The Annals of Statistics
1991, Vol. 19, No. 1, 314–328

SIMULTANEOUS ESTIMATION IN DISCRETE MULTIVARIATE EXPONENTIAL FAMILIES

BY JINE-PHONE CHOU

Academia Sinica

Let X have a discrete density of the form $f(x) = t(x)\xi(\theta)\theta_1^{x_1} \cdots \theta_p^{x_p}$, where $t(x)$ is nonzero on some infinite subset of Z^p. Consider simultaneous estimation of the θ_i under the loss $\mathscr{L}_m(\theta, \delta) = \sum_{i=1}^p \theta_i^{-m}(\theta_i - \delta_i)^2$, $m \geq 0$. For integers $m \geq 1$, estimators are found which improve on the maximum likelihood estimator or uniformly minimum variance unbiased estimator. The improved estimators are distinguished by the property that they do not depend on m for "large values" of the observed vector. On the other hand, we prove admissibility of a class of estimators, including the MLE and UMVUE, for some discrete densities of the indicated form under squared error loss ($m = 0$).

1. Introduction. Let $X = (X_1, \dots, X_p)$ be a random vector with a discrete exponential density of the form

$$f(x) = t(x)\xi(\theta)\theta_1^{x_1} \cdots \theta_p^{x_p}. \tag{1.1}$$

Here $t(x)$ has a support E which is an infinite subset of $Z^p \equiv \{\dots, -2, -1, 0, 1, 2, \dots\}^p$. We consider the problem of estimating $\theta = (\theta_1, \dots, \theta_p)$ with loss function

$$\mathscr{L}_m(\theta, \delta) = \sum_{i=1}^{p} \theta_i^{-m}(\theta_i - \delta_i)^2, \tag{1.2}$$

where m is a nonnegative integer and $\delta = (\delta_1, \dots, \delta_p)$ is an estimator.

Since Stein (1956) first showed that the uniformly minimum variance unbiased estimator of a p-dimensional normal mean is inadmissible for $p \geq 3$, a number of authors have considered improved estimators for parameters of continuous and discrete exponential families. The work on discrete exponential families has concentrated on distributions with independent marginals and the improved estimators depend critically on the loss functions. Recent results in this field may be found in Hwang (1982a), Ghosh, Hwang and Tsui (1983) and Tsui (1984).

In this paper we prove inadmissibility results under (1.2) with integers $m \geq 1$ and some admissibility results when $m = 0$. Our inadmissibility results apply to densities with dependent marginals as well as independent marginals,

Received 1984; revised January 1990.
AMS 1980 *subject classifications*. 62C15, 62F10, 62H99.
Key words and phrases. Discrete multivariate exponential families, admissibility, uniformly minimum variance unbiased estimators, maximum likelihood estimators, negative multinomial distribution, quota fulfillment.

and the improved estimators are independent of m for large x. Here are examples of the settings we have in mind.

EXAMPLE 1.1. $X = (X_1, \dots, X_p)$ has independent Poisson marginals: $X_i \sim P(\lambda_i)$.

EXAMPLE 1.2. $X = (X_1, \dots, X_p)$ has independent negative binomial marginals: $X_i \sim \mathrm{NB}(r, \theta_i)$.

EXAMPLE 1.3. $X = (X_1, \dots, X_p)$, $p \geq 2$, has the negative multinomial distribution with density

$$(1.3) \quad f(x) = \frac{(x_1 + \cdots + x_p + r - 1)!}{(x_1)! \cdots (x_p)!(r-1)!} \theta_1^{x_1} \cdots \theta_p^{x_p} (1 - \theta_1 - \cdots - \theta_p)^r,$$

where $x_i = 0, 1, 2, \dots, r \geq 1$, $0 < \theta_i$ and $\sum_{i=1}^p \theta_i < 1$. This distribution arises frequently in capture–recapture models relating to several types of objects, if in the recapture stage one samples until r objects of a particular type are captured. For capture–recapture models, see, for instance, Seber (1962, 1982) and Smith (1988). Some applications and properties of the negative multinomial are discussed in Sibuya, Yoshimura and Shimizu (1964).

EXAMPLE 1.4 (Quota fulfillment distribution). Here is a coupon collector's generalization of the negative multinomial. For this let there be objects of type $1, \dots,$ type p and suppose objects are sampled with replacement until at least $r_i \geq 1$ of type i have been observed, $i = 1, \dots, p$. X_i denotes the number of type i counted when sampling ceases. With $\theta = (\theta_1, \dots, \theta_p)$, θ_i the proportion of type i, the density of $X = (X_1, \dots, X_p)$ is given by

$$(1.4) \qquad f(x) = \big(t_1(x) + \cdots + t_p(x)\big) \theta_1^{x_1} \cdots \theta_p^{x_p},$$

where $0 < \theta_i < 1$, $\sum_{i=1}^p \theta_i = 1$ and, with $A_j = \{x \colon x_j = r_j,\ x_i \geq r_i,\ i = 1, \dots, p\}$ for $r_i \geq 1$,

$$t_j(x) = \frac{(x_1 + \cdots + x_p - 1)!}{(x_1)! \cdots (x_{j-1})!(x_j - 1)!(x_{j+1})! \cdots (x_p)!} 1_{A_j}(x).$$

Now, we extend the previous distribution further by allowing for partial fulfillment of quotas, and no quota restriction on some types. Conveniently, define $r_i = 0$ if there is no quota restriction on type i. Let X_i denote the number of type i observed when precisely a of the inequalities $X_i \geq r_i \geq 1$ are satisfied, $1 \leq a \leq \#\{i,\ r_i \neq 0,\ i = 1, \dots, p\}$. The distribution of $X = (X_1, \dots, X_p)$ is denoted by $Q(a, r, \theta)$, where $r = (r_1, \dots, r_p)$, r_i a nonnegative integer. Some properties of this distribution are discussed in Young (1961) and Anderson, Sobel and Uppuluri (1982).

EXAMPLE 1.5 (Gas station distribution). Let N be a nonnegative integer-valued random variable and suppose that given $N = n$, $X = (X_1, \dots, X_p)$ is multinomial with parameters $n, \theta_1, \ldots, \theta_p$. Accordingly the density is

$$(1.5) \qquad f(x) = g(x_1 + \cdots + x_p) \frac{(x_1 + \cdots + x_p)!}{(x_1)! \cdots (x_p)!} \theta_1^{x_1} \cdots \theta_p^{x_p},$$

where $x_i = 0, 1, \ldots, 0 < \theta_i$, $\Sigma_{i=1}^p \theta_i = 1$ and $\Sigma_{n=0}^\infty g(n) = 1$. This distribution arises naturally in many situations including the following example: Suppose that there are p different types of gasoline, and the total number of sales over a fixed period of time is random with p.d.f. g. Let X_i denote the number of purchases of gasoline to type i when time expires, and θ_i denote the true proportion of customers who use gasoline of type i. The density of $X = (X_1, \ldots, X_p)$ is given by (1.5).

We begin in Section 2 with those distributions having support $E = \{x \geq x^0\}$ for a point x^0 in Z^p (here $\geq$ is the usual partial ordering in $\mathbb{R}^p$). In such a case, it is easy to derive an unbiased estimator $D(x)$ of the difference between two risk functions $R(\theta, \delta + g)$ and $R(\theta, \delta)$ under $\mathscr{L}_m$, $m \geq 1$, where $\delta + g$ and δ are arbitrary estimators of θ with finite risk, i.e., $ED(X) = R(\theta, \delta + g) - R(\theta, \delta)$. By choosing g such that $D(x) \leq 0$ for all x and with strict inequality for some x, an improved estimator can be constructed. This technique traces back to Stein (1973). It is used in Section 2 to construct estimators which improve upon the UMVUE and, more generally, upon estimators which are larger than a "semi tail upper bound" [see Hwang (1982b)]. Applications are made to the distributions in Examples 1.1, 1.2 and 1.3. For these distributions, the improved estimators are of Clevenson–Zidek type (1975). For independent negative binomials under $\mathscr{L}_m$, $m = 2, 3, \ldots$, the improved estimators have simpler form than those previously found, for example, by Ghosh, Hwang and Tsui (1983). For the negative multinomial distribution with density (1.3), all estimators $\delta(x)$ with ith component $x_i/(x_1 + \cdots + x_p + r + a)$, a class which contains the UMVUE and MLE, can be improved by a class of estimators under $\mathscr{L}_m$, $m = 1, 2, \ldots$, provided p is larger than a specific bound depending on a (see Example 2.5).

There is also a discussion pertinent to the quota fulfillment distribution in Section 2. Here the difference inequality technique is not directly applicable. However, by decomposing the density suitably, one may use the previous result to produce a class of estimators which improve on the MLE of θ under $\mathscr{L}_m$, $m \geq 1$, provided p is large enough.

Section 3 contains admissibility results under $\mathscr{L}_0$. For the negative multinomial distribution and for any p, we demonstrate the admissibility of a class of estimators including the UMVUE and the MLE. This contrasts sharply with the results of Hwang (1982a) who showed the inadmissibility of the UMVUE for $p \geq 3$, when the marginal distributions of the X_i with density (1.1) are independent. Admissibility results, relating to the gas station and quota fulfillment distributions, are also provided.

2. Construction.

2.1. *Improving upon usual estimators by solving difference inequalities.* Let e_i denote the vector with 1 as ith component and 0 for others. Now suppose X has a p-variate discrete density of the form (1.1) which satisfies the conditions:

(2.1) Support $E = \{x \geq x^0\}$ for some x^0 in Z^p.

(2.2) For $m = 1, 2, \dots$, there exists $\tilde{x}$ in E such that

$$t(x + (m-1)e_i) \geq t(x + (m-1)e_j)$$

whenever $x \geq \tilde{x}$ and $x_i \leq x_j$.

(2.3) For some a with $a + \sum_{i=1}^{p} \tilde{x}_i + m - 1 > 0$, b with $b + \min\{\tilde{x}_i, i = 1, \dots, p\} + m - 1 \geq 0$ and a nonnegative number α,

$$\frac{t(x - e_i)}{t(x)} \geq \frac{x_i + b}{(x_1 + \cdots + x_p + a)^{\alpha}} > 0$$

whenever $x \geq \tilde{x}$ and $x_i \geq \tilde{x}_i + m$.

For example, if $X = (X_1, \dots, X_p)$ is a random vector with independent Poissons or negative binomial marginals, or has a negative multinomial distribution, the conditions are satisfied. Now we are going to construct estimators which improve upon the usual estimators of the parameter $\theta = (\theta_1, \dots, \theta_p)$ in (1.1) by solving a difference inequality. We use the convention $0/0 \equiv 0$, and for $x = (x_1, \dots, x_p)$ we define $s(x) = s = \sum_{i=1}^{p} x_i$ throughout.

The following lemma generalizes the identity of Hudson (1978) and Hwang (1982a) to the dependent case.

Lemma 2.1. *Let X have density* (1.1) *and satisfy condition* (2.1). *For a real-valued function g and an integer m, if $E|g(X)| < \infty$ and $g(x) = 0$ whenever $x_i < m + x_i^0$, then*

$$E\frac{1}{\theta_i^m} g(X) = E\frac{t(X + me_i)}{t(X)} g(X + me_i).$$

Proof. The proof follows by a change of variables. □

From Lemma 2.1 with $m = -1$ and $g(x) \equiv 1$, it follows that $\delta^0(X) = (\delta_i^0(X))$, $\delta_i^0(X) = t(X - e_i)/t(X)$, is an unbiased estimator of θ. Moreover, it is the UMVUE if the parameter space contains an open set in R^p. Let δ be an estimator of θ with $E|\delta|^2 < \infty$. For an integer m, assume that for all $i = 1, \dots, p$, $E\phi_i^2(X) < \infty$ and $\phi_i(x) = 0$ whenever $x_i < m + x_i^0$. Then under the loss function $\mathscr{L}_m$, from Lemma 2.1,

$$\begin{aligned} R(\theta, \delta + \phi) &- R(\theta, \delta) \\ &= ED(X) \\ &= E\sum_{i=1}^{p} \left\{ \frac{t(X + me_i)}{t(X)} \left[\left(\phi_i^2(X + me_i) + 2\delta_i(X + me_i)\phi_i(X + me_i) \right) \right] \right. \\ &\qquad \left. -2\frac{t(X + (m-1)e_i)}{t(X)} \phi_i(X + (m-1)e_i) \right\}. \end{aligned} \tag{2.4}$$

In particular, if $\delta(x) = \delta^0(x)$, the function $D(x)$ reduces to

$$\begin{aligned} \mathscr{D}(X) = \sum_{i=1}^{p} \left[\frac{t(X + me_i)}{t(x)} \phi_i^2(X + me_i) + 2\frac{t(X + (m-1)e_i)}{t(X)} \right. \\ \left. \times \left(\phi_i(X + me_i) - \phi_i(X + (m-1)e_i) \right) \right]. \end{aligned} \tag{2.5}$$

2.1.1. *Improving upon* δ^0. If we find $\phi(x)$ such that $\mathscr{D}(x) \le 0$ for all x, with strict inequality for some x in the support, then $\delta^0(X) + \phi(X)$ improves upon $\delta^0(X)$.

Suppose that there exist $\tilde{x}$, a, b, α and m such that conditions (2.2) and (2.3) hold. For these $\tilde{x}$, a, b, α and m, define $\phi(X) = (\phi_i(X))$ by

$$\phi_i(x) = \frac{-h(s)(x_i + b)}{(s(x) + c)^{\alpha + 1}} 1_{\Omega_i}(x), \qquad i = 1, \dots, p, \tag{2.6}$$

where $\Omega_i = (x\colon x \ge \tilde{x},\ x_i \ge \tilde{x}_i + m\}$, $c \ge a$,

$$\begin{aligned} \frac{2(c + m)}{m + b} - 2(\alpha + 1) > 0 \quad &\text{if } m + b > 0, \\ c + m \ge 0 \quad &\text{if } m + b \le 0, \end{aligned} \tag{2.7}$$

and $h(\cdot)$ is a nondecreasing function such that $h \not\equiv 0$,

$$\begin{aligned} &0 \le h(\cdot) \le \min\{2p - 2(\alpha + 1), \\ &\qquad\qquad 2(c + m)/(m + b) - 2(\alpha + 1)\} \quad \text{if } m + b > 0, \\ &0 \le h(\cdot) \le 2p - 2(\alpha + 1) \qquad\qquad\qquad\quad \text{if } m + b \le 0. \end{aligned} \tag{2.8}$$

For example, if $X = (X_1, \dots, X_p)$ is a random vector of independent Poissons, we may choose $\tilde{x} = 0$, $a = b = 0$, $\alpha = 0$, m a positive integer, and define $\phi_i(x) = [-h(s(x))x_i/(s(x) + c)]1_{(x_i \ge m)}(x)$, where $c > 0$, $h(\cdot)$ is nondecreasing, $h \not\equiv 0$ and $0 \le h(\cdot) \le \min\{2p - 2, 2c/m\}$.

THEOREM 2.1. *Assume that* X *has the density* (1.1) *and satisfies conditions* (2.1)–(2.3). *For estimating* θ *under the loss function* $\mathscr{L}_m$, $m = 1, 2, \dots$, $\delta^0(X)$ *is dominated by* $\delta^0(X) + \phi(X)$, *where* $\phi(X)$ *is given in* (2.6), *provided* $p > \alpha + 1$.

PROOF. See the Appendix. □

Note that the improved estimator does not depend on m if x is large.

EXAMPLE 2.1 (Independent Poissons). For estimating $\lambda = (\lambda_1, \dots, \lambda_p)$ under $\mathscr{L}_m$, $m = 1, 2, \dots$, the UMVUE $\delta^0(X) = X$ is dominated by $\delta^m(X) = \delta(X) + \phi(X)$, provided $p > \alpha + 1$. Here $\phi(x)$ has the form (2.6) with α being any nonnegative number, $\tilde{x}$ being any nonnegative point in Z^p, b satisfying $1 - m - \min\{\tilde{x}_i,\ i = 1, 2, \dots, p\} \le b$ for $\alpha > 0$ or $1 - m - \min\{\tilde{x}_i,\ i = 1, 2, \dots, p\} \le b \le 0$ for $\alpha = 0$, and a satisfying $a + s(x) + m - 1 > 0$ and $x_i \ge (x_i + b)/(s(x) + a)^\alpha$ if $x_i \ge m + \tilde{x}_i$. If we choose $\alpha = 0$, $\tilde{x} = 0$ and $b = 0$, then the improved estimators can be written $\delta^m(X) = (X_i) - (h_m(s)X_i(s + c)^{-1}1_{(X_i \ge m)}(X))$, where $c > 0$, h_m is nondecreasing, $h_m \not\equiv 0$ and $0 \le h_m \le \min\{2p - 2, 2c/m\}$. Note that for $\mathscr{L}_1$, the improved estimator proposed by Tsui and Press (1982) is $\delta^1(X)$. Moreover, for any $\mathscr{L}_m$, $m = 1, 2, \dots$, if we require $c \ge p - 1$, then these δ^m are Clevenson and Zidek (1975) estimators. Furthermore, they have some better results for this problem.

EXAMPLE 2.2 (Independent negative binomials). For estimating $\theta = (\theta_1, \dots, \theta_p)$ under $\mathscr{L}_m$, $m = 1, 2, \dots$, the UMVUE $\delta^0(X) = (X_i/(X_i + r - 1))$ is dominated by $\delta^0(X) + \phi(X)$, provided $p > \alpha + 1$, $\alpha > 0$. Here $\phi(x)$ is given in (2.6) with α being any positive number, and $\tilde{x}$, a, b satisfy the same regularities as that in Example 2.1 except $x_i \ge (x_i + b)/(s(x) + a)^\alpha$, is replaced by $x_i/(x_i + r - 1) \ge (x_i + b)/(s(x) + a)^\alpha$. In particular, if we choose $\alpha = 1$, $\tilde{x} = 0$ and $b = 0$ in (2.6), then $\phi_i(x)$ has the simple form, $-h(s)x_i(s + c)^{-2}1_{(x_i \ge m)}(x)$, where $c \ge r - 1$, $c > m$, $h(\cdot)$ is nondecreasing, $h \not\equiv 0$ and $0 \le h \le \min\{2p - 4, 2c/m - 2\}$. Different improved estimators have been constructed by many authors, for instance, Ghosh, Hwang and Tsui (1983).

EXAMPLE 2.3 (Negative multinomial distribution). For estimating θ under $\mathscr{L}_m$, $m = 1, 2, \dots$, the UMVUE $\delta^0(X) = (X_i/(s + r - 1))$ is dominated by $\delta^0(X) + \phi(X)$. Here $\phi(x)$ is given in (2.6), provided $p > \alpha + 1$, $\alpha > 0$. Note that α given in (2.6) is an arbitrary positive number and $\tilde{x}$ is any nonnegative point in Z^p. If we choose $\tilde{x} = 0$, $\alpha = 1$ and $b = 0$, then $\delta^0(X)$ is dominated by $\delta^0(X) - (h(s)X_i(s + c)^{-2}1_{(X_i \ge m)}(X))$, where $c \ge r - 1$, $c > m$ and $h(\cdot)$ is nondecreasing with $h \not\equiv 0$ and $0 \le h \le \min\{2p - 4, 2c/m - 2\}$.

2.1.2. *Improving other estimators.* Under the independence assumption, Hwang (1982b) defined a semi tail upper bound [an estimator is a STUB if every estimator $\delta(X)$ which is larger than it for sufficiently large X is inadmissible] and constructed STUBs for various cases. In Theorem 2.2 we extend these results by finding a class of STUBs without the independence assumption.

THEOREM 2.2. *Let X have density* (1.1) *and satisfy* (2.1)–(2.3). *For estimating θ under $\mathscr{L}_m$, $m = 1, 2, \dots$, the estimator $\delta^0(X)(1 - k/(s(X) + \beta))$ is a STUB, provided $k \geq 0$ and $p > \alpha + 1 + k$. In particular, every estimator $\delta(X) \geq \delta^0(X)(1 - k/(s(X) + \beta))$ for sufficiently large X is inadmissible and dominated by $\delta(X) + \phi(X)$, where ϕ is given in* (2.6) *with some specific c, h and Ω_i.*

PROOF. See the Appendix. □

REMARK 2.1. From Theorem 2.2 and its proof, it is obvious that under $\mathscr{L}_m$, $m = 1, 2, \dots$, every estimator $\delta(X) \geq \delta^0(X)$ for large X is dominated by $\delta(X) + \phi(X)$ where $\phi(X)$ is given in (2.6) with $\tilde{x}$ replaced by some large x, provided $p > \alpha + 1$. Some applications of Theorem 2.2 follow.

EXAMPLE 2.4 (Independent negative binomials). For estimating $\theta = (\theta_1, \dots, \theta_p)$ under $\mathscr{L}_m$, $m = 1, 2, \dots$, all estimators $\delta(X) = (X_i/(X_i + r + l))$ are inadmissible, if $p \geq 3$ and $l \leq -1$.

EXAMPLE 2.5 (Negative multinomial distribution). For estimating $\theta = (\theta_1, \dots, \theta_p)$ under $\mathscr{L}_m$, $m = 1, \dots$, the MLE $\delta(X) = (X_i/(s + r))$ is inadmissible if $p \geq 4$. Moreover, all estimators of the form $\delta^a(X) = (X_i/(s + r + a))$ are inadmissible under $\mathscr{L}_m$, $m = 1, 2, \dots$, provided $p \geq p(a)$, where $p(a) = 4$ if $a \leq 0$ and $p(a) = 4 + a$ if $a > 0$.

2.2. *Improving upon the MLE for $Q(a, r, \theta)$.* Let X be $Q(a, r, \theta)$ (see Example 1.4). For simultaneously estimating θ_i, we are going to improve upon the MLE. Note first that if $p \geq 2$, the support of X is not $\{x \geq x^0\}$ for any x^0 in Z^p, hence the results of Theorem 2.1 and 2.2 cannot be directly applied to solve the problem here. However, we can decompose the density into several parts and then think of each part as a "modified" negative multinomial density. For example, the density in (1.4) is $f(x) = f_1(x) + \cdots + f_p(x)$, where, for $j = 1, \dots, p$ with A_j given in (1.4),

$$f_j(x) = \frac{(x_1 + \cdots + r_j - 1 + \cdots + x_p)!}{(x_1)! \cdots (r_j - 1)! \cdots (x_p)!} \theta_1^{x_1} \cdots \theta_j^{r_j} \cdots \theta_p^{x_p} 1_{A_j}(x).$$

THEOREM 2.3. *Let $X = (X_1, \dots, X_p)$ be $Q(a, r, \theta)$ with $a + \#\{i, r_i = 0\} \geq 5$. For estimating θ under $\mathscr{L}_m$, $m = 1, 2, \dots$, the MLE $\delta(X) = (X_i/(X_1 + \cdots + X_p))$ is dominated by $\delta(X) + \phi(X)$, where $\phi(X) = (\phi_i(X))$ is* *given by*

$$\phi_i(x) = \frac{-(x_i - m)}{(x_1 + \cdots + x_p)^2} 1_{(x_i \geq r_i + m)}(x), \qquad i = 1, \dots, p.$$

PROOF. See the Appendix. □

REMARK 2.2. There are many possible choices for $\phi(x)$. For example, $\phi_i(x)$ can be chosen as follows:

$$\phi_i(x) = \frac{-h(s)(x_i - m)}{(x_1 + \cdots + x_p + c)^2} 1_{(x_i \geq r_i + m)}(x),$$

where $-1 \leq c \leq 0$, $h(\cdot)$ is nondecreasing such that $0 \leq h(\cdot) \leq 2(a + \#\{i, r_i = 0\} - 3)$ and $h(\cdot) \not\equiv 0$.

3. Admissibility of usual estimators under squared error loss. For some random vectors X with density (1.1), we can prove admissibility of standard estimators by using an induction argument on p together with the "local" admissibility property of a "modified" Bayes rule w.r.t. some prior and subset of the support of X, and a "modified" Blyth method [Berger (1980)]. For simplicity, $(\beta)_+$, β a real number, denotes $\max\{0, \beta\}$.

THEOREM 3.1 (Negative multinomial distribution). *Let X have density (1.3). For estimating θ under the loss function $\mathscr{L}(\theta, b) = \sum_{j \in S}(\theta_j - b_j)^2$, S a subset of $\{1, \dots, p\}$, all estimators*

$$\delta_a^\alpha(X) = \frac{\{X_i - a_i\}_+}{s(X) + r - \alpha},$$

where $a = (a_1, \dots, a_p)$, $a_i \geq 0$, $\sum_{i=1}^p a_i - r < \alpha \leq 1$, are admissible for all p.

PROOF. See the Appendix. □

It follows from Theorem 3.1 that for all p, the MLE $\delta_i(X) = X_i/(X_1 + \cdots + X_p + r)$ and the UMVUE $\delta_i^1(X) = X_i/(X_1 + \cdots + X_p + r - 1)$ of θ_i are admissible under the squared error loss function. In contrast, these two estimators are inadmissible for all loss functions $\mathscr{L}_m$, $m = 1, 2, \dots$, and $p \geq 4$.

THEOREM 3.2 (Gas station distribution). *Let X have the density* (1.5). *For estimating $\theta = (\theta_1, \dots, \theta_p)$ under the loss function given in Theorem* 3.1, *all estimators*

$$\delta_a(X) = \frac{\{X_i - a_i\}_+}{\{s(X) - s(a)\}_+},$$

where $a = (a_1, \dots, a_p)$, $a_i \geq 0$, $\sum_{i=1}^p a_i < 1$, are admissible for all $p \geq 2$.

PROOF. The proof is similar to that of Theorem 3.1, and is omitted. □

It is a consequence of Theorem 3.2 that the MLE $X_i/(X_1 + \cdots + X_p)$ is admissible.

REMARK 3.1. In Theorem 3.2, g can be an arbitrary discrete distribution with finite or countably infinite subset of $N \cup \{0\}$ as support. If g is degenerate at a positive point n, X is multinomial or binomial (if $p = 2$). Thus for simultaneously estimating θ_i, the MLE $X_i/(X_1 + \cdots + X_p)$ is admissible under squared error loss, or the loss function given in Theorem 3.1, for all $p \geq 2$ and n. However, this result was subsumed under Johnson (1971).

THEOREM 3.3 (Quota fulfillment distribution). *Let* $X = (X_1, \dots, X_p)$ *be* $Q(a, r, \theta)$. *For estimating* θ *under squared error loss, the MLE* $(X_i/(X_1 + \cdots + X_p))$ *is admissible for all* $p \geq 2$.

PROOF. The proof is similar to that of Theorem 3.1, and is omitted. □

Note that although, for all p, the MLE is admissible for $\mathscr{L}_0$, it is inadmissible for all $\mathscr{L}_m$, $m = 1, 2, \dots$, and for large p.

REMARK 3.2. In Theorem 3.3, if $p = 2$, $a = \#\{i, r_i \neq 0\} = 1$, X is a negative binomial distribution and obviously the MLE is admissible.

REMARK 3.3. As before, in Theorem 3.3 the loss function can be replaced by that given in Theorem 3.1, and the class of admissible estimators includes all estimators $(\{X_i - a_i\}_+/\{s(X) - s(a)\}_+)$, $a_i \geq 0$, $s(a) < 1$.

APPENDIX

PROOF OF THEOREM 2.1. It is sufficient to prove that $\phi(x)$ is a solution of $\mathscr{D}(x) \leq 0$ for all x, with strict inequality for some x, where $\mathscr{D}(x)$ is given in (2.5). For $x \ngeq \tilde{x}$, there exists some j_0 such that $x_{j_0} < \tilde{x}_{j_0}$, so $\phi_i(x + me_i) = \phi_i(x + (m-1)e_i) = 0$ for all $i = 1, \dots, p$. Therefore $\mathscr{D}(x) = 0$ for all $x \ngeq \tilde{x}$.

Now suppose $x \geq \tilde{x}$. Because $t(x - e_i)/t(x) \geq (x_i + b)/(s + c)^\alpha$ whenever $x_i \geq m + \tilde{x}_i$, and $h(\cdot)$ is nondecreasing,

$$\mathscr{D}(x) \leq \frac{1}{t(x)} \sum_{i=1}^{p} t(x + (m-1)e_i) \Bigg[\frac{h^2(s+m)(x_i + m + b)}{(s + m + c)^{\alpha+2}} + 2\Bigg(\frac{-h(s+m)(x_i + m + b)}{(s + m + c)^{\alpha+1}} + \frac{h(s+m)(x_i + m - 1 + b)}{(s + m + c - 1)^{\alpha+1}} \Bigg) \Bigg].$$

Define

$$d(x_i|x) = \frac{h^2(s+m)(x_i+m+b)}{(s+m+c)^{\alpha+2}} + 2\left(\frac{-h(s+m)(x_i+m+b)}{(s+m+c)^{\alpha+1}} + \frac{h(s+m)(x_i+m-1+b)}{(s+m+c-1)^{\alpha+1}}\right),$$

then

$$\begin{aligned}\sum_{i=1}^{p} d(x_i|x) &= \frac{h(s+m)}{B^{\alpha+2}(B-1)^{\alpha+1}}\big[h(s+m)A(B-1)^{\alpha+1} \\ &\qquad -2AB(B-1)^{\alpha+1} + 2(A-p)B^{\alpha+2}\big] \\ &\le \frac{h(s+m)}{(B-1)^{\alpha+1}B^{\alpha+2}}\big[h(s+m)AB^{\alpha+1} \\ &\qquad +2(\alpha+1)AB^{\alpha+1} - 2pB^{\alpha+2}\big] \\ &= \frac{h(s+m)}{(B-1)^{\alpha+1}B}\{(h(s+m)+2(\alpha+1)-2p)s \\ &\qquad +p[(m+b)(h(s+m)+2(\alpha+1))-2(c+m)]\} \le 0,\end{aligned}$$

where $A = s + mp + bp$, $B = s + m + c$. Note that these two inequalities follow from the facts, $(B-1)^{\alpha+1} \ge B^{\alpha+1} - (\alpha+1)B^{\alpha}$ for all $\alpha \ge 0$, (2.7) and (2.8); the first inequality is strict for those x satisfying $h(s(x)) \ne 0$. Since it is impossible that the $d(x_i|x)$ are positive for all $i = 1, \dots, p$, define $x_{i_0} = \max\{x_i, i = 1, \dots, p, d(x_i|x) \le 0\}$. Because $t(x + (m-1)e_i) \ge t(x + (m-1)e_j)$ and $d(x_i|x) \le d(x_j|x)$ if $x_i \le x_j$, it follows that

$$\begin{aligned}\mathscr{D}(x) &\le \frac{1}{t(x)}\sum_{i=1}^{p} t(x+(m-1)e_i)d(x_i|x) \\ &= \frac{1}{t(x)}\Bigg[\sum_{i,\, x_i \le x_{i_0}} t(x+(m-1)e_i)d(x_i|x) \\ &\qquad + \sum_{i,\, x_i > x_{i_0}} t(x+(m-1)e_i)d(x_i|x)\Bigg] \\ &\le \frac{1}{t(x)}t(x+(m-1)e_{i_0})\sum_{i=1}^{p} d(x_i|x).\end{aligned}$$

From the property of $\sum_{i=1}^{p} d(x_i|x)$, the proof is complete. □

PROOF OF THEOREM 2.2. Let $\phi = (\phi_i)$ be as given in (2.6) with $\Omega_i = \{x: x \ge \tilde{\tilde{x}}, x_i \ge \tilde{\tilde{x}}_i + m\}$, and for a given $\tilde{k}$ with $\tilde{k} \ge k$ and $p - (\alpha+1) - \tilde{k} > 0$,

$c \geq a$, h and $\tilde{\tilde{x}}$ satisfy the conditions:

(i) $\tau = 2(m + c)/(m + b) - 2(\alpha + 1) - 2\tilde{k} > 0$ if $m + b > 0$, or $m + c \geq 0$ if $m + b \leq 0$;

(ii) $\tilde{\tilde{x}} \geq \tilde{x}$, and for all $x^{*} \geq \tilde{\tilde{x}} + m + s(x) > 0$, $\delta(x) \geq \delta^0(x)(1 - k/(s(x) + \beta))$ and $k(s(x) + m + c)/(s(x) + m + \beta) \leq \tilde{k}$;

(iii) $h(\cdot)$ is nondecreasing with $h \not\equiv 0$, $0 \leq h \leq \min\{2p - 2(\alpha + 1) - 2\tilde{k}, \tau\}$ if $m + b > 0$, or $0 \leq h \leq 2p - 2(\alpha + 1) - 2\tilde{k}$ if $m + b \leq 0$.

From (2.4) and (2.5), $R(\theta, \delta + \phi) - R(\theta, \delta) = ED(X)$, $R(\theta, \delta^0 + \phi) - R(\theta, \delta^0) = E\mathscr{D}(X)$ and

$$D(x) = \mathscr{D}(x) + \sum_{i=1}^{p} 2\frac{t(x + (m - 1)e_i)}{t(x)}\left(\frac{\delta_i(x + me_i)}{\delta_i^0(x + me_i)} - 1\right)\phi_i(x + me_i).$$

For $x \not\geq \tilde{\tilde{x}}$, we can show $D(x) = 0$ as in Theorem 2.1. Now suppose $x \geq \tilde{\tilde{x}}$, and observe first that

$$\mathscr{D}(x) \leq \frac{1}{t(x)}\sum_{i=1}^{p} t(x + (m - 1)e_i)d(x_i|x),$$

where $d(x_i|x)$ is defined in Theorem 2.1. Hence

$$D(x) \leq \frac{1}{t(x)}\sum_{i=1}^{p} t(x + (m - 1)e_i)$$

$$\times\left[d(x_i|x) + 2\frac{k}{(s + m + \beta)}\,\frac{h(s + m)(x_i + m + b)}{(s + m + c)^{\alpha+1}}\right].$$

Let

$$d^*(x_i|x) = d(x_i|x) + 2\frac{kh(s + m)(x_i + m + b)}{(s + m + \beta)(s + m + c)^{\alpha+1}}.$$

Then with A, B given in Theorem 2.1,

$$\sum_{i=1}^{p} d^*(x_i|x) \leq \frac{h(s + m)}{(B - 1)^{\alpha+1}B}\left(h(s + m)A + 2(\alpha + 1)A - 2pB + \frac{2k(s + mp + bp)(s + m + c)}{s + m + \beta}\right)$$

$$\leq \frac{h(s + m)}{(B - 1)^{\alpha+1}B}(h(s + m)A + 2(\alpha + 1)A - 2pB + 2\tilde{k}A) \leq 0.$$

Note that these inequalities follow from the facts

$$\sum_{i=1}^{p} d(x_i|x) \leq \frac{h(s + m)}{(B - 1)^{\alpha+1}B}(h(s + m)A + 2(\alpha + 1)A - 2pB),$$

and (i)–(iii). By the same arguments as those in Theorem 2.1,

$$D(x) \le \frac{1}{t(x)} t\big(x + (m-1)e_{i_0}\big) \sum_{i=1}^{p} d^*(x_i|x) \le 0,$$

where i_0 is the index of $\max\{x_i,\ d^*(x_i|x) \le 0\}$. □

PROOF OF THEOREM 2.3. We only give proofs for the cases (i) $a = p$, $r_i \neq 0$ for all i, and (ii) $a = p - 2$, $\#\{i,\ r_i \neq 0\} = p - 1$, since the proofs for the other cases are similar.

CASE (i). $a = p$, $r_i \neq 0$ for all i. For this case X has the density (1.4). We find $\Delta R = R(\delta + \phi, \theta) - R(\delta, \theta) = \sum_{j=1}^{p} R_j$, where, for $j = 1, \dots, p$,

$$R_j = \sum_x \left(\sum_{\substack{i=1\\ i\neq j}}^{p} \left(\frac{(\delta_i(x) + \phi_i(x) - \theta_i)^2}{\theta_i^m} - \frac{(\delta_i(x) - \theta_i)^2}{\theta_i^m} \right) t_j(x) \theta_1^{x_1} \cdots \theta_p^{x_p} \right).$$

The reason for $i \neq j$ in the second summation is $\phi_j(x) = 0$ whenever $x_j = r_j$. Consider a $p - 1$ variate negative multinomial Y with the density

$$f(y) = \frac{(y_1 + \cdots + y_{p-1} + r_p - 1)!}{(y_1)! \cdots (y_{p-1})!(r_p - 1)!} \theta_1^{y_1} \cdots \theta_{p-1}^{y_{p-1}} (1 - \theta_1 - \cdots - \theta_{p-1})^{r_p}.$$

By Theorem 2.2 if $p \ge 5$, the MLE, $\tilde{\delta}(Y) = Y_i/(Y_1 + \cdots + Y_{p-1} + Y_p)$ is dominated by $\tilde{\delta}(Y) + \tilde{\phi}(Y)$ under $\mathscr{L}_m$, $m > 0$, where $\tilde{\phi}_i(y) = \phi_i(y) 1_{(y_p = r_p)}$, $i = 1, \dots, p - 1$. And we know that the difference between $R_Y(\theta, \tilde{\delta})$ and $R_Y(\theta, \tilde{\delta} + \tilde{\phi})$ is exactly equal to R_p. Therefore $R_p < 0$. Similarly we can prove $R_j < 0$ for $j = 1, \dots, p$. So $\Delta R < 0$.

CASE (ii). $a = p - 2$, $\#\{i,\ r_i \neq 0\} = p - 1$. Without loss of generality, we assume $r_1 = 0$, $r_i \neq 0$, $i = 2, \dots, p$. For simplicity, let $A_{j,k} = \{x\colon x_j = r_j, x_k < r_k, x_i \ge r_i, 1 \le i \le p, i \neq k\}$. The density of X is

$$f(x) = \sum_{\substack{j,k=2\\ j\neq k}}^{p} \frac{(x_1 + \cdots + x_p - 1)!}{(x_1)! \cdots (x_j - 1)! \cdots (x_p)!} 1_{A_{j,k}}(x) \theta_1^{x_1} \cdots \theta_p^{x_p}.$$

We find

$$\Delta R = R(\delta + \phi, \theta) - R(\delta, \theta) = \sum_{\substack{j,k=2\\ j\neq k}}^{p} R_{j,k},$$

where, for example,

$$R_{p,p-1} = \sum_{x_{p-1}=0}^{r_{p-1}-1} \frac{(x_{p-1} + r_p - 1)!}{(x_{p-1})!(r_p - 1)!} \frac{\theta_{p-1}^{x_{p-1}} \theta_p^{r_p}}{(1 - \theta_1 - \cdots - \theta_{p-2})^{x_{p-1} + r_p}} d(x_{p-1}),$$

and, defining $\Omega = \{x\colon x_p = r_p, x_i \geq r_i, 1 \leq i \leq p-2\}$,

$$d(x_{p-1}) = \sum_{\Omega} \left(\sum_{i=1}^{p-2} \left[\left(\frac{(\delta_i(x) + \phi_i(x) - \theta_i)^2}{\theta_i^m} - \frac{(\delta_i(x) - \theta_i)^2}{\theta_i^m} \right) \times \frac{(x_1 + \cdots + x_{p-2} + x_{p-1} + r_p - 1)!}{(x_1)! \cdots (x_{p-2})!(x_{p-1} + r_p - 1)!} \times \theta^{x_1} \cdots \theta_{p-2}^{x_{p-2}} (1 - \theta_1 - \cdots - \theta_{p-2})^{x_{p-1} + r_p} \right] \right).$$

By the same argument for $R_p < 0$ in case (i), $d(x_{p-1}) < 0$. Hence $R_{p,p-1} < 0$. Similarly, $R_{j,k} < 0$ for all $j, k, j \neq k$, so $\Delta R < 0$. □

In order to prove Theorem 3.1, we need two inequalities. The first follows from Holder's inequality: For the ordinary Gamma function Γ,

(A.1) $\qquad \Gamma(x + a) \leq x^a \Gamma(x)$ for all a, $0 < a < 1$, and all $x > 0$.

A straightforward calculation establishes the second:

(A.2) $$\sum_{\Omega} \frac{1}{x_1 \cdots x_p} \leq \frac{1}{n} [2(\log n + 1)]^{p-1},$$

where $\Omega = \{(x_1, \dots, x_p)\colon x_1 + \cdots + x_p = n, x_i \geq 1, 1 \leq i \leq p\}$.

Proof of Theorem 3.1. We only give proofs for the cases (i) $S = \{1, \dots, p\}$, $a_i = 0$, $i = 1, \dots, p$, $-r < \alpha < 1$ and (ii) $S = \{1, \dots, p\}$, $a_i = 0$, $i = 1, \dots, p$, $\alpha = 1$. The proofs for the other cases are similar.

Case (i). Let Θ denote the parameter space $\{(\theta_1, \dots, \theta_p)\colon 0 < \theta_i, \sum_{i=1}^p \theta_i < 1\}$, and let $\tilde{\Theta} = \{(\theta_1, \dots, \theta_p)\colon 0 \leq \theta_i, \sum_{i=1}^p \theta_i < 1\}$. Note that $R(\theta, \delta^\alpha)$, where $\delta_i^\alpha = x_i/(x_1 + \cdots + x_p + r - \alpha)$, can be defined and is continuous in $\tilde{\Theta}$. Then the admissibility of δ^α in Θ is equivalent to the admissibility of δ^α in $\tilde{\Theta}$.

When $p = 1$ it is clear that δ^α is the unique Bayes estimator w.r.t. the prior π_α with $d\pi_\alpha = \theta^{-1}(1 - \theta)^{-(1+\alpha)} 1_{(0,1)}(\theta)\, d\theta$, hence it is admissible. Now assume the theorem is true for $p - 1$. Suppose there exists an estimator $\delta'(X)$ such that, for all θ in $\tilde{\Theta}$, $R(\theta, \delta') \leq R(\theta, \delta^\alpha)$. With $\theta_j = 0$ in both sides of the inequality, from the property of the density (1.3) and the induction hypothesis, we have $\delta_i'(x) = \delta_i(x)$ for all $i = 1, \dots, p$, if $x_j = 0$. Therefore $\tilde{R}(\theta, \delta') \leq \tilde{R}(\theta, \delta^\alpha)$. Here $\tilde{R}(\theta, \delta) = \sum_{\Omega} (\sum_{i=1}^p (\delta_i(x) - \theta_i)^2 f(x|\theta))$, where $\Omega = \{x\colon x_i \geq 1, 1 \leq i \leq p\}$. Define the modified Bayes risk $B(\gamma, \delta) = \int \tilde{R}(\theta, \delta)\, d\gamma$. For any x with $x_i \geq 1$, $\delta_i^\alpha(x) = (\int \theta_i f(x|\theta)\, d\pi_\alpha)/(\int f(x|\theta)\, d\pi_\alpha)$, where $d\pi_\alpha =$

$1/[\theta \dots \theta_p(1 - \theta_1 - \cdots - \theta_p)^{1+\alpha}]1_{\Theta}(\theta)\,d\theta$. By (A.1) and (A.2)

$$B(\pi_\alpha, \delta^\alpha)$$
$$\leq \frac{\Gamma(r-\alpha)}{\Gamma(r)} \sum_n (n + r - \alpha)^{\max\{\alpha, 0\}} \frac{1}{n + r - \alpha} \frac{1}{n} (2(\log n + 1))^{p-1} < \infty.$$

Thus $\delta^\alpha(X)$ is the unique modified Bayes estimator w.r.t. π_α, and $R(\theta, \delta') \leq R(\theta, \delta^\alpha)$ for all θ implies $\delta'(X) = \delta^\alpha(X)$.

CASE (ii). We also prove this by induction. For $p = 1$, the distribution is negative binomial and it is well known that $\delta(X) = X/(X + r - 1)$ is admissible. Assume that the theorem is true for $p - 1$. If there is an estimator $\delta'(X)$ such that $R(\theta, \delta') \leq R(\theta, \delta)$ for all θ, and $R(\theta, \delta') < R(\theta, \delta) - \varepsilon_0$ for some $\varepsilon_0 > 0$ and all θ in some open set O, we can similarly show that $\delta_i'(x) = \delta_i(x)$ for all i, if $x_j = 0$. Hence $\tilde{R}(\theta, \delta') \leq \tilde{R}(\theta, \delta)$ for all θ and $\tilde{R}(\theta, \delta') < \tilde{R}(\theta, \delta) - \varepsilon_0$ for all θ in O. Consider the modified Bayes risk of δ', δ and δ^α, $0 < \alpha < 1$ w.r.t. the prior π_α, where δ^α and π_α are given in case (i). Calculation, by (A.1) and (A.2), gives us $B(\pi_\alpha, \delta) - B(\pi_\alpha, \delta^\alpha) \leq (1 - \alpha)^2 M$, M some constant. And $B(\pi_\alpha, \delta') - B(\pi_\alpha, \delta) < -\varepsilon_0 A$, where $A = \text{area}(O) > 0$. Therefore $0 \leq B(\pi_\alpha, \delta') - B(\pi_\alpha, \delta^\alpha) < -\varepsilon_0 A + (1 - \alpha)^2 M$, which converges to $-\varepsilon_0 A$ as $\alpha \to 1$. We thus have a contradiction and the theorem is proved. □

Acknowledgments. I would like to thank Professor D. Ylvisaker and Professor J. T. Hwang for their patient supervision and assistance. I would also like to thank the associate editor and referees for countless helpful comments, suggestions and corrections.

REFERENCES

ANDERSON, K., SOBEL, M. and UPPULURI, V. R. R. (1982). Quota fulfillment times. *Canad. J. Statist.* **10** 73–88.

BERGER, J. O. (1980). *Statistical Decision Theory: Foundations, Concepts, and Methods.* Springer, New York.

CLEVENSON, M. L. and ZIDEK, J. V. (1975). Simultaneous estimation of the means of independent poisson laws. *J. Amer. Statist. Assoc.* **70** 698–705.

GHOSH, M., HWANG, J. T. and TSUI, K. W. (1983). Construction of improved estimators in multiparameter estimation for discrete exponential families. *Ann. Statist.* **11** 351–367.

HUDSON, H. M. (1978). A natural identity for exponential families with applications in multiparameter estimation. *Ann. Statist.* **6** 473–484.

HWANG, J. T. (1982a). Improving upon standard estimators in discrete exponential families with applications to Poisson and negative binomial cases. *Ann. Statist.* **10** 857–867.

HWANG, J. T. (1982b). Semi tail upper bounds on the class of admissible estimators with applications to Poisson and negative binomial families. *Ann. Statist.* **10** 1137–1147.

JOHNSON, B. MCK. (1971). On the admissible estimators for certain fixed sample binomial problems. *Ann. Math. Statist.* **42** 1579–1587.

SEBER, G. A. F. (1962). The multi-sample single recapture census. *Biometrika* **49** 339–349.

SEBER, G. A. F. (1982). *The Estimation of Animal Abundance*, 2nd ed. Macmillan, New York.

SIBUYA, M., YOSHIMURA, I. and SHIMIZU, R. (1964). Negative multinomial distribution. *Ann. Inst. Statist. Math.* **16** 409–426.

SMITH, P. J. (1988). Bayesian methods for multiple capture–recapture surveys. *Biometrics* **44** 1177–1189.

STEIN, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 197–206. Univ. California Press, Berkeley.

STEIN, C. (1973). Estimation of the mean of a multivariate normal distribution. Technical Report 48, Stanford Univ.

TSUI, K. W. and PRESS, S. J. (1982). Simultaneous estimation of several Poisson parameters under k-normalized squared error loss. *Ann. Statist.* **10** 93–100.

TSUI, K. W. (1984). Robustness of Clevenson–Zidek type estimators. *J. Amer. Statist. Assoc.* **79** 152–157.

YOUNG, D. H. (1961). Quota fulfillment using unrestricted random sampling. *Biometrika* **48** 333–342.

INSTITUTE OF STATISTICAL SCIENCE
ACADEMIA SINICA
TAIPEI, TAIWAN 11529
REPUBLIC OF CHINA