MIXTURES OF POISSON DISTRIBUTIONS. I.
SOME SIMPLE CASES

By Lucien Le Cam*

University of California, Berkeley

1. Introduction. A numerical random variable X is said to be distributed according to a mixture of Poisson distributions if there is a probability measure F on the positive real line such that

$$\text{Prob}(X=k) = \int_0^\infty e^{-\lambda} \frac{\lambda^k}{k!} F(d\lambda)$$

for every nonnegative integer k.

Mixtures of Poisson distributions occur in many different contexts. A general type of circumstances leading to such mixtures has been described in [1] as follows.

Let $\mathcal{X}$ be an arbitrary space carrying a σ-field $\mathcal{A}$. Let μ be a positive measure on $\mathcal{A}$. Let $\mathcal{B}$ be the subring of $\mathcal{A}$ formed by the sets which have finite measure. Assume that for each $B \in \mathcal{B}$ and each integer m there is a $B_m \in \mathcal{B}$ such that $B \subset B_m$ and $m \leq \mu(B_m) < \infty$.

Consider a random mechanism by which points are selected in $\mathcal{X}$ in such a way that if $A \in \mathcal{B}$ the number N(A) of points falling in A is almost surely finite. Assume in addition that, given

---

*The author is a research fellow of the A. P. Sloan Foundation. This research was prepared with the partial support of the Office of Ordnance Research, U. S. Army under Contract DA-04-200-ORD-171, Task Order 3.

$N(A) = n$ the position of the n points $x_1, x_2, \cdots, x_n$ in A have the same joint distribution as if the x_j were selected at random independently of each other with the distribution

$$\text{Prob}(x_j \in S) = \frac{\mu(S \cap A)}{\mu(A)}.$$

Under these assumptions, for every $A \in \mathcal{B}$ the distribution of the variable $N(A)$ is a mixture of Poisson distributions.

This can be easily proved as follows, using the operational notations of [2]. Let Δ be the measure giving mass (-1) to zero and mass (+1) to the point unity of the line. Let Q be the distribution of $N(A)$. For any $\mathcal{B} \supset A$ let Q_n be the distribution of $N(A)$ given that $N(B) = n$. Finally let F_B be the distribution of

$$\lambda = \frac{\mu(A \cap B)}{\mu(B)} N(B).$$

According to the assumptions made

$$Q_n = [I + p\Delta]^n$$

with $p = p_B = \mu(A \cap B)/\mu(B)$.

Hence (see [3]),

$$\|Q_n - \exp[np_B\Delta]\| \leq 4p_B$$

$$\|Q - \int \exp(\lambda\Delta) F_B(d\lambda)\| \leq 4p_B.$$

If $\mu(B)$ tends to infinity, p_B tends to zero, hence

$$\int \exp(\lambda\Delta) F_B(d\lambda)$$

converges to the limit Q, in the sense of the L_1-norm. As we shall prove in section 2 such a convergence implies that F_B tends in the ordinary sense to a limit F and that $Q = \int \exp(\lambda\Delta)F(\alpha\lambda)$.

It has been shown by Greenwood and Yule [4] that certain empirical distributions of accidents are fairly well represented by Poisson mixtures. These authors assume that the mixing measure F is a Gamma distribution.

In accident studies, the occurrence of a mixture can often be attributed to a lack of homogeneity of the population. More precisely, consider a population of n individuals subject to "accidents" and observed for a given length of time. Let X_j be the number of accidents sustained by the jth member of the population. It is often assumed that X_j has a Poisson distribution with expectation $\lambda_j = EX_j$. The number λ_j is called the proneness of the jth individual. A lack of homogeneity is reflected in the variation of λ_j from member to member. If it is assumed that the individuals have been taken at random independently from some very large population, one may consider that the λ_j are results of independent trials on a population characterized by the measure F. Under these assumptions the X_j's are independent variables distributed according to a Poisson mixture.

Neyman and Scott [5], have derived tests of homogeneity which are in particular applicable to the case where F is assumed to be a Gamma distribution.

It is clear that the Gamma distribution has been selected in such studies because of its general shape and because of its general tractability. For this reason it may be interesting to investigate the performance of the tests developed in [5] for other mixing distributions. The present paper represents a study of some particular cases in which such investigations can be carried out rather easily.

2. Mixing distributions and corresponding mixtures. The purpose of the present section is to show that the correspondence between a mixing measure and the resulting Poisson mixture is one to one. Further, the correspondence is bicontinuous for suitably selected topologies. The existence of consistent estimates of the mixing measure follows from this.

Whenever convenient, we shall employ operation notations. For instance, let $\mathcal{E}$ be the space of bounded numerical functions on the space of all integers. The space $\mathcal{E}$ is a Banach space for the norm defined by

$$|u| = \sup_{S} |u(x)|.$$

The finite signed measures on the integers form, under the convolution operation, a cumulative Banach algebra $\mathcal{M}$ identifiable to a subspace of the dual of $\mathcal{E}$. The measure Δ which assigns mass (-1) to zero and mass (+1) to the point unity is an element of $\mathcal{M}$. The Poisson distribution having expectation λ is simply $\exp\{\lambda\Delta\}$. A Poisson mixture Q with mixing measure F can be written

$$Q = \int \exp\{\lambda\Delta\} F(d\lambda).$$

Let u be a complex-valued function defined on the nonnegative integers and such that

$$\sum_{k=0}^{\infty} |u(k)| \frac{\lambda^k}{k!} < \infty,$$

for every nonnegative λ. Let $\hat{u}$ be the transform

$$\hat{u}(\lambda) = \sum_k u(k) \frac{\lambda^k}{k!}.$$

By definition

$$\int u(k)Q(dk) = \sum_k u(k)\int e^{-\lambda}\frac{\lambda^k}{k!}F(d\lambda).$$

Therefore, whenever $\int|u(k)|Q(dk) < \infty$ one can write

$$\int\hat{u}(\lambda)e^{-\lambda}F(d\lambda) = \int u(k)Q(dk).$$

It follows that Q determines all integrals of the type $\int v(\lambda)F(d\lambda)$ where $v(\lambda) = \lambda^m e^{-\alpha\lambda}$, $m \geq 0$, $\alpha > 0$.

The Stone-Weierstrass theorem shows that Q determines all integrals of the type $\int v(\lambda)F(d\lambda)$ where v is continuous and vanishes at infinity. Therefore the correspondence $F \longleftrightarrow Q_F$ is one to one.

Consider two random variables X and Y which take only nonnegative values. One can write

$$E\, e^{Y\Delta} - E\, e^{X\Delta} = E(e^{(Y-X)\Delta} - 1)e^{X\Delta}.$$

Suppose that $\text{Prob}\{|Y-X| > \varepsilon\} < \delta$.

Let $f(X,Y) = 1$ if $|Y-X| \leq \varepsilon$ and let $f = 0$ otherwise. Then

$$\|E[e^{Y\Delta} - e^{X\Delta}]\| \leq 2\delta + E\|f(X,Y)e^{X\Delta}[e^{(Y-X)\Delta} - 1]\|.$$

For every real number u one can write

$$e^{u\Delta} - 1 = \int_0^u e^{\xi\Delta}\,\Delta\, d\xi.$$

It follows that

$$\|E[e^{Y\Delta} - e^{X\Delta}]\| \leq 2\delta + E\|\Delta e^{X\Delta}\|\,\{f(X,Y)e^{2|Y-X|}|Y-X|\}$$

$$\leq 2\delta + 2\varepsilon e^{2\varepsilon}$$

when X takes only very large values, better bounds may be obtained as follows.

If $\text{Prob}[X < \lambda] < \delta_1$ then

$$\| E[e^{Y\Delta} - e^{X\Delta}] \| \leq 2\delta + 2\delta_1 + \varepsilon e^{2\varepsilon} E \| \Delta e^{Z\Delta} \|$$

where Z is a random variable taking only values larger than λ. For $Z \geq \lambda \geq 1$ it is easily seen that $\|\Delta e^{Z\Delta}\| \leq 2/\sqrt{\lambda}$. Finally one can conclude as follows.

Prohorov

Theorem 1. Let F and G be two measures on the positive real axis. Suppose that the Paul Levy distance of F and G is inferior to ε. Let Q_F and Q_G denote the corresponding Poisson mixtures. Then

$$\|Q_F - Q_G\| \leq 2\varepsilon[1 + e^{2\varepsilon}]$$

If in addition $F\{[0, \lambda)\} < \delta_1$ for $\lambda \geq 1$ then

$$\|Q_F - Q_G\| \leq 2\varepsilon + 2\delta_1 + \frac{2\varepsilon e^{2\varepsilon}}{\sqrt{\lambda}}.$$

This theorem implies in particular that if $F_n \longrightarrow F$ in the ordinary sense then $Q_{F_n} \longrightarrow Q_F$ in the sense of the norm. In addition, the proof shows that it may happen that $\|Q_{F_n} - Q_{G_n}\| \longrightarrow 0$ even though the Lévy distance between F_n and G_n does not tend to zero. This may happen when for every $\lambda > 0$ the measure $|F_n - G_n|\{[0, \lambda)\}$ tends to zero.

To obtain a result in the reverse direction note that if $\mathcal{P}_\lambda$ denotes the Poisson distribution $P_\lambda = \exp\{\lambda\Delta\}$ then

$$P_\lambda\{(m, \infty)\} = \int_0^\lambda e^{-u} \frac{u^m}{m!} \, du.$$

If Q_F denotes the mixture corresponding to F this gives

$$Q_F\{(m,\infty)\} = \int_0^\infty F(d\lambda)\int_0^\lambda e^{-u}\frac{u^m}{m!}\,du$$
$$\geq \int_t^\infty F(d\lambda)\int_0^\lambda e^{-u}\frac{u^m}{m!}\,du$$
$$\geq P_t\{(m,\infty)\}\,F\{(t,\infty)\}.$$

It follows from this that for a family $\{Q_F, F \in \mathcal{F}\}$ to be relatively compact for the topology induced by Paul Lévy's distance it is necessary and sufficient that the family $\{F: F \in \mathcal{F}\}$ have the same property. As a consequence, if $Q_{F_n} \longrightarrow Q_F$ in the sense of Paul Lévy's distance, then $F_n \longrightarrow F$ in the same sense.

The continuity of the map $Q_F \longrightarrow F$ implies that whenever consistent estimates of Q_F are available consistent estimates of F are also available. For instance, suppose that N_j; $j=1,2,\cdots,n,\cdots$ are independent random variables having distribution Q_F. Let G_n be the empirical cumulative distribution of the first n variables. Take for estimate of F any distribution F_n such that

$$\sup_m\left|[1 - G_n(m)] - Q_{F_n}\{(m,\infty)\}\right| \leq \inf_H\,\sup_m\left|[1-G_n(m)] - Q_H(m,\infty)\right| + \frac{1}{n},$$

then $\hat{F}_n$ converges almost surely in the sense of Paul Lévy's distance to the "true" distribution F.

3. Contiguity and asymptotic normality. Let $\{P_n\}$ and $\{Q_n\}$ be sequences of probability measures on σ-fields $\{\mathcal{A}_n\}$. By the norm $\|P_n-Q_n\|$ of the difference P_n-Q_n will be meant the quantity

$$\|P_n - Q_n\| = 2\sup_{A\in\mathcal{A}_n}|P_n(A_n) - Q_n(A_n)|.$$

The two sequences $\{P_n\}$ and $\{Q_n\}$ will be called "equivalent" if $\|P_n - Q_n\|$ tends to zero as n tends to infinity.

Let $\{T_n\}$ denote a sequence of $\{\mathcal{A}_n\}$-measurable functions taking their values in a fixed Euclidean space. The equivalence of $\{P_n\}$ and $\{Q_n\}$ implies that the sequences $\{T_n\}$ having a limiting distribution $\mathcal{L}(T)$ under $\{P_n\}$ (resp. Q_n) have also the same limiting distribution $\mathcal{L}(T)$ under $\{Q_n\}$ (resp. P_n). Conversely the preceding property implies the equivalence of $\{P_n\}$ and $\{Q_n\}$.

An equivalence relation, weaker than the preceding, has been introduced in [6] under the name of "contiguity." Specifically, the sequences $\{P_n\}$ and $\{Q_n\}$ are called contiguous if the sequences $\{T_n\}$ tending to zero in probability are the same for $\{P_n\}$ and $\{Q_n\}$.

Let $\mu_n = P_n + Q_n$ and let h be the Radon-Nikodym derivative $h = (dQ_n/d\mu_n)$. Let $\Lambda[Q_n; P_n]$ denote $\log(h/1-h)$. This measurable function which takes values in $[-\infty, +\infty]$ will be called the logarithm of the likelihood ratio of Q_n to P_n. In the sequel we shall have to add quantities of the form $\Lambda[Q_n; P_n]$. This can be done with obvious conventions (such as $\infty - \infty = -\infty$) concerning the addition of infinite values. Another procedure avoiding infinities altogether is given in [6].

Letting $\chi_n = \Lambda[Q_n; P_n]$, the contiguity of $\{P_n\}$ and $\{Q_n\}$ is equivalent to the compactness requirement that for every $\varepsilon > 0$ there is an $n(\varepsilon)$ and a $b(\varepsilon) < \infty$ such that $n \geq n(\varepsilon)$ implies $(P_n + Q_n)\{|\chi_n| > b(\varepsilon)\} < \varepsilon$. The equivalence of $\{P_n\}$ and $\{Q_n\}$ corresponds to the requirement that for every ε there is an $n(\varepsilon)$ such that $n \geq n(\varepsilon)$ implies $(P_n + Q_n)\{|\chi_n| > \varepsilon\} < \varepsilon$.

Contiguity will often enter in our arguments by the following proposition. Suppose that $\{T_n\}$ is a sequence of $\{\mathcal{A}_n\}$-measurable functions such that the distributions $\mathcal{L}\{e^{X_n}, T_n|P_n\}$ converge to a limit $\mathcal{L}[e^X, T|$ in the usual sense of convergence of integrals of bounded continuous functions. Assume that for every $\varepsilon > 0$ there is an $n(\varepsilon)$ and a $b(\varepsilon)$ such that $Q_n\{X_n > b_n(\varepsilon)\} < \varepsilon$ for $n > n(\varepsilon)$. Then, $\mathcal{L}[e^{X_n}, T_n|Q_n]$ converges to the distribution $e^X\mathcal{L}[e^X, T]$ which has density e^X with respect to $\mathcal{L}[e^X, T]$.

In the following sections we shall be concerned with sequences $\{P_n\}$ and $\{Q_n\}$ such that $\mathcal{L}[X_n|P_n]$ converges, as n tends to infinity, to a normal distribution $\mathcal{N}(\mu, \sigma^2)$. The contiguity assumption is then equivalent to the relation $2\mu = -\sigma^2$.

Furthermore, we shall be particularly interested in the case where both P_n and Q_n are product measures describing the distribution of independent variables $X_{n,j}$; $j=1,2,\cdots,\nu_n$. Let $p_{n,j} = \mathcal{L}[X_{n,j}|P_n]$ and let $q_{n,j} = \mathcal{L}[X_{n,j}|Q_n]$. The differences $p_{n,j} - q_{n,j}$ will be called asymptotically negligible if $\sup_j\|p_{n,j} - q_{n,j}\|$ converges to zero as n tends to infinity.

In [6] we have stated the following theorem.

<u>Proposition</u> 1. <u>Let the differences</u> $\|q_{n,j} - p_{n,j}\|$ <u>be asymptotically negligible and let</u> $X_{n} = \Lambda[q_{n,j}; p_{n,j}]$. <u>Let</u> $X_n = \Sigma_j X_{n,j}$ <u>and let</u> $S_n = \Sigma_j [e^{X_{n,j}} - 1]$.

<u>The following conditions are equivalent</u>:

(a) $\mathcal{L}(X_n)$ <u>converges to</u> $\mathcal{N}[-\frac{1}{2}\sigma^2, \sigma^2]$.

(b) $\mathcal{L}(S_n)$ <u>converges to</u> $\mathcal{N}[0, \sigma^2]$.

<u>Proof</u>. This result is closely related to a classical theorem of Bobrov (see [7], p.144). Since the reduction to Bobrov's theorem

requires some argument which is not given in [6] we shall give the details of the proof.

Note first that both (a) and (b) imply that for every $\varepsilon > 0$, say $\varepsilon = 10^{-2}$, the sum

$$\sum_j P_n\{|\chi_{n,j}| > \varepsilon\}$$

tends to zero as $n \longrightarrow \infty$. Since (a) implies the contiguity of $\{P_n\}$ and $\{Q_n\}$ it follows that under (a) the sum

$$\sum_j Q_n\{|\chi_{n,j}| > \varepsilon\}$$

also tends to zero. Under (b), let $A_{n,j}$ be the set where $|\chi_{n,j}| \leqq \varepsilon$. Then, according to the central limit theorem

$$\sum_j \int_{A_{n,j}} [e^{\chi_{n,j}} - 1] dp_{n,j} = \sum_j [q_{n,j}(A_{n,j}) - p_{n,j}(A_{n,j})]$$

also tends to zero, therefore $\Sigma_j Q_n[|\chi_{n,j}| > \varepsilon]$ tends to zero. Let $p'_{n,j}(B)$ be defined by

$$p'_{n,j}(B) = \frac{p_{n,j}(B \cap A_{n,j})}{p_{n,j}(A_{n,j})}$$

and similarly for $q'_{n,j}$. Clearly $\|p'_{n,j} - p_{n,j}\| \leqq 4p_{n,j}(A^c_{n,j})$. If P'_n is the product of the $p'_{n,j}$ then $\|P'_n - P_n\|$ tends to zero. Similarly for Q'_n. Replacing the sequences distributions $\{P_n\}$ and $\{Q_n\}$ by the equivalent sequences $\{P'_n\}$ and $\{Q'_n\}$ if necessary one can assume that $|\chi_{n,j}| \leqq \varepsilon = 10^{-2}$ for every u and j. With this supplementary assumption one has always

$$E\{e^{\chi_{n,j}} - 1|P_n\} = 0.$$

Therefore, if S_n is asymptotically normal, the limiting distribution must have expectation zero. Let $S_{n,j} = e^{\chi_{n,j}} - 1$

and assume (b) satisfied. Then

$$\chi_{n,j} = \log(1 + S_{n,j}) = S_{n,j} - \frac{1}{2} S_{n,j}^2 + \theta|S_{n,j}|^3$$

with $|\theta| < 1$. According to the central limit theorem $\Sigma_j S_{n,j}^2$ converges in probability to σ^2 and

$$\sum_j E|S_{n,j}|^3 \leqq [\sum_j E\, S_{n,j}^2] \sup_j E|S_{n,j}|$$

converges to zero. Therefore (a) holds.

Conversely, if (a) holds, let $\mu_{n,j} = -E\, \chi_{n,j}$. Expanding the exponential gives

$$S_{n,j} = \chi_{n,j} + \frac{1}{2} \chi_{n,j}^2 + \theta|\chi_{n,j}|^3.$$

Hence

$$|\mu_{n,j} - \frac{1}{2} E\, \chi_{n,j}^2| \leqq E|\chi_{n,j}|^3.$$

For n large enough this implies

$$|\mu_{n,j} - \frac{1}{2} E\, \chi_{n,j}^2| \leqq \frac{1}{4} E|\chi_{n,j}|^2,$$

and finally

$$0 \leqq \mu_{n,j} \leqq \frac{3}{4} E\, \chi_{n,j}^2.$$

The convergence of $\Sigma_j \mu_{n,j}$ to $(1/2)\sigma^2$ implies that

$$\sum_j \mu_{n,j}^2 \leqq \sup_j |\mu_{n,j}| \sum_j |\mu_{n,j}|$$

converges to zero. Hence both $\Sigma_j E\, \chi_{n,j}^2$ and $\Sigma_j E[\chi_{n,j} + {}_{n,j}]^2$ converge to the limit σ^2 as $n \longrightarrow \infty$. This, in turn, implies that $\Sigma[S_{n,j} - \chi_{n,j} - \frac{1}{2} \chi_{n,j}^2]$ converges in probability to zero and (b) follows by application of the central limit theorem.

The preceding proposition 1 will simplify some of the arguments of the next sections. The situation considered there can be described in its full generality as follows.

Let $\{T_{n,j}\}$; $j=1,2,\cdots,\nu_n$ be a double sequence of positive numbers. For each n, let $\{X_{n,j}\}$, $j=1,2,\cdots,\nu_n$ be a set of <u>independent</u> random variables. It will be assumed that for each value of n there is a probability measure F_n on the interval $(0, \infty)$ such that, when k is a nonnegative integer,

$$\text{Prob}[X_{n,j} = k] = \frac{1}{k!}\int e^{-\lambda T_{n,j}}(\lambda T_{n,j})^k \, dF_n(\lambda).$$

The problem is to construct optimal asymptotically similar tests of the hypothesis that the F_n are degenerate distributions giving mass unity to a point of $(0, \infty)$ against the hypothesis that F_n is not degenerate.

To apply the result of [6] we may consider a particular sequence $\{P_n\}$ of simple hypotheses for which the measures F_n are degenerate at λ_n and compare it to a particular sequence $\{Q_n\}$ of alternatives for which the distribution G_n of (λ/λ_n) is nondegenerate. The results of [6] depend on the conditions

(A) the sequences $\{P_n\}$ and $\{Q_n\}$ are contiguous,

(B) the sequence $\mathcal{L}(\chi_n) = \mathcal{L}\{\Lambda(Q_n; P_n)|P_n\}$ converges to a normal distribution.

Let $a_{n,j} = \lambda_n T_{n,j}$ and let $s_n = \Sigma_j a_{n,j}$. It will be assumed throughout that the following condition is satisfied:

(C) The sequence $\{s_n\}$ tends to infinity.

Let $p_{n,j}$ and $q_{n,j}$ be the distributions of $X_{n,j}$ under P_n and Q_n respectively. It will be assumed that

(D) the differences $p_{n,j} - q_{n,j}$ are asymptotically negligible. That is,

$$\sup_j \|p_{n,j} - q_{n,j}\| \longrightarrow 0.$$

Let ψ_n be the function defined by

$$\psi_n(x,a) = \int_0^{\infty} e^{-a(\xi-1)} \xi^x \, dG_n(\xi).$$

The likelihood ratio of $q_{n,j}$ to $p_{n,j}$ takes the form $\psi_n(X_{n,j}, a_{n,j})$. The function ψ_n and its logarithm are convex functions of the pair (x,a). For a given value of a, the function $k \longrightarrow \psi_n(k,a)$ defined on the set of nonnegative integers determines G_n entirely. Furthermore, for a fixed value of a, there is a nonempty interval of nonnegative integers k such that $\psi_n(k,a) < 1$ unless G_n is entirely concentrated at the point $\xi = 1$.

If X is a Poisson variable having expectation a then $E\,\psi_n(X,a) = 1$ and Variance $\psi_n(x,a) = E(e^{aU} - 1)$ where U has the same distribution as $(\xi-1)(\xi'-1)$ for variables ξ and ξ' distributed independently according to G_n. This variance is a convex increasing function of a.

The condition that the differences $p_{n,j} - q_{n,j}$ be asymptotically negligible is equivalent to the asymptotic negligibility of the variables $\psi_n(X_{n,j}, a_{n,j}) - 1$, for variables $X_{n,j}$ which are Poisson variables such that $EX_{n,j} = a_{n,j}$.

Lemma 1. *Let* $\rho(a) = \min(a, \sqrt{a})$. *Let* X_n *be a Poisson variable such that* $EX_n = a_n$. *In order that* $\psi_n(X_n, a_n)$ *converges in probability to unity as* $n \longrightarrow \infty$ *it is necessary and sufficient that when* $\mathcal{L}(\xi_n) = G_n$ *the variables* $\rho(a_n)|\xi_n - 1|$ *converge in probability to zero.*

Proof. The sufficiency of the condition is an immediate consequence of theorem 1. To prove the necessity of the condition it is sufficient to consider three cases: (1) $a_n \longrightarrow 0$,

(2) $a_n \longrightarrow a \neq 0$, (3) $a_n \longrightarrow \infty$. In the first case $\text{Prob}(x_n = 0) \longrightarrow 1$. For the convergence of ψ_n to unity it is then necessary and sufficient that

$$\psi_n(0,a_n) = \int_0^\infty e^{-a_n(\xi-1)} dG_n(\xi) \longrightarrow 1 .$$

Since $-a_n(\xi-1) < a_n$ the condition is also equivalent to the convergence in probability to zero of $a_n|\xi_n-1|$.

For the other cases, let $Z_n = \sqrt{a_n}\,[(X_n/a_n) - 1]$ and write ψ_n in the form

$$\psi_n^*(Z_n) = \int_0^\infty \exp\{-a_n(\xi-1-\log\xi) + Z_n\sqrt{a_n}\,\log\xi\} dG_n(\xi)$$

$$= \int e^{Z_n t} dM_n(t),$$

where $t = \sqrt{a_n}\,\log\xi$ and where M_n is an appropriate measure. The possible values of Z_n range from $-\sqrt{a_n}$ to $+\infty$ by equidistant steps. Since ψ_n^* is a convex function of Z_n, convergence to unity at points α, β, γ such that $\alpha < \beta < \gamma$ implies uniform convergence to unity in the interval $[\alpha, \gamma]$. In both cases Z_n has a limiting distribution. From this it follows that convergence to unity, in probability, of ψ_n implies that $\psi_n^*(x) = \int e^{xt} dM_n(t)$ converges to unity for every x in some interval $[\alpha, \infty)$ with $\alpha < 0$. Furthermore, the convergence is uniform on bounded subintervals. In particular

$$\|M_n\| = \psi_n^*(0) = \int_0^\infty \exp\{-a_n(\xi-1-\log\xi)\} dG_n(\xi) \leq 1$$

must converge to unity. This implies the convergence in probability $a_n|\xi_n-1|^2$ to zero.

The foregoing lemma 1 leads to a condition (D') which is equivalent to the condition (D) stated previously.

<u>Lemma</u> 2. <u>Let</u> $a_n = \sup_j a_{n,j}$ <u>and let</u> $\rho(a) = \min(a, \sqrt{a})$. <u>The differences</u> $p_{n,j} - q_{n,j}$ <u>are asymptotically negligible if and only if</u> $\rho(a_n)|\xi_n - 1|$ <u>converges to zero in probability</u>.

The condition (B) which expresses the convergence to a normal distribution of $\Sigma_j\{\psi_n[X_{n,j}, a_{n,j}] - 1\}$ is a more stringent condition than condition (D). Before investigating it in particular cases note the following.

Let $S_{n,j} = \psi_n[X_{n,j}, a_{n,j}] - 1$ and assume that $X_{n,j}$ is Poisson with expectation $a_{n,j}$. For every sequence $\{\alpha_n\}$ of positive numbers we may write

$$S_{n,j} = S'_{n,j} + S''_{n,j}$$

with

$$S'_{n,j} = \int_{|\xi-1|\leq\alpha_n}\left\{\exp[-a_{n,j}(\xi-1)]\xi^{X_{n,j}}-1\right\}dG_n(\xi)$$

$$S''_{n,j} = \int_{|\xi-1|>\alpha_n}\left\{\exp[-a_{n,j}(\xi-1)]\xi^{X_{n,j}}-1\right\}dG_n(\xi).$$

Suppose that there exist normal distributions $\mathcal{N}'_n$ and $\mathcal{N}''_n$ such that the Levy distances

$$\text{dis}\{\mathcal{N}'_n, \ \mathcal{L}[\sum_j S'_{n,j}]\}$$

and

$$\text{dist}\{\mathcal{N}''_n, \ \mathcal{L}[\sum_j S''_{n,j}]\}$$

both tend to zero, and suppose that the sequences $\{\mathcal{N}'_n\}$ and $\{\mathcal{N}''_n\}$ are relatively compact sequences. Then the sequence of joint distributions $\{\mathcal{L}[\Sigma\, S'_{n,j}], \mathcal{L}\Sigma\, S''_{n,j}\}$ is also relatively compact. Furthermore, every convergent subsequence tends to a joint normal distribution.

From this remark we conclude that, in many cases it will be sufficient to investigate the limiting distributions of $\Sigma_j\, S'_{n,j}$ and $\Sigma_j\, S''_{n,j}$. Of course the convergence of these distributions to normal limits may not be necessary for the validity of condition (B).

In many derivations it is important to be able to truncate the variables $X_{n,j}$ and the variables ξ_n. In this connection, the following inequalities are relevant.

Assume $a \leqq 1$. Let $\eta_n = \xi_n$ if $\xi_n \geqq \beta_n$ and let $\eta_n = \beta_n$ if $\xi_n < \beta_n$ with $\beta_n < 1$. Then

$$\rho_n = E\left[e^{\xi_n a\Delta} - e^{\eta_n a\Delta}\right]$$
$$= \int_{[0,\beta_n)} e^{\xi a\Delta}\left[1 - e^{(\beta_n - \xi_n)a\Delta}\right] dG_n(\xi).$$

Hence

$$\|\rho_n\| \leqq \int_{[0,\beta_n)} \|1 - e^{(\beta_n - \xi)a\Delta}\| \, dG_n(\xi)$$
$$\leqq 2\int_{[0,\beta_n)} \left[1 - e^{-(\beta_n - \xi)a}\right] dG_n(\xi)$$
$$\leqq 2a \int_{[0,\beta_n)} (\beta_n - \xi)\, dG_n(\xi).$$

If $a > 1$, the same inequalities give

$$\|\rho_n\| \leqq 2\int_{[0,\beta_n)} dG_n(\xi).$$

Similarly, let $\eta_n = \xi_n$ if $\xi_n \leq \gamma_n$ and let $\eta_n = \gamma_n$ otherwise. Take $\gamma_n > 1$. One can write

$$\rho_n = \int_{(\gamma_n,\infty)} e^{\gamma_n a\Delta}\left[e^{-(\xi-\gamma_n)a\Delta} - 1\right] dG_n(\xi)$$

For all values of a this gives:

$$\|\rho_n\| \leq 2\int_{(\gamma_n,\infty)} \left[1 - e^{-(\xi-\gamma_n)a}\right] dG_n(\xi)$$

$$\leq 2\int_{(\gamma_n,\infty)} \min\left[(\xi-\gamma_n)a,\ 1\right] dG_n(\xi).$$

For large values of a it may be possible to obtain better bounds. Let $\varepsilon_{n,j}$ be defined by

$$\varepsilon'_{n,j} = \min\left\{a_{n,j}\int_{[0,\beta_n)} (\beta_n-\xi)\, dG_n(\xi),\ \int_{[0,\beta_n)} dG_n(\xi)\right\}$$

$$\varepsilon''_{n,j} = \int_{(\gamma_n,\infty)} \min\left[(\xi-\gamma_n)a_{n,j}, 1\right] dG_n(\xi)$$

$$\varepsilon_{n,j} = \varepsilon'_{n,j} + \varepsilon''_{n,j}.$$

Let Q'_n be the distribution obtained when ξ_n is replaced by the variable ξ'_n defined by

$$\xi'_n = \begin{cases} \beta_n & \text{if } \xi_n \leq \beta_n < 1 \\ \xi_n & \text{if } \beta_n \leq \xi_n \leq \gamma_n \\ \gamma_n & \text{if } \gamma_n \leq \xi_n \end{cases}$$

If $\Sigma_j\ \varepsilon_{n,j} \longrightarrow 0$ then $\|Q'_n - Q_n\| \longrightarrow 0$.

In the following sections we shall almost always substitute Q'_n with Q_n.

For the variables $X_{n,j}$, note that for Poisson distributions

$$P_\lambda\{(m,\infty)\} = \int_0^\lambda e^{-u}\frac{u^m}{m!}\,du \leq e^{-\lambda}\frac{\lambda^{m+1}}{m!} \leq \frac{\lambda^m}{m!}e^{-\lambda}$$

$$m! \geq \sqrt{2\pi m}\; e^{-m}\, m^m .$$

Consequently, for variables ξ_n such that $\xi_n \leq \gamma_n$

$$\sum_j \operatorname{Prob}[X_{n,j} \geq m_n] \leq \frac{\gamma_n}{\sqrt{2\pi m_n}}\sum_j a_{n,j}\, e^{-\gamma_n a_{n,j}}\left[a_{n,j}\gamma_n e\right]^{m_n}$$

$$\leq \frac{1}{\sqrt{2\pi m_n}}\left[\frac{\gamma_n e}{m_n}\right]^{m_n}\left[\sum_j a_{n,j}^{m_n}\right]$$

4. Mixtures with bounded expectations. As a particular case of the general situation described in section 3, consider the case where $a_n = \sup_j a_{n,j}$ stays bounded. We shall be particularly interested in the situation where in addition both a_n and the ratio a_n'/a_n with $a_n' = \inf_j a_{n,j}$, stays bounded away from zero. To refer to this in a simple manner, let us say that condition (E_1) is satisfied if a_n and a_n' stay bounded away from zero and infinity.

As previously stated it will also be assumed that $s_n = \Sigma_j a_{n,j}$ tends to infinity.

In addition, throughout the present section, it will be assumed that the following condition holds.

(F) There is a sequence $\{\alpha_n\}$ a number b and a positive integer k such that

1) $\alpha_n \longrightarrow \infty$.
2) If $\tau_n = \alpha_n \log \xi_n$ then $|\tau_n| \leq b$.
3) $s_n\,\alpha_n^{-2k}$ remains bounded.

In the remainder of the present section, it will be convenient to denote by k an arbitrary fixed integer satisfying condition (F).

Let $X_{n,j}$ have a Poisson distribution with expectation $a_{n,j}$. The moment $E[X_{n,j} - a_{n,j}]^r$ is a polynomial of degree at most r whose term of lowest degree is $a_{n,j}$. Therefore, if a_n stays bounded there exists coefficients C_k such that

$$\sum_j E[X_{n,j} - a_{n,j}]^{2k} \leq C_k s_n.$$

For every $\varepsilon > 0$ there exists an $n(\varepsilon)$ such that $n \geq n(\varepsilon)$ implies $\varepsilon\alpha_n > a_n \exp(b/\alpha_n) + 1$. For such values of n the inequality $|X_{n,j} - a_{n,j}| \leq \varepsilon\alpha_n$ is equivalent to the one-sided inequality $X_{n,j} \leq a_{n,j} + \varepsilon\alpha_n$.

Taking $\gamma_n = \exp(b/\alpha_n)$ one can apply the truncation inequalities given in the preceding section. These inequalities imply that, for Q_n as well as for P_n, one has

$$\sum_j \text{Prob}[X_{n,j} > \varepsilon\alpha_n] \leq \frac{C}{\sqrt{\alpha_n \varepsilon}} s_n \left(\frac{3a_n}{\varepsilon\alpha_n}\right)^{\varepsilon\alpha_n}$$

as soon as $e \exp(b/\alpha_n) \leq 3$. For n large this implies

$$\sum_j \text{Prob}[X_{n,j} > \varepsilon\, \alpha_n] \leq \frac{C}{\sqrt{\varepsilon\alpha_n}} s_n e^{-\varepsilon\alpha_n}.$$

According to the boundedness condition $\sup_n s_n \alpha_n^{-2k} < \infty$ the above sum tends to zero as $n \longrightarrow \infty$. Let P_n', (resp (Q_n') be the distributions obtained from P_n (resp. Q_n) by replacing $X_{n,j}$ by zero whenever $X_{n,j} > \varepsilon\, \alpha_n$. The differences $\|P_n - P_n'\|$ and $\|Q_n - Q_n'\|$ tend to zero as n increases. Consequently, it will be sufficient to investigate the behavior of the measures P_n' and Q_n'.

Let $Y_{n,j} = 1/\alpha_n\, [X_{n,j} - a_{n,j}]$. Let $U_{n,j} = Y_{n,j}$ if $|Y_{n,j}| \leq \varepsilon$ and let $U_{n,j} = 0$ otherwise.

Let H_n be the distribution of τ_n corresponding to $G_n = \mathcal{L}(\xi_n)$. Let $\beta_{n,j}(\tau) = \log E\, e^{\tau U_{n,j}}$, the expectation being taken under the assumption that the distribution of $U_{n,j}$ is given by P_n'. Let $M_{n,j}$ be the measure defined by

$$M_{n,j}(S) = \int_S \exp[-\beta_{n,j}(\tau)]\, dH_n(\tau).$$

The density of Q_n' with respect to P_n' is the product of the functions $\psi_{n,j}$ defined by

$$\psi_{n,j}(u) = \int \exp \tau u \; dM_{n,j}(\tau).$$

Let $\psi_{n,j} = \psi_{n,j}(U_{n,j})$ where $U_{n,j}$ has the distribution corresponding to the truncated Poisson P_n'. We shall now investigate the limiting distribution of

$$S_n = \sum_j [\psi_{n,j} - 1]$$

$$= \sum_j \int \left\{ \exp[\tau U_{n,j}] - E \exp[\tau\, U_{n,j}] \right\} dM_{n,j}(\tau).$$

Let

$$B_{n,m,j} = \int \tau^m \; dM_{n,j}(\tau),$$

$$V_{n,m,j} = \left[U_{n,j}^m - E\, U_{n,j}^m \right] B_{n,m,j}.$$

According to Taylor's formula

$$[\psi_{n,j} - 1] = \sum_{m=1}^{2k} \frac{1}{m!} V_{n,m,j}$$

$$+ \frac{1}{(2k-1)!} \int_0^1 (1-v)^{2k-1} dv \int \tau^{2k}\; W_{n,j}(\tau,v)\, dM_{n,j}(\tau),$$

with

$$W_{n,j}(\tau,v) = U_{n,j}^{2k}\left[e^{\tau v U_{n,j}} - 1\right] - E\, U_{n,j}^{2k}\left[e^{\tau v U_{n,j}} - 1\right].$$

Let

$$B_{n,2} = \int \tau^2 \, dH_n(\tau).$$

Lemma 3. Assume that (F) is satisfied and that a_n stays bounded. Furthermore, assume that the sequence of distributions

$$\mathcal{L}\left\{B_{n,2} \sum_j \left[U_{n,j}^2 - E\, U_{n,j}^2\right]\right\}$$

is relatively compact. Then

$$S_n - \sum_j V_{n,1,j} - \frac{1}{2} B_{n,2} \sum_j \left[U_{n,j}^2 - E\, U_{n,j}^2\right]$$

converges to zero in P_n' probability. Furthermore, if S_n has a limiting distribution, it is a normal distribution having expectation, equal to zero.

Proof. Let us show first that assumption (F) and the boundedness of a_n imply that

$$S_n - \sum_j \sum_{m=1}^{k} \frac{1}{m!} V_{n,m,j}$$

converges in probability to zero. For this purpose note that

$$\sum_j E\, U_{n,j}^{2k} \leq C_k\, s^k\, \alpha_n^{-2k}$$

remains bounded. Furthermore, for every positive δ the sum $\Sigma_j \operatorname{Prob}\{|U_{n,j}| > \delta\}$ converges to zero. It follows that the integral terms of the Taylor expansion of S_n converges in probability to zero. Consider now a term of the form

$$V_{n,m} = \sum_j V_{n,m,j}$$

for $m > k$. Since the coefficients $B_{n,m,j}$ are bounded, the variance of $V_{n,m}$ is bounded by an expression of the type

$$\sum_j E(U_{n,j}^{2m}) \leq \frac{C_m s_n}{\alpha_n^{2m}}.$$

Since this converges to zero and since the $V_{n,m,j}$ are bounded variables having expectation equal to zero the sum $V_{n,m}$ converges in probability to zero. Therefore,

$$S_n - \sum_j \sum_{m=1}^{k} \frac{1}{m!} V_{n,m,j}$$

converges in probability to zero.

Taking $k \geq 2$ one can write

$$\sum_j E\, Y_{n,j}^2 \; I[|Y_{n,j}| > \delta] \leq \frac{C_k s_n \alpha_n^{-2k}}{\delta^{2k-2}},$$

for $Y_{n,j} = 1/\alpha_n (X_{n,j} - a_{n,j})$. It follows that

$$B_{n,2}^2 \text{ var } \sum_j [Y_{n,j}^2 - E\, Y_{n,j}^2]$$

and variance $V_{n,2}$ are simultaneously bounded or unbounded. Consequently, if $\{\mathcal{L}(V_{n,2})\}$ is relatively compact then

$$\frac{1}{\alpha_n^4} B_{n,2}^2 s_n$$

must remain bounded. This implies that

$$\begin{aligned}\sum_j E\, V_{n,m,j}^2 &\leq \left[\sup_j B_{n,m,j}^2\right] \sum_j E\, U_{n,j}^{2m} \\ &\leq \left[\sup_j B_{n,m,j}^2\right] C_m s_n \alpha_n^{-2m} \\ &\leq C_m \left[b^{m-2}\right]^2 \left[B_{n,2}^2 s_n \alpha_n^{-4}\right] \alpha_n^{-2(m-2)},\end{aligned}$$

converges to zero for every $m > 2$. Finally

$$S_n - V_{n,1} - \frac{1}{2} V_{n,2}$$

converges in probability to zero as claimed. The limit points of the sequence $\{\mathcal{L}[V_{n,2}]\}$ are necessarily normal distributions with expectation zero.

To evaluate the limiting distribution of $V_{n,1}$ note that such a limit can exist only if $\Sigma_j |B_{n,1,j}|^2 E[U_{n,j} - E U_{n,j}]^2$ remains bounded. In this case the uniform asymptotic negligibility of the variables $U_{n,j}$ implies again that the limiting distribution is normal with expectation zero. This completes the proof of the lemma.

One may inquire whether the sum

$$\sum_j \sum_{m=1}^{k} \frac{1}{m!} V_{n,m,j}$$

may have a limiting distribution in cases where $V_{n,2}$ does not. The following argument shows that in the situation covered by (E_1) such a possibility is highly improbable. If n is large the covariance matrix of the sums $V_{n,m}$ is not singular. Let $\sigma^2_{n,m}$ = variance $V_{n,m}$ and let $\sigma_n = \sup_m \sigma_{n,m}$. If σ_n tends to infinity the joint distribution of the variables

$$\left\{\frac{1}{\sigma_n} V_{n,m}\right\}$$

can be approximated by a normal distribution with expectation equal to zero. This normal distribution is either nonsingular or is such that some coordinates are zero while the other coordinates have a nonsingular distribution. In any event the sum

$$\sum_j \sum_m \frac{1}{\sigma_n m!} V_{n,m,j}$$

is approximately normal and nondegenerate. Therefore, the original sum cannot be expected to have a limiting distribution.

To complete this section let us show that, under the conditions of lemma 3 we may substitute to the variables $U_{n,j}$ the nontruncated variables $Y_{n,j}$ without affecting the limiting behavior of S_n. Let m be an integer $m \leq 2$ and let k_1 be larger than or equal to 2. Then

$$E\{|Y_{n,j}|^m I[|Y_{n,j}| > \varepsilon]\} \leq \varepsilon^{m-2k_1} E|Y_{nkj}^{2k_1}|$$

$$\leq \varepsilon^{m-2k_1} \alpha_n^{-2k_1} C_{k_1} a_{n,j}.$$

Therefore,

$$\sum_j E\{|Y_{n,j}|^m I[|Y_{n,j}| > \varepsilon]\} \leq C_{k_1} \varepsilon^{m-2k_1} s_n \alpha_n^{-2k_1}$$

taking $k_1 > k$ tives the desired result.

Collecting the results just established one obtains the following theorem.

Theorem 2. Assume that $\sup a_n < \infty$ and that condition (F) is satisfied. Furthermore, assume that $B_{n,2}^2 s_n \alpha_n^{-4}$ stays bounded. Let $\omega_{n,j}$ be the density of $q_{n,j}$ with respect to $p_{n,j}$ considered as a random variable for the Poisson measure P_n.

Let $W_n = \Sigma_j(\omega_{n,j}-1)$. Then

$$\sum_j(\omega_{n,j}-1) - \sum_j B_{n,1,j} Y_{n,j}$$

$$- \frac{1}{2} B_{n,2} \sum_j [Y_{n,j}^2 - E\, Y_{n,j}^2]$$

converges in probability to zero. The sequence of distributions $\mathcal{L}(W_n)$ is relatively compact if and only if $1/\alpha_n^2 \Sigma_j B_{n,1,j}^2 a_{n,j}$

stays bounded. Under this supplementary restriction, the cluster points of the sequence $\mathcal{L}(W_n)$ are normal distributions whose expectation is zero and the sequences $\{P_n\}$ and $\{Q_n\}$ are contiguous.

Note 1. The preceding theorem applies in particular to the cases where $s_n \alpha_n^{-4}$ stays bounded, but this condition is not necessary.

Note 2. The domain of validity of theorem 2 can be substantially extended by means of truncation procedures applied to ξ_n as indicated in section 3.

Note 3. Under the conditions of theorem 2, assume that W_n has a limiting distribution. Then there are numbers c_n such that

$$\log \frac{dQ_n}{dP_n} - c_n - W_n$$

converges in probability to zero.

The expression of $\Sigma[\omega_{n,j}-1]$ can be simplified further as follows. Let $B_n = \int \tau dH_n(\tau)$. Assume that the conditions of theorem 2 are satisfied and that $\{\mathcal{L}(W_n)\}$ is a relatively compact sequence. Consider the sum

$$\sum_j [B_{n,1,j} - B_n] Y_{n,j}.$$

This sum has expectation zero and variance

$$\sum_j a_{n,j} \frac{1}{\alpha_n^2} [B_{n,1,j} - B_n]^2.$$

The difference $B_n - B_{n,1,j}$ may be written

$$B_n - B_{n,1,j} = \int \tau\left[1 - e^{-\beta_{n,j}(\tau)}\right] dH_n(\tau)$$

$$= \int e^{-\beta_{n,j}(\tau)} \tau\left[e^{\beta_{n,j}(\tau)} - 1\right] dH_n(\tau)$$

Consequently there is a constant K such that

$$\left|B_n - B_{n,1,j}\right|^2 \leq K\int \tau^2 \, dH_n(\tau)\int \left[e^{\beta_{n,j}(\tau)} - 1\right]^2 dH_n(\tau)$$

$$\leq K\, B_{n,2}\int \left[e^{\beta_{n,j}(\tau)} - 1\right]^2 dH_n(\tau).$$

The exponential $e^{\beta_{n,j}(\tau)}$ is, by definition, equal to $E\, e^{\tau U_{n,j}}$ where $U_{n,j}$ is a bounded variable.

Taylor's formula gives

$$\exp\left[\beta_{n,j}(\tau)\right] = 1 + \tau\, E\, U_{n,j} + \frac{\tau^2}{2}(E\, U_{n,j}^2)\varphi_{n,j}(\tau)$$

where $\varphi_{n,j}(\tau)$ is bounded by some constant K_1. Therefore,

$$\left[e^{\beta_{n,j}(\tau)} - 1\right]^2 \leq \tau^2(E\, U_{n,j})^2 + |\tau|^3(E\, U_{n,j})E(U_{n,j}^2)K_1 + \frac{\tau^4}{4} K_1^2 \left[E(U_{n,j}^2)\right]^2.$$

Reverting to the Poisson variables it is easy to show that $|EU_{n,j}|$ is smaller than α_n^{-2} multiplied by a term which decreases like $e^{-\varepsilon\alpha_n}$. Finally

$$|B_n - B_{n,j}|^2 \leq \delta_n\, B_{n,2}^2\, \alpha_n^{-2}$$

where $\delta_n \longrightarrow 0$. Since $B_{n,2}^2\, s_n\, \alpha_n^{-4}$ is bounded, the sum

$$\sum_j \left[B_{n,1,j} - B_n\right] Y_{n,j}$$

converges in probability to zero and $\Sigma_j\left[\omega_{n,j} - 1\right]$ differs from

$$T_n = \frac{B_n}{\alpha_n}\sum_j \left[X_{n,j} - a_{n,j}\right] + \frac{B_{n,2}}{2\alpha_n^2}\sum \left[Z_{n,j}^2 - a_{n,j}\right]$$

(with $Z_{n,j} = X_{n,j} - a_{n,j}$) by a quantity which tends to zero in probability.

5. Applications to the construction of tests of homogeneity. In this section it will be assumed that the conditions of Theorem 2 are satisfied and that $\{\mathcal{L}(W_n)\}$ is a relatively compact sequence of distributions.

Under these conditions the random part of $\log dQ_n/dP_n$ is equivalent to

$$W_n' = R_n \sum_j X_{n,j} + R_n' \sum_j (X_{n,j} - a_{n,j})^2$$

where R_n and R_n' are constants such that $R_n' \geq 0$.

It follows from this that asymptotically optimum tests of the hypothesis (P_n) against the alternative (Q_n) can be constructed by rejecting P_n if W_n' is too large.

The difficulty in the application of this result to tests of homogeneity lies in the fact that the hypothesis of homogeneity does not specify the values of the $a_{n,j}$ entirely.

In the situation described in section 3 where the $a_{n,j}$ are proportional to some λ_n multiplied by the length of the period of observation one may consider that the ratios of the $a_{n,j}$ to their sum are given. This will be assumed here.

To obtain some results in such circumstances, let $\{a_{n,j}\}$ be a fixed sequence of numbers and introduce the following variables

1) $Z_{n,j} = X_{n,j} - a_{n,j}$.

2) $U_n = \frac{1}{\sqrt{s_n}} \sum_j Z_{n,j}$.

3) $V_n = \frac{1}{\sqrt{s_n}} \sum_j [Z_{n,j}^2 - a_{n,j}]$.

Under the assumptions made here if $X_{n,j}$ has a Poisson distribution with expectation $a_{n,j}$ then U_n and V_n have a joint distribution whose distance to a centered normal distribution tends to zero as $n \longrightarrow \infty$.

Consider the following local problem. Under the hypothesis tested the variables $X_{n,j}$ have Poisson distributions with expectations $\lambda_n a_{n,j}$ such that $\lambda_n \Sigma a_{n,j} = s_n + w\sqrt{s_n}$ with w finite.

Under the alternatives the $X_{n,j}$ have distributions which are mixtures of Poisson distributions corresponding to measures H_n which satisfy the conditions of Theorem 2 and the requirement that $\{\mathcal{L}(W_n)\}$ be relatively compact.

Let P_n be the distribution corresponding to the Poisson sequence $\{a_{n,j}\}$ itself. The log of the likelihood ratio, $\log dQ_n/dP_n$, of another one of the measures considered here to P_n can be written as

$$T_n = u_n U_n + v_n V_n + c_n + \varepsilon_n$$

where ε_n tends to zero in probability and c_n is nonrandom. For instance, the case of Poisson variables with expectations $\lambda_n a_{n,j}$ such that $\lambda_n \Sigma a_{n,j} = s_n + w\sqrt{s_n}$ corresponds to $u_n = w$ and $v_n = 0$. The whole family of approximate likelihood ratios can then be indexed by the two parameters u and v, coefficients of U_n and V_n in T_n.

It follows that the family of distributions considered here is asymptotically normal in the sense of [6]. An asymptotically best asymptotically similar test of the hypothesis of homogeneity corresponds to a test that $v = 0$ against $v > 0$ for the normal approximation.

Let ρ_n be the regression coefficient of V_n on U_n. The best similar test of the hypothesis, under normal assumptions would be to reject the hypothesis if $V_n - \rho_n U_n \geq \gamma_n$ for a suitably selected number γ_n. Consequently, the same test is asymptotically best for the local problem envisaged here. The test statistic indicated by this reasoning is

$$\theta_n = V_n - U_n = \frac{1}{\sqrt{s_n}} \Sigma \left[Z_{n,j}^2 - a_{n,j} - Z_{n,j}\right]$$

$$= \frac{1}{\sqrt{s_n}} \Sigma \left[(X_{n,j} - a_{n,j})^2 - X_{n,j}\right].$$

As the last expression indicates θ_n can be computed only if the $a_{n,j}$ are actually known. To be able to test the global hypothesis instead of the local one suppose that S_n is an estimate of s_n and take $\hat{a}_{n,j} = (S_n/s_n) a_{n,j}$ as estimate of $a_{n,j}$. When the ratios $a_{n,j}/s_n$ are known the estimate $\hat{a}_{n,j}$ can actually be computed.

Let

$$\theta_n^* = \frac{1}{\sqrt{S_n}} \sum_j \left\{[X_{n,j} - \hat{a}_{n,j}]^2 - X_{n,j}\right\}$$

and let $\hat{\theta}_n = \sqrt{S_n/s_n}\, \theta_n^*$. Then

$$\hat{\theta}_n - \theta_n = \frac{1}{\sqrt{s_n}} \left[\frac{S_n - s_n}{s_n}\right]^2 \Sigma\, a_{n,j}^2$$

$$- \frac{s}{\sqrt{s_n}} \frac{S_n - s_n}{s_n} \sum_j a_{n,j} Z_{n,j}$$

the term $\Sigma\, a_{n,j} Z_{n,j}$ has expectation zero and variance $\Sigma_j\, a_{n,j}^3 \leq a_n^2 \sum_j a_{n,j} = a_n^2 s_n$. Therefore,

$$|\hat{\theta}_n - \theta_n| \leq \frac{[S_n - s_n]^2}{s_n\sqrt{s_n}} a_n + 2 \left|\frac{S_n - s_n}{s_n}\right| a_n |\xi_n|$$

where $E|\xi_n|^2 \leq 1$. Since a_n stays bounded the difference $\hat{\theta}_n - \theta_n$ will tend to zero in probability provided that $(S_n - s_n)/s_n^{3/4}$ converges in probability to zero. Such a condition is satisfied for instance if S_n is taken equal to $\Sigma_j X_{n,j}$.

Summarizing, let $r_{n,j} = a_{n,j} s_n^{-1}$ and let S_n be an estimate of s_n such that $[S_n - s_n] s_n^{-3/4}$ converges in probability to zero. For instance S_n may be taken equal to $\Sigma_j X_{n,j}$. Let θ_n^* be equal to

$$\theta_n^* = \frac{1}{\sqrt{S_n}} \sum_j \{[X_{n,j} - S_n r_{n,j}]^2 - X_{n,j}\}.$$

Then $\theta_n^* - \theta_n$ tends in probability to zero as $n \longrightarrow \infty$.

This convergence property has been proved above for the Poisson measures having expectations $a_{n,j}$. Because of the contiguity properties insured by theorem 2 the convergence to zero of $[S_n - s_n] s_n^{-3/4}$ and $\theta_n^* - \theta_n$ will also take place for all mixtures corresponding to measures H_n which satisfy all the requirements of theorem 2.

Under the hypothesis tested the distribution of θ_n is approximately normal with expectation zero and variance $1 + 2(\Sigma\, a_{n,j}^2) s_n^{-1}$. It will be convenient to use instead of θ_n^* the statistic T_n^* defined by

$$T_n^* = \frac{\theta_n^*}{\sqrt{1 + 2S_n \Sigma\, r_{n,j}^2}}.$$

Under the hypothesis tested T_n^* is asymptotically normal with mean zero and variance unity. Finally the results just established can be summarized as follows.

Theorem 3. Let $\{\{X_{n,j}\}$ be a double sequence of random variables. Assume that under the hypothesis tested the $X_{n,j}$ are independent Poisson variables with expectations $a_{n,j} = s_n r_{n,j}$ where the $r_{n,j}$ are known and s_n is unknown but tends to infinity. Assume also that $\sup_{n,j} s_n r_{n,j} < \infty$.

Consider alternative hypotheses where the $X_{n,j}$ are independent but have distributions $q_{n,j}$ which are Poisson mixtures

$$q_{n,j} = \int e^{\xi a_{n,j}\Delta} \, dG_n(\xi)$$

such that assumption (F) of section 4 is satisfied for some sequence α_n. Let

$$u_n = \sqrt{s_n} \int [\log \xi] \, dG_n(\xi)$$

$$v_n = \sqrt{s_n} \int [\log \xi]^2 dG_n(\xi)$$

$$S_n = \sum_j X_{n,j}$$

$$T_n^* = [S_n + 2S_n^2 \Sigma \, r_{n,j}^2]^{-1/2} \sum_j \; [X_{n,j} - S_n r_{n,j}]^2 - X_{n,j} \; .$$

Assume that for each sequence $\{G_n\}$ considered the parameters u_n and v_n stay bounded. Let γ be such that

$$\frac{1}{\sqrt{2\pi}} \int_\gamma^\infty e^{-\frac{u^2}{2}} du = \delta,$$

the test which rejects the Poisson hypothesis if $T_n^* \geq \gamma$ is asymptotically similar of size δ. Furthermore, this test is asymptotically most powerful among asymptotically similar tests.

To complete the statement of the theorem it remains to compute an expression for the power function of the test. For this purpose note that, according to the theorems proved in [6], if θ_n is asymptotically normal with mean zero and variance σ^2 for the Poisson variables then θ_n is asymptotically normal with the same variance σ^2 under the contiguous alternatives. The expectation of the limiting distribution of θ_n is equal to the limiting covariance of θ_n and $T_n = \log dQ_n/dP_n$. In the present case $T_n = u_n U_n + v_n V_n$. Hence the limiting covariance of T_n and θ_n is equal to the limit of $2v_n(\Sigma\, a_{n,j}^2)s_n^{-1}$.

Let $K_n = (\Sigma\, a_{n,j}^2)s_n^{-1}$. It follows that under the alternatives considered the distance between $\mathcal{L}(T_n^*)$ and the normal distribution with expectation

$$\mu_n = \frac{2\, v_n\, K_n}{\sqrt{1 + 2K_n}}$$

tends to zero as $n \longrightarrow \infty$. The power of the test is approximated by

$$\frac{1}{\sqrt{2\pi}} \int_{\gamma-\mu_n} e^{-\frac{u^2}{2}}\, du.$$

This expression is increasing in μ_n. Furthermore, for a fixed value of v_n, the expectation μ_n is increasing in K_n. This is understandable if one notes that

$$K_n = \sum_j \left[\frac{a_{n,j}}{s_n}\right] a_{n,j}$$

is an average of the value of the $a_{n,j}$.

REFERENCES

[1] J. DUBOURDIEU, Theorie Mathematique des Assurances, Paris, Gauthier-Villars (1952), 266 pp.

[2] L. LE CAM, "An approximation theorem for the Poisson-Binomial distribution," to be published.

[3] YU. V. PROHOROV, "Asymptotic behavior of the Binomial distribution," (Russian), Uspekhii Matematicheskiikh Nauk, Vol. 8, No. 3, Moscow, 1953, pp. 135-142.

[4] M. GREENWOOD and G. U. YULE, "An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attack of disease or of repeated accidents," J. Roy. Stat. Soc., Vol. 83 (1920), pp. 255-279.

[5] J. NEYMAN and E. L. SCOTT, "Test of the hypothesis that all the individuals of a given category are equally prone to accidents," Report on Air Force Contract AF.18(600)-456, "Care of the Flier-Accident Proneness," Part III, September, 1957.

[6] L. LE CAM, "Locally asymptotically normal families of distributions," University of Calif. Publ. in Statist. (in press).

[7] B. V. GNEDENKO and A. N. KOLMOGOROV, "Limit distributions for sums of independent random variables," translated by K. L. Chung, Addison-Wesley, Cambridge, Massachusetts, 1954 (p. 144).