delta = 2$. This means, since a δ computed for known σ always gives a lower bound for the corresponding δ with σ unknown, that $\delta(\mathfrak{S}_A, \mathfrak{S}_B) = 2$ also when σ is unknown. □

A consequence of Proposition 2.1 is that the Δ-distance is 2 between experiments given by $X_1, \cdots, X_n$ independent and normally distributed $\operatorname{Var} X_i = \sigma^2$, $EX_i = \alpha + \beta t_i$, $i = 1, \cdots, n$ and $Y_1, \cdots, Y_n$ independent and normally distributed $\operatorname{Var} Y_i = \sigma^2$, $EY_i = \alpha + \beta t_i + \gamma t_i^2$, $i = 1, \cdots, n$ whether σ is known or not. Hence the Δ-distance is of no help if we want to study the effect of variation in the regression variables $t_1, \cdots, t_n$ for estimation of α and β in the two experiments.

From now on suppose row $[B'] \subset$ row $[A']$. Let

$$r_A = \text{rank}\,[A], \qquad r_B = \text{rank}\,[B]$$

$$S_A^2 = \inf_\beta (X - A'\beta)'(X - A'\beta), \qquad S_B^2 = \inf_\beta (X - B'\beta)'(X - B'\beta)$$

where X is the vector of observations in $\mathcal{E}_A$ and Y is the vector of observations in $\mathcal{E}_B$. Then there exist nonzero linear functionals

$$\theta_i = C_i'\beta, \qquad i = 1, \cdots, r_A$$

of β such that $\theta_1, \cdots, \theta_{r_A}$ are estimable in $\mathcal{E}_A$ and $\theta_1, \cdots, \theta_{r_B}$ are estimable in $\mathcal{E}_B$.

These functionals may be constructed as follows. Let V' be a $r_A \times k$ matrix such that the first r_B rows of V' span row $[B']$ while all the rows of V' span row $[A']$, i.e., row $[V']$ = row $[A']$. Then there is a $n_A \times r_A$ matrix S' of rank r_A and a $n_B \times r_A$ matrix T' of rank r_B such that

$$A' = S'V' \qquad \text{and} \qquad B' = T'V'.$$

Let $\Delta_1, \cdots, \Delta_{r_A}$ be the characteristic roots counting multiplicites of $(SS')^{-1}(TT')$. Without loss of generality we may assume that they are arranged so that $\Delta_{r_B+1} = \cdots = \Delta_{r_A} = 0$. By the theorem [2] on simultaneous reduction of two quadratic forms there is a $r_A \times r_A$ nonsingular matrix F so that

$$F'SS'F = I \qquad \text{and} \qquad F'TT'F = \Delta$$

where Δ is the diagonal matrix whose ith diagonal element is Δ_i. Denote the ith row of F^{-1} by d_i. Then we may put

$$C_i = Vd_i, \qquad i = 1, \cdots, r_A.$$

From linear normal theory it now follows that if $\hat{X}_1, \cdots, \hat{X}_{r_A}$ are the UMVU estimators of $\theta_1, \cdots, \theta_{r_A}$ in $\mathcal{E}_A$ and $\hat{Y}_1, \cdots, \hat{Y}_{r_B}$ are the UMVU estimators of $\theta_1, \cdots, \theta_{r_B}$ in $\mathcal{E}_B$, then S_A, $\hat{X}_1, \cdots, \hat{X}_{r_A}$ are independent in $\mathcal{E}_A$ and S_B, $\hat{Y}_1, \cdots, \hat{Y}_{r_B}$ are independent in $\mathcal{E}_B$, with the following distributions

$$\hat{X}_i \sim N(\theta_i, \sigma^2), \quad i = 1, \cdots, r_A, \qquad S_A^2/\sigma^2 \sim \chi^2_{n_A - r_A}$$

and

$$\hat{Y}_i \sim N(\theta_i, \sigma^2/\Delta_i), \quad i = 1, \cdots, r_B, \qquad S_B^2/\sigma^2 \sim \chi^2_{n_B - r_B}.$$

Furthermore for σ unknown (known) S_A, $\hat{X}_1, \cdots, \hat{X}_{r_A}(\hat{X}_1, \cdots, \hat{X}_{r_B})$ and S_B, $\hat{Y}_1, \cdots, \hat{Y}_{r_B}(\hat{Y}_1, \cdots, \hat{Y}_{r_B})$ are sufficient in $\mathcal{E}_A$ and $\mathcal{E}_B$ respectively.

From now on we drop the $\hat{}$. Thus $\mathcal{E}_A$ are given by S_A, $X_1, \cdots, X_{r_A}$ and $\mathcal{E}_B$ by S_B, $Y_1, \cdots, Y_{r_B}$ with both sets of random variables having the properties described above.

3. The case of known σ. Put $l = r_B$ = rank B and let $\Delta_1, \cdots, \Delta_{r_A}$ be as described in the preceding section.

PROPOSITION 3.1. *If row $[B'] \subset$ row $[A']$,*

$$\delta(\mathcal{E}_A, \mathcal{E}_B) = E|1 - \Pi_{\Delta_i > 1} \Delta_i^{\frac{1}{2}} \exp(-(\Delta_i - 1)Z_i^2/2)|$$

where $Z_1, \cdots, Z_{r_A}$ are independent and identically $N(0, 1)$ distributed.

PROOF. By the results in Section 2 $\mathcal{E}_A$ is given by $X_1, \cdots, X_{r_A}$ independent and normally distributed $EX_i = \theta$, $\operatorname{Var} X_i = \sigma^2$, $i = 1, \cdots, r_A$ and $\mathcal{E}_B$ is given by $Y_1, \cdots, Y_l$ independent and normally distributed with $EY_i = \theta_i$, $\operatorname{Var} Y_i = \sigma^2/\Delta_i$, $i = 1, \cdots, l$.

By invariance considerations similar to those of the proof of Proposition 2.1 in Hansen and Torgersen [3], it may be shown that $X_{l+1}, \cdots, X_{r_A}$ may be deleted in $\mathcal{E}_A$.

Now $\mathcal{E}_A$ and $\mathcal{E}_B$ are translation experiments for addition in $\mathbb{R}^l$. Since this is a commutative operation, and $\mathcal{E}_A$ and $\mathcal{E}_B$ both are dominated, $\delta(\mathcal{E}_A, \mathcal{E}_B)$ may be found by the method described in Section 1. Let P_θ, $\theta \in \mathbb{R}^l$ be the probability measure defined by $X_1, \cdots, X_l$ and let Q be the probability measure defined by $Y_1, \cdots, Y_l$ for the parameter value $\theta = 0$. Then the least favorable distribution μ_0 for testing

$$H : P_\theta, \quad \theta \in \mathbb{R}^l \qquad \text{against} \qquad K : Q$$

is given by the independent variables $U_1, \cdots, U_l$ where $U_i = 0$ with probability 1 if $\Delta_i \geqslant 1$ and U_i is $N(0, \sigma^2(\Delta_i^{-1} - 1))$ distributed if $\Delta_i < 1$. Hence $\delta(\mathcal{E}_A, \mathcal{E}_B) = \| \mu_0 * P_0 - Q \|$. But $\mu_0 * P_0$ has density

$$\Pi_{\Delta_i \geqslant 1} \sigma^{-1} \phi(x_i/\sigma) \Pi_{\Delta_i < 1} \Delta_i^{\frac{1}{2}} \phi\left(\Delta_i^{\frac{1}{2}} x_i/\sigma\right)$$

with respect to the Lebesgue measure. □

Now suppose $r_A = k$ and $BB' \geqslant AA'$, i.e., $BB' - AA'$ is positive semidefinite. If F has the same meaning as in Section 2, then

$$Z'(\Delta - I)Z = Z'F'F'^{-1}(\Delta - I)F^{-1}FZ = W'(BB' - AA')W$$

where $W = FZ$. Furthermore

$$EWW' = EFZZ'F' = FF' = (AA')^{-1}$$

and

$$\det[BB'](\det[AA'])^{-1} = (\det[F'BB'F])(\det[F'AA'F])^{-1} = \Delta_1 \cdots \Delta_k$$

so we may write

$$\delta(\mathcal{E}_A, \mathcal{E}_B) = E|(\det[BB'])^{\frac{1}{2}}(\det[AA'])^{-\frac{1}{2}} \exp\left(-\frac{1}{2} W'(BB' - AA')W\right) - 1|$$

where W is multivariate normal with mean 0 and covariance matrix $(AA')^{-1}$. This is the result given by Le Cam in [5].

COROLLARY 3.1. *If σ^2 is known, then for any pair $\mathcal{E}_A, \mathcal{E}_B$ of linear normal experiments such that $BB' \geqslant AA'$ and A is of full rank*

$$\delta(\mathcal{E}_A, \mathcal{E}_B) = \|N\left(0, (AA')^{-1}\right) - N\left(0, (BB')^{-1}\right)\|.$$

If $\mathcal{E}_A$ and $\mathcal{E}_B$ are given by X, Y multivariate normal with known covariance matrices Σ_A, Σ_B and $EX = A'\beta$, $EY = B'\beta$ then after a linear transformation we may consider equivalent experiments given by independent normally distributed variables. Thus $\delta(\mathcal{E}_A, \mathcal{E}_B)$ may be computed also in this case.

4. The case of unknown σ. Some of the notation we will use in this section are: If $(\mathfrak{X}, \tau)$ is a topological space, let $\mathfrak{B}(\mathfrak{X}) = \sigma(\{B : B \in \tau\})$ be the Borel sets in $\mathfrak{X}$.

Let $P_{l,n_1,\beta_1,\cdots,\beta_l,\sigma}(Q_{l,n_2,\beta_1,\cdots,\beta_l,\sigma})$ be the probability measure on $(\mathbb{R}^l \times \mathbb{R}^+, \mathfrak{B}(\mathbb{R}^l \times \mathbb{R}^+))$ given by the independent random variables $X_1, \cdots, X_l, S_A(Y_1, \cdots, Y_l, S_B)$ where $X_1, \cdots, X_l(Y_1, \cdots, Y_l)$ are normally distributed with $EX_i = \beta_i$, $\operatorname{Var} X_i = \sigma^2(EY_i = \beta_i$, $\operatorname{Var} Y_i = \sigma^2/\Delta_i)i = 1, \cdots, l$ and $S_A^2/\sigma^2(S_B^2/\sigma^2)$ is distributed as $\chi^2_{n_1}(\chi^2_{n_2})$. $\Delta_1, \cdots, \Delta_l$ are known positive constants. $P'_{l,\beta_1,\cdots,\beta_l,\sigma}$ and $Q'_{l,\beta_1,\cdots,\beta_l,\sigma}$ are the marginal distributions on $(\mathbb{R}^l, \mathfrak{B}(\mathbb{R}^l))$ of $P_{l,n_1,\beta_1,\cdots,\beta_l,\sigma}$ and $Q_{l,n_2,\beta_1,\cdots,\beta_l,\sigma}$ respectively.

$\#C$ is the number of elements in the set C.

Let $\Delta_1, \cdots, \Delta_{r_A}$ have the same meaning as in Section 2. Let $m = \#\{i : 0 < \Delta_i < 1\}$. Without loss of generality we may assume that $\Delta_1, \cdots, \Delta_{r_B - m} \geq 1$, $1 > \Delta_{r_B - m+1}, \cdots, \Delta_{r_B} > 0$ and $\Delta_{r_B+1} = \cdots = \Delta_{r_A} = 0$.

It follows by the reduction in Section 2 and by a reduction similar to that of the proof of Proposition 2.1 in Hansen and Torgersen [3], that if row $[B'] \subset$ row $[A']$ and if $l = r_B = \operatorname{rank}[B]$, we may assume

$$\mathcal{E}_A = \Big(\big(\mathbb{R}^l \times \mathbb{R}^+, \mathfrak{B}(\mathbb{R}^l \times \mathbb{R}^+)\big), \big(P_{l,n_A - r_A,\theta_1,\cdots,\theta_l,\sigma} : (\theta_1, \cdots, \theta_l, \sigma) \in \mathbb{R}^l \times \mathbb{R}^+\big)\Big)$$

if $n_A > r_A$ and with P' instead of P if $n_A = r_A$. Similarly, we may assume

$$\mathcal{E}_B = \Big(\big(\mathbb{R}^l \times \mathbb{R}^+, \mathfrak{B}(\mathbb{R}^l \times \mathbb{R}^+)\big), \big(Q_{l,n_B - l,\theta_1,\cdots,\theta_l,\sigma} : (\theta_1, \cdots, \theta_l, \sigma) \in \mathbb{R}^l \times \mathbb{R}^+\big)\Big)$$

if $n_B > r_B$ and with Q' instead of Q if $n_B = r_B$.

We will first treat the cases where there is no estimator for σ in $\mathcal{E}_A$, but one in $\mathcal{E}_B$ (Proposition 4.1) and the case where there is no estimator for σ neither in $\mathcal{E}_A$ nor in $\mathcal{E}_B$ (Proposition 4.2).

PROPOSITION 4.1. *If $n_A = r_A$ and $n_B > l$, then*

$$\delta(\mathcal{E}_A, \mathcal{E}_B) = 2.$$

PROOF. Let the group G be given by

$$\begin{aligned} g(x_1, \cdots, x_l) &= (g_0 x_1 + g_1, \cdots, g_0 x_l + g_l) \\ g(y_1, \cdots, y_l, z) &= (g_0 y_1 + g_1, \cdots, g_0 y_l + g_l, g_0 z) \\ g(\theta_1, \cdots, \theta_l, \sigma) &= (g_0 \theta_1 + g_1, \cdots, g_0 y_l + g_l, g_0 \sigma) \end{aligned}$$

where $(g_1, \cdots, g_l, g_0) \in \mathbb{R}^l \times \mathbb{R}^+$. It can be verified that the assumptions (i)–(iii) in Section 1 are satisfied so that it is enough to consider invariant kernels. Furthermore, it is not difficult to see that every invariant kernel has ϕ as the exceptional set.

Now suppose $\delta(\mathcal{E}_A, \mathcal{E}_B) < 2$. Then if $\varepsilon + \delta < 2$, there exists an invariant Markov kernel such that

$$\| P'_{l,\theta_1,\cdots,\theta_l,\sigma} M - Q_{l,n_B - l,\theta_1,\cdots,\theta_l,\sigma} \| < \varepsilon + \delta, (\theta_1, \cdots, \theta_l, \sigma) \in \mathbb{R}^l \times \mathbb{R}^+.$$

Let K be a compact set in $\mathbb{R}^l \times \mathbb{R}^+$. Choose $B_i \in \mathfrak{B}(\mathbb{R})$, $i = 1, \cdots, l$ and $B_0 \in \mathfrak{B}(\mathbb{R}^+)$ such that $K \subset B_1 \times \cdots \times B_l \times B_0$. Then

$$\begin{aligned} M(B_1 \times \cdots \times B_l \times B_0 | x_1, \cdots, x_l) \\ = M(g_0 B_0 + g_1 \times \cdots \times g_0 B_l + g_l \times g_0 B_0 | g_0 x_1 + g_1, \cdots, g_0 x_l + g_l), \\ (g_1, \cdots, g_l, g_0) \in \mathbb{R}^l \times \mathbb{R}^+ . \end{aligned}$$

Let $g_0 \to 0$. Then $\mathbb{R}^l \times \mathbb{R}^+ \cap g_0 B_1 + g_1 \times \cdots \times g_0 B_l + g_l \times g_0 B_0 \to \varnothing$ so that $M(B_1 \times \cdots \times B_l \times B_0 | x_1, \cdots, x_l) = 0$. Hence $M(\cdot | x_1, \cdots, x_l)$ vanishes on every compact. Since $M(\cdot | x_1, \cdots, x_l)$ is a probability measure and $\mathbb{R}^l \times \mathbb{R}^+$ is σ-compact, this gives a contradiction. □

PROPOSITION 4.2. *If $n_A = r_A$ and $n_B = l$*

$$\begin{aligned} \delta(\mathcal{E}_A, \mathcal{E}_B) &= \| P'_{l,0,\cdots,0,1} - Q'_{l,0,\cdots,0,1} \| \\ &= E | 1 - \Pi_{\Delta_i > 0} \Delta_i^{\frac{1}{2}} \exp[-(\Delta_i - 1) Z_i^2 / 2] | \end{aligned}$$

where $Z_1, \cdots, Z_{r_A}$ are independent $N(0, 1)$ distributed.

PROOF. The proof is analogous to a part of the proof of Proposition 2.1 in Hansen and Torgersen [3]. Let G be the group given by

$$\begin{aligned} g(x_1, \cdots, x_l) &= (g_0 x_1 + g_1, \cdots, g_0 x_l + g_l) \\ g(y_1, \cdots, y_l) &= (g_0 y_1 + g_1, \cdots, g_0 y_l + g_l) \\ g(\theta_1, \cdots, \theta_l, \sigma) &= (g_0 \theta_1 + g_1, \cdots, g_0 \theta_l + g_l, g_0 \sigma) \end{aligned}$$

where $(g_1, \cdots, g_l, g_0) \in \mathbb{R}^l \times \mathbb{R}^+$. It is easy to verify that

$$\begin{aligned} \delta(\mathcal{E}_A, \mathcal{E}_B) &= \inf_{M \in \mathfrak{M}_G} \sup_{(\theta_1, \cdots, \theta_l, \sigma)} \| P'_{l, \theta_1, \cdots, \theta_l, \sigma} M - Q'_{l, \theta_1, \cdots, \theta_l, \sigma} \| \\ &= \inf_{M \in \mathfrak{M}_G} \| P'_{l,0,\cdots,0,1} M - Q'_{l,0,\cdots,0,1} \| . \end{aligned}$$

Consider $M \in \mathfrak{M}_G$. Since $M(\cdot | x_1, \cdots, x_l)$ is a probability measure on a σ-compact space, there exists for every $\varepsilon > 0$ a K compact so that

$$M(K | x_1, \cdots, x_l) > 1 - \varepsilon.$$

Choose a_i, b_i, $i = 1, \cdots, l$ so that

$$\{x_1, \cdots, x_l\} \cup K \subset \Pi_{i=1}^l [a_i, b_i].$$

Then

$$\begin{aligned} M(\Pi_{i=1}^l [a_i, b_i] | x_1, \cdots, x_l) &= M(\Pi_{i=1}^l g_0 [a_i, b_i] + g_i | g_0 x_1 + g_1, \cdots, g_0 x_l + g_l) \\ &= M(\Pi_{i=1}^l g_0 ([a_i, b_i] - x_i) + x_i | x_1, \cdots, x_l) \\ &> 1 - \varepsilon \end{aligned}$$

by inserting $g_i = x_i - g_0 x_i$. Now, let $g_0 \to 0$. Then

$$M(\{x_1, \cdots, x_l\} | x_1, \cdots, x_l) > 1 - \varepsilon$$

so that the only $M \in \mathfrak{M}_G$ is the Markov kernel given by $M(B|x_1, \cdots, x_l) = I_B(x_1, \cdots, x_l)$, $B \in \mathfrak{B}(\mathbb{R}^l)$. □

From now on in this section we will, unless explicitly otherwise stated, assume $n_A > r_A$ and $n_B > r_B$. Let $x = (x_1, \cdots, x_l, x_0), y = (y_1, \cdots, y_l, y_0) \in \mathbb{R}^l \times \mathbb{R}^+$ and define a group operation by

$$xy = (y_1 + y_0 x_1, \cdots, y_l + y_0 x_l, y_0 x_0).$$

Then $\mathbb{R}^l \times \mathbb{R}^+$ is solvable and consequently has an invariant mean. With the standard topology for $\mathbb{R}^l \times \mathbb{R}^+$ the operation just defined is continuous. Hence $\mathbb{R}^l \times \mathbb{R}^+$ is a locally compact group which is Hausdorff and satisfies the second axiom of countability.

Let $Z_1, \cdots, Z_l, W$ be independent random variables, $Z_i \sim N(0, 1)$, $i = 1, \cdots, l$ and $W \sim \chi^2_{n_A - r_A}$. Then

$$P_{l, n_A - r_A, \theta_1, \cdots, \theta_l, \sigma}(B)$$
$$= P_{l, n_A - r_A, 0, \cdots, 0, 1}((Z_1, \cdots, Z_l, W)(\theta_1, \cdots, \theta_l, \sigma) \in B)$$
$$= P_{l, n_A - r_A, 0, \cdots, 0, 1}\big(B(\theta_1, \cdots, \theta_l, \sigma)^{-1}\big) \quad \text{where} \quad B \in \mathfrak{B}(\mathbb{R}^l \times \mathbb{R}^+).$$

Similarly

$$Q_{l, n_A - l, \theta_1, \cdots, \theta_l, \sigma}(B)$$
$$= Q_{l, n_B - l, 0, \cdots, 0, 1}\big(B(\theta_1, \cdots, \theta_l, \sigma)^{-1}\big) \quad \text{where} \quad B \in \mathfrak{B}(\mathbb{R}^l \times \mathbb{R}^+)$$

so that $\mathcal{E}_A$ and $\mathcal{E}_B$ are translation experiments. Since $(P_{l, n_A - r_A, \theta_1, \cdots, \theta_l, \sigma} : (\theta_1, \cdots, \theta_l, \sigma) \in \mathbb{R}^l \times \mathbb{R}^+)$ is dominated the method described in Section 1 may be applied.

If $B \in \mathfrak{B}(\mathbb{R}^l \times \mathbb{R}^+)$, then

$$P''_{l, n_A - r_A, \theta_1, \cdots, \theta_l, \sigma}(B)$$
$$= P_{l, n_A - r_A, 0, \cdots, 0, 1}\big((\theta_1, \cdots, \theta_l, \sigma)^{-1} B\big)$$
$$= P_{l, n_A - r_A, 0, \cdots, 0, 1}\big((Z_1, \cdots, Z_l, W) \in (\theta_1, \cdots, \theta_l, \sigma)^{-1} B\big)$$
$$= P_{l, n_A - r_A, 0, \cdots, 0, 1}((\theta_1, \ldots, \theta_l, \sigma)(Z_1, \cdots, Z_l, W) \in B)$$
$$= P_{l, n_A - r_A, 0, \cdots, 0, 1}((Z_1 + \theta_1 W, \cdots, Z_l + \theta_l W, \sigma W) \in B)$$
$$= \int I_B(z_1 + \theta_1 w, \cdots, z_l + \theta_l w, w\sigma) \textstyle\prod_{i=1}^l \phi(z_i) \tilde{\gamma}_{n_A - r_A}(w)\, dw\, dz_1 \cdots dz_l$$
$$= \int_B \sigma^{-1} \tilde{\gamma}_{n_A - r_A}(w/\sigma) \textstyle\prod_{i=1}^l \phi(z_i - \theta_i w/\sigma)\, dz_1 \cdots dz_l dw.$$

Thus $P''_{l, n_A - r_A, \theta_1, \cdots, \theta_l, \sigma}$ has density $\sigma^{-1} \tilde{\gamma}_{n_A - r_A}(w/\sigma) \prod_{i=1}^l \phi(z_i - \theta_i w/\sigma)$ with respect to the Lebesgue measure. Here ϕ is the density of the standard normal distribution, and $\tilde{\gamma}_{n_A - r_A}$ is the density of S where S^2 is $\chi^2_{n_A - r_A}$ distributed.

Similarly, $Q_{l, n_B - l, 0, \cdots, 0, 1}$ has density $\tilde{\gamma}_{n_A - l}(w) \prod_{i=1}^l \Delta_i^{\frac{1}{2}} \phi(z_i \Delta_i^{\frac{1}{2}})$.

PROPOSITION 4.3. *If* $m = \#\{i : 0 < \Delta_i < 1\} = 0$ *and* $1 \leqslant n_A - r_A < n_B - l$, *then*

$$\delta(\mathcal{E}_A, \mathcal{E}_B) = \| P''_{l, n_A - r_A, 0, \cdots, 0, [(n_B - l)/(n_A - r_A)]^{\frac{1}{2}}} - Q_{l, n_B - l, 0, \cdots, 0, 1} \|$$

$$= E\left|1 - (n_B - l)\gamma_{n_B - l}(W)\left[(n_A - r_A)\gamma_{n_A - r_A}((n_A - r_A)W/(n_B - l))\right]^{-1} \times \Pi_{\Delta_i > 1}\Delta_i^{\frac{1}{2}}\exp(-(\Delta_i - 1)Z_i^2/2)\right|$$

where $Z_1, \cdots, Z_{r_A}$ *and* W *are independent,* $Z_i \sim N(0,1)$, $i = 1, \cdots, r_A$ *and* $(n_A - r_A)W/(n_B - l) \sim \chi^2_{n_A - r_A}$. *Here* $\gamma_{n_B - l}$ *denotes the density of the* $\chi^2_{n_B - l}$ *distribution.*

PROOF. Let $n_A - r_A = n_1$, $n_B - r_B = n_2$, $\nu = (n_2/n_1)^{\frac{1}{2}}$ and $\delta_x(B) = I_B(x)$. We must show that $\mu_0 = \delta_0 \times \cdots \times \delta_0 \times \delta_\nu$ is a least favorable distribution for testing

$$H : P''_{l, n_1, \theta_1, \cdots, \theta_l, \sigma}, (\theta_1, \cdots, \theta_l, \sigma) \in \mathbb{R}^l \times \mathbb{R}^+ \quad \text{against} \quad K : Q$$

at all levels α. Then the proposition will follow from the results given in Section 1.

The strongest α-level test δ_{μ_0} for H_{μ_0} against Q is given by

$$\delta_{\mu_0}(z_1, \cdots, z_l, w) = 1$$

$$\Leftrightarrow \Pi_{i=1}^l \Delta_i^{\frac{1}{2}}\phi(z_i\Delta_i^{\frac{1}{2}})\tilde{\gamma}_{n_2}(w) > C\Pi_{i=1}^l\phi(z_i)\tilde{\gamma}_{n_1}(w/\nu)\nu^{-1}$$

$$\Leftrightarrow \exp(-\Sigma_{i=1}^l(\Delta_i - 1)z_i^2/2)w^{n_2 - n_1}\exp(-w^2(1 - n_1/n_2)/2) > C^1$$

$$\Leftrightarrow (z_1, \cdots, z_l, w) \in K$$

where

(i) $\alpha = P''_{l, n_1, 0, \cdots, 0, \nu}(K)$

(ii) $K_w = \{(z_1, \cdots, z_l) : (z_1, \cdots, z_l, w) \in K\}$
$= \{(z_1, \cdots, z_l) : \Sigma_{i=1}^l(\Delta_i - 1)z_i^2/2 < -\log C^1 + \log[w^{n_2 - n_1}\exp(-(1 - n_1/n_2)w^2/2)]\}$

is an ellipse which may be degenerate since $\Delta_i = 1$ is possible.

(iii) $K_{z_1, \cdots, z_l} = \{w : (z_1, \cdots, z_l, w) \in K\} =]k_1(z_1, \cdots, z_l), k_2(z_1, \cdots, z_l)[$ where

$$k_1(z_1, \cdots, z_l)^{n_2 - n_1}\exp(-(1 - n_1/n_2)k_1^2(z_1, \cdots, z_l)/2) = k_2(z_1, \cdots, z_l)^{n_2 - n_1}\exp(-(1 - n_1/n_2)k_2^2(z_1, \cdots, z_l)/2).$$

Then if $k_3 = \max_w \log[w^{n_2 - n_1}\exp(-(1 - n_1/n_2)w^2/2)]$,

$$P''_{l, n_1, 0, \cdots, 0, \nu}(K) = \int_{\Sigma(\Delta_i - 1)z_i^2/2 < -\log C^1 + k_3}\int_{k_1(z_1, \cdots, z_l)}^{k_2(z_1, \cdots, z_l)}\Pi_{i=1}^l\phi(z_i)\nu^{-1}\gamma_{n_1}\left(\frac{w}{\nu}\right)dw\,dz_1 \cdots dz_l.$$

Let $E_{\theta_1, \cdots, \theta_l, \sigma}$ be the expectation taken relative to $P''_{l, n_1, \theta_1, \cdots, \theta_l, \sigma}$. Then

$$P''_{l, n_1, \theta_1, \cdots, \theta_l, \sigma}(K) = E_{\theta_1, \cdots, \theta_l, \sigma}\left[I_K(X_1, \cdots, X_l, W)\right] = E_{\theta_1, \cdots, \theta_l, \sigma}E_{\theta_1, \cdots, \theta_l, \sigma}\left[I_K(X_1, \cdots, X_l, W)|W\right].$$

But $E_{\theta_1,\cdots,\theta_l,\sigma}[I_K(X_1,\cdots,X_l,W)|W]$ is a function of $(X_1,\cdots,X_l,W)$ only through W. Thus the distribution is independent of $(\theta_1,\cdots,\theta_l)$. Consequently,

$$P''_{l,n_1,\theta_1,\cdots,\theta_l,\sigma}(K) = E_{0,\cdots,0,\sigma}E_{\theta_1,\cdots,\theta_l,\sigma}\left[I_K(X_1,\cdots,X_l,W)|W\right].$$

Furthermore,

$$E_{\theta_1,\cdots,\theta_l,\sigma}\left[I_K(X_1,\cdots,X_l,W)|W\right] \leqslant E_{0,\cdots,0,\sigma}\left[I_K(X_1,\cdots,X_l,W)|W\right]$$

since K_w is an ellipse with center in $(0,\cdots,0) \in \mathbb{R}^l$, and the probability for $(Z_1,\cdots,Z_l) \in K_w$ where $Z_1,\cdots,Z_l$ are independent random variables and $Z_i \sim N(\theta_i w/\sigma, 1)$, $i = 1,\cdots,l$ is maximized when the center of the ellipse and the center of the distribution coincide [1]. Thus

$$P''_{l,n_1,\theta_1,\cdots,\theta_l,\sigma}(K) \leqslant P''_{l,n_1,0,\cdots,0,\sigma}(K), \qquad (\theta_1,\cdots,\theta_l,\sigma) \in \mathbb{R}^l \times \mathbb{R}^+$$

and if we show that

$$P''_{l,n_1,0,\cdots,0,\sigma}(K) \leqslant P''_{l,n_1,0,\cdots,0,\nu}(K) = \alpha \qquad \text{for } \sigma > 0$$

it follows from Theorem 3.7 in [9] that μ_0 is a least favorable distribution as claimed above.

Let

$$\begin{aligned}\alpha(\sigma) &= P''_{l,n_1,0,\cdots,0,\sigma}(K)\\ &= \int_{\sum_{i=1}^l(\Delta_i-1)z_i^2/2<k_3-\log C^1}\Pi_{i=1}^l\phi(x_i)\left[\tilde{\Gamma}_{n_1}(\sigma^{-1}k_2(z_1,\cdots,z_l))\right.\\ &\qquad \left.-\tilde{\Gamma}_{n_2}(\sigma^{-1}k_1(z_1,\cdots,z_l))\right]dz_1\cdots dz_l\end{aligned}$$

where $\tilde{\Gamma}_{n_1}$ is the cumulative distribution function for W where W^2 is distributed as $\chi^2_{n_1}$. $\{P''_{l,n_1,0,\cdots,0,\sigma} : \sigma \in \mathbb{R}^+\}$ is an exponential family of distributions and

$$\alpha(\sigma) = \int I_K(z_1,\cdots,z_l,w)P''_{l,n_1,0,\cdots,0,\sigma}(dz_1,\cdots,dz_l,dw).$$

Hence by Theorem 2.9 in [9] differentiation with respect to σ under the integral sign is permitted, and

$$\begin{aligned}\alpha'(\sigma) &= \int_{\sum_{i=1}^l(\Delta_i-1)z_i^2<k_3+\log C^1}\Pi_{i=1}^l\phi(z_i)\sigma^{-2}\left[k_1(z_1,\cdots,z_l)\tilde{\gamma}_{n_1}(\sigma^{-1}k_1(z_1,\cdots,z_l))\right.\\ &\qquad \left.-k_2(z_1,\cdots,z_l)\tilde{\gamma}_{n_1}(\sigma^{-1}k_2(z_1,\cdots,z_l))\right]dz_1\cdots dz_l.\end{aligned}$$

By (iii) above $\alpha'(\nu) = 0$ so that α has an extremal point at ν.

Consider $f(t) = \tilde{\Gamma}(k_2/t) - \tilde{\Gamma}(k_1/t)$, $k_2 > k_1 > 0$, $t > 0$. $f(t)$ can have only one extremal point t_0. Since $f > 0$ and $f(t) \to 0$ as $t \to \infty$ or $t \to 0$, this must be a maximum point and hence $f'(t) < 0$ when $t > t_0$ and $f'(t) > 0$ when $t < t_0$. These results applied to the integrand in the expression for $\alpha'(\sigma)$ show that ν must be a maximum point. □

PROPOSITION 4.4. *If* $0 \leqslant n_B - l + \#\{i : 0 < \Delta_i < 1\} \leqslant n_A - r_A$, *then*

$$\begin{aligned}\delta(\mathcal{E}_A, \mathcal{E}_B) &= \| P'_{l-m,0,\cdots,0,1} - Q'_{l-m,0,\cdots,0,1}\|\\ &= E\left|1 - \Pi_{\Delta_i>1}\Delta_i^{\frac{1}{2}}\exp(-(\Delta_i-1)Z_i^2/2)\right|\end{aligned}$$

where $Z_1,\cdots,Z_{r_A}$ *are independent and* $N(0,1)$ *distributed random variables.*

REMARK. The following proof uses the fact that $n_B > r_B$ so that $\mathcal{E}_A$ and $\mathcal{E}_B$ are translation experiments. However, even if $n_B = r_B$, i.e., σ^2 is not estimable in $\mathcal{E}_B$, the result is true. This might be seen by using a similar argument to the one given below to obtain an upper bound for $\delta(\mathcal{E}_A, \mathcal{E}_B)$ and fixing σ and using the result from Section 3 to obtain a lower bound.

PROOF. By Proposition 2.1 in Hansen and Torgersen [3] the deficiency of the experiment given by $X_{l-m+1}, \cdots, X_l, S_A$ with respect to the experiment given by $Y_{l-m+1}, \cdots, Y_l, S_B$ is 0, i.e., the first one is more informative than the second. Hence there exists an invariant Markov kernel M so that

$$P_{m,n_A-r_A,\theta_{l-m+1},\cdots,\theta_l,\sigma}M = Q_{m,n_B-l,\theta_{l-m+1},\cdots,\theta_l,\sigma}$$

for all $(\theta_{l-m+1}, \cdots, \theta_l, \sigma) \in \mathbb{R}^m \times \mathbb{R}^+$. Since M is invariant, it may be represented by a probability measure μ_2 and

$$P_{m,n_A-r_A,\theta_{l-m+1},\cdots,\theta_l}M = \mu_2 * P_{m,n_A-r_A,\theta_{l-m+1},\cdots,\theta_l,\sigma}$$

for all $(\theta_{l-m+1}, \cdots, \theta_l, \sigma) \in \mathbb{R}^m \times \mathbb{R}^+$. Hence

$$\begin{aligned} Q_{m,n_B-l,0,\cdots,0,1}(B) &= \mu_2 * P_{m,n_A-r_A,0,\cdots,0,1}(B) \\ &= \mu_2 \times P_{m,n_A-r_A,0,\cdots,0,1}(\{(x,y) : xy \in B\}) \\ &= \int P_{m,n_A-r_A,0,\cdots,0,1}(x^{-1}B)\mu_2(dx) \\ &= \int P''_{m,n_A-r_A,\theta_{l-m+1},\cdots,\theta_l,\sigma}(B)\mu_2(d\theta_{l-m+1}, \cdots, d\theta_l, d\sigma). \end{aligned}$$

Now, let $\mu_1 = \delta_0 \times \cdots \times \delta_0$ be the probability measure on $\mathbb{R}^{l-m}$ with all mass in 0, and let $\mu_0 = \mu_1 \times \mu_2$. Then the strongest α-level test for H_{μ_0} against Q has rejection region of the form

$$\{(z_1, \cdots, z_l, w) : \exp(-\Sigma_{i=1}^{l-m}(\Delta_i - 1)z_i^2/2 > k\}.$$

The projection into $\mathbb{R}^{l-m}$ is an ellipse and by using an argument with conditioning on W, as in the proof of the preceding proposition, it follows that μ_0 is at least favorable distribution for all levels α. □

It still remains to consider the case $1 \leqslant n_A - r_A < n_B - r_B + \#\{i : 0 < \Delta_i < 1\}$ and $\#(\Delta_i : 0 < \Delta_i < 1\} > 0$. We have not been able to find an expression for $\delta(\mathcal{E}_A, \mathcal{E}_B)$ in this case.

Consider now the situation treated by Le Cam for σ known, i.e., AA' nonsingular and $BB' - AA'$ positive semidefinite so that $F'AA'F = I$ and $F'BB'F = \Delta$ for a nonsingular $r_A \times r_A$ matrix F. Since BB' is positive semidefinite and since $\Delta_1, \cdots, \Delta_{r_A}$ are the solutions of $\det[BB' - \Delta AA'] = 0$, $\Delta_1, \cdots, \Delta_{r_A} \geqslant 1$. By noticing that AA' nonsingular implies BB' nonsingular, $\delta(\mathcal{E}_A, \mathcal{E}_B)$ may now be found by Proposition 4.1 if $r_A = n_A < n_B$, by Proposition 4.3 if $r_A < n_A < n_B$ and by Proposition 4.4 if $r_A \leqslant n_B \leqslant n_A$.

COROLLARY 4.1. *If σ^2 is unknown, then for any pair $\mathcal{E}_A, \mathcal{E}_B$ of linear normal experiments such that $BB' \geqslant AA'$ and A is of full rank*

$$\begin{aligned}\delta(\mathcal{E}_A, \mathcal{E}_B) &= \|N(0,(AA')^{-1}) - N(0,(BB')^{-1})\| \quad \text{if } r_A \leqslant n_B \leqslant n_A \\ &= \|N(0,(AA')^{-1}) \times \Gamma_{n_A - r_A, (n_B - r_B)/(n_A - r_A)} \\ &\quad - N(0,(BB')^{-1}) \times \Gamma_{n_B - r_B, 1}\| \quad \text{if } r_A < n_A < n_B\end{aligned}$$

where $\Gamma_{n,t}$ is the measure corresponding to S if S/t is χ^2_n distributed.

5. Applications and examples. Let $\mathfrak{M}_i$ and $\mathfrak{M}_a$ denote the least informative and the totally informative experiments respectively. Then $\delta(\mathfrak{M}_i, \mathcal{E}_A)$ and $\delta(\mathcal{E}_A, \mathfrak{M}_a)$ are absolute measures of the information in the experiment $\mathcal{E}_A$. Unfortunately, for translation experiments on the real line both of these deficiencies are equal to 2, as shown by Torgersen [20]. Hence $\delta(\mathfrak{M}_i, \mathcal{E}_A) = \delta(\mathcal{E}_A, \mathfrak{M}_i) = 2$ for the case where σ is known and consequently also for σ unknown.

However, if an experiment is given by the independent, identically distributed observations $X_1, \cdots, X_n$, deficiencies may be used to compute the information in an additional observation.

In the experiments considered in this paper the observations are not identically distributed. The question then arises whether deficiencies may be of help to determine the regression coefficients for additional coefficients.

Let $A = (a_{ij})$ be an $k \times n_A$ matrix which we suppose is of rank k. Let $B = [A \,|\, t]$ be the $k \times (n_A + 1)$ matrix having the columns of A as the first n_A columns and the vector $t = (t_1, \cdots, t_k)'$ as the $(n_A + 1)$th column. Then

$$\begin{aligned}\det[BB' - \Delta AA'] &= \det[AA' + tt' - \Delta AA'] \\ &= \det[tt' - (\Delta - 1)AA'].\end{aligned}$$

Since $\operatorname{rank}[tt'] = 1$, the solutions of $\det[tt' - \Delta AA'] = 0$ are all equal to zero, except one. Hence, if Δ_0 is the nonzero solution,

$$\begin{aligned}\det[tt' - \Delta(AA')] &= \det[AA'] \det[tt'(AA')^{-1} - \Delta I] \\ &= \det[AA'](-\Delta)^{k-1}(\Delta_0 - \Delta) \\ &= [(-1)^k \Delta^k + (-1)^{k-1} \Delta_0 \Delta^{k-1}] \det[AA'].\end{aligned}$$

On the other hand

$$\begin{aligned}\det[tt'(AA')^{-1} - \Delta I] &= (-\Delta)^k + \operatorname{tr}[tt'(AA')^{-1}](-\Delta)^{k-1} \\ &\quad + \text{factors of lower order in } \Delta.\end{aligned}$$

Here tr denotes trace. Hence $\Delta_0 = \operatorname{tr}(tt')(AA')^{-1}$, and consequently $1 + \operatorname{tr}[(tt')(AA')^{-1}], 1, \cdots, 1$ are the k solutions of $\det[BB' - \Delta AA'] = 0$.

Now, let $Z \sim N(0,1)$. Then by noticing that $\operatorname{tr}[(tt')(AA')^{-1}] = t'(AA')^{-1}t$,

$$\begin{aligned}\delta(\mathcal{E}_A, \mathcal{E}_B) &= E|1 - \left[t'(AA')^{-1}t + 1\right]^{\frac{1}{2}} \exp\left(-t'(AA')^{-1}tZ^2/2\right)| \\ &= 4\Big[\Phi\Big(\Big[\big(1 + t'(AA')^{-1}t\big)\big(t'(AA')^{-1}t\big)^{-1}\log\big(1 + t'(AA')^{-1}t\big)\Big]^{\frac{1}{2}}\Big) \\ &\qquad - \Phi\Big(\Big[\big(t'(AA')^{-1}t\big)^{-1}\log\big(1 + t'(AA')^{-1}t\big)\Big]^{\frac{1}{2}}\Big)\Big]\end{aligned}$$

when σ is known. In the expression above we have written the integrand f and used that $\int|f| = 2\int f^+$.

If Z and W are independent, $Z \sim N(0,1)$, $(n-k)W/(n-k+1) \sim \chi^2_{n-k}$, then

$$\delta(\mathcal{E}_A, \mathcal{E}_B) = E\left|1 - \left[\frac{n-k+1}{n-k}\right]^{(n-k)/2} \frac{\Gamma\left(\frac{n-k}{2}\right)}{\Gamma\left(\frac{n-k+1}{2}\right)} \left[\left(1 + t'(AA')^{-1}t\right)\frac{W}{2}\right]^{\frac{1}{2}} \times \exp\left(-t'(AA')^{-1}t\frac{Z}{2} - \frac{W}{2(n-k+1)}\right)\right|$$

when σ is unknown.

Another possible use of the deficiencies might be in optimal experimental design. Here various plans are compared in terms of such functions as $\det(AA')$ and $\operatorname{tr}(AA')$ which are functions of the design matrix A'. In comparison of experiments there are two different experiments involved and a possible choice as reference design might be one with a design matrix having the upper $k \times k$ matrix equal to the identity matrix and the lower $(n-k) \times k$ matrix consisting only of zeroes. This would correspond to a situation where $n - k$ observations are taken with the regression coefficients in a fixed state, and taking the k remaining with one regression coefficient at a time differing from the fixed value.

By comparing this experiment with one with design matrix A' it will be the eigenvalues of AA' that determine the deficiency. The dependence on the eigenvalues are, however, much more complicated than for functions such as $\det(AA')$ and $\operatorname{tr}(AA')$.

Finally as an example let us consider regression with two variables, i.e., $A = (a_{ij})$ and $B = (b_{ij})$ are $2 \times n_A$ and $2 \times n_B$ matrices respectively with $a_{1j} = 1$, $a_{2j} = t_j$, $j = 1, \cdots, n_A$ and $b_{1j} = 1$, $b_{2j} = s_j$, $j = 1, \cdots, n_B$. Assume $\operatorname{rank}[A] = 2$ so that not all $t_1, \cdots, t_{n_A}$ are equal. If $M_A = \Sigma_{i=1}^{n_A}(t_i - \bar{t})^2$, $M_B = \Sigma_{i=1}^{n_B}(s_i - \bar{s})^2$, $\det[BB' - \Delta AA'] = 0$ has two solutions given by

$$[2n_A M_A]^{-1}\Big[n_B M_A + n_A M_B + n_A n_B(\bar{s} - \bar{t})^2 \pm \Big[\big(n_B M_A - n_A M_B + n_A n_B(\bar{t} - \bar{s})^2\big)^2 - 4n_A n_B M_A M_B\Big]^{\frac{1}{2}}\Big].$$

$\delta(\mathcal{E}_A, \mathcal{E}_B)$ may now be found for σ known, and for σ unknown, except when $0 < n_A - 2 < n_B - \text{rank}[B] + \#\{I : 0 < \Delta_i < 1\}$ and $\#\{i : 0 < \Delta_i < 1\} > 0$.

Acknowledgment. This work is based on the author's master thesis at the University of Oslo, 1975. The author wants to thank Professor Erik N. Torgersen for suggesting the problem and for valuable suggestions that led to improvements in the presentation and simplified several of the proofs.

REFERENCES

[1] ANDERSON, T. W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6** 170–176.

[2] ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.

[3] HANSEN, O. H. and TORGERSEN, E. N. (1974). Comparison of linear normal experiments. *Ann. Statist.* **2** 367–373.

[4] LE CAM. L. (1964). Sufficiency and approximate sufficiency. *Ann. Math. Statist.* **35** 1419–1455.

[5] LE CAM, L. (1975). Distances between experiments. In: *A Survey of Statistical Design and Linear Models* (J. N. Srivastava, ed.), pages 385–396. North-Holland, Amsterdam.

[6] LINDQVIST, B. (1977). How fast does a Markov chain forget the initial state? A decision theoretic approach. *Scand. J. Statist.* **4** 145–152.

[7] LINDQVIST, B. (1978a). On the loss of information incurred by lumping states of a Markov chain. *Scand. J. Statist.* **5** 92–98.

[8] LINDQVIST, B. (1978b). A decision theoretical characterization of weak ergodicity. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **44** 155–158.

[9] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

[10] TORGERSEN, E. N. (1972). Comparison of translation experiments. *Ann. Math. Statist.* **43** 1383–1399.

[11] TORGERSEN, E. N. (1979). Measures of information based on comparison with total information and with total ignorance. I & II. Unpublished manuscript.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720