The Annals of Statistics
1975, Vol. 3, No. 1, 165-188

CENTRAL LIMIT THEOREMS FOR MULTINOMIAL SUMS

BY CARL MORRIS

The Rand Corporation, Santa Monica

Let $(N_1, \cdots, N_k)$ be a multinomial vector with $n = \sum N_i$ and with parameter $(p_1, \cdots, p_k)$, $\sum p_i = 1$. Let $f_1, \cdots, f_k$ be real-valued functions defined on the integers $\{0, 1, \cdots, n\}$, and let $S_k = \sum_{i=1}^k f_i(N_i)$. Suppose $k \to \infty$, and as $k \to \infty$, that $n \to \infty$ and $\max_{1 \leqq i \leqq k} p_i \to 0$. Conditions on the $\{f_i\}$ are given which guarantee that S_k, suitably centered and scaled, has a normal limit in law. An application shows that if $\min_{1 \leqq i \leqq k} (np_i)$ is bounded away from zero and the $\{f_i\}$ are polynomials of bounded degree as $k \to \infty$, that S_k is asymptotically normal provided only that a "uniformly asymptotically negligible" (uan) condition on the $\{f_i\}$ holds.

For testing the specified simple hypothesis $p_i = p_i^0$ for all $1 \leqq i \leqq k$, Pearson's "chi-square" statistic and the likelihood ratio statistic can be written in the form of S_k. It is shown that these two statistics are asymptotically normal as $k \to \infty$ provided they satisfy simple conditions which are equivalent to their respective uan conditions.

1. Introduction. Let $N = (N_1, \cdots, N_k)$ be a multinomial vector with parameter $p = (p_1, \cdots, p_k)$ such that $\sum N_i = n$ (an integer) and $\sum p_i = 1$. Let $f_1, \cdots, f_k$ be real-valued functions, each having domain on the set of integers $\{0, 1, \cdots, n\}$, and let $S_k = \sum_{i=1}^k f_i(N_i)$. We are interested in determining conditions which insure that S_k, suitably scaled and centered, has approximately a normal distribution for sufficiently large k. The difficulty in proving such a theorem is that S_k is a sum of dependent random variables.

Statistics of the form S_k are of interest because they arise in practice. For example, several well-known test statistics for testing simple hypotheses in a multinomial distribution have such a form, most notably Pearson's "chi-square" statistic [5] and the likelihood ratio statistic [4]. These particular statistics are well known to have the same asymptotic chi-square sampling distribution in the "standard case" where k is fixed and n is large. The results of this paper provide asymptotically normal sampling distributions for these and other statistics when k is large. When n/k is moderate, the chi square and likelihood ratio statistics have different asymptotic normal distributions. These distributions for large k would be needed, for example, in determining a consistent sequence of tests, based on either Pearson's chi-square statistic or the likelihood ratio statistic, that a given sample comes from a specified probability distribution.

Special cases of the above central limit theorem have been considered earlier. Tumanyan [9] demonstrated that Pearson's chi-square statistic is asymptotically

Received May 1972; revised December 1973.

AMS 1970 *subject classifications.* Primary 62E20; Secondary 62F05.

Key words and phrases. Central limit theorem for multinomial sums, hypothesis testing with many parameters, likelihood ratio statistic for the multinomial distribution, multinomial distribution, Pearson's chi-square statistics.

normal when both $\min_{1 \le i \le k} np_i \to \infty$ and $k \to \infty$ and the null hypothesis holds. Under these conditions, the usual chi-square limit holds, but the degrees of freedom are large and so the chi-square distribution is nearly normal. Tumanyan's method does not lend itself to generalization. Steck [7] showed that Pearson's chi-square statistic is asymptotically normal for large k under the null and alternative hypotheses. His approach to the problem is described in the next section. It is Steck's approach that was used by the author [3] to show that S_k has a normal limit if for every i, f_i is a polynomial of the second degree, and provided only that

$$\max_{1 \le i \le k} p_i = o(1)\,, \tag{1.1}$$

$$\min_{1 \le i \le k} np_i \quad \text{is bounded away from zero as} \quad k \to \infty \tag{1.2}$$

and the "uan (uniformly asymptotically negligible) condition"

$$\max_{1 \le i \le k} \operatorname{Var} f_i(N_i) / \textstyle\sum_{i=1}^{k} \operatorname{Var} f_i(N_i) = o(1) \tag{1.3}$$

hold.

In this paper, Steck's approach is used again to develop a fundamental lemma (Lemma 2.2) where conditions are given which insure that S_k has the same limit in law as S_k^*,

$$S_k^* \equiv \textstyle\sum f_i(X_i)\,, \tag{1.4}$$

whenever S_k^* has a normal limit, X_i being Poisson with the same mean as N_i, and $\{X_i\colon 1 \le i \le k\}$ being independent. That is, the fundamental lemma reduces the problem to consideration of sums of independent random variables. Then, in Section 3, certain useful lemmas concerning the Poisson–Charlier representations for multinomial sums are developed before turning to the main results contained in Sections 4 and 5.

In Section 4, the basic central limit theorem for multinomial sums (Theorem 4.1) gives conditions for asymptotic normality of S_k in terms of the coefficients of the Poisson–Charlier expansion of the $\{f_i\}$. Then Theorem 4.2 gives sufficient conditions for Theorem 4.1. Corollary 4.1 to Theorem 4.2 then shows that if the $\{f_i\}$ are polynomials with bounded degree as $k \to \infty$ that very simple conditions guarantee the central limit theorem.

Applications to the distribution of Pearson's chi-square statistic [5] and the likelihood ratio test statistic (for multinomials) [4] are considered in Section 5. It is shown in Theorems 5.1 and 5.2 that Pearson's chi-square statistic and the likelihood ratio statistic are asymptotically normal as $k \to \infty$ provided only that (1.1), (1.2) and the uan condition hold. The uan condition for these two statistics always holds under (1.1) and (1.2) if the "null hypothesis" is true.

2. Fundamental lemma for asymptotic normality of multinomial sums. The purpose of this section is to establish notation and to present Lemma 2.2 for proving asymptotic normality of multinomial sums. The following notation will be used throughout this paper. The well-known "choice function" is denoted by $\binom{n}{x_1,\cdots,x_k} = n!/\prod_{i=1}^k x_i!$ if $\sum_i x_i = n$ and if the x_i are all nonnegative integers. We define $\binom{n}{x_1,\cdots,x_k}$ for any other $\{x_i\}$ to be zero. The choice function $\binom{n}{x}$ is defined to be $\binom{n}{x_1, n-x_1}$ and is similarly extended to be zero when x_1 is not an integer in the $[0, n]$ interval. To simplify notation, we adopt the convention that when the limits of a sum are not specified, the sum is assumed to be over all nonzero terms. That is, $\sum_i x_i$ indicates the sum over all values of x_i which are defined and for which $x_i \neq 0$. The symbol $\mathscr{L}$ is used to denote the "law" or "distribution" of a random variable. The vector $N = (N_1, \cdots, N_k)$ has a multinomial distribution with parameters $p = (p_1, \cdots, p_k)$, k and n, denoted $\mathscr{L}(N) = \mathscr{M}(p, k, n)$, if

$$P(N = (x_1, \cdots, x_k)) = \binom{n}{x_1 \cdots x_k} \prod_{i=1}^k p_i^{x_i} .$$

We will assume throughout that $p_i > 0$. Due to the conventions just adopted for the choice function, N must always be a vector of nonnegative integers with $\sum N_i = n$. We indicate that X has the binomial distribution with parameters $n \geqq 1$ and p, by $\mathscr{L}(X) = B(n, p)$ if $P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ for integers $0 \leqq x \leqq n$, and we indicate that X has the Poisson distribution with mean $\lambda \geqq 0$ by $\mathscr{L}(X) = \text{Poisson}(\lambda)$ if $P(X = x) = \lambda^x e^{-\lambda}/x!$ for nonnegative integers x. If $X = (X_1, \cdots, X_k)$ has the k-variate normal distribution with parameters $\mu \equiv EX$ and $\Sigma = EX'X - \mu'\mu$ it will be denoted as $\mathscr{L}(X) = \mathscr{N}_k(\mu, \Sigma)$. If $k = 1$, we may delete the subscript, writing $\mathscr{N}_1(\mu, \Sigma) = \mathscr{N}(\mu, \Sigma)$.

Steck's argument [7] will now be outlined. Let $\mathscr{L}(N) = \mathscr{M}(p, k, n)$ with $p = (p_1, \cdots, p_k)$ given. Let $X_1, \cdots, X_k$ be independent random variables with $\mathscr{L}(X_i) = \text{Poisson}(\lambda_i)$, $\lambda_i \equiv np_i$, and denote $X = (X_1, \cdots, X_k)$. It is well known that

$$\mathscr{L}(N) = \mathscr{L}(X \mid \textstyle\sum_{i=1}^k X_i = n) , \tag{2.1}$$

the latter notation representing the conditional law of X given $\sum X_i = n$.

Let $f_1, \cdots, f_k$ be real-valued functions, each having domain $\{0, \cdots, n\}$. Let f_i^* be any function on the nonnegative integers which agrees with f_i on $\{0, \cdots, n\}$. (We note that f_i^* may be taken to be a polynomial of degree at most n by defining $f_i^*(x) = \sum_{j=0}^n \Delta^j f_i(0) x^{(j)}/j!$ with $x^{(j)} \equiv \prod_{i=1}^j (x - i + 1)$ if $j \geqq 1$ $x^{(0)} = 1$ and Δ the finite difference operator: $\Delta f(x) \equiv f(x+1) - f(x)$. This fact is discussed further at the end of Section 4.) Define the function g_i on the nonnegative integers by

$$g_i(x) = f_i^*(x) - Ef_i^*(X_i) - \gamma(x - \lambda_i) \tag{2.2}$$

with

$$\gamma \equiv \frac{1}{n} \textstyle\sum_{i=1}^k \text{Cov}\,(f_i^*(X_i), X_i) . \tag{2.3}$$

Define

$$s_k^2 \equiv \textstyle\sum_{i=1}^k \text{Var}\, g_i(X_i) = \sum_{i=1}^k \text{Var}\, f_i^*(X_i) - n\gamma^2 \tag{2.4}$$

and

$$U_k(X) = (1/s_k) \textstyle\sum_{i=1}^k g_i(X_i) , \qquad V_k(X) = 1/n^{\frac{1}{2}} \sum_{i=1}^k (X_i - \lambda_i) . \tag{2.5}$$

We are interested in the distribution of $S_k = \sum_{i=1}^k f_i(N_i)$. Since $0 \leqq N_i \leqq n$ and $\sum_i (N_i - \lambda_i) = 0$, we have

$$S_k = \sum_{i=1}^k f_i^*(N_i) = \sum g_i(N_i) + \sum_i Ef_i^*(X_i) \tag{2.6}$$
$$= s_k U_k(N) + \sum_i Ef_i^*(X_i) .$$

From the expression (2.1) we have, since $\sum \lambda_i = n$,

$$\mathscr{L}(U_k(N)) = \mathscr{L}(U_k(X) \mid V_k(X) = 0) .$$

Hence, to study the distribution of $\sum f_i(N_i)$, it suffices to study the conditional distribution of $U_k(X)$ given $V_k(X)$, U_k and V_k each being sums of *independent* random variables. It is easy to see that

$$EU_k(X) = EV_k(X) = 0 , \qquad \text{Var } U_k(X) = \text{Var } V_k(x) = 1 \quad \text{and}$$
$$\text{Cov } (U_k(X), V_k(X)) = 0 ,$$

so conditions guaranteeing that

$$\mathscr{L}((U_k, V_k)) \to \mathscr{N}_2((0, 0), (\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix})) \qquad \text{as} \quad k \to \infty$$

are well known. If the conditional law $\mathscr{L}(U_k \mid V_k)$ converges to the conditional $\mathscr{N}(0, 1)$ law of the limiting distribution, then

$$\mathscr{L}(U_k \mid V_k = 0) \to \mathscr{N}(0, 1) \qquad \text{as} \quad k \to \infty ,$$

and hence

$$\mathscr{L}((1/s_k)\{\sum f_i(N_i) - \sum Ef_i^*(X_i)\}) \to \mathscr{N}(0, 1) \qquad \text{as} \quad k \to \infty .$$

Steck's condition in [7] for convergence of conditional distributions, requiring uniform equicontinuity of the conditional characteristic functions follows. We recall that a set of functions $\psi_k(\cdot)$ are uniformly equicontinuous on bounded sets if for every given $\varepsilon > 0$ and bounded set C, there exists $\delta > 0$ such that for all k,

$$|\psi_k(v + h) - \psi_k(v)| < \varepsilon$$

whenever $|h| < \delta$, and $v + h$, v are in C.

THEOREM 2.1. (Steck [7], page 241.) *Let (U_k, V_k) be a sequence of random variables with limits in law denoted by (U, V), so $\mathscr{L}(U_k, V_k) \to \mathscr{L}(U, V)$ as $k \to \infty$. Let*

$$\psi_k(v; t) = E\{\exp(itU_k) \mid V_k = v\}$$

be a version of the conditional characteristic function. If for every fixed t, the family of functions

$$\{\psi_k(\cdot; t) : k = 1, 2, \cdots\}$$

is uniformly equicontinuous on bounded sets, then

$$\mathscr{L}(U_k \mid V_k = v) \to \mathscr{L}(U \mid V = v) \qquad \textit{as} \quad k \to \infty .$$

The lemma we prove for convergence to the multivariate normal distribution is the following.

LEMMA 2.1. *A central limit lemma for the multivariate normal distribution* $\mathscr{N}_p(0, I)$.

Let $S_k = (S_{1k}, \cdots, S_{pk}) = \sum_{i=1}^k X_{ik}$ *with* $X_{ik} = (X_{i1k}, \cdots, X_{ipk})$ *a p-vector and* $\{X_{1k}, \cdots, X_{kk}\}$ *be independent. Suppose* $EX_{ijk} = 0$ *for all* i, j, k *and* $ES_{jk}S_{j'k} = \delta_{jj'}$ *(Kronecker delta) for all* k, j, j'. *Suppose all coordinates* $S_{jk} = \sum_i^k X_{ijk}$ *satisfy the* uan *condition, this is* $\max_{1\le i\le k} \mathrm{Var}\,(X_{ijk}) = o(1)$ *as* $k \to \infty$ *for each* $j = 1, \cdots, p$, *and suppose* $\mathscr{L}(S_{jk}) \to \mathscr{N}(0, 1)$ *for each* $j = 1, \cdots, p$. *Then* $\mathscr{L}(S_k) \to \mathscr{N}_p(0, I)$ *as* $k \to \infty$.

PROOF. Since each coordinate satisfies the uan condition and is asymptotically normal, it satisfies the Lindeberg condition ([2], page 280),

$$\delta_{kj}(\varepsilon) = \sum_{i=1}^k EX_{ijk}^2 I_{[\varepsilon,\infty)}(X_{ijk}^2) = o(1)$$

as $k \to \infty$ for each j, for every fixed $\varepsilon > 0$, I being the indicator function.

It will be shown that the Lindeberg condition holds for $\sum_{j=1}^p a_j X_{ijk}$ for every vector $a = (a_1, \cdots, a_p)$ satisfying $\sum a_j^2 = 1$. Since

$$(\textstyle\sum_j a_j X_{ijk})^2 \leqq \sum_j X_{ijk}^2 \leqq p \max_j X_{ijk}^2 ,$$

we have

$$\begin{aligned}\delta_k(\varepsilon) &\equiv \textstyle\sum_{i=1}^k E(\sum_{j=1}^p a_j X_{ijk})^2 I_{[\varepsilon,\infty)}((\sum_{j=1}^p a_j X_{ijk})^2)\\ &\leqq \textstyle\sum_i Ep(\max_j X_{ijk}^2) I_{[\varepsilon,\infty)}(p \max_j X_{ijk}^2)\\ &\leqq p \textstyle\sum_i \sum_j EX_{ijk}^2 I_{[\varepsilon/p,\infty)}(X_{ijk}^2)\\ &= p \textstyle\sum_{j=1}^p \delta_{kj}(\varepsilon/p) = o(1) \qquad \text{as} \quad k \to \infty .\end{aligned}$$

This proves the lemma.

The results stated above are now restated in the form of a fundamental lemma for asymptotic normality of multinomial sums.

LEMMA 2.2. *Let* $\mathscr{L}(N_k) = \mathscr{L}(N_{1k}, \cdots, N_{kk}) = \mathscr{M}(p_k = (p_{1k}, \cdots, p_{kk}), k, n_k)$ *be given.*

For given k, let the real-valued functions $\{f_{ik}: 1 \leqq i \leqq k\}$ *be given, each having as domain the nonnegative integers. With* $\lambda_{ik} \equiv n_k p_{ik}$, *let* $\{X_{ik}: i = 1, \cdots, k\}$ *be independent random variables,* $\mathscr{L}(X_{ik}) = \mathrm{Poisson}\,(\lambda_{ik})$. *Suppose* $Ef_{ij}(X_{ik}) = 0$ *for each i and* $\mathrm{Cov}\,(\sum_i f_{ik}(X_{ik}), \sum_i X_{ik}) = \sum_i \mathrm{Cov}\,(f_{ik}, X_{ik}) = 0$: *(The last two requirements of the* $\{f_{ik}\}$ *can always be accomplished by a transformation similar to that of* (2.2).)

Define

$$\sigma_{ik}^2 = \mathrm{Var}\, f_{ik}(X_{ik}) , \qquad s_k^2 = \textstyle\sum_{i=1}^k \sigma_{ik}^2 . \tag{2.7}$$

Suppose as $k \to \infty$ *that* $n_k \to \infty$, $\max_{1\le i\le k} p_{ik} = o(1)$ *and the* uan *condition*

$$(1/s_k^2) \max_{1\le i\le k} \sigma_{ik}^2 = o(1) \tag{2.8}$$

holds. Assume that

$$\lim_{h\to 0} \sup_k \sup_v (1/s_k^2) E(\textstyle\sum_{i=1}^k \{f_{ik}(L_{ik} + M_{ik}) - f_{ik}(L_{ik})\})^2 = 0 , \tag{2.9}$$

$L_k = (L_{1k}, \cdots, L_{kk})$ *and* $M_k = (M_{1k}, \cdots, M_{kk})$ *being independent multinomial vectors* *with* $\mathscr{L}(L_k) = \mathscr{M}(p_k, k, n_k + v_k n_k^{\frac{1}{2}})$ *and* $\mathscr{L}(M_k) = \mathscr{M}(p_k, k, hn_k^{\frac{1}{2}})$. h *and* v_k *are such that* $l_k \equiv n_k + v_k n_k^{\frac{1}{2}}$ *and* $m_k \equiv hn_k^{\frac{1}{2}}$ *are nonnegative integers and* v_k *is bounded as* $k \to \infty$.

Suppose further that

$$\mathscr{L}((1/s_k) \textstyle\sum_i f_{ik}(X_{ik})) \to \mathscr{N}(0, 1) \quad \text{as} \quad k \to \infty . \tag{2.10}$$

Then $\mathscr{L}((1/s_k) \sum f_{ik}(N_{ik})) \to \mathscr{N}(0, 1)$ *as* $k \to \infty$.

Proof. Let $V_k = (1/n_k^{\frac{1}{2}}) \sum (X_{ik} - \lambda_{ik})$. Then $EV_k = 0$, $\operatorname{Var} V_k = 1$. V_k is a sum of independent random variables and the Liapounov condition ([2], page 275) is

$$\textstyle\sum_i E(X_{ik} - \lambda_{ik})^4/n_k^2 = \sum \{\lambda_{ik} + 3\lambda_{ik}^2\}/n_k^2 \leqq 1/n_k + 3 \max_{1 \leqq i \leqq k} p_{ik} = o(1) \quad \text{as } k \to \infty .$$

Hence $\mathscr{L}(V_k) \to \mathscr{N}(0, 1)$ and V_k satisfies the uan condition.

Let $U_k = (1/s_k) \sum_i f_{ik}(X_{ik})$. Then $U_k \to \mathscr{N}(0, 1)$ and satisfies the uan condition (2.8). We have $\operatorname{Cov}(U_k, V_k) = 0$, and hence from Lemma 2.1 $(U_k, V_k) \to \mathscr{N}_2(0, (\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}))$.

When $V_k = v_k$, $\sum X_{ik} = n_k + v_k n_k^{\frac{1}{2}}$, and so

$$\mathscr{L}(U_k \mid V_k = v_k) = \mathscr{L}((1/s_k) \textstyle\sum f_{ik}(L_{ik})) .$$

Similarly, when $V_k = v_k + h$, $\sum X_{ik} = n_k + v_k n_k^{\frac{1}{2}} + hn_k^{\frac{1}{2}}$ we have $\mathscr{L}(U_k \mid V_k = v_k + h) = \mathscr{L}((1/s_k) \sum f_{ik}(L_{ik} + M_{ik}))$. Letting $\phi_k(v_k; t) = E\{\exp(itU_k) \mid V_k = v_k\}$, then

$$\begin{aligned} &|\phi_k(v_k + h; t) - \phi_k(v_k; t)| \\ &\quad = |E \exp(i(t/s_k) \textstyle\sum_i f_{ik}(L_{ik} + M_{ik})) - E \exp(i(t/s_k) \sum_i f_{ik}(L_{ik}))| \\ &\quad \leqq (|t|/s_k) E|\textstyle\sum_i f_{ik}(L_{ik} + M_{ik}) - \sum_i f_{ik}(L_{ik})| . \end{aligned}$$

From (2.9) and the Schwarz inequality, it follows that

$$\lim_{h \to 0} \sup_k \sup_{v_k} |\phi_k(v_k + h; t) - \phi_k(v_k; t)| = 0$$

which is the equicontinuity condition of Theorem 2.1. Hence, $\mathscr{L}(U_k \mid V_k = 0) \to \mathscr{N}(0, 1)$ as $k \to \infty$. Since $\mathscr{L}(U_k \mid V_k = 0) = \mathscr{L}((1/s_k) \sum f_{ij}(N_{ik}))$, the conclusion follows.

3. Useful lemmas concerning Poisson-Charlier polynomials. This section contains results concerning the Poisson-Charlier polynomials used in the central limit theorems of Section 4.

We define the polynomial $x^{(j)}$ as

$$x^{(j)} = \textstyle\prod_{i=1}^{j} (x - i + 1) \quad \text{for } j \geqq 1 \quad \text{and} \quad x^{(0)} = 1 . \tag{3.1}$$

x may take any real value, and j must be a nonnegative integer.

The addition formula

$$(x + y)^{(j)} = \textstyle\sum_{i=0}^{j} \binom{j}{i} x^{(i)} y^{(j-i)} \tag{3.2}$$

is easy to remember from its binomial expansion analogy and is equivalent to the Vandermonde convolution ([6], page 8). There is also a multiplication formula:

$$x^{(i)}x^{(j)} = \sum_{k=0}^{i} \binom{i}{k} j^{(i-k)} x^{(j+k)} = \sum_k \binom{i}{k}\binom{j}{k} k!\, x^{(i+j-k)}, \tag{3.3}$$

which is easily proved by induction.

The Poisson–Charlier polynomials (Szego [8]) are for $\lambda > 0$:

$$\Pi_j(x, \lambda) \equiv \sum_i \binom{j}{i} (-\lambda)^{j-i} x^{(i)} \tag{3.4}$$

and

$$\varphi_j(x, \lambda) \equiv (j!)^{-\frac{1}{2}} \lambda^{-j/2} \Pi_j(x, \lambda). \tag{3.5}$$

The reader will find that Π_j is easy to remember if he compares it to the binomial expansion of $(x - \lambda)^j$.

Let $\mathscr{L}(X) =$ Poisson (λ). Then

$$E\Pi_j(X, \lambda) = E\varphi_j(X, \lambda) = \delta_{j0} \quad \text{(the Kronecker delta)}, \tag{3.6}$$

$$E\varphi_i(X, \lambda)\varphi_j(X, \lambda) = \delta_{ij}. \tag{3.7}$$

(See [8], pages 33–34.)

The inversion of formula (3.4) is

$$x^{(j)} = \sum_{i=0}^{j} \binom{j}{i} \lambda^{j-i} \Pi_i(x, \lambda). \tag{3.8}$$

The addition formula for Π_j is

$$\Pi_j(x + y, \lambda) = \sum_{i=0}^{j} \binom{j}{i} x^{(i)} \Pi_{j-i}(y, \lambda)$$

which is proved by using (3.4) and (3.2) and interchanging the order of summation. Hence for φ_j the addition formula is

$$\varphi_j(x + y, \lambda) = \sum_{i=0}^{j} \binom{j}{i}^{\frac{1}{2}} \frac{\lambda^{-i/2}}{(i!)^{\frac{1}{2}}} x^{(i)} \varphi_{j-i}(y, \lambda). \tag{3.9}$$

The differentiation formula for Π_j is easily established by induction to be

$$\partial^i \Pi_j(x, \lambda)/\partial\lambda^i = (-1)^i j^{(i)} \Pi_{j-i}(x, \lambda). \tag{3.10}$$

The multiplication formula for the polynomials of (3.5) is as follows.

Lemma 3.1. *With*

$$\varphi_i(x, \lambda)\varphi_j(x, \lambda) \equiv \sum_{h=0}^{i+j} c_{ijh}(\lambda)\varphi_h(x, \lambda)$$

then

$$\begin{aligned} c_{ijh}(\lambda) &= 0 \qquad \text{if} \quad h < |i-j|, \\ &= \left(\frac{h!}{i!\,j!}\right)^{\frac{1}{2}} \sum_t \lambda^{-\frac{1}{2}(2t+h-i-j)} t! \binom{i}{t}\binom{j}{t}\binom{t}{2t+h-i-j} \\ &\qquad\qquad \text{if} \quad |i-j| \leqq h \leqq i+j. \end{aligned} \tag{3.11}$$

Equation (3.11) is a polynomial of degree $r_{ijh} = \max\{0, \min(i - j + h, j - i + h, j + i - h)\}$ in the variable $\theta = \lambda^{-\frac{1}{2}}$. Obviously $c_{ij0}(\lambda) = \delta_{ij}$. The limits of summation are over integer values of t in the interval

$$(i + j - h)/2 \leqq t \leqq \min\{i, j, i + j - h\}.$$

PROOF. We use (3.4), (3.3) and then (3.8) to establish

$$\begin{aligned}\Pi_i(x, \lambda)\Pi_j(x, \lambda) &= \sum_r \sum_s \binom{i}{r}\binom{j}{s}(-\lambda)^{i+j-r-s} \sum_t \binom{r}{t}\binom{s}{t}t!\, x^{(r+s-t)} \\ &= (-1)^{i+j} \sum_h \Pi_h(x, \lambda)\lambda^{-h} \sum_t \binom{i}{t}\binom{j}{t}t!\, \lambda^{i+j-t}B_{i-t,j-t}(h, t)\end{aligned}$$

with

$$\begin{aligned}B_{i-t,j-t}(h, t) &\equiv \sum_{r=t}^{i} \sum_{s=t}^{j} \binom{i-t}{r-t}\binom{j-t}{s-t}\binom{r+s-t}{h}(-1)^{r+s} \\ &= \sum_{r+s} \binom{i+j-2t}{r+s-2t}\binom{r+s-t}{h}(-1)^{r+s} \\ &= (-1)^{i+j}\binom{t}{h-i-j+2t}\end{aligned}$$

first using the Vandermonde convolution and then [6], page 11. Substituting φ for π through (3.5) yields (3.11). That $c_{ijh}(\lambda) = 0$ if $h < |i-j|$ follows from $0 \leqq 2t + h - i - j$ and $t \leqq \min(i, j)$. The limits of summation are determined from the values of t which make all choice functions positive. This completes the proof of Lemma 3.1.

LEMMA 3.2. *Let* $\mathscr{L}(N) = B(n, p)$ *and define* $\mu_j = E\varphi_j(N, np)$. *Then* $\mu_0 = 1$, $\mu_1 = 0$ *and for* $j \geqq 2$,

$$\mu_j = -(j-1)\left(\frac{p}{nj}\right)^{\frac{1}{2}} \mu_{j-1} - p\left(\frac{j-1}{j}\right)^{\frac{1}{2}} \mu_{j-2}\,. \tag{3.12}$$

PROOF. Using $EN^{(i)} = n^{(i)}p^i$, it is trivially shown that $E\pi_j(N, np) = p^j\pi_j(n, n)$. Now $\pi_j(n, n)$ satisfies the recurrence

$$\pi_j(n, n) = -(j-1)[\pi_{j-1}(n, n) + n\pi_{j-2}(n, n)] \tag{3.13}$$

from which (3.12) follows easily. To prove (3.13) note that the right-hand side may be written as $-(j-1) \sum \binom{j-2}{i-1}(-n)^{j-i-1}n^{(i)}$ using Pascal's formula. Then substitute

$$-(j-1)\binom{j-2}{i-1} = [(n-i) - n]\binom{j-1}{i}$$

to get

$$\sum \binom{j-1}{i}(-n)^{j-i}n^{(i)} + \sum \binom{j-1}{i}(-n)^{j-i-1}n^{(i+1)}\,.$$

The result follows from substituting $i - 1$ for i in the second term only, and using Pascal's formula again to recombine the two sums.

The author is grateful to the referee who suggested this substantial improvement of the original proof.

LEMMA 3.3. $c_{ijk}(\lambda)$ *defined in* (3.11) *satisfies for some appropriately large constant* $A > 0$ (*independent of* i, j, k):

$$0 \leqq c_{ijk}(\lambda) \leqq A^{i+j}\left\{1 + \left(\frac{i+j}{\lambda}\right)^{(i+j)/4-|k/2-(i+j)/4|}\right\} \tag{3.14}$$

and

$$0 \leqq c_{ijk}(\lambda) \leqq A^i\left\{1 + \left(\frac{i}{\lambda}\right)^{i/4}\right\} A^j\left\{1 + \left(\frac{j}{\lambda}\right)^{j/4}\right\}. \tag{3.15}$$

PROOF. Clearly $c_{ijk}(\lambda) \geqq 0$. Let $l \equiv \frac{1}{2}(i + j - k)$, $u \equiv \min(i, j, i + j - k)$.

If $k < |i-j|$, (3.14) holds and otherwise $l \leqq u$ and

$$c_{ijk}(\lambda) = \left(\frac{k!}{i!\,j!}\right)^{\frac{1}{2}} \sum_{t \geq l}^{u} \lambda^{-\frac{1}{2}(2t+k-i-j)} t! \binom{i}{t}\binom{j}{t}\binom{t}{2t+k-i-j} . \tag{3.16}$$

$$\frac{k!}{i!\,j!} = \frac{\binom{i+j}{k}}{\binom{i+j}{k}(2l)!} \leqq \frac{2^{i+j}}{(2l)!} \leqq 2^{i+j}\left(\frac{e}{2l}\right)^{2l} \tag{3.17}$$

using Stirling's inequality and interpreting $0^0 = 1$. We have

$$l \leqq u \leqq 2l \leqq i+j . \tag{3.18}$$

Making frequent use of inequalities like $\binom{i}{t} \leqq 2^i$ and of (3.18), (3.16) and (3.17) gives

$$\begin{aligned} c_{ijk}(\lambda) &\leqq A_1^{i+j}\left(\frac{\lambda}{l}\right)^l \sum_{t=l}^{u} \lambda^{-t} t! \leqq A_2^{i+j} \sum_{m=0}^{u-l} \lambda^{-m} m! \\ &\leqq A^{i+j}\left\{1 + \left(\frac{i+j}{\lambda}\right)^{u-l}\right\} \end{aligned} \tag{3.19}$$

for appropriate A_1, A_2, A.

The result (3.14) follows by noting

$$u - l \leqq \min\left\{\frac{k}{2}, l\right\} = \frac{i+j}{4} - \left|\frac{k}{2} - \frac{i+j}{4}\right| . \tag{3.20}$$

To prove (3.15), use (3.14) and note the result is trivial if $i + j \leqq \lambda$. If $i + j > \lambda$ then (3.14) gives for approximate A_1

$$\begin{aligned} c_{ijk}(\lambda) &\leqq A_1^{i+j}\left(\frac{i+j}{\lambda}\right)^{(i+j)/4} \\ &= A_1^{i+j}\left(\frac{i}{\lambda}\right)^{i/4}\left(\frac{j}{\lambda}\right)^{j/4} \cdot \left(\frac{i+j}{i}\right)^{i/4}\left(\frac{i+j}{j}\right)^{j/4} . \end{aligned} \tag{3.21}$$

The function $f(p) = p^p(1-p)^{1-p}$ is minimized for $0 \leqq p \leqq 1$ at $p = \frac{1}{2}$ and $f(\frac{1}{2}) = \frac{1}{2}$. Hence the last two terms of (3.21) are bounded by $2^{(i+j)/4}$ (letting $p = i/(i+1)$), and so (3.21) yields (3.15). The proof is complete.

LEMMA 3.4. *Let φ_j be defined as in* (3.5) *and $\mathscr{L}(N) = B(n, p)$, $\lambda \equiv np$. Suppose $\alpha > 0$ satisfies $|(\lambda - \alpha)/(\lambda p)^{\frac{1}{2}}| \leqq B$ for a given constant $B < \infty$. Then there exists a constant A (depending only on B) such that*

$$E\varphi_j^2(N, \alpha) \leqq A^j\left\{1 + \left(\frac{j}{n}\right)^j\right\} . \tag{3.22}$$

PROOF. Using Lemma 3.2, we first prove by induction that

$$|\mu_j| \equiv |E\varphi_j(N, \lambda)| \leqq (4p)^{j/2}\{1 + (j/n)^{j/2}\} . \tag{3.23}$$

Note that (3.23) holds for $j = 0, 1$. (3.12) yields for $j \geqq 2$,

$$|\mu_j| \leqq \left(\frac{jp}{n}\right)^{\frac{1}{2}} |\mu_{j-1}| + p|\mu_{j-2}| . \tag{3.24}$$

The proof proceeds by induction, taking two cases. For the case $j \leqq n$, assume $|\mu_i| \leqq (4p)^{i/2}$ for all $i \leqq j-1$ and use (3.24) to prove $|\mu_j| < (4p)^{j/2}$. When $j > n$, assume $|\mu_i| \leqq (4pi/n)^{i/2}$ for $i = j-1$ and $j-2$ and show this holds for $i = j$ by using (3.24).

We now prove if $|(\lambda - \alpha)/(\lambda p)^{\frac{1}{2}}| \leqq B < \infty$ (we take $B \geqq 1$), then

$$|E\varphi_j(N, \alpha)| \leqq (Cp)^{j/2}\left\{1 + \left(\frac{j}{n}\right)^{j/2}\right\} \tag{3.25}$$

for appropriately large C (depending only on B).

Using (3.10), expand $\varphi_j(N, \alpha)$ in its Taylor series about $\alpha = \lambda$ to get

$$\varphi_j(N, \alpha) = \sum_{i=0}^{j} \binom{j}{i}^{\frac{1}{2}} \frac{(p^{\frac{1}{2}}\theta)^i}{(i!)^{\frac{1}{2}}} \varphi_{j-i}(N, \lambda), \tag{3.26}$$

where $\theta \equiv (\lambda - \alpha)/(\lambda p)^{\frac{1}{2}}$. Hence, using (3.23) and $|\theta| \leqq B$,

$$\begin{aligned} |E\varphi_j(N, \alpha)| &\leqq \sum_{i=0}^{j} \binom{j}{i}^{\frac{1}{2}} \frac{|\theta|^i p^{i/2}}{(i!)^{\frac{1}{2}}} (4p)^{(j-i)/2}\left\{1 + \left(\frac{j-i}{n}\right)^{(j-i)/2}\right\} \\ &\leqq (8pB^2)^{j/2} \sum_{i=0}^{j} \left\{1 + \left(\frac{j}{n}\right)^{(j-i)/2}\right\} \end{aligned}$$

which yields (3.25) for an appropriately large C.

To establish (3.22), we use Lemma 3.1, Lemma 3.3, (3.14), and then (3.25). Hence, for sufficiently large D given by Lemma 3.3, and assuming $C \geqq 1$,

$$\begin{aligned} E\varphi_j^2(N, \alpha) &\leqq \sum_{k=0}^{2j} c_{jjk}(\lambda)|E\varphi_k(N, \alpha)| \\ &\leqq D^j C^{j/2} \sum_{k=0}^{j} \left\{1 + \left(\frac{2j}{np}\right)^{k/2}\right\}\left\{1 + \left(\frac{k}{n}\right)^{k/2}\right\} p^{k/2} \\ &\quad + D^j C^{j/2} \sum_{k=j+1}^{2j} \left\{1 + \left(\frac{2j}{np}\right)^{j-k/2}\right\}\left\{1 + \left(\frac{k}{n}\right)^{k/2}\right\} p^{k/2}. \end{aligned}$$

Note that the preceding bound is increased by setting $p = 1$. The result (3.22) then follows from simple inequalities and choosing A large enough.

LEMMA 3.5. *Let $\mathscr{L}(N) = \mathscr{M}(p, k, n)$ and $r \geqq 0$ be an integer. Then there exists a constant $A < \infty$ such that*

$$EN_i^{(r)}N_j^{(r)} \leqq r!\,(Anp_i)^r\delta_{ij} + (An^2p_ip_j)^r. \tag{3.27}$$

PROOF. If $i \neq j$, then $EN_i^{(r)}N_j^{(r)} = n^{(2r)}p_i^rp_j^r$ which satisfies (3.27). If $i = j$, use (3.3) to get

$$\begin{aligned} EN_i^{(r)2} &= \sum_k \tbinom{r}{k}^2 k!\, EN_i^{(2r-k)} \\ &= \sum_k \tbinom{r}{k}^2 k!\, n^{(2r-k)}p_i^{2r-k} \\ &\leqq 2^{2r}(np_i)^{2r} \sum_{k=0}^{r} \frac{k!}{(np_i)^k} \\ &\leqq A^r(np_i)^{2r}\left\{1 + \frac{r!}{(np_i)^r}\right\} \end{aligned}$$

if A is large enough. This is equivalent to (3.27).

4. General central limit theorems for multinomial sums.

THEOREM 4.1. *Basic central limit theorem for multinomial sums.*

Let $\mathscr{L}(N_k) = \mathscr{M}(p_k, k, n_k)$. *Suppose as* $k \to \infty$ *that* $n_k \to \infty$ *and* $\max_{1 \leq i \leq k} p_{ik} = o(1)$.

Let $\{f_{ik}: 1 \leq i \leq k\}$ *be real-valued functions with domain the nonnegative integers, and express* f_{ik} *as*

$$f_{ik}(x) = \sum_{j=0}^{\infty} \alpha_{ijk} \varphi_j(x, \lambda_{ik})$$

where $\lambda_{ik} \equiv n_k p_{ik}$. *Here,* $\alpha_{ijk} = Ef_{ik}(X_{ik})\varphi_j(X_{ik}, \lambda_{ik})$, $\mathscr{L}(X_{ik})$ *being Poisson* (λ_{ik}) *and* $\{X_{ik}: 1 \leq i \leq k\}$ *are independent. The* $\{\varphi_j\}$ *are the Poisson–Charlier polynomials of* (3.5). *Assume* $Ef_{ik}(X_{ik}) = 0$, *i.e.,* $\alpha_{i0k} = 0$, *and* $\sum_i \mathrm{Cov}\,(f_{ik}(X_{ik}), X_{ik}) = 0$, *i.e.,* $\sum_i (p_{ik})^{\frac{1}{2}} \alpha_{i1k} = 0$.

Define

$$\sigma_{ik}^2 = \mathrm{Var}\, f_{ik}(X_{ik}) = \sum_{j=1}^{\infty} \alpha_{ijk}^2 \quad \textit{and} \quad s_k^2 = \sum_{i=1}^{k} \sigma_{ik}^2 . \tag{4.1}$$

Assume that as $k \to \infty$, *the* uan *condition*

$$\frac{1}{s_k^2} \max_{1 \leq i \leq k} \sigma_{ik}^2 = o(1) \tag{4.2}$$

holds, and assume that for every fixed A, $0 < A < \infty$

$$\delta_k(A) \equiv \frac{1}{s_k^2} \sum_{i=1}^{k} \sum_{j=1}^{\infty} \alpha_{ijk}^2 A^j \left\{1 + \left(\frac{j}{n_k}\right)^j\right\} = O(1) \tag{4.3}$$

is bounded as $k \to \infty$. *This will be called "condition-*δ*."*

Suppose

$$\mathscr{L}\left(\frac{1}{s_k} \sum_{i=1}^{k} f_{ik}(X_{ik})\right) \to \mathscr{N}(0, 1) \quad \textit{as} \quad k \to \infty . \tag{4.4}$$

Note this is a sum of independent random variables, appropriately scaled and centered.

Then

$$\mathscr{L}\left(\frac{1}{s_k} \sum_{i=1}^{k} f_{ik}(N_{ik})\right) \to \mathscr{N}(0, 1) \quad \textit{as} \quad k \to \infty . \tag{4.5}$$

PROOF. It is only necessary to verify condition (2.9) of Lemma 2.2. The notation of Lemma 2.2 will be used. Denote $l_k = n_k + v_k n_k^{\frac{1}{2}}$, $m_k = hn_k^{\frac{1}{2}}$, and assume $h \leq 1$.

$$\frac{1}{s_k^2} E(\sum_{i=1}^{k} \{f_{ik}(L_{ik} + M_{ik}) - f_{ik}(L_{ik})\})^2$$

$$= \frac{1}{s_k^2} E(\sum_{i=1}^{k} \sum_{j=1}^{\infty} \alpha_{ijk}\{\varphi_j(L_{ik} + M_{ik}, \lambda_{ik}) - \varphi_j(L_{ik}, \lambda_{ik})\})^2$$

$$= \frac{1}{s_k^2} E\left(\sum_{i=1}^{k} \sum_{j=1}^{\infty} \alpha_{ijk} \sum_{s=1}^{j} \binom{j}{s}^{\frac{1}{2}} \frac{\lambda_{ik}^{-s/2}}{(s!)^{\frac{1}{2}}} M_{ik}^{(s)} \varphi_{j-s}(L_{ik}, \lambda_{ik})\right)^2$$

using (3.9),

$$= \frac{1}{s_k^2} E\left(\sum_{s=1}^{\infty} \sum_{i=1}^{k} \frac{\lambda_{ik}^{-s/2} M_{ik}^{(s)}}{(s!)^{\frac{1}{2}}} B_{iks}\right)^2 \tag{4.6}$$

where we define

$$B_{iks} \equiv \sum_{j=s}^{\infty} \alpha_{ijk}\binom{j}{s}^{\frac{1}{2}}\varphi_{j-s}(L_{ik}, \lambda_{ik}) .$$

Note that from the Schwarz inequality

$$\begin{aligned} EB_{iks}^2 &\leqq (1/2^{s-1}) \sum_{j=s}^{\infty} \alpha_{ijk}^2\binom{j}{s}2^j E\varphi_{j-s}^2(L_{ik}, \lambda_{ik}) \\ &\leqq \sum_{j=s}^{\infty} \alpha_{ijk}^2 4^j A_1^{j-s}\left\{1 + \left(\frac{j-s}{l_k}\right)^{j-s}\right\} , \end{aligned}$$

A_1 being guaranteed by Lemma 3.4 since the hypothesis of that lemma

$$\left|\frac{l_k p_{ik} - n_k p_{ik}}{l_k^{\frac{1}{2}} p_{ik}}\right| = |v_k| \left(\frac{n_k}{n_k + v_k n_k^{\frac{1}{2}}}\right)^{\frac{1}{2}} = O(1) \tag{4.7}$$

holds, $|v_k|$ being bounded by assumption.

Define

$$\varepsilon_{ik}(A) \equiv \sum_{j=1}^{\infty} \alpha_{ijk}^2 A^j \left\{1 + \left(\frac{j}{n_k}\right)^j\right\} .$$

Since $l_k/n_k = 1 + o(1)$ as $k \to \infty$, there exists a sufficiently large A_2 such that

$$EB_{iks}^2 \leqq \varepsilon_{ik}(A_2) . \tag{4.8}$$

Then (4.6) is dominated by

$$\begin{aligned} &\frac{1}{s_k^2} \sum_{s=1}^{\infty} \frac{2^s}{s!} E\left(\sum_{i=1}^k \lambda_{ik}^{-s/2} M_{ik}^{(s)} B_{iks}\right)^2 \\ &\quad = \frac{1}{s_k^2} \sum_{s=1}^{\infty} \frac{2^s}{s!} \sum_{i=1}^k \sum_{t=1}^k \lambda_{ik}^{-s/2}\lambda_{tk}^{-s/2} EM_{ik}^{(s)} M_{tk}^{(s)} EB_{iks} B_{tks} \\ &\quad \leqq \frac{1}{s_k^2} \sum_{s=1}^{\infty} \frac{2^s}{s!} \sum_{i=1}^k \sum_{t=1}^k \lambda_{ik}^{-s/2}\lambda_{tk}^{-s/2} (A_3 m_k^2 p_{ik} p_{jk})^s E^{\frac{1}{2}} B_{iks}^2 E^{\frac{1}{2}} B_{tks}^2 \\ &\qquad + \frac{1}{s_k^2} \sum_{s=1}^{\infty} 2^s \sum_{i=1}^k \lambda_{ik}^{-s} (A_3 m_k p_{ik})^s EB_{iks}^2 \end{aligned}$$

using A_3 guaranteed by Lemma 3.5,

$$\begin{aligned} &= \frac{1}{s_k^2} \sum_{s=1}^{\infty} \frac{2^s}{s!} \left(\sum_{i=1}^k A_3^{s/2} \left(\frac{m_k^2}{n_k}\right)^{s/2} p_{ik}^{s/2} E^{\frac{1}{2}} B_{iks}^2\right)^2 \\ &\quad + \frac{1}{s_k^2} \sum_{s=1}^{\infty} 2^s \sum_{i=1}^k \left(\frac{A_3 m_k}{n_k}\right)^s EB_{iks}^2 \\ &\leqq \frac{1}{s_k^2} \sum_{s=1}^{\infty} \frac{1}{s!} \left(\frac{2A_3 m_k^2}{n_k}\right)^s \sum_{i=1}^k p_{ik}^s \sum_{i=1}^k EB_{iks}^2 \\ &\quad + \frac{1}{s_k^2} \sum_{s=1}^{\infty} \left(\frac{2A_3 m_k}{n_k}\right)^s \sum_{i=1}^k \varepsilon_{ik}(A_2) . \end{aligned} \tag{4.9}$$

Since

$$\frac{1}{s_k^2} \sum_{i=1}^k \varepsilon_{ik}(A_2) = \delta_k(A_2) , \qquad m_k^2 = h^2 n_k ,$$

and

$$2A_3 m_k/n_k = h \cdot (2A_3/n_k^{\frac{1}{2}}) < h/2$$

for sufficiently large k, (4.9) is dominated for sufficiently large k by

$$h^2 \exp(2A_3)\delta_k(A_2) + h\delta_k(A_2) \sum_{s=1}^{\infty} \frac{1}{2^s} = o(1) \quad \text{as} \quad h \to 0 . \tag{4.10}$$

The proof of the theorem is complete.

We now give a condition, called "condition-η," which guarantees both condition-δ and the asymptotic normality of

$$\frac{1}{s_k} \sum_{i=1}^{k} f_{ik}(X_{ik})$$

of Theorem 4.1.

THEOREM 4.2. *With the notation of Theorem* 4.1, *suppose that as* $k \to \infty$

$$n_k \to \infty , \qquad \max_{1 \le i \le k} p_{ik} = o(1) \tag{4.11}$$

and that the uan *condition*

$$\frac{1}{s_k^2} \max_{1 \le i \le k} \sigma_{ik}^2 = o(1)$$

holds.

Then

$$\mathscr{L}\left(\frac{1}{s_k} \sum_{i=1}^{k} f_{ik}(N_{ik})\right) \to \mathscr{N}(0, 1)$$

provided the following condition, condition-η, holds.

For every fixed $A < \infty$,

$$\eta_k(A) \equiv \frac{1}{s_k^2} \sum_{i=1}^{k} \sum_{j=1}^{\infty} \alpha_{ijk}^2 A^j \left\{1 + \left(\frac{j}{n_k p_{ik}}\right)^j\right\} = O(1) \tag{4.12}$$

as $k \to \infty$.

PROOF. Clearly condition-η implies condition-δ. It therefore remains to show that (4.4) of Theorem 4.1 holds. We check the Liapounov condition, that with $\mathscr{L}(X_{ik}) = \text{Poisson}(\lambda_{ik})$, $\lambda_{ik} \equiv n_k p_{ik}$,

$$\frac{1}{s_k^4} \sum_{i=1}^{k} Ef_{ik}^4(X_{ik}) = o(1) . \tag{4.13}$$

$$\begin{aligned} Ef_{ik}^4(X_{ik}) &= \sigma_{ik}^4 + \operatorname{Var}\left(\sum_{j=1}^{\infty} \sum_{l=1}^{\infty} \alpha_{ijk}\alpha_{ilk}\varphi_j(X_{ik}, \lambda_{ik})\varphi_l(X_{ik}, \lambda_{ik})\right) \\ &= \sigma_{ik}^4 + \operatorname{Var}\left(\sum_j \sum_l \alpha_{ijk}\alpha_{ilk} \sum_{r=0}^{j+l} c_{jlr}(\lambda_{ik})\varphi_r(X_{ik}, \lambda_{ik})\right) \end{aligned}$$

using Lemma 3.1

$$\begin{aligned} &= \sigma_{ik}^4 + \operatorname{Var}\left(\sum_{r=0}^{\infty} d_{ikr}\varphi_r(X_{ik}, \lambda_{ik})\right) \\ &= \sigma_{ik}^4 + \sum_{r=1}^{\infty} d_{ikr}^2 \end{aligned}$$

where we have defined $d_{ikr} = \sum_j \sum_l \alpha_{ijk}\alpha_{ilk}c_{jlr}(\lambda_{ik})$, the region of summation being $j, l \geqq 1, j + l \geqq r$.

From (3.15), $|d_{ikr}| \leqq \sum_j \sum_l q_{ijk}^{(A)} q_{ilk}^{(A)}$ with $q_{ijk}^{(A)} \equiv |\alpha_{ijk}|A^j\{1 + (j/\lambda_{ik})^{j/4}\}$. Hence

$$\begin{aligned} \sum_{r=1}^{\infty} d_{ikr}^2 &\leqq \left(\sum_j q_{ijk}^{(A)}\right)^2 \sum_j \sum_l \sum_{r=1}^{j+l} q_{ijk}^{(A)} q_{ilk}^{(A)} \\ &\leqq \left(\sum_{j=1}^{\infty} q_{ijk}(B)\right)^4 \leqq \left(\sum_{j=1}^{\infty} q_{ijk}^2(B2^{\frac{1}{2}})\right)^2 \end{aligned}$$

using the Schwarz inequality and with B enough larger than A to account for $\sum_{r=1}^{j+l} 1 = j + l$. The left side of (4.13) is therefore bounded by

$$\frac{1}{s_k^4} \sum_i \sigma_{ik}^4 + \frac{1}{s_k^4} \sum_i (\sum_j \alpha_{ijk}^2)(\sum_j q_{ijk}^4(B2^{\frac{1}{2}})/\alpha_{ijk}^2)$$

$$\leqq \frac{\max \sigma_{ik}^2}{s_k^2} + \frac{4}{s_k^4} \sum_i \sigma_{ik}^2 \sum_j \alpha_{ijk}^2 (4B^4)^j \{1 + (j/\lambda_{ik})\}^j \tag{4.14}$$

$$\leqq \frac{\max \sigma_{ik}^2}{s_k^2} \{1 + \eta_k(4B^4)\} = o(1) \qquad \text{as} \quad k \to \infty .$$

This completes the proof.

COROLLARY 4.1. *If, with the notation of Theorem* 4.1, *the* $\{f_{ik}\}$ *are polynomials of degree at most* r *as* $k \to \infty$ *(*r *independent of* k*), if there exists* $\varepsilon > 0$ *such that* $n_k p_{ik} \geqq \varepsilon$ *for all* i, k, *and if* $\max_{1 \leqq i \leqq k} p_{ik} = o(1)$ *as* $k \to \infty$, *then the* uan *condition*

$$\frac{1}{s_k^2} \max_{1 \leqq i \leqq k} \sigma_{ik}^2 = o(1) \qquad \text{as} \quad k \to \infty \tag{4.15}$$

is sufficient to guarantee that

$$\mathscr{L}((1/s_k) \textstyle\sum_{i=1}^k f_{ik}(N_{ik})) \to \mathscr{N}(0, 1) \qquad \text{as} \quad k \to \infty .$$

PROOF. We have $n_k \to \infty$ since $\eta_k p_{ik} \geqq \varepsilon$, and from (4.12), assuming $A \geqq 1$,

$$\eta_k(A) \leqq A^r \left\{2 + \left(\frac{r}{\varepsilon}\right)^r\right\} = O(1) \qquad \text{as} \quad k \to \infty .$$

The hypothesis of Theorem 4.2 therefore holds, thus establishing this corollary.

It may be useful to note that the functions f_{ik} are always equivalent to polynomials of degree n so that they may be expressed as a finite sum

$$f_{ikn}(x) = \textstyle\sum_{j=0}^n \alpha_{ijkn}(\lambda_{ik}) \varphi_j(x, \lambda_{ik}) . \tag{4.16}$$

This follows from the fact that every function $f(x)$ is equal to the polynomial ($\Delta f(x) \equiv f(x+1) - f(x)$, $\Delta^r f(x) \equiv \Delta(\Delta^{r-1} f(x))$ if $r \geqq 2$).

$$f_n(x) \equiv \textstyle\sum_{r=0}^n \Delta^r f(0) x^{(r)}/r! \tag{4.17}$$

at the values $x = 0, 1, \cdots, n$ ([6], page 201).

The Poisson–Charlier coefficients α_{ijkn} of (4.16) are calculated as

$$\alpha_{ijkn}(\lambda) = \lambda^{j/2} (j!)^{-\frac{1}{2}} \textstyle\sum_{s=0}^{n-j} \{\Delta^{j+s} f_{ik}(0)\} \lambda^s / s! , \tag{4.18}$$

if $j \leqq n$ and $\alpha_{ijkn}(\lambda) = 0$ if $j > n$. To prove this, we suppress the i, k subscripts. Then if $j \leqq n$, from (4.17) we have

$$\alpha_{jn}(\lambda) = E \textstyle\sum_{r=0}^n \Delta^r f(0) X^{(r)} \varphi_j(X, \lambda)/r!$$

$$= \frac{1}{(j!)^{\frac{1}{2}} \lambda^{j/2}} \sum_{r=0}^n \frac{\Delta^r f(0)}{r!} \sum_l \binom{r}{l} \lambda^{r-l} E \Pi_l(X, \lambda) \Pi_j(X, \lambda)$$

using (3.8)

$$= \frac{1}{(j!)^{\frac{1}{2}}\lambda^{j/2}} \sum_{r=0}^{n} \frac{\Delta^r f(0)}{r!} \sum_l \binom{r}{l} \lambda^{r-l} j!\, \lambda^j \delta_{l,j}$$

$$= (j!)^{\frac{1}{2}}\lambda^{j/2} \sum_{r=j}^{n} \frac{\Delta^r f(0)}{r!} \binom{r}{j} \lambda^{r-j} ,$$

which, upon substituting $r = s + j$, is (4.18). Clearly $\alpha_{jn}(\lambda) = 0$ if $j > n$ since f_n is a polynomial of degree at most n. This completes the proof of (4.18).

If $\mathscr{L}(N_k) = \mathscr{M}(p_k, k, n_k)$, then

$$\sum_{i=1}^{k} f_{ik}(N_{ik}) = \sum_{i=1}^{k} f_{ikn}(N_{ik})$$

with f_{ikn} as in (4.16) and $\alpha_{ijkn}(\lambda_{ik})$ as in (4.18) since N_{ik} is an integer between 0 and n. When this representation is used, we have $\alpha_{ijk}(= \alpha_{ijkn}(\lambda_{ik})) = 0$ if $j > n$ and therefore condition-δ (4.3) simplifies to a "condition-δ^*," this condition being that for every $0 < A < \infty$,

(4.19) $$\delta_k^*(A) \equiv (1/s_k^2) \sum_{i=1}^{k} \sum_{j=1}^{n} \alpha_{ijk}^2 A^j = O(1) \quad \text{as} \quad k \to \infty .$$

Furthermore, the $\alpha_{ijkn}(\lambda_{ik})$ given in (4.18) always exist, but the α_{ijk} of Theorem 4.1 may not exist if f_{ik} does not have sufficiently many Poisson moments. (Of course this latter condition can always be corrected by altering the definition of $f_{ik}(x)$ for $x > n$.)

5. Central limit theorems for Pearson's chi-square statistic and the likelihood ratio statistic. Let $\mathscr{L}(N_k) = \mathscr{M}(p_k, k, n_k)$ and suppose $p_k{}^0 = (p_{1k}^0, \cdots, p_{kk}^0)$ is a specified vector. Of the many statistics that have been proposed for testing the simple null hypothesis

$$H_0 : p_{ik} = p_{ik}^0 \qquad \text{for every} \quad i = 1, \cdots, k$$

against all possible alternatives, Pearson's chi-square test in [5] and the likelihood ratio test in [4] are the best known and most often used.

Pearson's test is to reject H_0 if and only if

(5.1) $$\sum_{i=1}^{k} (N_{ik} - n_k p_{ik}^0)^2 / n_k p_{ik}^0 \geqq c_1 .$$

The likelihood ratio test is to reject H_0 if and only if

(5.2) $$2 \sum_{i=1}^{k} N_{ik} \log\left(\frac{N_{ik}}{n_k p_{ik}^0}\right) \geqq c_2 .$$

c_1 and c_2 are determined by the H_0 distribution of their statistics and the desired level of significance.

The asymptotic distributions of the statistics in (5.1) and (5.2) are well known to be the chi-square distribution under H_0 and the noncentral chi-square distribution under the alternative hypothesis, provided k is fixed, $n_k p_{ik}^0 \to \infty$, and $n_k p_{ik} \to \infty$. The distribution of (5.1) as $k \to \infty$ has been considered by Tumanyan [9] and Steck [7]. The distribution of (5.2) as $k \to \infty$ has not been considered before, although an example of Stein's [1] assumed the asymptotic normality of (5.2) in considering the power properties of the likelihood ratio test as $k \to \infty$. Here, we will determine the distribution of both statistics as $k \to \infty$, beginning with Pearson's test statistic (5.1).

These distributional results for the tests (5.1) and (5.2) provide a basis for efficiency comparisons. In the examples of Stein [1] and of the author's [3] it is shown that Pearson's chi-square test is the uniformly best test in a class which contains the likelihood ratio test, assuming near alternatives and $k \to \infty$. Unlike the familiar case with k fixed, (5.1) and (5.2) are not equivalent tests under these circumstances if n_k/k is moderate, so that (5.1) is superior. The author in unpublished work has obtained numerical results which show how much Pearson's chi-square test dominates the likelihood ratio test.

THEOREM 5.1. *Asymptotic normality of Pearson's chi-square statistic when* $k \to \infty$.

Let $\mathscr{L}(N_k) = \mathscr{M}(p_k, k, n_k)$. *Let* $\{p^0_{ik}: 1 \leqq i \leqq k\}$ *be given with* $p^0_{ik} > 0$, $\sum_i p^0_{ik} = 1$. *Suppose*

$$\max_{1\leq i\leq k} p_{ik} = o(1) \quad \text{as} \quad k \to \infty \tag{5.3}$$

and that there exists $\varepsilon > 0$ *such that*

$$n_k p_{ik} \geqq \varepsilon \quad \text{for all} \quad i, \quad k. \tag{5.4}$$

Denote

$$\mu_k \equiv \sum_i \frac{p_{ik}}{p^0_{ik}} + n_k \sum_i \frac{(p_{ik} - p^0_{ik})^2}{p^0_{ik}}, \tag{5.5}$$

$$\gamma_k \equiv \sum_i \left(\frac{1}{n_k p^0_{ik}} + 2\frac{p_{ik}}{p^0_{ik}}\right) p_{ik}, \tag{5.6}$$

$$\sigma^2_{ik} \equiv 2\frac{p^2_{ik}}{p^{02}_{ik}} + n_k \left(\frac{1}{n_k p^0_{ik}} + 2\frac{p_{ik}}{p^0_{ik}} - \gamma_k\right)^2 p_{ik} \tag{5.7}$$

and

$$s_k^{\,2} = \sum_i \sigma^2_{ik}. \tag{5.8}$$

Suppose the uan *condition,*

$$\frac{\max_{1\leq i\leq k} \sigma^2_{ik}}{s_k^{\,2}} = o(1) \quad \text{as} \quad k \to \infty \tag{5.9}$$

holds.

Then

$$\mathscr{L}\left(\frac{1}{s_k}\left\{\sum_i \frac{(N_{ik} - n_k p^0_{ik})^2}{n_k p^0_{ik}} - \mu_k\right\}\right) \to \mathscr{N}(0, 1) \quad \text{as} \quad k \to \infty \tag{5.10}$$

Define

$$\alpha_k \equiv \sum \left(\frac{p_{ik} - p^0_{ik}}{p^0_{ik}}\right) p_{ik} = \sum \frac{(p_{ik} - p^0_{ik})^2}{p^0_{ik}}$$

and

$$\theta_{ik} \equiv \frac{p_{ik} - p^0_{ik}}{p^0_{ik}} - \alpha_k.$$

Then s_k^2 *is asymptotically of the exact order of*

$$k + k\alpha_k^2 + n_k \textstyle\sum_{i=1}^k \theta_{ik}^2 p_{ik} \,, \tag{5.11}$$

and the uan *condition* (5.9) *is equivalent to the condition that*

$$\frac{\max_{1 \leq i \leq k} n_k \theta_{ik}^2 p_{ik}}{k + k\alpha_k^2 + n_k \sum \theta_{ik}^2 p_{ik}} = o(1) \qquad as \quad k \to \infty \,. \tag{5.12}$$

When the "null hypothesis" $p_{ik} = p_{ik}^0$ *for every* i, *condition* (5.12) *is trivially met and so* (5.10) *holds provided only that* (5.3) *and* (5.4) *are valid.*

PROOF. With Π_j defined as in (3.4) and with $\lambda_{ik} \equiv n_k p_{ik}$, and using $\sum_i \Pi_1(N_{ik}, \lambda_{ik}) = 0$, it follows that

$$\textstyle\sum_i \dfrac{(N_{ik} - n_k p_{ik}^0)^2}{n_k p_{ik}^0} - \mu_k = \sum_i f_{ik}(N_{ik}) \tag{5.13}$$

with

$$f_{ik}(N_{ik}) \equiv \frac{1}{n_k p_{ik}^0} \Pi_2(N_{ik}, \lambda_{ik}) + \left(\frac{1}{n_k p_{ik}^0} + 2\frac{p_{ik}}{p_{ik}^0} - \gamma_k\right) \Pi_1(N_{ik}, \lambda_{ik}) \,. \tag{5.14}$$

Let $\{X_{ik}, \cdots, X_{kk}\}$ be independent, $\mathscr{L}(X_{ik}) = \text{Poisson}\,(\lambda_{ik})$. We have $Ef_{ik}(X_{ik}) = 0$, $\sum_i \text{Cov}\,(X_{ik}, f_{ik}(X_{ik})) = 0$ and $\sigma_{ik}^2 = \text{Var}\, f_{ik}(X_{ik})$, using (3.5)–(3.7). The other conditions of Corollary 4.1 are met, and since the $\{f_{ik}\}$ are polynomials of the second degree,

$$\mathscr{L}\left(\frac{1}{s_k} \sum f_{ik}(N_{ik})\right) \to \mathscr{N}(0, 1) \qquad \text{as} \quad k \to \infty \tag{5.15}$$

provided only that the uan condition (4.15) holds. This is assumed in (5.9), and so (5.10) follows from (5.13) and (5.15).

We now establish the asymptotic equivalence of (5.8) and (5.11). The symbol O_e will be used to denote "exact order of" as $k \to \infty$. That is, $a_k = O_e(b_k)$ if and only if $a_k = O(b_k)$ and $b_k = O(a_k)$ as $k \to \infty$.

Define

$$D_k^2 = \sum \frac{p_{ik}^2}{p_{ik}^{02}} \,, \qquad \beta_{ik} \equiv \frac{1}{n_k p_{ik}^0} - \sum \frac{p_{ik}}{n_k p_{ik}^0} \,, \qquad B_k^2 \equiv n_k \sum \beta_{ik}^2 p_{ik}$$

and

$$T_k^2 = n_k \sum \theta_{ik}^2 p_{ik} \,.$$

Then

$$\begin{aligned} s_k^2 &= 2D_k^2 + B_k^2 + 4T_k^2 + 4n_k \textstyle\sum \beta_{ik}\theta_{ik} p_{ik} \\ &= 2D_k^2 + B_k^2 + 4T_k^2 + O(B_k T_k) \end{aligned} \tag{5.16}$$

using the Schwarz inequality. But

$$B_k^2 \leqq \frac{2}{n_k} \sum \frac{p_{ik}}{p_{ik}^{02}} + \frac{2}{n_k}\left(\sum \frac{p_{ik}}{p_{ik}^0}\right)^2 \leqq \frac{2}{\varepsilon} \sum \frac{p_{ik}^2}{p_{ik}^{02}} + \frac{2k}{n_k} \sum \frac{p_{ik}^2}{p_{ik}^{02}} = O(D_k^2) \,. \tag{5.17}$$

Using (5.16), (5.17) and that $s_k^2 \geqq D_k^2$ always, from (5.7),

$$s_k^2 = O_e(D_k^2 + T_k^2) \,. \tag{5.18}$$

Note $D_k^2 = \sum (1 + \alpha_k + \theta_{ik})^2 = O(k + k\alpha_k^2 + \sum \theta_{ik}^2)$. Since $\sum \theta_{ik}^2 \leqq n_k/\varepsilon \times \sum p_{ik}\theta_{ik}^2 = O(T_k^2)$, then $D_k^2 + T_k^2 = O(k + k\alpha_k^2 + T_k^2)$. This must be the exact order, for if not, both T_k^2 and D_k^2 are $o(k + k\alpha_k^2)$ implying both $\sum \theta_{ik}^2$ and D_k^2 are $o(k + k\alpha_k^2)$, which contradicts the fact that $D_k^2 = \sum (1 + \alpha_k + \theta_{ik})^2$. With (5.18), $s_k^2 = O_e(k + k\alpha_k^2 + T_k^2)$. Thus (5.11) is established.

To establish the equivalence of (5.9) and (5.12), note that

$$\max_{1\leqq i\leqq k} \sigma_{ik}^2 = \max_i \{2(1 + \alpha_k + \theta_{ik})^2 + n_k p_{ik}(\beta_{ik} + 2\theta_{ik})^2\}$$

will vanish relative to (5.11) if (5.12) holds and if $n_k p_{ik}\beta_{ik}^2$ vanishes relative to (5.11). But

$$\begin{aligned}\frac{1}{s_k^2} n_k \beta_{ik}^2 p_{ik} &\leqq \frac{2}{n_k}\left\{\frac{p_{ik}}{p_{ik}^{02}} + \left(\sum \frac{p_{ik}}{p_{ik}^0}\right)^2 p_{ik}\right\} \leqq \frac{2}{\varepsilon}\left(\frac{p_{ik}}{p_{ik}^0}\right)^2 + 2p_{ik}\frac{k}{n_k}\sum \frac{p_{ik}^2}{p_{ik}^{02}} \\ &= \frac{1}{s_k^2} O((1 + \alpha_k + \theta_{ik})^2) + p_{ik} O(D_k^2))/s_k^2 \\ &= \frac{1}{s_k^2} O(1 + \alpha_k^2) + \frac{1}{s_k^2} O(n_k p_{ik}\theta_{ik}^2) + \frac{1}{s_k^2} p_{ik} O(D_k^2) = o(1)\end{aligned}$$

from (5.11), (5.3) and (5.18). Therefore $\max_i \sigma_{ik}^2 = o(s_k^2) + \max_i (\mathrm{np}_{ik}\theta_{ik}^2)$, giving the result. The proof of this theorem is complete.

Discussion. The conditions (5.9) or (5.12) are probably the weakest possible for asymptotic normality of the chi-square statistic when $n_k p_{ik} \geqq \varepsilon$ and $\max_{1\leqq i\leqq k} p_{ik} = o(1)$. Steck's conditions [7], that as $k \to \infty$,

$$p_{ik}\,, \qquad p_{ik}^0 \leqq \frac{c}{k}\,, \qquad \frac{k^2}{n_k} \to \infty\,, \qquad n_k(p_{ik} - p_{ik}^0)^2 \to 0\,,$$

and that

$$(\min_{1\leqq i\leqq k} p_{ik}^0) \sum_{i=1}^k \frac{p_{ik}}{p_{ik}^0} \geqq \varepsilon \geqq 0\,,$$

are weaker than those of Theorem 5.1 only when (5.4) fails, and otherwise easily imply (5.9). When (5.4) holds, our condition for the null distribution ($p_{ik} = p_{ik}^0$) of the chi-square statistic is considerably weaker than Steck's condition which requires constants c_1 and c_2 such that $c_1 \leqq kp_{ik}^0 \leqq c_2$. Our conditions are also much more general for the alternative hypothesis.

The likelihood ratio statistic (LRS) (5.2) will now be considered. Theorem 4.1 can be applied, but condition-δ of that theorem holds for the LRS if and only if $\max_{1\leqq i\leqq k} n_k p_{ik}$ is bounded. Instead, we will use Lemma 2.2 directly to get a general central limit theorem for the LRS, not requiring $n_k p_{ik}$ to be bounded. As with the chi-square statistic, the essential requirement of the ensuing theorem is that p_{ik} and p_{ik}^0 be sufficiently close to one another for the uan condition to hold. When $n_k p_{ik} \geqq \varepsilon$ and $\max p_{ik} = o(1)$, the condition of Theorem 5.2 is probably necessary as well as sufficient.

THEOREM 5.2. *Central limit theorem for the likelihood ratio statistic.*

Let $\mathscr{L}(N_k) = \mathscr{M}(p_k, k, n_k)$, $p_k{}^0 = (p^0_{1k}, \cdots, p^0_{kk})$. *Define*

$$\theta_{ik} = \log\left(\frac{p_{ik}}{p^0_{ik}}\right) - \sum_i p_{ik} \log\left(\frac{p_{ik}}{p^0_{ik}}\right). \tag{5.19}$$

Define the function $I(\cdot, \cdot)$ *for* $x \geqq 0$, $\lambda > 0$ *by*

$$I(x, \lambda) = x \log\left(\frac{x}{\lambda}\right) - x + \lambda \quad \textit{if} \quad x > 0 \quad \textit{and} \quad I(0, \lambda) = \lambda . \tag{5.20}$$

($I(\cdot, \cdot)$ *is the Kullback–Liebler information kernel for the Poisson distribution.*) *Letting*

$$\mathscr{L}(X_{ik}) = \text{Poisson}(\lambda_{ik}), \qquad \lambda_{ik} \equiv n_k p_{ik},$$

define

$$\gamma_k = \frac{1}{n_k} \sum_{i=1}^k \text{Cov}\,(I(X_{ik}, \lambda_{ik}), X_{ik}), \tag{5.21}$$

$$\sigma^2_{ik} = \text{Var}\,\{I(X_{ik}, \lambda_{ik}) + X_{ik}(\theta_{ik} - \gamma_k)\}, \tag{5.22}$$

and

$$s_k{}^2 = \sum_{i=1}^k \sigma^2_{ik}. \tag{5.23}$$

Suppose

$$\max_{1 \leqq i \leqq k} p_{ik} = o(1) \quad \textit{as} \quad k \to \infty, \tag{5.24}$$

and

$$n_k p_{ik} \geqq \varepsilon \quad \textit{for all} \quad i, \quad k, \quad \textit{some} \quad \varepsilon > 0 \quad \textit{fixed.} \tag{5.25}$$

Suppose

$$\frac{\max_{1 \leqq i \leqq k} n_k p_{ik} \theta^2_{ik}}{k + n_k \sum_{i=1}^k p_{ik} \theta^2_{ik}} = o(1) \quad \textit{as} \quad k \to \infty. \tag{5.26}$$

(5.26) *is equivalent to the* uan *condition*

$$\frac{1}{s_k{}^2} \max_{1 \leqq i \leqq k} \sigma^2_{ik} = o(1) \quad \textit{as} \quad k \to \infty. \tag{5.27}$$

Then

$$\mathscr{L}\left(\frac{1}{s_k}\left\{\sum_{i=1}^k N_{ik} \log\left(\frac{N_{ik}}{n_k p^0_{ik}}\right) - \sum_{i=1}^k EI(X_{ik}, \lambda_{ik}) - n_k \sum_{i=1}^k I(p_{ik}, p^0_{ik})\right\}\right) \to \mathscr{N}(0, 1) \quad \textit{as} \quad k \to \infty. \tag{5.28}$$

When the hypothesis $p_{ik} = p^0_{ik}$ *holds for every* i, k,

$$\mathscr{L}\left(\frac{1}{s_k}\left\{\sum_{i=1}^k N_{ik} \log\left(\frac{N_{ik}}{n_k p_{ik}}\right) - \sum_{i=1}^k EI(X_{ik}, \lambda_{ik})\right\}\right) \to \mathscr{N}(0, 1) \quad \textit{as} \quad k \to \infty \tag{5.29}$$

provided only that (5.24) *and* (5.25) *hold.*

Before proceeding to the proof of Theorem 5.2, we first list some of the properties of the Kullback–Liebler information kernel $I(\cdot, \cdot)$.

LEMMA 5.1. *The function $I(\cdot, \cdot)$ defined in* (5.20) *satisfies*

$$0 \leqq I(x, \lambda) \leqq \frac{(x-\lambda)^2}{\lambda} \tag{5.30}$$

$$|I(x+y, \lambda) - I(x, \lambda)| \leqq \frac{y^2}{\lambda} + yh(x, \lambda) \qquad \text{if} \quad x, y = 0, 1, 2, \cdots \tag{5.31}$$

where

$$h(x, \lambda) \equiv |x - \lambda| \left\{\frac{1}{\lambda} + \frac{2}{x+1}\right\}. \tag{5.32}$$

Letting $\mathscr{L}(X) =$ Poisson (λ),

$$EI^4(X, \lambda) \leqq C\left\{\frac{1}{\lambda^3} + 1\right\} \tag{5.33}$$

for some constant C,
and

$$\alpha_2 \equiv EI(X, \lambda)\varphi_2(X, \lambda) \geqq \frac{1}{2^{\frac{1}{2}}} \frac{\lambda}{\lambda + 2}, \tag{5.34}$$

φ_2 defined in (3.5).

PROOF. (5.30) is proved by noting $g(x) \equiv (x-\lambda)^2/\lambda - I(x, \lambda)$ is concave on $[0, \lambda/2]$ and thereafter convex. Since $g(0) = g(\lambda) = 0$ and $g'(\lambda) = 0$, g must be nonnegative.

To establish (5.31), the mean value theorem gives $|I(x+y, \lambda) - I(x, \lambda)| = y|\log((x+ty)/\lambda)|$ for some $0 < t < 1$. Assuming $x \geqq 1$, the two cases $x + ty \geqq \lambda$ and $x + ty < \lambda$ must be considered separately, and each time the inequality $\log(z) \leqq z - 1$ is used. If $x = 0$ and $y \geqq 1$, then $|I(y, \lambda) - I(0, \lambda)| = y|\log(y/\lambda) - 1| \leqq y\{y/\lambda + \lambda/y + 1\}$ yields the result. The case $x = 0$ and $y = 0$ is trivial.

Inequality (5.33) is a simple, direct computation from (5.30).

Inequality (5.34) uses (3.5) and then (3.4) to get

$$\begin{aligned}
\alpha_2 &= \frac{1}{\lambda 2^{\frac{1}{2}}} EI(X, \lambda)\Pi_2(X, \lambda), \qquad \Pi_2 \quad \text{as in (3.4)} \\
&= \frac{1}{\lambda 2^{\frac{1}{2}}} \textstyle\sum_{i=0}^2 EI(X, \lambda)\binom{2}{i}(-\lambda)^{2-i}X^{(i)} \\
&= \frac{1}{\lambda 2^{\frac{1}{2}}} \textstyle\sum_{i=0}^2 \binom{2}{i}(-\lambda)^{2-i} \sum_{x=i}^\infty \frac{1}{(x-i)!} \lambda^x e^{-\lambda} I(x, \lambda) \\
&= \frac{\lambda^2}{\lambda 2^{\frac{1}{2}}} \textstyle\sum_{i=0}^2 \binom{2}{i}(-1)^{2-i} \sum_{y=0}^\infty \frac{1}{y!} \lambda^y e^{-\lambda} I(y+i, \lambda) \\
&= \frac{\lambda}{2^{\frac{1}{2}}} E\Delta^2 I(X, \lambda).
\end{aligned}$$

From applying the mean value theorem twice, and for every $x \geqq 0$ (since $I(\cdot, \lambda)$ is continuous on $[0, \infty)$), for some $0 < \theta = \theta(x, \lambda) < 1$,

$$\Delta^2 I(x, \lambda) = I''(x + 2\theta, \lambda) = 1/(x + 2\theta).$$

Therefore

$$\alpha_2 \geqq \frac{\lambda}{2^{\frac{1}{2}}} E \frac{1}{X+2}$$

$$\geqq \frac{\lambda}{2^{\frac{1}{2}}} \frac{1}{\lambda + 2} \qquad \text{from Jensen's inequality.}$$

The proof of the lemma is complete.

PROOF OF THEOREM 5.2. We use Lemma 2.2 and its notation. Define $\lambda_{ik}^0 = n_k p_{ik}^0$ and define

$$f_{ik}(x) = I(x, \lambda_{ik}) - EI(X_{ik}, \lambda_{ik}) + (x - \lambda_{ik})(\theta_{ik} - \gamma_k) . \tag{5.35}$$

Then the statistic of (5.28) can be expressed as

$$\frac{1}{s_k}\left\{\sum N_{ik} \log\left(\frac{N_{ik}}{n_k p_{ik}^0}\right) - \sum EI(X_{ik}, \lambda_{ik}) - n_k \sum I(p_{ik}, p_{ik}^0)\right\} \tag{5.36}$$

$$= \frac{1}{s_k} \sum f_{ik}(N_{ik}) .$$

The $\{f_{ik}\}$, so defined, satisfy the requirements of Lemma 2.2,

$$Ef_{ik}(X_{ik}) = 0 , \qquad \sum_i \operatorname{Cov}(f_{ik}(X_{ik}), X_{ik}) = 0 ,$$

while σ_{ik}^2 of (5.22) is the variance of $f_{ik}(X_{ik})$ and $s_k^2 = \sum \sigma_{ik}^2$ as required. Conditions (5.24) and (5.25) assure that $n_k \to \infty$ as $k \to \infty$ and $\max_{1 \leqq i \leqq 1} p_{ik} = o(1)$.

We first show that the variance s_k^2 and $k + \zeta_k^2$ where

$$\zeta_k^2 \equiv n_k \sum p_{ik} \theta_{ik}^2 \tag{5.37}$$

are of the same order as $k \to \infty$. We have

$$s_k^2 = \sum_{i=1}^k \operatorname{Var} f_{ik}(X_{ik})$$

$$\geqq \sum E^2 f_{ik}(X_{ik}) \varphi_2(X_{ik}, \lambda_{ik})$$

$$= \sum E^2 I(X_{ik}, \lambda_{ik}) \varphi_2(X_{ik}, \lambda_{ik}) \qquad \text{from (5.35) and (3.7)}$$

$$\geqq \tfrac{1}{2} \sum \left(\frac{\lambda_{ik}}{\lambda_{ik} + 2}\right)^2 \qquad \text{using (5.34)}$$

so

$$s_k^2 \geqq c_1 k \qquad \text{with} \quad c_1 \equiv \frac{1}{2}\left(\frac{\varepsilon}{\varepsilon + 2}\right)^2 > 0 . \tag{5.38}$$

Define

$$\beta_{ik} = \frac{1}{\lambda_{ik}} \operatorname{Cov}(I(X_{ik}, \lambda_{ik}), X_{ik}) - \gamma_k . \tag{5.39}$$

Let $T_{ik}(X_{ik}) \equiv I(X_{ik}, \lambda_{ik}) - X_{ik}/\lambda_{ik} \operatorname{Cov}(I(X_{ik}, \lambda_{ik}), X_{ik})$, so that $\operatorname{Cov}(T_{ik}(X_{ik}), X_{ik}) = 0$. Then

$$s_k^2 = \sum \sigma_{ik}^2 = \sum \operatorname{Var} T_{ik}(X_{ik}) + \sum \lambda_{ik}(\beta_{ik} + \theta_{ik})^2 \tag{5.40}$$

$$\geqq 2 \sum \lambda_{ik} \beta_{ik} \theta_{ik} + \sum \lambda_{ik} \theta_{ik}^2 . \tag{5.41}$$

However,

$$
\begin{aligned}
n_k \sum p_{ik}\beta_{ik}^2 + n_k\gamma_k^2 &= n_k \sum p_{ik}\frac{1}{\lambda_{ik}^2}\operatorname{Cov}^2(I(X_{ik}, \lambda_{ik}), X_{ik}) \\
&\leqq n_k \sum p_{ik}\frac{1}{\lambda_{ik}^2} EI^2(X_{ik}, \lambda_{ik}) \operatorname{Var} X_{ik} \\
&\leqq \sum E(X_{ik} - \lambda_{ik})^4/\lambda_{ik}^2 \qquad \text{using (5.30)} \\
&= \textstyle\sum_{i=1}^k (\lambda_{ik} + 3\lambda_{ik}^2)/\lambda_{ik}^2 < k\left(\frac{1}{\varepsilon} + 3\right).
\end{aligned} \tag{5.42}
$$

Hence,

$$
n_k\gamma_k^2 = O(k) \qquad \text{and} \qquad n_k \sum p_{ik}\beta_{ik}^2 = O(k)\,. \tag{5.43}
$$

The Schwarz inequality applied to (5.41) with (5.43) yields

$$
s_k^2 \geqq O(k^{\frac{1}{2}}\zeta_k) + \zeta_k^2\,. \tag{5.44}
$$

(5.38) and (5.44) together guarantee the existence of a constant $c > 0$ such that

$$
s_k^2 \geqq c(k + \zeta_k^2) = c(k + n_k \sum p_{ij}\theta_{ik}^2)\,. \tag{5.45}
$$

From (5.40), we also have

$$
\begin{aligned}
s_k^2 &= \sum \{\operatorname{Var} I(X_{ik}, \lambda_{ik}) - \lambda_{ik}(\beta_{ik} + \gamma_k)^2\} + n_k \sum p_{ik}(\beta_{ik} + \theta_{ik})^2 \\
&\leqq \sum E\frac{(X_{ik} - \lambda_{ik})^4}{\lambda_{ik}^2} + 2n_k \sum p_{ik}\beta_{ik}^2 + 2n_k \sum p_{ik}\theta_{ik}^2 \\
&\leqq k\left(\frac{1}{\varepsilon} + 3\right) + O(k) + 2\zeta_k^2 \qquad \text{using (5.42) and (5.43).}
\end{aligned} \tag{5.46}
$$

(5.45) and (5.46) together imply that

$$
s_k^2 = O_e(k + n_k \sum p_{ik}\theta_{ik}^2)\,. \tag{5.47}
$$

We now turn to the condition (2.9) of Lemma 2.2. Note that the function h of (5.32) satisfies

$$
\begin{aligned}
h^2(x, \lambda) &\leqq \frac{2(x - \lambda)^2}{\lambda^2} + 8\frac{(x - \lambda)^2}{(x + 1)^2} \\
h^2(x, \lambda) &\leqq \frac{2(x - \lambda)^2}{\lambda^2} + 16\frac{(x - \lambda)^2}{(x + 1)(x + 2)} \qquad \text{since } x \geqq 0 \\
&= \frac{2(x - \lambda)^2}{\lambda^2} + 16\left\{1 - \frac{3 + 2\lambda}{x + 1} + \frac{4 + 4\lambda + \lambda^2}{(x + 1)(x + 2)}\right\}.
\end{aligned} \tag{5.48}
$$

Let $\{L_{ik}\}$ and $\{M_{ik}\}$ and other notation be as in Lemma 2.2. From (5.48) and the easily verified inequality that $E1/(x + i)^{(i)} \leqq 1/(n + i)^{(i)}p^{(i)} < 1/(np)^i$ when $X \sim Bin(n, p)$, then

$$
\begin{aligned}
Eh^2(L_{ik}, \lambda_{ik}) \leqq{} & \frac{2}{\lambda_{ik}^2}\{l_k p_{ik}(1 - p_{ik}) + (l_k p_{ik} - \lambda_{ik})^2\} \\
& + 16\left\{1 - \frac{3 + 2\lambda_{ik}}{l_k p_{ik}} + \frac{4 + 4\lambda_{ik} + \lambda_{ik}^2}{l_k^2 p_{ik}^2}\right\}.
\end{aligned}
$$

Since $l_k = n_k + v_k n_k^{\frac{1}{2}}$ with $v_k = O(1)$, then

$$\frac{l_k}{n_k} = O(1) \quad \text{and} \quad \frac{n_k}{l_k} = O(1) . \tag{5.49}$$

Obvious inequalities using (5.49) and the fact that

$$1 - 2\lambda_{ik}/l_k p_{ik} + \lambda_{ik}^2/l_k^2 p_{ik}^2 = (l_k - n_k)^2/l_k^2 = v_k^2 n_k/l_k^2$$

yield

$$Eh^2(L_{ik}, \lambda_{ik}) = O(1/\lambda_{ik}) . \tag{5.50}$$

$$\begin{aligned}
&\frac{1}{s_k^2} E(\textstyle\sum_{i=1}^k \{f_{ik}(L_{ik} + M_{ik}) - f_{ik}(L_{ik})\})^2 \\
&\quad \leqq \frac{2}{s_k^2} E(\textstyle\sum_{i=1}^k \{I(L_{ik} + M_{ik}, \lambda_{ik}) - I(L_{ik}, \lambda_{ik})\})^2 \\
&\qquad + \frac{2}{s_k^2} E(\textstyle\sum_{i=1}^k M_{ik}(\theta_{ik} - \gamma_k))^2 \\
&\quad \leqq \frac{2}{s_k^2} E\left(\textstyle\sum_{i=1}^k \frac{M_{ik}^{(2)}}{\lambda_{ik}} + M_{ik}\left\{h(L_{ik}, \lambda_{ik}) + \frac{1}{\lambda_{ik}}\right\}\right)^2 \\
&\qquad + \frac{2}{s_k^2} E(\textstyle\sum_i M_{ik}\theta_{ik} - m_k\gamma_k)^2 \\
&\quad \leqq \frac{4}{s_k^2} \textstyle\sum_{i=1}^k \sum_{j=1}^k \frac{1}{\lambda_{ik}\lambda_{jk}} EM_{ik}^{(2)} M_{jk}^{(2)} \qquad (5.51) \\
&\qquad + \frac{4}{s_k^2} E\left(\textstyle\sum_i M_{ik}\left\{h(L_{ik}, \lambda_{ik}) + \frac{1}{\lambda_{ik}}\right\}\right)^2 \\
&\qquad + \frac{4}{s_k^2} E(\textstyle\sum_i M_{ik}\theta_{ik})^2 + \frac{4m_k^2\gamma_k^2}{s_k^2} \\
&\quad \leqq \frac{4}{s_k^2} \textstyle\sum_i \sum_j \frac{1}{\lambda_{ik}\lambda_{jk}} \{m_k^{(4)} p_{ik}^2 p_{jk}^2 + \delta_{ij}(4m_k^{(3)} p_{ik}^3 + 2m_k^{(2)} p_{ik}^2)\} \\
&\qquad + \frac{4m_k^2}{s_k^2} \textstyle\sum_i E\frac{M_{ik}}{m_k} E\left\{h(L_{ik}, \lambda_{ik}) + \frac{1}{\lambda_{ik}}\right\}^2 \\
&\qquad + \frac{4m_k^2}{s_k^2} \textstyle\sum_i E\frac{M_{ik}}{m_k} \theta_{ik}^2 + \frac{4m_k^2\gamma_k^2}{s_k^2} .
\end{aligned}$$

Using (5.50) together with obvious identities and inequalities, the assumption $m_k^2 = O(h^2)n_k$, and finally (5.43) and (5.47) shows (5.51) is $O(h)$ and hence that condition (2.9) is satisfied.

Conditions (2.8) and (2.10) can be checked simultaneously by verifying Liapounov's theorem in the form

$$\frac{1}{s_k^4} \textstyle\sum_{i=1}^k Ef_{ik}^4(X_{ik}) = o(1) \quad \text{as} \quad k \to \infty . \tag{5.52}$$

Referring to (5.35),

$$\begin{aligned}
Ef_{ik}^4(X_{ik}) &\leqq 8EI^4(X_{ik}, \lambda_{ik}) + 8(\theta_{ik} - \gamma_k)^4 E(X_{ik} - \lambda_{ik})^4 \\
&\leqq O(1) + 8(\theta_{ik} - \gamma_k)^4(\lambda_{ik} + 3\lambda_{ik}^2) = O(1)\{1 + (\theta_{ik}^4 + \gamma_k^4)\lambda_{ik}^2\} ,
\end{aligned}$$

using (5.33). Hence (5.52) is bounded by

$$\frac{1}{s_k^4} O(1)\{k + n_k^2 \sum p_{ik}^2 \theta_{ik}^4 + n_k^2 \gamma_k^4 \sum p_{ik}^2\}$$

$$\leqq O(1)\left\{\frac{1}{k} + \frac{\max_{1\leqq i\leqq k}(n_k p_{ik}\theta_{ik}^2)}{s_k^2} + \max_{1\leqq i\leqq k} p_{ik}\right\} = o(1)$$

from (5.43), (5.47) and then is $o(1)$ as $k \to \infty$ from assumptions (5.24), (5.26). This establishes (2.8) and (2.10).

Thus (5.26) $\Rightarrow$ (5.27), and it only remains to show the converse. But $\max_i \lambda_{ik}\beta_{ik}^2 = 0(1)$ and $\max_i \operatorname{Var}(T_{ik}) = O(1)$ by repeating the inequalities leading to (5.42) and (5.46) respectively, excluding the summation signs. Therefore

$$\begin{aligned}\sigma_{ik}^2/s_k^2 &= \{\operatorname{Var}(T_{ik}) + \lambda_{ik}(\beta_{ik} + \theta_{ik})^2\}/s_k^2 \\ &= o(1) + O_e(n_k p_{ik}\theta_{ik}^2/(k + \sum n_k p_{ik}\theta_{ik}^2))\end{aligned}$$

from (5.47). Thus (5.27) $\Rightarrow$ (5.26). The proof of this theorem is complete.

REFERENCES

[1] HOEFFDING, W. (1965). Asymptotically optimal tests for multinomial distributions. *Ann. Math. Statist.* **36** 369–401.

[2] LOÈVE, *M.* (1960). *Probability Theory*, (2nd ed.). Van Nostrand, Princeton.

[3] MORRIS, C. (1966). Admissible Bayes procedures and classes of epsilon Bayes procedures for testing hypotheses in a multinomial distribution. Technical Report No. 55, Department of Statistics, Stanford Univ.

[4] NEYMAN, J. and PEARSON, E. S. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference. *Biometrika* **20-A** 175–247, 264–299.

[5] PEARSON, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philos. Mag.* **50** 157–175.

[6] RIORDAN, J. (1968). *Combinatorial Identities.* Wiley, New York.

[7] STECK, G. P. (1957). Limit theorems for conditional distributions. *Univ. California Publ. Statist.* **2** No. 12, 237–284.

[8] SZEGÖ, G. (1939). *Orthogonal Polynomials.* American Mathematical Society Colloquium Publication, **23**, New York.

[9] TUMANYAN, S. H. (1954). On the asymptotic distribution of the chi-square criterion. *Dokl. Akad. Nauk. SSSR* **94** 1011–1012.

THE RAND CORPORATION
1700 MAIN STREET
SANTA MONICA, CALIFORNIA 90406