The Annals of Mathematical Statistics
1969, Vol. 40, No. 5, 1809–1823

ASYMPTOTICALLY NEARLY EFFICIENT ESTIMATORS OF MULTIVARIATE LOCATION PARAMETERS[1]

By D. S. Moore

Purdue University

1. Introduction and Summary. Common estimates of multivariate location parameters have the property that each component of the parameter is estimated using only the corresponding component of the observations. This is true of the sample mean, sample median and the vector of medians of averages (studied in [1]), as well as of the rank-order statistics often applied to testing for location. In some cases, particularly the multivariate normal, such estimators achieve asymptotic efficiency, but in general information is lost. This paper presents three methods of estimating multivariate location parameters which use more information than is available in the marginal distributions. These classes of estimators are *asymptotically nearly efficient* (ANE), in the sense that for every $\epsilon > 0$ there is an estimator in the class with asymptotic efficiency $> 1 - \epsilon$ (if efficiency is measured by a comparison of the asymptotic covariance matrix to the inverse of the information matrix).

Our ANE estimators are motivated by those of Ogawa [6] for univariate location parameters. Ogawa obtained the asymptotically minimum-variance asymptotically unbiased estimator (ABLUE) for location or scale from a chosen set of sample quantiles. It was soon observed (Tischendorf [10]) that the reciprocal of the asymptotic variance of Ogawa's estimator (properly normalized) is essentially a Riemann sum for the information integral for the parameter being estimated. Thus under mild regularity conditions the ABLUE approaches asymptotic efficiency as larger sets of more closely spaced quantiles are chosen for use. Ogawa's estimators are therefore ANE for univariate location parameters.

In the present paper we describe three classes of ANE estimators for multivariate location parameters. The first two consist of linear estimators, and represent multivariate generalizations of Ogawa's ANE class. Our three classes are as follows: (1) Choose a set of marginal sample quantiles in each direction from a continuous r-variate location parameter distribution. These quantiles generate a random partition of Euclidean r-space R_r, and for $r > 1$ the observed cell frequencies contain additional information. We obtain the ABLUE's in terms of the sample quantiles and the observed cell frequencies for $r = 2$ and show that they are ANE. (2) For all $r > 1$, linear ANE estimators are obtained by choosing a single sample quantile in each direction and partitioning R_r by mark-

Received 25 March 1968; revised 29 January 1969.

[1] This is a revision of part of the author's Ph.D. dissertation at Cornell University. Research supported in part by the Office of Naval Research under Contract No. NONR 401(50). Reproduction in whole or in part is permitted for any purpose of the United States Government.

ing off fixed distances from these. The ABLUE's in terms of the r chosen quantiles and the observed cell frequencies are ANE. These estimators have much simpler coefficients than do those of class (1). (3) Finally, ANE estimators can be obtained by exploiting analytic properties of RBAN estimators (Neyman [5]) for a sequence of multinomial problems related to the given location parameter family. These estimators are usually not expressed in closed form.

Ogawa proceeded by applying least squares theory to the asymptotic distribution of his chosen set of sample quantiles. The ABLUE's of classes (1) and (2) are here derived by the same method, but establishing the joint asymptotic distribution of the marginal sample quantiles and the observed cell frequencies is non-trivial. Our method is to reduce the problem to one involving the multinomial distribution. A similar idea was used by Weiss [11] to obtain the joint asymptotic distribution of the quantiles alone, but the present problem requires more elaborate arguments.

Section 2 contains a preliminary result for the multinomial distribution. The ANE classes (1) and (2) are discussed in Section 3, while Section 4 presents the third class. Estimators of all three classes will require use of a computer if the distribution function cannot be expressed in closed form, and may therefore be of limited practical usefulness. Section 5, however, contains an example for which estimators of classes (1) and (2) can be computed with relative ease. For this example, a bivariate logistic distribution, the performance of our estimators is compared with that of the sample mean and median and of Ogawa's univariate ANE estimators.

Throughout, K denotes a generic positive constant, $\mathfrak{L}\{X\}$ is the probability law of the random variable X, and $\mathfrak{L}\{X_n\} \to \mathfrak{L}\{X\}$ designates convergence in law. $N(\mu, \Sigma)$ is the normal law with mean μ and covariance matrix Σ (which may be 1×1).

2. A Preliminary Lemma. We require a limit theorem for the conditional distribution of multinomial random variables given certain linear restraints. This result has often been used in the literature on conditional chi-square tests, but that literature does not seem to contain a proof. The lemma below can be shown to follow from the very general theorems of Section 2 of Steck [9], but we will give a more direct proof based on the uniform local limit theorem.

Let $\{n_i : i = 0, \cdots, M\}$ be multinomial random variables with parameters $(n, \{P_i\})$, where $0 < P_i < 1$ for all i, and set $Q_i = n^{\frac{1}{2}}(n_i/n - P_i)$. $\{Z_i : i = 0, \cdots, M\}$ are random variables having the jointly normal asymptotic distribution of the Q_i. $\{b_i\}^*$ will denote the set of objects b_i having indices in a stated subset of $\{0, \cdots, M\}$.

LEMMA. *Let $a_{1n}, \cdots, a_{mn}$ be constants such that $a_{in} \to a_i$ for each $1 \leqq i \leqq m$ and*

$$P[\textstyle\sum_{A_1} Q_i = a_{1n}, \cdots, \sum_{A_m} Q_i = a_{mn}] > 0$$

for all n, where $A_1, \cdots, A_m$ are subsets of the index set $\{0, \cdots, M\}$. Denote by

$\{Q_i\}^*$ a set of $M - m$ Q_i's which have a non-degenerate distribution given the m specified linear restraints. Then

$$(2.1) \quad \mathcal{L}\{\{Q_i\}^* \mid \textstyle\sum_{A_1} Q_i = a_{1n}, \cdots, \sum_{A_m} Q_i = a_{mn}\} \rightarrow \mathcal{L}\{\{Z_i\}^* \mid \sum_{A_1} Z_i = a_1, \cdots, \sum_{A_m} Z_i = a_m\}.$$

PROOF. For the sake of notational simplicity we give the proof for the case of a single linear restraint, $\sum_{i=0}^k Q_i = a_n$, where $a_n \rightarrow a$. $\{Q_i\}^*$ will be chosen to be $\{Q_1, \cdots, Q_{M-1}\}$. It is well known that the multinomial n_i have the distribution of $M + 1$ independent Poisson random variables with means nP_i, conditional on $\sum_{i=0}^M n_i = n$. Regarding the n_i as Poisson allows the use of a univariate local limit theorem. We must find the limit of

$$\mathcal{L}\{\{Q_i\}^* \mid \textstyle\sum_{i=0}^k Q_i = a_n, \sum_{i=0}^M Q_i = 0\} = \mathcal{L}\{\{Q_i\}^* \mid R\}$$

where R denotes the linear restraints.

Let $p_0(x), \cdots, p_M(x)$ be the probability functions of the Q_i, and $n_0(x), \cdots, n_M(x)$ the densities of the limiting $N(0, P_i)$ distributions. The results we use, describing the convergence of $n^{\frac{1}{2}}p_i(x)$ to $n_i(x)$ as $n \rightarrow \infty$ and also $x \rightarrow \infty$, are familiar for the binomial case (see Feller [2], Chapter VII). They are proved for more general random variables in Richter [7].

First, for fixed $q_1, \cdots, q_{M-1}$,

$$\begin{aligned} P[\{Q_i\}^* = \{q_i\}^* \text{ and } R] &= p_0(a_n - \textstyle\sum_{i=1}^k q_i) \cdot \prod_{i=1}^{M-1} p_i(q_i) \cdot p_M(-a_n - \sum_{i=k+1}^{M-1} q_i) \\ &\sim n^{-\frac{1}{2}(M+1)} n_0(a - \textstyle\sum_{i=1}^k q_i) \cdot \prod_{i=1}^{M-1} n_i(q_i) \cdot n_M(-a - \sum_{i=k+1}^{M-1} q_i). \end{aligned}$$

Let now $r = n^{\epsilon}$ for $0 < \epsilon < \frac{1}{6}$, so that $r \rightarrow \infty$ but $r = o(n^{1/6})$. Define the square D_r in R_{M-1} by

$$D_r = \{x: -r \leqq x_i \leqq r, i = 1, \cdots, M - 1\}.$$

Then by the local limit theorem,

$$\begin{aligned} P[\{Q_i\}^* \not\in D_r] &\leqq \textstyle\sum_{i=1}^{M-1} P[Q_i \not\in [-r, r]] \\ &\sim K(1 - \Phi(r)) \end{aligned}$$

where Φ is a normal df. The last expression tends to 0 exponentially as $n \rightarrow \infty$ by a standard estimate for the tails of the normal distribution (Feller [2], page 166). Therefore

$$\begin{aligned} P[R] &= \textstyle\sum_{\{q_i\}^*} P[\{Q_i\}^* = \{q_i\}^* \text{ and } R] \\ &\sim \textstyle\sum_{D_r} P[\{Q_i\}^* = \{q_i\}^* \text{ and } R] \\ &\sim n^{-\frac{1}{2}(M+1)} \textstyle\sum_{D_r} \{n_0(a - \sum_{i=1}^k q_i) \cdot \prod_{i=1}^{M-1} n_i(q_i) \cdot n_M(-a - \sum_{i=k+1}^{M-1} q_i)\} \end{aligned}$$

Again as a consequence of the uniform local limit theorem, n times the last expression above is asymptotic to

$$\int \cdots \int_{D_r} n_0(a - \textstyle\sum_{i=1}^k z_i) \cdot \prod_{i=1}^{M-1} n_i(z_i) \cdot n_M(-a - \sum_{i=k+1}^{M-1} z_i)\, dz_1 \cdots dz_{M-1},$$

which converges to the corresponding integral (say $P_\infty[R]$) over R_{M-1} as $n \to \infty$. So $nP[R] \sim P_\infty[R]$.

Combining these results, we have for the probability function of $\{Q_i\}^*$ conditional on R that uniformly in $|q_i| < n^\epsilon$ (by Theorem 3 of [7])

$$\begin{aligned} n^{\frac{1}{2}(M-1)} P[\{Q_i\}^* &= \{q_i\}^* \text{ and } R]/P[R] \\ &\sim n_0(a - \textstyle\sum_{i=1}^k q_i) \cdot \prod_{i=1}^{M-1} n_i(q_i) \\ &\quad \cdot n_M(-a - \textstyle\sum_{i=k+1}^{M-1} q_i)/P_\infty[R]. \end{aligned} \tag{2.2}$$

The right side of (2.2) is the density of $Z_1^*, \cdots, Z_{M-1}^*$ given the linear restraints $\sum_{i=0}^k Z_i^* = a$ and $\sum_{i=0}^M Z_i^* = 0$, where $Z_0^*, \cdots, Z_M^*$ are independent $N(0, P_i)$ random variables. But the conditional distribution of $Z_0^*, \cdots, Z_M^*$ given $\sum_{i=0}^M Z_i^* = 0$ is the distribution of $Z_0, \cdots, Z_M$. Therefore the right side of (2.2) is the density of the right side of (2.1) for $m = 1$, and (2.2) implies (2.1).

3. Two linear ANE estimators. A natural r-variate analog of Ogawa's estimator is the ABLUE from chosen sets of sample quantiles in each direction and the observed cell frequencies of the partition of R_r generated by the quantiles. Unfortunately, the coefficients of these linear estimators are very complicated. Theorem 1 is therefore stated only for the bivariate case.

For $0 < \alpha_1 < \cdots < \alpha_L < 1$, denote by ξ_i, $i = 1, \cdots, L$, the sample α_i-quantiles based on the x-components of n independent observations from a population with continuous bivariate location parameter cdf $F(x - \theta_1, y - \theta_2)$. The corresponding marginal population quantiles are $x_i = u_i + \theta_1$, where u_i is the population α_i-quantile for $\theta_1 = 0$. For given $0 < \beta_1 < \cdots < \beta_M < 1$, the sample and population quantiles from the y-component are denoted by $\zeta_1, \cdots, \zeta_M$ and $y_1, \cdots, y_M$. Here $y_j = v_j + \theta_2$, where v_j is the marginal population β_j-quantile for $\theta_2 = 0$. The LM sample quantiles partition the plane into $(L + 1)(M + 1)$ cells. Let N_{ij} denote the number of observations falling in the open cell with "northeast corner" (ξ_i, ζ_j). (The cells not indexed by this scheme are redundant when the sample quantiles are given.) Define for $i = 1, \cdots, L + 1$ and $j = 1, \cdots, M + 1$ the probabilities

$$P_{ij} = F(u_i, v_j) - F(u_i, v_{j-1}) - F(u_{i-1}, v_j) + F(u_{i-1}, v_{j-1}) \tag{3.1}$$

with the conventions

$$u_0 = v_0 = -\infty, \qquad u_{L+1} = v_{M+1} = +\infty. \tag{3.2}$$

We abbreviate (3.1) as $P_{ij} = \Delta_{ij}F$, thereby defining difference operators Δ_{ij}.

The P_{ij} can be thought of as asymptotic cell probabilities. Theorem 1 states that the ABLUE's in terms of the sample quantiles and the $Q_{ij} = N_{ij}/n - P_{ij}$ are ANE. We first introduce some notation.

Let F_1 and F_2 denote the first partial derivatives of $F(x, y)$ with respect to x and y, respectively, and define the quantities

$$\delta_{ij} = F_1(u_i, v_j) - F_1(u_i, v_{j-1}), \kappa_{ij} = F_2(u_i, v_j) - F_2(u_{i-1}, v_j)$$
$$a_i = \sum_{j=1}^{M+1} \delta_{ij}(\Delta_{ij}F_1/P_{ij} - \Delta_{i+1,j}F_1/P_{i+1,j})$$
$$b_j = \sum_{i=1}^{L+1} \kappa_{ij}(\Delta_{ij}F_2/P_{ij} - \Delta_{i,j+1}F_2/P_{i,j+1})$$
$$c_i = \sum_{j=1}^{M} (\delta_{ij}\kappa_{ij}/P_{ij} - \delta_{i,j+1}\kappa_{ij}/P_{i,j+1}$$
$$-\delta_{ij}\kappa_{i+1,j}/P_{i+1,j} + \delta_{i,j+1}\kappa_{i+1,j}/P_{i+1,j+1})$$
$$d_j = \sum_{i=1}^{L} (\text{summand as in } \quad c_i)$$
$$s_{ij} = -\Delta^* F_1, r_{ij} = -\Delta^* F_2$$

where

$$\Delta^* H = \frac{\Delta_{ij}H}{P_{ij}} - \frac{\Delta_{i,M+1}H}{P_{i,M+1}} - \frac{\Delta_{L+1,j}H}{P_{L+1,j}} + \frac{\Delta_{L+1,M+1}H}{P_{L+1,M+1}}.$$

Any term above containing a factor $1/P_{rs}$ for $P_{rs} = 0$ is interperted to be 0. Note that the conventions (3.2) give $F_1(u_{L+1}, y) = 0$, $F_2(u_{L+1}, y) = f_Y(y)$ (the marginal density of Y), and $F_1(u_0, y) = F_2(u_0, y) = 0$, etc. Define the 2×2 matrix $I^* = |I_{kr}^*|$ by

$$I_{kr}^* = \sum_{i=1}^{L+1} \sum_{j=1}^{M+1} (\Delta_{ij}F_k)(\Delta_{ij}F_r)/P_{ij}.$$

$I = |I_{kr}|$ will denote the information matrix,

$$I_{kr} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [f_k(x, y)f_r(x, y)/f(x, y)]\, dx\, dy.$$

Finally, let

$$\mu_1 = \sum_{i=1}^{L} a_i(\xi_i - u_i) + \sum_{j=1}^{M} d_j(\zeta_j - v_j) + \sum_{i=1}^{L} \sum_{j=1}^{M} s_{ij}Q_{ij}$$
$$\mu_2 = \sum_{i=1}^{L} c_i(\xi_i - u_i) + \sum_{j=1}^{M} b_j(\zeta_j - v_j) + \sum_{i=1}^{L} \sum_{j=1}^{M} r_{ij}Q_{ij}$$

and define $(\theta_1^*, \theta_2^*)' = (I^*)^{-1}(\mu_1, \mu_2)'$, where prime denotes transpose.

THEOREM 1. *Suppose that the density $f(x, y)$ of $F(x, y)$ is continuous in the plane. Then (θ_1^*, θ_2^*) are the* ABLUE's *for (θ_1, θ_2) in terms of the sample quantities and the N_{ij}. If (θ_1, θ_2) is true,*

$$\mathcal{L}\{n^{\frac{1}{2}}(\theta_1^* - \theta_1), n^{\frac{1}{2}}(\theta_2^* - \theta_2)\} \to N(0, (I^*)^{-1}).$$

I^* is the information matrix for (θ_1, θ_2) from the joint asymptotic distribution of $\{n^{\frac{1}{2}}(\xi_i - x_i), n^{\frac{1}{2}}(\zeta_j - y_j), n^{\frac{1}{2}}Q_{ij}: i = 1, \cdots, L$ and $j = 1, \cdots, M\}$.

If the information integrals I_{kr} are finite and the derivatives f_1 and f_2 of f are continuous in the plane, each I_{kr}^* can be made as close as desied to I_{kr} by choosing α_1 and β_1 sufficiently near 0, *α_L and β_M sufficiently near* 1 *and L, M sufficiently large with* $\max_i |\alpha_i - \alpha_{i-1}|$ *and* $\max_j |\beta_j - \beta_{j-1}|$ *sufficiently small.*

The proof differs only in detail from that of Theorem 2 below. Since it is computationally more complicated, we omit it.

Let now $F(x_1 - \theta_1, \cdots, x_r - \theta_r)$ be a continuous r-variate location parameter family. For $j = 1, \cdots, r$, denote by ξ_j the sample α_j-quantile from the jth components of n independent observations on F. Partition the x_j — axis into $M_j + 2$ intervals by marking off known distances from ξ_j. These intervals are indexed $0, \cdots, M_j + 1$, with those indexed 0 and $M_j + 1$ being the half-infinite intervals at the tail, and those indexed $0, \cdots, K_j$ lying to the left of ξ_j. The Cartesian product of these partitions is a partition of Euclidean r-space R_r. We index the cells of this partition by attaching index $(i_1, \cdots, i_r)$ to the product of the i_jth interval in the x_j direction, for $j = 1, \cdots, r$. Usually we use σ as an abbreviation for $(i_1, \cdots, i_r)$. Denote by N_σ the number of n independent observations on F falling in the σth cell. We take all cells to be open.

Let ν_j be the $\theta = 0$ population α_j-quantile in the x_j direction, and mark off the given set of fixed distances from ν_j. This scheme yields a nonrandom partition of R_r. The probability when $\theta = 0$ of an observation on F falling in the σth cell of this partition is $P_\sigma = \Delta_\sigma F$, which defines the difference operator Δ_σ. (Explicit definitions may be found in many texts.) We also need the $r \times r$ information matrices I and I^* having entries

$$I_{ks} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} [f_k(x_1, \cdots, x_r) f_s(x_1, \cdots, x_r)/f(x_1, \cdots, x_r)]\, dx_1 \cdots dx_r$$

$$I_{ks}^* = \sum_\sigma (\Delta_\sigma F_k) \cdot (\Delta_\sigma F_s)/P_\sigma$$

for $k, s = 1, \cdots, r$, where $F_k = \partial F/\partial x_k$, etc. In the definition of I_{ks}^*, terms for which $P_\sigma = 0$ are regarded as themselves being 0. The same convention applies to the $c_{k\sigma}$ defined below.

The cells of the partition are linearly dependent when the sample quantiles are given. We therefore omit as redundant the following $r + 1$ cells: those indexed $(M_1 + 1, \cdots, M_{j-1} + 1, 0, M_{j+1} + 1, \cdots, M_r + 1)$ for $j = 1, \cdots, r$ and that indexed $(M_1 + 1, M_2 + 1, \cdots, M_r + 1)$. $\sum^*$ will denote summation over all indices except those of the omitted cells. Denote by $A_{k_1 \cdots k_s}$ the collection of indices $\sigma = (i_1, \cdots, i_r)$ such that all $i_j > K_j$ except $i_{k_m} \leqq K_{k_m}$ for $m = 1, \cdots, s$. A_0 is the collection of indices with all $i_j > K_j$. Δ_k^* and P_k^* will denote Δ_σ and P_σ for that σ with all $i_j = M_j + 1$ except $i_k = 0$, and Δ^* and P^* the corresponding quantities for $\sigma = (M_1 + 1, \cdots, M_r + 1)$. Finally, let ξ be the $r \times 1$ vector with jth component $\xi_j - \nu_j$, $j = 1, \cdots, r$.

With this considerable weight of notation, we can define a vector θ^{**} of estimators of $(\theta_1, \cdots, \theta_r) = \theta'$ by $\theta^{**} = \xi - (I^*)^{-1} Q$ where Q is the $r \times 1$ vector with jth component $\sum^* c_{j\sigma}(N_\sigma/n - P_\sigma)$ and the coefficients $c_{1\sigma}, \cdots, c_{r\sigma}$ are given by

$$\begin{aligned} c_{j\sigma} &= \Delta_\sigma F_j/P_\sigma - \Delta^* F_j/P^* && \sigma \in A_0 \\ &= \Delta_\sigma F_j/P_\sigma - \textstyle\sum_{m=1}^{s} \Delta_{k_m}^* F_j/P_{k_m}^* + (s-1)\Delta^* F_j/P^*, && \sigma \in A_{k_1 \cdots k_s}, s = 1, \cdots, r. \end{aligned}$$

THEOREM 2. *Let $F(x_1 - \theta_1, \cdots, x_r - \theta_r)$ be a continuous location parameter family with continuous density $f(x_1 - \theta_1, \cdots, x_r - \theta_r)$. Then the components of θ^{**} are the* ABLUE's *of $(\theta_1, \cdots, \theta_r)$ in terms of the ξ_j and the N_σ. When θ is true,*

$$\mathfrak{L}\{n^{\frac{1}{2}}(\theta^{**} - \theta)\} \to N(0, (I^*)^{-1}).$$

I^* is the information matrix for θ from the asymptotic distribution of

$$\{n^{\frac{1}{2}}(\xi_j - \nu_j - \theta_j),\ n^{\frac{1}{2}}(N_\sigma/n - P_\sigma) : \text{all } j \text{ and } \sigma\}.$$

If the information integrals I_{ks} are finite and the derivatives $f_1, \cdots, f_r$ are continuous, each I_{ks}^* may be made as close as desired to I_{ks} by appropriately choosing the set of fixed distances used in defining the estimators.

Proof. For notational convenience, we give the proof for the bivariate case, $r = 2$. Write (x, y) and (L, M) for (x_1, x_2) and (M_1, M_2), and let (u_i, v_j) be the vertices of the partition of the plane obtained by marking off the given distances from the $\theta = 0$ population quantiles (ν_1, ν_2). Then $P_{ij} = \Delta_{ij}F$ is given by (3.1), with the conventions that $u_{-1} = v_{-1} = -\infty$ and $u_{L+1} = v_{M+1} = \infty$.

Form a non-random partition of the plane by marking off the given fixed distances from the point $(\nu_1 + \theta_1 + u/n^{\frac{1}{2}},\ \nu_2 + \theta_2 + v/n^{\frac{1}{2}})$. The probability of an observation on $F(x - \theta_1, y - \theta_2)$ falling in the (i, j)th open cell of this partition if P_{ij}^n, where P_{ij}^n is independent of (θ_1, θ_2) and $P_{ij}^n \to P_{ij}$ as $n \to \infty$.

Consider first the conditional distribution

$$\mathfrak{L}\{\{n^{\frac{1}{2}}(N_{ij}/n - P_{ij}\}^* \mid n^{\frac{1}{2}}(\xi_1 - \nu_1 - \theta_1) = u,\ n^{\frac{1}{2}}(\xi_2 - \nu_2 - \theta_2) = v\} \tag{3.3}$$

where $\{B_{ij}\}^*$ denotes the set of all quantities B_{ij} except those having the omitted indices $(0, M + 1)$, $(L + 1, 0)$ and $(L + 1, M + 1)$. Under the conditions of (3.3) define $\tau_1 = 0$ if the x-component of the observation on the line $y = \nu_2 + \theta_2 + v/n^{\frac{1}{2}}$ is $\geqq \nu_1 + \theta_1 + u/n^{\frac{1}{2}}$, and $\tau_1 = 1$ otherwise. Similarly, set $\tau_2 = 0$ if the y-component of the observation on $x = \nu_1 + \theta_1 + u/n^{\frac{1}{2}}$ is $\geqq \nu_2 + \theta_2 + v/n^{\frac{1}{2}}$, and $\tau_2 = 1$ otherwise.

There remain $n - 2$ observations falling into the open cells ($n - 1$ if ξ_1 and ξ_2 come from the same observation). The key to the proof is the observation that when (θ_1, θ_2) is true,

$$\mathfrak{L}\{\{N_{ij}\}^* \mid \xi_1 = \nu_1 + \theta_1 + u/n^{\frac{1}{2}},\ \xi_2 = \nu_2 + \theta_2 + v/n^{\frac{1}{2}},\ \tau_1, \tau_2\}$$

is the same as

$$\mathfrak{L}\{\{n_{ij}\}^* \mid \textstyle\sum_{i=0}^{K_1} \sum_{j=0}^{M+1} n_{ij} = [n\alpha_1] - \tau_1,\quad \sum_{i=0}^{L+1} \sum_{j=0}^{K_2} n_{ij} = [n\alpha_2] - \tau_2\}, \tag{3.4}$$

where $\{n_{ij}\}$ are multinomial with parameters $n - 2$ (or $n - 1$) and P_{ij}^n. This plausible fact is easily verified by calculations on the model of Siddiqui [8], who displays the joint density of $\xi_1, \xi_2, \tau_1, \tau_2$ and the N_{ij} for the case $L = M = 0$ (4 cells). Setting $Q_{ij} = n^{\frac{1}{2}}(n_{ij}/n - P_{ij}^n)$, Taylor's theorem shows that the condition of (3.4) is

$$R: \textstyle\sum_1 Q_{ij} = a_n, \qquad \sum_2 Q_{ij} = b_n$$

where Σ_1 and Σ_2 denote the summations appearing in (3.4) and $a_n = -f_X(\nu_1)u + o(1)$, $b_n = -f_Y(\nu_2)v + o(1)$.

If $Q'_{ij} = n^{\frac{1}{2}}(n_{ij}/n - P_{ij})$, then by Taylor's theorem

$$Q'_{ij} = Q_{ij} + (\Delta_{ij}F_1)u + (\Delta_{ij}F_2)v + o(1).$$

If $\{Z_{ij}\}$ are a set of random variables having the asymptotic distribution $N(0, \Sigma)$ of the Q_{ij} (omitting $Q_{L+1,M+1}$), then the limit of $\mathcal{L}\{\{Q'_{ij}\}^* \mid R\}$ is

$$\mathcal{L}\{\{Z_{ij} + \mu_{ij}\}^* \mid \textstyle\sum_1 Z_{ij} = -f_X(\nu_1)u, \ \sum_2 Z_{ij} = -f_Y(\nu_2)v\} \tag{3.5}$$

where $\mu_{ij} = (\Delta_{ij}F_1)\, u + (\Delta_{ij}F_2)v$. This is true because inspection of the expansion of $p_i(x)$ given by the local limit theorem shows that the proof of the Lemma of Section 2 is unchanged if the n_i have parameters P_i^n with $P_i^n \to P_i$. Since this holds for all values of τ_1 and τ_2, (3.5) is the limit of the distribution (3.3).

Let $c_1 = -1/f_X(\nu_1)$ and $c_2 = -1/f_Y(\nu_2)$. Then it is easy to show by the method of Weiss [11] that when (θ_1, θ_2) is true,

$$\mathcal{L}\{n^{\frac{1}{2}}(\xi_1 - \nu_1 - \theta_1), n^{\frac{1}{2}}(\xi_2 - \nu_2 - \theta_2)\} \to \mathcal{L}\{c_1 \textstyle\sum_1 Z_{ij}, c_2 \sum_2 Z_{ij}\}.$$

Combining these results, we have that the asymptotic joint distribution of

$$n^{\frac{1}{2}}(\xi_1 - \nu_1 - \theta_1), \qquad n^{\frac{1}{2}}(\xi_2 - \nu_2 - \theta_2), \qquad \{n^{\frac{1}{2}}(N_{ij}/n - P_{ij})\}^*$$

when (θ_1, θ_2) is true is just the joint distribution of

$$c_1 \textstyle\sum_1 Z_{ij}, \qquad c_2 \sum_2 Z_{ij}, \qquad \{Z_{ij} + c_1(\Delta_{ij}F_1) \sum_1 Z_{ks} + c_2(\Delta_{ij}F_2) \sum_2 Z_{ks}\}^*.$$

If $n(\cdot)$ is the density of the normal distribution with means μ_{ij} and covariance Σ, these last random variables can be computed to have density h given by

$$h(u, v, \{q_{ij}\}^*) = n(\{q_{ij}\}^*, -\textstyle\sum_{i=0}^{K_1} \sum_{j=0}^{M} q_{ij}, -\sum_{i=0}^{L} \sum_{j=0}^{K_2} q_{ij}),$$

where the last two entries are in the $(0, M + 1)$ and $(L + 1, 0)$ places, respectively.

Since the inverse of Σ is well known, it is a matter of routine arithmetic to establish that $h = Ke^{-\frac{1}{2}S}$, where

$$S = I_{11}^* u^2 + I_{22}^* v^2 + 2I_{12}^* uv + 2 \textstyle\sum^* c_{1ij}q_{ij}u + 2 \sum^* c_{2ij}q_{ij}v$$
$$+ \text{ terms not containing } u \text{ or } v.$$

Standard least squares theory now shows that $(\theta_1^{**}, \theta_2^{**})$ are the least squares estimators from this distribution, and hence the minimum-variance linear unbiased estimators. That $(I^*)^{-1}$ is the asymptotic covariance matrix of these estimators also follows from least squares theory.

It remains to show that I^* approximates I for appropriate choice of the set of fixed distances used in defining the estimators. Let D_{ij} be the (i, j)th cell of the partition with vertices $\{(u_k, v_s)\}$. First notice that by Schwarz's inequality,

$$I_{11}^* = \textstyle\sum_{i,j} (\Delta_{ij}F_1)^2/P_{ij} = \sum_{i,j} [\int \int_{D_{ij}} f_1(x, y)\, dx\, dy]^2 / \int \int_{D_{ij}} f(x, y)\, dx\, dy$$
$$\leqq \textstyle\sum_{i,j} \int \int_{D_{ij}} [f_1(x, y)]^2/f(x, y)\, dx\, dy = I_{11}.$$

Let D denote a compact set which is a union of closures of bounded cells D_{ij} in which $f(x, y)$ is bounded away from zero, and which is such that

$$I_{11} - \int \int_D [f_1(x, y)]^2/f(x, y)\, dx\, dy < \epsilon.$$

By choosing u_0, v_0, u_L and v_M appropriately and making the cells containing points at which $f = 0$ sufficiently small, such a D can be obtained for any given $\epsilon > 0$. Let $D' = \{(i, j) : D_{ij} \subset D\}$.

If A_{ij} is the area of D_{ij}, the mean value theorem for integrals implies that

$$\begin{aligned}\textstyle\sum_{D'} (\Delta_{ij}F_1)^2/P_{ij} &= \textstyle\sum_{D'} (\int\int_{D_{ij}} f_1(x, y)\, dx\, dy]^2/\int\int_{D_{ij}} f(x, y)\, dx\, dy \\ &= \textstyle\sum_{D'} [f_1(x^*, y^*)A_{ij}]^2/[f(x^{**}, y^{**})A_{ij}]\end{aligned}$$

for some (x^*, y^*) and (x^{**}, y^{**}) in D_{ij}. By Taylor's theorem this is

$$\textstyle\sum_{D'} [f_1(u_{i-1}, v_{j-1})A_{ij} + R_{ij}^*]^2/[f(u_{i-1}, v_{j-1})A_{ij} + R_{ij}^{**}]$$

where $R_{ij}^* = [f_1(x^*, y^*) - f_1(u_{i-1}, v_{j-1})]A_{ij} = o(A_{ij})$ as $\max_i |u_i - u_{i-1}| \to 0$ and $\max_j |v_j - v_{j-1}| \to 0$, since $f_1(x, y)$ is continuous. Similarly, $R_{ij}^{**} = o(A_{ij})$, and by compactness of D both of these are uniform in $(i, j) \varepsilon D'$. It is now easy to show that

$$\textstyle\sum_{D'} (\Delta_{ij}F_1)^2/P_{ij} = \sum_{D'} A_{ij}[f_1(u_{i-j}, v_{j-i})]^2/f(u_{i-1}, v_{j-1}) + o(1)$$

as $\max_{D'} (|u_i - u_{i-1}|, |v_j - v_{j-1}|) \to 0$. The sum on the right is a Riemann sum for

$$\int\int_D [f_1(x, y)]^2/f(x, y)\, dx\, dy.$$

This establishes that I_{11} can be approached as closely as desired by first choosing D, then refining the partition.

The proof for I_{22}^* is similar. For I_{12}^* use Schwarz's inequality as follows (where C is the set of $(i, j) \not\varepsilon D'$):

$$\left(\textstyle\sum_C \Delta_{ij}F_1 \cdot \Delta_{ij}F_2/P_{ij}\right)^2 \leqq [\textstyle\sum_C (\Delta_{ij}F_1)^2/P_{ij}] \cdot [\sum_C (\Delta_{ij}F_2)^2/P_{ij}].$$

It is a consequence of the proofs for I_{11}^* and I_{22}^* that the right side may be made as small as desired by choosing an appropriate D. The sum over $(i, j) \varepsilon D'$ is then treated as in the other cases. This completes the proof of Theorem 2.

4. Use of RBAN estimators. Our third ANE method of estimation for multivariate location parameters is based on Neyman's [5] theory of regular best asymptotically normal (RBAN) estimators for multinomial problems. Suppose $n_1, \cdots, n_m$ are multinomial random variables with parameters n and $\{P_i\}$. If $P_i = \pi_i(\theta)$, where π_i is a known function and $\theta = (\theta_1, \cdots, \theta_r)$ is an unknown parameter, we can estimate θ from the n_i. If $q_i = n_i/n$ are the observed cell frequencies, a function $\varphi_j(q_1, \cdots, q_m)$ is a RBAN estimator for θ_j if φ_j is continuously differentiable with respect to each q_i and $n^{\frac{1}{2}}(\varphi_i - \theta_j)$ is asymptotically normal with mean zero and variance equal to the Cramér-Rao lower bound.

Neyman presents three methods of obtaining such estimators, including the maximum likelihood method, but we will be concerned only with certain analytic properties of the functions φ_j. Let G be the $r \times r$ matrix with entries

$$G_{ks} = \sum_{i=1}^{m} \frac{1}{\pi_i} \frac{\partial \pi_i}{\partial \theta_k} \cdot \frac{\partial \pi_i}{\partial \theta_s}.$$

Neyman shows that any function φ_j such that $\varphi_j(q_1, \cdots, q_m)$ is RBAN must satisfy

$$\varphi_j(\pi_1(\theta), \cdots, \pi_m(\theta)) \equiv \theta_j \tag{4.1}$$

and

$$\frac{\partial \varphi_j}{\partial q_k}\bigg|_{\{q_i\}=\{\pi_i\}} = \sum_{s=1}^{r} G^{js} \frac{1}{\pi_k} \frac{\partial \pi_k}{\partial \theta_s} \tag{4.2}$$

where $|G_j^s| = G^{-1}$.

Suppose $F(x_1 - \theta_1, \cdots, x_r - \theta_r)$ is a continuous r-variate location parameter family. For $0 < \alpha_{i1} < \cdots < \alpha_{iM_i} < 1$ and $i = 1, \cdots, r$, denote by ξ_{ij} the marginal sample α_{ij}–quantile in the x_i–direction. These quantiles partition R_r into cells which we index by $\sigma = (i_1, \cdots, i_r)$. Let N_σ be the number of observations falling in the σth open cell, and $q_\sigma = N_\sigma/n$. For given observed values z_{ij} of the ξ_{ij} the partition is non-random. Let $\pi_\sigma(\{z_{ij} - \theta_i\})$ be the probability that an observation on $F(x_1 - \theta_1, \cdots, x_r - \theta_r)$ falls in the σth cell of this partition. For example, if $r = 2$, π_{ks} is defined by

$$\pi_{ks}(\{z_{ij} - \theta_i\}) = F(z_{1k} - \theta_1, z_{2s} - \theta_2) - F(z_{1k} - \theta_1, z_{2,s-1} - \theta_2) - F(z_{1,k-1} - \theta_1, z_{2s} - \theta_2) + F(z_{1,k-1} - \theta_1, z_{2,s-1} - \theta_2)$$

for $k = 1, \cdots, M_1 + 1$ and $s = 1, \cdots, M_2 + 1$ with conventions that $z_{i0} = -\infty$ and $z_{iM_i+1} = \infty$.

The proposed methof of estimation is as follows: for the observed values z_{ij} of the ξ_{ij} the functions $\pi_\sigma(\{z_{ij} - \theta_i\})$ may be taken as cell probabilities for a multinomial problem to which Neyman's theory applies. Let $\varphi_k(\{z_{ij}\}, \{\cdot\})$ be a RBAN estimator for θ_k in this multinomial problem. We will estimate the location parameter θ_k by $\varphi_k(\{\xi_{ij}\}, \{q_\sigma\})$, where q_σ are the observed cell frequencies from the random partition formed by the ξ_{ij}.

Let $P_\sigma^n = \pi_\sigma(\{\xi_{ij} - \theta_i\})$ where the ξ_{ij} arise from n independent observations on F. Clearly

$$P_\sigma^n \to P_\sigma = \pi_\sigma(\{\nu_{ij}\})$$

$$\frac{\partial P_\sigma^n}{\partial \theta_k} \to a_{k\sigma} \text{ (say)}$$

in probability, where ν_{ij} are the $\theta = 0$ population quantiles. These limits are independent of θ. Define $r \times r$ matrices G_n and G_* by

$$(G_n)_{ks} = \sum_\sigma \frac{1}{P_\sigma^n} \frac{\partial P_\sigma^n}{\partial \theta_k} \cdot \frac{\partial P_\sigma^n}{\partial \theta_s}$$

$$(G_*)_{ks} = \sum_\sigma \frac{a_{k\sigma} a_{s\sigma}}{P_\sigma}$$

so that G_* is the limit in probability of G_n.

We have supposed that the RBAN estimators for the multinomial scheme $\pi_\sigma(\{z_{ij} - \theta_i\})$ could be written as fixed functions φ_k of the z_{ij} and the q_σ. This is not unreasonable, since RBAN estimators are typically defined implicitly as solutions of systems of equations in the π_σ and q_σ. Existence of solutions φ_k having the properties required by theorem 3 will typically be guaranteed by the implicit function theorem. For the three RBAN methods discussed by Neyman, this is the case if G_n is non-singular. This in turn is shown by him to be a consequence of an assumption that the parameters $\theta_1, \cdots, \theta_r$ are not dependent for the multinomial problem. Thus our assumptions on the functional form of φ_k are essentially the requirement that we use the same RBAN method of estimation for each of the sequence of multinomial schemes presented by the observed values of the ξ_{ij} as $n \to \infty$.

THEOREM 3. *Let the density $f(x_1 - \theta_1, \cdots, x_r - \theta_r)$ be continuous. Suppose there exist functions $\varphi_k(\{z_{ij}\}, \{q_\sigma\})$, continuous in each z_{ij} and continuously differentiable in each q_σ, such that $\varphi_k(\{z_{ij}\}, \{\cdot\})$ is a* RBAN *estimator for θ_k from the multinomial problem with cell probabilities $\pi_\sigma(\{z_{ij} - \theta_i\})$. Let $\varphi = (\varphi_1(\{\xi_{ij}\}, \{q_\sigma\}), \cdots, \varphi_r(\{\xi_{ij}\}, \{q_\sigma\}))$. If G_n is non-singular for n sufficiently large, then when θ is true*

$$\mathfrak{L}\{n^{\frac{1}{2}}(\varphi - \theta)\} \to N(0, G_*^{-1}).$$

If the partial derivatives $f_1, \cdots, f_r$ of f are continuous and the information integrals I_{ks} exist, each $(G_*)_{ks}$ can be made as close as desired to I_{ks} by choosing a sufficient number of sufficiently closely spaced quantiles in each direction.

PROOF. Abbreviate $\pi_\sigma(\{z_{ij} - \theta_i\})$ by π_σ and let τ_σ be the $r \times 1$ vector with kth component $(\pi_\sigma)^{-1}\partial\pi_\sigma/\partial\theta_k$. Then from (4.1) and (4.2) it follows that for any fixed $\{z_{ij}\}$ the Taylor's series about $\{q_\sigma\} = \{\pi_\sigma\}$ is

$$\varphi(\{z_{ij}\}, \{q_\sigma\}) = \theta + \textstyle\sum_\sigma (G^{-1}\tau_\sigma)(q_\sigma - \pi_\sigma) + R(\{q_\sigma\}, \{\pi_\sigma\})$$

where the kth component of R is

$$\sum_\sigma \left(\frac{\partial \varphi_k}{\partial q_\sigma}\bigg|_{q_\sigma^*} - \frac{\partial \varphi_k}{\partial q_\sigma}\bigg|_{\pi_\sigma} \right)(q_\sigma - \pi_\sigma)$$

for some q_σ^* between q_σ and π_σ. If therefore q_σ are the observed cell frequencies from the random partition and $\rho_{n\sigma}$ is the $r \times 1$ vector with kth component $(P_\sigma^n)^{-1}\, \partial P_\sigma^n/\partial\theta_k$, we have that

$$n^{\frac{1}{2}}(\varphi - \theta) = \textstyle\sum (G_n^{-1}\rho_{n\sigma})n^{\frac{1}{2}}(q_\sigma - P_\sigma^n) + n^{\frac{1}{2}}R(\{q_\sigma\}, \{P_\sigma^n\}).$$

Now $\{n^{\frac{1}{2}}(q_\sigma - P_\sigma^n)\}$ is asymptotically $N(0, \Sigma)$, where

$$\textstyle\sum_{\sigma\sigma} = P_\sigma(1 - P_\sigma), \qquad \sum_{\sigma\tau} = -P_\sigma P_\tau, \qquad \sigma \neq \tau.$$

This important result is proved by arguments which differ only in detail from those used in the proof of Theorem 2. From this result and the continuity properties of φ_k it is immediate that $n^{\frac{1}{2}}R(\{q_\sigma\}, \{P_\sigma^n\}) \to 0$ in probability. Thus $n^{\frac{1}{2}}(\varphi - \theta)$ is asymptotically distributed as

$$\textstyle\sum_\sigma (G_*^{-1}\rho_\sigma)Z_\sigma, \tag{4.3}$$

where $\{Z_\sigma\}$ are $N(0, \Sigma)$ and ρ_σ has kth component $a_{k\sigma}/P_\sigma$. That the r random variables (4.3) have distribution $N(0, G_*^{-1})$ follows after some calculation.

The final statement of the theorem follows as in Theorem 2, after writing the $a_{k\sigma}$ as difference operators acting on the derivative F_k.

As was remarked above, RBAN estimators are usually implicitly defined by systems of equations which can only rarely be explicitly solved. There is a large literature on methods for obtaining RBAN estimators (iteratively or otherwise) for multinomial problems. See Ferguson [3] for reference to some such methods. Since our φ_k are obtained from RBAN estimators for certain multinomial schemes, the problem of computing them should be approached by reference to this literature.

5. Example: The bivariate logistic distribution. In this section we apply the estimators θ^* and θ^{**} to the problem of estimating the parameters of a bivariate logistic distribution. This is a bivariate location parameter family with

$$F(x, y) = [1 + e^{-x} + e^{-y}]^{-1} \qquad -\infty < x, y < \infty$$

described in detail by Gumbel [4]. The marginal distributions of F are logistic, and we have the relation $F(x, y) = F(y, x)$ which simplifies many of the calculations below.

We compare the performance of our estimators with that of the vector $\bar{X}$ of sample means and the vector M of sample medians. The asymptotic distributions of $n^{\frac{1}{2}}(\bar{X} - \theta)$ and $n^{\frac{1}{2}}(M - \theta)$ are bivariate normal with zero means and covariances given by Theorem 3.1 of [1]. For the bivariate logistic case, $n^{\frac{1}{2}}(\bar{X} - \theta)$ has asymptotic covariance matrix σ with entries

$$\sigma_{11} = \sigma_{22} = \pi^2/3, \qquad \sigma_{12} = \sigma_{21} = \pi^2/6$$

and $n^{\frac{1}{2}}(M - \theta)$ has asymptotic covariance matrix τ with entries

$$\tau_{11} = \tau_{22} = 4, \qquad \tau_{12} = \tau_{21} = \tfrac{4}{3}.$$

One could also estimate the bivariate location parameter θ by using Ogawa's univariate estimators to estimate each component of θ from the corresponding components of the observations. Call the resulting estimator $\hat{\theta}$. We expect θ^* to attain greater asymptotic efficiency than $\hat{\theta}$ using the same set of quantiles, since θ^* utilizes additional information from the cell frequencies.

We will measure the asymptotic efficiency of an estimator T such that

$$\mathcal{L}\{n^{\frac{1}{2}}(T - \theta)\} \to N(0, \beta)$$

by

$$e(T) = (\|I^{-1}\|/\|\beta\|)^{\frac{1}{2}}$$

where I is the information matrix and $\|\alpha\|$ is the determinant of the matrix α. This means that the relative asymptotic efficiency of two estimators is the inverse ratio of the sample sizes required to reach equal "generalized variance" (see section 4 of [1]). The information matrix for the bivariate logistic family is

$$I_{11} = I_{22} = \tfrac{1}{2}, \qquad I_{12} = I_{21} = -\tfrac{1}{4}$$

with $\|I^{-1}\| = \frac{16}{3}$. We have therefore that $e(\bar{X}) = .810$ and $e(M) = .612$.

Our three types of ANE estimators in the bivariate case involve random partitions of the plane. In θ^* and φ the partition is formed by a set of sample quantiles in each direction, while in θ^{**} it arises from a single quantile in each direction with fixed distances marked off from it. These random partitions converge in probability to a fixed limiting partition as the sample size $n \to \infty$. (These limiting partitions were used in the proofs of Theorems 1, 2 and 3.) Inspection of the asymptotic covariance matrices of θ^*, θ^{**} and φ yields the important observation that *when the limiting partitions are the same, all three estimators have the same asymptotic distribution.* For a given set of sample quantiles (or their equivalent fixed distances in θ^{**}) the estimators have identical asymptotic efficiency, and the choice among them may be made on the basis of computational convenience.

Let us compute θ^{**} for the case of 16 cells obtained by marking off ± 1 from the sample median in each component. The limiting partition is formed by partitioning each axis at -1, 0, $+1$. The difference operators Δ_{ij} refer to the cells of this partition. Values of $F(x, y)$ are given in Table 1 of [4]. Computation of $P_{ij} = \Delta_{ij}F$ is simplified by the fact that $P_{ij} = P_{ji}$. From $F_1(x, y) = F_2(y, x)$ we have $\Delta_{ij}F_1 = \Delta_{ji}F_2$, again reducing the required computations. We obtain

$$I_{11}^* = I_{22}^* = 0.4028, \qquad I_{12}^* = I_{21}^* = -0.1674,$$

so $\|(I^*)^{-1}\| = 7.452$ and $e(\theta^{**}) = .846$. Thus θ^{**} for 16 cells is already more efficient than the sample mean. The estimator itself is

$$\begin{aligned}\theta_1^{**} = \xi_1 &+ 3.06q_{00} + 4.71q_{10} - 0.44q_{20} + 2.30q_{01} \\ &+ 4.22q_{11} - 0.37q_{21} + 0.58q_{31} + 0.20q_{02} \\ &+ 2.15q_{12} - 2.08q_{22} - 0.62q_{32} + 1.95q_{13} \\ &- 1.97q_{23} - 1.04\end{aligned}$$

where $q_{ij} = N_{ij}/n$ and ξ_1 is the sample median from the x-component. Similarly

$$\begin{aligned}\theta_2^{**} = \xi_2 &+ 3.06q_{00} + 2.30q_{10} + 0.20q_{20} + 4.71q_{01} \\ &+ 4.22q_{11} + 2.15q_{21} + 1.95q_{31} - 0.44q_{02} \\ &- 0.37q_{12} - 2.08q_{22} - 1.97q_{32} + 0.58q_{13} \\ &- 0.62q_{23} - 1.04.\end{aligned}$$

Estimators θ^* with the same performance are obtained by choosing the $(1 + e)^{-1}$, $\frac{1}{2}$ and $(1 + e^{-1})^{-1}$ sample quantiles in each direction (so the $\theta = 0$ population quantiles are -1, 0, $+1$). The estimators θ^* are much more laborious to compute, even though in this example we have $a_j = b_j$, $c_j = d_j$ and $\kappa_{ij} = \delta_{ji}$.

The results are (in the notation of Theorem 1)

$$\theta_1^* = 0.36\xi_1 + 0.36\xi_2 + 0.28\xi_3 - 0.06\zeta_1 + 0.02\zeta_2 + 0.04\zeta_3$$
$$+ 3.06q_{11} + 2.76q_{21} + 1.54q_{31} + 1.72q_{12} + 1.69q_{22}$$
$$+ 1.02q_{32} + 0.83q_{13} + 0.82q_{23} + 0.52q_{33} - 1.07,$$
$$\theta_2^* = -0.06\xi_1 + 0.01\xi_2 - 0.04\xi_3 + 0.36\zeta_1 + 0.36\zeta_2 + 0.28\zeta_3$$
$$+ 3.06q_{11} + 1.72q_{21} + 0.83q_{31} + 2.76q_{12} + 1.69q_{22}$$
$$+ 0.82q_{32} + 1.54q_{13} + 1.02q_{23} + 0.52q_{33} - 1.07.$$

The estimators φ are defined implicitly and are therefore not useful in this example, where θ^* and θ^{**} can be obtained routinely. In cases where $F(x, y)$ and its derivatives cannot be obtained in closed form, numerical solution of the equations defining φ offers an alternative to computation of $\Delta_{ij}F$ by numerical integration of the density.

Increasing the number of cells used increases the efficiency of these estimators in accordance with our Theorems. In this example, estimators with a limiting partition of 100 cells divided at the integers from -4 to 4 in each direction have asymptotic efficiency $e = .944$.

Ogawa's estimators, calculated using the same sample quantiles as in θ_1^*, are

$$\hat{\theta}_1 = 0.315\xi_1 + 0.370\xi_2 + 0.315\xi_3$$
$$\hat{\theta}_2 = 0.315\zeta_1 + 0.370\zeta_2 + 0.315\zeta_3$$

by (5.3) of [6] ($K_3 = 0$ by symmetry). The asymptotic variance of each component of $n^{\frac{1}{2}}(\hat{\theta} - \theta)$ is 3.205, and the asymptotic covariance is 1.388, so that $e(\hat{\theta}) = .799$. This compares with $e = .846$ for any of θ^*, θ^{**}, φ. The result illustrates the additional information contained in the cell frequencies.

Acknowledgment. I am happy to acknowledge the advice and encouragement of Professor J. Kiefer in the preparation of this paper.

REFERENCES

[1] Bickel, Peter J. (1964). On some alternative estimates for shift in the p-variate one-sample problem. *Ann. Math. Statist.* **35** 1079–1090.

[2] Feller, W. (1957). *An Introduction to Probability Theory and Its Applications* **1** 2nd Ed. Wiley, New York.

[3] Ferguson, T. A. (1958). A method of generating best asymptotically normal estimates with application to the estimation of bacterial densities. *Ann. Math. Statist.* **29** 1046–1062.

[4] Gumbel, E. J. (1961). Bivariate logistic distributions. *J. Amer. Statist. Assoc.* **56** 335–349.

[5] Neyman, J. (1949). Contributions to the theory of the χ^2 test. *Proc. Berkeley Symp. Math. Statist. Prob.* 239–273. Univ. of California Press.

[6] OGAWA, J. (1951). Contributions to the theory of systematic statistics I. *Osaka Math. J.* **3** 175–213.

[7] RICHTER, WOLFGANG (1957). Local limit theorems for large deviations. *Theor. Probability Appl.* **2** 206–219.

[8] SIDDIQUI, M. M. (1960). Distribution of quantiles in samples from a bivariate population. *J. Res. Nat. Bur. Standards* **64B** 145–150.

[9] STECK, GEORGE P. (1957). Limit theorems for conditional distributions. *Univ. California Publ. Statist.* **2** 237–284.

[10] TISCHENDORF, JOHN A. (1955). Linear estimation techniques using order statistics. Unpublished dissertation, Purdue Univ.

[11] WEISS, LIONEL (1964). On the asymptotic joint normality of quantiles from a multivariate distribution. *J. Res. Nat. Bur. Standards* **68B** 65–66.