ON A LOCALLY MOST POWERFUL BOUNDARY RANDOMIZED SIMILAR TEST FOR THE INDEPENDENCE OF TWO POISSON VARIABLES[1, 2]

By Mohamad Salahuddin Ahmed[3]

University of California, Berkeley

0. Summary. By definition (X, Y) is a bivariate Poisson vector if $(X, Y) = (X^* + U, Y^* + U)$ where X^*, Y^* and U are three independent Poisson variables with, say, respective expectations a, b and d.

Let $(X_n, Y_n)n = 1, 2, \cdots, N$ be independent observations on a bivariate Poisson vector (X, Y). It is shown that no test for the independence of X and Y can be both boundary randomized similar in a and b and also uniformly most powerful. However, a test of the form

$$\varphi_0(Z \mid s, t) = \begin{cases} 1 \\ \gamma_0 \\ 0 \end{cases} \text{according as} \quad Z = \sum_{n=1}^{N} X_n Y_n \gtreqless k_0(N, s, t)$$

given $S = s$ and $T = t$, where,

$$\sum_{n=1}^{N} X_n = S, \qquad \sum_{n=1}^{N} Y_n = T,$$

is boundary randomized similar and locally most powerful. Using a lemma on the convergence to a Normal probability distribution function of the conditional probability distribution function of $\sum_{n=1}^{N} (X_n - sN^{-1})(Y_n - tN^{-1})$ given $S = s$ and $T = t$, asymptotic formulae for the values of the $k_0(N, s, t)$ corresponding to a given level of significance are derived. In addition, it is shown that the asymptotic power of the test can be obtained from an approximation to a Normal probability function and that, in case instead of $k_0(N, s, t)$ its value calculated from the asymptotic formulae is used, the modified test is asymptotically locally most powerful in the sense of Definition 2.

To extend the domain of application of this test, we replace U in the definition of the bivariate Poisson vector by another random variable W also taking non-negative integral values according to the probability function

$$P\{W = w \mid \sigma \geqq 0\} = \int_0^\infty e^{-\sigma t} \frac{(\sigma t)^w}{w!} f(t)\, dt$$

Received January 27, 1958; revised November 22, 1960.

[1] This paper was prepared with the support of the Office of Ordnance Research, U. S. Army under Contract DA-04-200-ORD-171.

[2] This paper forms part of a doctoral thesis submitted to the Department of Statistics, University of California, Berkeley, California. Some of the results contained in this paper were announced at the Annual Membership meeting of the Institute of Mathematical Statistics, held in Berkeley, December, 1954.

[3] Now at College of Science, University of Baghdad, Baghdad, Iraq.

where $f(t)$ is any continuous probability function satisfying

$$\frac{\partial}{\partial\sigma}\int_0^\infty e^{-\sigma t}\frac{(\sigma t)^w}{w!}f(t)\,dt\bigg|_{\sigma=0}=\int_0^\infty\frac{\partial}{\partial\sigma}\left\{e^{-\sigma t}\frac{(\sigma t)^w}{w!}\right\}\bigg|_{\sigma=0}f(t)\,dt$$

$\sigma \geqq 0$, $w = 0, 1, 2, \cdots$ and $\int_0^\infty tf(t)\,dt < \infty$. In this way a class of bivariate probability functions is obtained such that for every member of this class also, regarded as a probability function of the random vector (X, Y), the locally most powerful boundary randomized similar test for the independence of the two random variables X and Y is the same as the one given in the Poisson case.

1. Introduction. In recent years the application of stochastic processes to problems in biology, physics, etc., has created a demand for new tests of independence. The present study is an attempt in this direction.

Under the hypothesis of independence we assume that the random variables X and Y are Poisson variables. The first problem is to define a bivariate Poisson vector (X, Y) when the Poisson variables X and Y are not independent. In Section 2 we give a definition of a bivariate Poisson vector. Equivalent definitions of a bivariate Poisson vector have been given earlier by several authors following different lines of attack (see Teicher [10], p. 2 and Loève [5], p. 84). Also, we consider some characteristics of the bivariate Poisson vector. In particular it is shown that the two Poisson variables X and Y are independent if and only if they are uncorrelated. In Section 3 a test for the independence of two Poisson variables is obtained and furthermore, the properties of the test, already enumerated in the Summary, are proved. Section 4 discusses the extension of the domain of application of the test of independence for two Poisson variables. An example is given for which the locally most powerful boundary randomized similar test of independence is not of the form obtained in the case of a bivariate Poisson vector.

2. The bivariate Poisson vector. The simplest definition of the bivariate Poisson vector is

DEFINITION 1. A bivariate random vector (X, Y) is a bivariate Poisson vector if

$$(X, Y) = (X^* + U, Y^* + U),$$

where X^*, Y^* and U are independent Poisson variables.

If the Poisson variables X^*, Y^* and U have respective expectations a, b and d, then the probability function of the bivariate Poisson vector (X, Y) is given by

$$(1)\qquad P\{X = x, Y = y\} = e^{-(a+b+d)}\,a^x\,b^y\sum_{u=0}^{u^*}\frac{(d/ab)^u}{(x-u)!u!(y-u)!},$$

where $u^* = \min(x, y)$. We shall refer to the system of parameters used in (1) as the system (a, b, d). A direct consequence of Definition 1 is

THEOREM 1. *If (X, Y) is a bivariate Poisson vector, then,*

(i) *the marginals X and Y are Poisson variables*;

(ii) *the correlation between X and Y is nonnegative*; *and*

(iii) *the Poisson variables X and Y are independent if and only if the correlation between X and Y is zero.*

Let $(X_n, Y_n)n = 1, 2, \cdots, N$ be independent observations on a bivariate Poisson vector (X, Y). Henceforth we shall use the notation

$$\chi = (X_1, Y_1, X_2, Y_2, \cdots, X_n, Y_n), \qquad S = \sum_{n=1}^{N} X_n, \qquad T = \sum_{n=1}^{N} Y_n.$$

We also introduce the system (a, b, c), where $d = abc$. With this notation we have

THEOREM 2. *For the system (a, b, c),*

(i) *the vector statistic (S, T) is a vector sufficient statistic for (a, b) whatever be c*; *and*

(*ii*) *if $c = 0$, the vector sufficient statistic (S, T) is also a vector complete sufficient statistic for (a, b).*

PROOF. The probability function of the bivariate Poisson vector (X, Y) corresponding to the system (a, b, c) is easily obtained by substituting $d = abc$ in (1). Accordingly, the joint conditional probability function of $(X_n, Y_n)n = 1, 2, \cdots, N$, given $S = s$, $T = t$, is

$$q_c(\chi \mid s, t) = \prod_{n=1}^{N} \sum_{u=0}^{U_n^*} \frac{c^u}{(X_n - u)!u!(Y_n - u)!} \div \sum_{S=s,\, T=t} \prod_{n=1}^{N} \sum_{u=0}^{U_n^*} \frac{c^u}{(X_n - u)!u!(Y_n - u)!}, \tag{2}$$

which is independent of a and b. This proves the sufficiency in (i) and (ii).

The proof of the completeness in (ii) depends on the fact that, when $c = 0$, the random variables S and T are independent Poisson variables with respective expectations Na and Nb. Now consider

$$\sum_{s=0}^{\infty} \sum_{t=0}^{\infty} g(s, t)e^{-Na-Nb} \frac{(Na)^s}{s!} \frac{(Nb)^t}{t!} \equiv 0,$$

for all nonnegative a and b. Viewing this series as a convergent power series in a and b, we see that $g(s, t) = 0$ almost everywhere. This proves completeness (see Lehmann and Scheffé [4], p. 311).

3. Test for independence. In this section we shall obtain a locally most powerful boundary randomized similar (LMP BRS) test for the independence of two Poisson variables and prove the nonexistence of the uniformly most powerful boundary randomized similar (UMP BRS) test. Also, we shall obtain the asymptotic distribution theory of the test statistic under the hypothesis and prove certain asymptotic properties of the test. Finally we shall discuss an application of this test.

3.1. LMP BRS test for independence. In the system (a, b, c) the test for the independence of two Poisson variables is equivalent to testing the hypothesis $c = 0$ against the alternative $c > 0$, whatever be a and b. Let $(X_n, Y_n)n = 1, 2, \cdots, N$ be independent observations on a bivariate Poisson vector (X, Y) with parameters (a, b, c). Finally, let $p_0(\chi)$ and $p_c(\chi)$ denote the likelihood functions respectively under the hypothesis and the alternative.

In order to get rid of the unknown (nuisance) parameters a and b occurring in $p_0(\chi)$ and $p_c(\chi)$ we shall use instead of these the conditional likelihood functions $q_0(\chi \mid s, t)$ and $q_c(\chi \mid s, t)$ respectively, when $S = s$, and $T = t$, which are obtainable from (2). The geometrical picture is of surfaces for each set of values of (S, T). To obtain the LMP similar test for the independence of X and Y, the method is to find a test function $\varphi(\chi \mid s, t)$, $0 \leqq \varphi(\chi \mid s, t) \leqq 1$, defined on the hypersurface given by $S = s$, $T = t$, such that whatever be a and b,

$$\sum \varphi(\chi \mid s, t) q_0(\chi \mid s, t) = \alpha, \qquad 0 < \alpha < 1$$

and

$$\frac{\partial}{\partial c} \sum \varphi(\chi \mid s, t) q_c(\dot{\chi} \mid s, t)|_{c=0}$$

is a maximum, where α is the preassigned size of the test. The summation is over the sample points on the hypersurface given by $S = s$, $T = t$. The meaning of this test function is that when χ is observed on the hypersurface defined by $S = s$, $T = t$, the hypothesis of independence is rejected with probability $\varphi(\chi \mid s, t)$. This procedure to obtain the LMP similar test is valid due to the conclusions obtained in Theorem 2 (also see Lehmann and Scheffé [4], pp. 311, 318). The solution to this problem is given (see Neyman and Pearson [7], p. 10) by

$$\varphi(\chi \mid s, t) = \begin{cases} 1 \\ \gamma \\ 0 \end{cases} \text{according as} \quad \frac{\partial}{\partial c} \log q_c(\chi \mid s, t)|_{c=0} \gtreqless k_1 .$$

Substituting the expression for $q_c(\chi \mid s, t)$ from (2) and after simplification of the required differentiation we obtain that the condition

$$\frac{\partial}{\partial c} \log q_c(\chi \mid s, t)\ |_{c=0} \gtreqless k_1$$

is equivalent to the condition

$$Z = \sum_{n=1}^{N} X_n Y_n \gtreqless k_0 = k_0(N, s, t). \tag{3}$$

In particular we notice that the test $\varphi(\chi \mid s, t)$ is a function of χ through Z only. Therefore, the test will be denoted by $\varphi_0(Z \mid s, t)$.

Since the sample space of the random variable Z is discrete, to obtain similar test it is sufficient to use randomization only on the critical value k_0 of Z; hence the name BRS test. We remark also that with this agreement there is a one-to-one correspondence between the BRS test and the BRS region. Henceforth we shall use these terms interchangeably. Finally we have,

THEOREM 3. *If* $(X_n, Y_n)n = 1, 2, \cdots, N$ *are independent observations on a bivariate Poisson vector* (X, Y) *with parameters* (a, b, c), *then the* LMP BRS *test for the independence of the two Poisson variables* X *and* Y *is given by the rule*:

$$\varphi_0(Z \mid s, t) = \begin{cases} 1 \\ \gamma_0 \\ 0 \end{cases} \text{according as} \quad Z \gtreqless k_0 = k_0(N, s, t). \tag{4}$$

The values of k_0 *and* γ_0 *are determined in such a way that*

$$\sum \varphi_0(Z \mid s, t) q_0(Z \mid s, t) = \alpha. \tag{5}$$

3.2. Nonexistence of the UMP BRS test. Let φ_0 and φ respectively be the LMP BRS and the UMP BRS tests of size α for testing $c = 0$ against $c > 0$. Let $\beta(\varphi_0; a, b, c)$ and $\beta(\varphi; a, b, c)$ respectively be the power of the tests φ_0 and φ against the alternative (a, b, c). Knowing that (S, T) is a vector sufficient statistic for (a, b) we can write for the test φ_0,

$$\beta(\varphi_0; a, b, c) = \sum_{s,t} \beta(\varphi_0; s, t, c) r(s, t; a, b, c) \tag{6}$$

where $\beta(\psi; s, t, c)$ is the conditional power of the test ψ on the hypersurface defined by $S = s$, $T = t$ against the alternative c, and $r(s, t; a, b, c)$ is the probability function of the vector statistic (S, T) when the parameters have the value (a, b, c). Similarly for the test φ we have

$$\beta(\varphi; a, b, c) = \sum_{s,t} \beta(\varphi; s, t, c) r(s, t; a, b, c). \tag{7}$$

THEOREM 4. *There exists no* UMP BRS *test for testing the independence of two Poisson variables* X *and* Y *based on the independent observations* $(X_n, Y_n)n = 1, 2, \cdots, N$ *made on the bivarite Poisson vector* (X, Y).

PROOF. Since a UMP test is also a LMP test

$$\frac{\partial}{\partial c} \beta(\varphi_0; a, b, c) \mid_{c=0} = \frac{\partial}{\partial c} \beta(\varphi; a, b, c)|_{c=0}. \tag{8}$$

But we can rewrite (8) using (6) and (7) as follows:

$$\begin{aligned} &\sum_{s,t} \{\beta'(\varphi_0; s, t) + \beta(\varphi_0; s, t, 0) g(s, t; a, b)\} r(s, t; a, b, 0) \\ &\qquad = \sum_{s,t} \{\beta'(\varphi; s, t) + \beta(\varphi; s, t, 0) g(s, t; a, b)\} r(s, t; a, b, 0) \end{aligned} \tag{9}$$

where

$$\beta'(\psi; s, t) = \frac{\partial}{\partial c} \beta(\psi; s, t, c)|_{c=0},$$

and

$$g(s, t; a, b) = \frac{\partial}{\partial c} \log r(s, t; a, b, c) \,|_{c=0} .$$

Furthermore, since the tests φ_0 and φ have Neyman structure, we have, for all possible values of s and t,

$$\beta(\varphi_0 ; s, t, 0) = \beta(\varphi; s, t, 0) = \alpha.$$

Therefore, we obtain from (9) that

$$\sum_{s,t} \{\beta'(\varphi_0 ; s, t) - \beta'(\varphi; s, t)\} r(s, t; a, b, 0) = 0.$$

This fact, together with the completeness of the vector sufficient statistic (S, T) for (a, b) whenever $c = 0$, proved in Theorem 2(ii), implies that

$$h(\varphi_0 , \varphi; s, t) = \beta'(\varphi_0 ; s, t) - \beta'(\varphi; s, t) = 0$$

almost everywhere. In fact, more is true, viz., $\varphi_0 = \varphi$ almost everywhere. The proof is by contradiction. Simple calculation shows that

$$(10) \qquad h(\varphi_0 , \varphi; s, t) = A^{-2}\{\sum \varphi_0 P(ZA - B) - \sum \varphi P(ZA - B)\},$$

where the summation is taken over the sample points on the hypersurface defined by $S = s$, $T = t$ and

$$Z = \sum_{n=1}^{N} X_n Y_n , \qquad P = \prod_{n=1}^{N} (X_n! Y_n!)^{-1}, \qquad A = \sum_{S=s,\, T=t} P$$

and

$$B = \sum_{S=s,\, T=t} PZ.$$

Using the form of the test φ_0 given in (4) we obtain from (10) that

$$h(\varphi_0 , \varphi; s, t) > A^{-2}(kA - B)(\sum \varphi_0 P - \sum \varphi P) = 0$$

which contradicts (8). Hence $\varphi_0 = \varphi$ almost everywhere.

Therefore, to show that there exists no UMP BRS test it is enough to show that φ_0 is not a UMP BRS test. We will show this numerically. For a specific alternative let $b = b$, $a \to 0$ and $c \to \infty$ in such a manner that in limit $a = 0$ and $abc = d$. Then using Definition 1 we can see that X_n and $Y_n - X_n$ are independent Poisson variables with respective expectations d and b. The most powerful conditional test against the alternative described above when $S = s$, $T = t$ is given by

$$\varphi(\chi \mid s, t) = \begin{cases} 1 \\ \gamma \\ 0 \end{cases} \text{according as } \prod_{n=1}^{N} Y_n!\{(Y_n - X_n)!\}^{-1} \gtreqless k(N, s, t).$$

Let $N = 2$, $s = 3$, $t = 4$ and $\alpha = \frac{1}{8}$. Then the test $\varphi_0(\chi \mid 3, 4)$ is 1 or 0 according as $\chi = (X_1, Y_1; X_2, Y_2)$ belongs to the set $\{(3, 4; 0, 0), (0, 0; 3, 4), (3, 3; 0, 1), (0, 1; 3, 3), (2, 4; 1, 0), (1, 0; 2, 4)\}$ or not. Also the test $\varphi(\chi \mid 3, 4)$ is such that it is equal to 1 if χ belongs to the set $\{(3, 4; 0, 0), (0, 0; 3, 4)\}$, it is equal to $\gamma = \frac{7}{16}$ if χ belongs to the set $\{(3, 3; 0, 1), (0, 1; 3, 3), (2, 3; 1, 1), (1, 1; 2, 3)\}$ and it is equal to zero for all other points on the hypersurface defined by $s = 3$, $t = 4$. We can easily check that the size of both these tests is $\frac{1}{8}$. Finally, simple calculation shows that the power of the tests φ_0 and φ respectively are $\frac{24}{96}$ and $\frac{38}{96}$, showing that the test φ_0 is not the UMP test. This completes the proof of the theorem.

3.3. Asymptotic distribution of the test statistic under the hypothesis. Define

$$S_N = N^{-\frac{1}{2}}\sum_{n=1}^{N} (X_n - a),$$

$$T_N = N^{-\frac{1}{2}}\sum_{n=1}^{N} (Y_n - b)$$

and

$$Z_N = N^{-\frac{1}{2}}\sum_{n=1}^{N} (X_n - a)(Y_n - b).$$

It follows from the central limit theorem for vectors (see Cramér [1], p. 316) that if $N \to \infty$ then

$$(11) \qquad Ee^{iuZ_N + i\xi S_N + i\eta T_N} \to e^{-\frac{1}{2}(abu^2 + a\xi^2 + b\eta^2)} = Ee^{iuZ_\infty + i\xi S_\infty + i\eta T_\infty},$$

for all u, ξ, η, where new random variables Z_∞, S_∞ and T_∞ are so defined that the equality sign holds. Using the notation,

$$E(e^{iuZ_k} \mid S_k = s, T_k = t) = \omega_k(u, s, t)$$

we obtain by rewriting (11)

$$(12) \qquad \int e^{i\xi s + i\eta t}\omega_N(u, s, t)\, dP_N(s, t) \to \int e^{i\xi s + i\eta t}\omega_\infty(u, s, t)\, dP_\infty(s, t)$$

for all u, ξ, η, where $P_k(s, t)$ is the probability distribution function of the bivariate random vector (S_k, T_k). Furthermore, if we define

$$Q_k(u, A) = \int_A \omega_k(u, s, t)\, dP_k(s, t),$$

then clearly $Q_k(u, A)$ is a complex measure satisfying

$$\sup_{k,A}|Q_k(u, A)| \leqq 1 \quad \text{for} \quad k = \infty, 1, 2, \cdots.$$

With this notation (12) can be written as

$$(13) \qquad \int e^{i\xi s + i\eta t}\, dQ_N(u, s, t) \to \int e^{i\xi s + i\eta t}\, dQ_\infty(u, s, t)$$

for all u, ξ, η.

If $f(s, t)$ is any bounded continuous function (see LeCam [3], p. 28) then (13) is equivalent to saying that

$$\int f(s, t)\, dQ_N(u, s, t) \to \int f(s, t)\, dQ_\infty(u, s, t)$$

for all u. Finally, reintroducing $\omega_k(u, s, t)$ we obtain that

$$\int f(s, t)\omega_N(u, s, t)\, dP_N(u, s, t) \to \int f(s, t)\, \omega_\infty(u, s, t)\, dP_\infty(u, s, t) \tag{14}$$

for all u. This indicates that $\omega_N(u, s, t)$ might tend pointwise to $\omega_\infty(u, s, t)$. Such is the case as will follow from Lemmas 1 and 2.

LEMMA 1. *If the distribution function of the random vector* (Z_N, S_N, T_N) *converges to the distribution function of the random vector* $(Z_\infty, S_\infty, T_\infty)$ *when* $N \to \infty$, *and if the family of the conditional characteristic functions* $\{\omega_N(u, s, t)\}$ *is uniformly bounded and is equicontinuous for all* u, s, t *and uniformly so for* u, s, t *bounded then,*

$$\omega_N(u, s, t) \to \omega_\infty(u, s, t)$$

as $N \to \infty$ *for all* u, s, t *and uniformly so for* u, s, t *bounded.*

PROOF. According to Ascoli's theorem (see Graves [2], p. 122) we can find a subsequence $\omega_{N_k}(u, s, t)$ and a continuous $\omega(u, s, t)$ such that

$$\omega_{N_k}(u, s, t) \to \omega(u, s, t) \tag{15}$$

for all u, s, t and uniformly so for u, s, t bounded. Replacing N by N_k in (14) we obtain

$$\int f(s, t)\omega_{N_k}(u, s, t)\, dP_{N_k}(s, t) \to \int f(s, t)\omega_\infty(u, s, t)\, dP_\infty(s, t) \tag{16}$$

for all u. Let us consider the equality

$$\begin{aligned} \int f(s, t)\omega_{N_k}(u, s, t)\, dP_{N_k}(s, t) &= \int f(s, t)\omega(u, s, t)\, dP_\infty(s, t) \\ &+ \int f(s, t)\omega(u, s, t)\, d\{P_{N_k}(s, t) - P_\infty(s, t)\} \\ &+ \int f(s, t)\{\omega_{N_k}(u, s, t) - \omega(u, s, t)\}\, dP_{N_k}(s, t). \end{aligned} \tag{17}$$

Using (15) and the fact that $f(s, t)$ is bounded, we can see that

$$\int f(s, t)\{\omega_{N_k}(u, s, t) - \omega(u, s, t)\}\, dP_{N_k}(s, t) \to 0$$

for all u when $N_k \to \infty$. Also due to the Helley-Bray theorem (see Loève [6], p. 182), since $P_{n_k}(s, t) \to P_\infty(s, t)$ for all values of s, t,

$$\int f(s, t)\omega(u, s, t)\, d\{P_{N_k}(s, t) - P_\infty(s, t)\} \to 0$$

for all u when $N_k \to \infty$. The above two conclusions, together with (17) implies that

$$\int f(s, t)\omega_{N_k}(u, s, t)\, dP_{N_k}(s, t) \to \int f(s, t)\omega(u, s, t)\, dP_\infty(s, t), \tag{18}$$

for all u. Comparing (16) and (18) we obtain that

$$\int f(s, t)\omega(u, s, t)\, dP_\infty(s, t) = \int f(s,t)\omega_\infty(u, s, t)\, dP_\infty(s, t)$$

for all values of u which implies that

$$\omega(u, s, t) = \omega_\infty(u, s, t) \text{ almost everywhere.} \tag{19}$$

The "almost everywhere" may be dropped since the functions under consideration are continuous. This proves that every convergent subsequence of $\{\omega_N(u, s, t)\}$ has the same limit. But more is true, viz.,

$$\omega_N(u, s, t) \to \omega(u, s, t) \tag{20}$$

for all u, s, t and uniformly so for u, s, t bounded. Suppose that

$$\omega_N(u, s, t) \nrightarrow \omega(u, s, t).$$

Then there exists subsequences $N_{k'}$ and $N_{k''}$ such that

$$\omega_{N_{k'}}(u, s, t) \to \omega'(u, s, t) = \limsup \omega_N(u, s, t)$$

and

$$\omega_{N_{k''}}(u, s, t) \to \omega''(u, s, t) = \liminf \omega_N(u, s, t)$$

for all u, s, t and uniformly so for u, s, t bounded, where $\omega'(u, s, t)$ and $\omega''(u, s, t)$ are not equal. In the proof given above if we replace the sequence N_k by the sequence $N_{k'}$ we obtain corresponding to (19) that

$$\omega'(u, s, t) = \omega(u, s, t)$$

for all u, s, t and similarly we obtain that

$$\omega''(u, s, t) = \omega(u, s, t),$$

contradicting our assumption that $\omega'(u, s, t)$ and $\omega''(u, s, t)$ are not equal. Therefore, (20) is true. This completes the proof of Lemma 1.

LEMMA 2. *The family $\{\omega_N(u, s, t)\}$ is uniformly bounded and is equicontinuous for all u, s, t and uniformly so for u, s, t bounded.*

PROOF. Since $\omega_N(u, s, t)$ are conditional characteristic functions it follows that

$$|\omega_N(u, s, t)| \leqq 1.$$

We remark that if $S_N = s$, then the X_n's, $n = 1, 2, \cdots, N$ are N-nomially distributed with probability of falling in any one of the N classes being N^{-1} and having parameter $k = N^{\frac{1}{2}}s + Na$ which is equal to the number of trials. Similarly, if $T_N = t$, then the Y_n's, $n = 1, 2, \cdots, N$ are N-nomially distributed with the same probability of falling in any one of the N classes and with parameter $l = N^{\frac{1}{2}}t + Nb$. If s and t are changed to s' and t' respectively, such that $s' > s$ and $t' > t$, then k and l will change respectively to say $k + k'$ and $l + l'$ and correspondingly X_n and Y_n will change to $X_n + X_n'$ and $Y_n + Y_n'$ where X_n' and Y_n' are completely independent N-nomials of the above type but with respective parameters k' and l'. Let

$$Z_N(k, l) = N^{-\frac{1}{2}}\sum_{n=1}^{N} (X_n - a)(Y_n - b).$$

It is easy to see that

$$\begin{aligned} Q_N &= Z_N(k + k', l + l') - Z_N(k, l) \\ &= N^{-\frac{1}{2}}\sum_{n=1}^{N} (X_n + X_n' - a)(Y_n + Y_n' - b) - N^{-\frac{1}{2}}\sum_{n=1}^{N} (X_n - a)(Y_n' - b) \\ &= N^{-\frac{1}{2}}\sum_{n=1}^{N} (X_n - a)Y_n' + N^{-\frac{1}{2}}\sum_{n=1}^{N} X_n'(Y_n - b) + N^{-\frac{1}{2}}\sum_{n=1}^{N} X_n'Y_n'. \end{aligned}$$

In order to prove the required equicontinuity, we shall first show that for an arbitrary $\epsilon > 0$, there exists an $\epsilon'(\epsilon, s, t, u, a, b)$ such that if

$$s' - s < \epsilon' \quad \text{and} \quad t' - t < \epsilon'$$

then for all N

$$|Q_N^*| < \tfrac{1}{2}\epsilon$$

where

$$\begin{aligned} |Q_N^*| &= |\, E\{e^{iuZ_N(k,l)}[Ee^{iuQ_N} - 1] \mid k, k', l, l'\}| \\ &\leqq |\, E\{[e^{iuQ_N} - 1] \mid k, k', l, l'\}| \\ &\leqq |\, E\left\{\left[iuQ_N - \tfrac{1}{2}u^2 Q_N^2 \int_0^1 e^{iuQ_N\xi}(1 - \xi)\, d\xi\right] \middle| k, k', l, l'\right\}| \\ &\leqq |u|\, |E\{Q_N \mid k, k', l, l'\}| + |u|^2\, |\, E\{Q_N^2 \mid k, k', l, l'\}|. \end{aligned}$$

By direct calculation it can be shown that

$$E\{Q_N \mid k, k', l, l'\} = N^{-\frac{1}{2}}(s't' - st)$$

and

$$\begin{aligned} E\{Q_N^2 \mid k, k', l, l'\} = (N - 1)N^{-\frac{3}{2}}\{b(s' - s) + a(t' - t)\} \\ + (N - 1)N^{-2}(s't' - st) + N^{-1}(s't' - st).^2 \end{aligned}$$

Therefore, with a suitable choice of $\epsilon' = \epsilon'(\epsilon, s, t, u, a, b)$

$$|Q_N^*| \leqq \tfrac{1}{2}\epsilon \quad \text{for all } N.$$

I.e., in terms of $\omega(u, s, t)$ we have shown that if

$$s' - s < \epsilon' \quad \text{and} \quad t' - t < \epsilon',$$

then

$$|\omega_N(u, s', t') - \omega_N(u, s, t)| \leqq \tfrac{1}{2}\epsilon \quad \text{for all } N.$$

For any two points (s_1, t_1) and (s_2, t_2) satisfying the conditions

$$|s_1 - s_2| < \epsilon' \quad \text{and} \quad |t_1 - t_2| < \epsilon'$$

we consider an auxillary point

$$(s, t) = (\min\{s_1, s_2\}, \min\{t_1, t_2\})$$

and then using the triangular inequality we obtain that

$$|\omega_N(u, s_1, t_1) - \omega_N(u, s_2, t_2)| \leqq |\omega_N(u, s_1, t_1) - \omega_N(u, s, t)| + |\omega_N(u, s, t) - \omega_N(u, s_2, t_2)| \leqq \tfrac{1}{2}\epsilon + \tfrac{1}{2}\epsilon = \epsilon \quad \text{for all } N.$$

This completes the proof of the equicontinuity of $\{\omega_N(u, s, t)\}$ for all u, s, t. It is easy to see that for bounded values of u, s, t the family $\{\omega_N(u, s, t)\}$ is uniformly continuous. The above type of reasoning was necessary to prove equicontinuity because only the sum of the multinomials with the same number of events is again a multinomial.

Therefore, in view of these lemmas, the needed asymptotic characteristic function is $\omega_\infty(u, s, t)$. We can easily check that

$$\omega_\infty(u, s, t) = \exp\{-abu^2/2\},$$

which is a characteristic function of a Normal variable with mean zero and variance ab. Finally, denoting the probability law of a random variable W by $\mathcal{L}(W)$ we can summarize the results obtained above as follows:

$$\mathcal{L}\{(ab)^{-\frac{1}{2}}Z_N \mid S_N = s, T_N = t\} = \mathcal{L}\{(\text{Nab})^{-\frac{1}{2}}\sum_{n=1}^{N}(X_n - a)(Y_n - b) \mid \sum_{n=1}^{N}(X_n - a) = N^{\frac{1}{2}}s, \sum_{n=1}^{N}(Y_n - b) = N^{\frac{1}{2}}t\} \rightarrow N(0, 1)$$

for all values of s, t and uniformly so for s, t bounded.

But to make use of this result we have to know the values of a and b. We recall that

$$\sum_{n=1}^{N} X_n = N\bar{X} \quad \text{and} \quad \sum_{n=1}^{N} Y_n = N\bar{Y}.$$

Using Slutsky's theorem (see Cramér [1], p. 254) it follows that the sequence

$$\{(N\bar{X}\bar{Y})^{-\frac{1}{2}}\sum_{n=1}^{N}(X_n - \bar{X})(Y_n - \bar{Y}) \mid \sum_{n=1}^{N}(X_n - a) = N^{\frac{1}{2}}s,$$
$$\sum_{n=1}^{N}(Y_n - b) = N^{\frac{1}{2}}t\}$$

also have the same limiting distribution $N(0, 1)$, for all values of s, t and for s, t bounded the convergence is uniform (see LeCam [3], p. 24). In what follows we shall use the notation

$$\tilde{Z}_N = (N\bar{X}\bar{Y})^{-\frac{1}{2}}\sum_{n=1}^{N}(X_n - \bar{X})(Y_n - \bar{Y}).$$

This gives the following

THEOREM 5. *When* $N \to \infty$,

$$\mathcal{L}(\tilde{Z}_N \mid s, t) \to N(0, 1)$$

for all s, t *and the convergence is uniform for* s, t *bounded.*

In order to use the test of independence developed in Section 3.1 we must find first the values of k_0 and γ_0 satisfying (5). In the absence of the knowledge of the exact distribution function of Z we propose to use instead the test statistic $\tilde{Z}_N$ and its limiting distribution obtained above. For a given level of significance α, first we find the value of k such that

$$(2\pi)^{-\frac{1}{2}}\int_{-\infty}^{k} e^{-\frac{1}{2}u^2}\,du = \alpha,$$

and then the modified test is given by the rule:

$$\tilde{\varphi}(\tilde{Z}_N \mid s, t) = \begin{cases} 1 \\ 0 \end{cases} \text{according as} \quad \tilde{Z}_N \gtrless k.$$

It is important to note that the condition of equicontinuity in the proof of Theorem 5 is not only sufficient but is also necessary in a certain sense. In applications, s and t are in fact functions of N say s_N, t_N, which will tend to s and t when $N \to \infty$. In the calculation of probabilities, etc., though we propose to use the asymptotic distribution but instead of s and t we have to use the values s_N and t_N respectively. Equicontinuity will ensure us that the values calculated with this substitution do not differ too much from the true values.

3.4. Asymptotic power function. The power of the modified test against the alternative (a, b, c) is

$$\beta(a, b, c) = \int P\{\tilde{Z}_N > k \mid s, t, a^*, b^*, d\}\, dP_N(s, t \mid a^*, b^*, d)$$

where

$$a^* = a + d, \qquad b^* = b + d, \qquad d = abc.$$

It is interesting to note ahead that we can compute the asymptotic power of the modified test without using the conditional probability distribution function of $\tilde{Z}_N$ given $S_N = s$, $T_N = t$. If we define

$$Z_N^* = N^{-\frac{1}{2}}\sum_{n=1}^{N}\{(X_n - a^*)(Y_n - b^*) - d\},$$

then it follows from the central limit theorem that under the alternative (a, b, c),

$$\mathfrak{L}(Z_N^*) \to N(0, \sigma^2) \tag{21}$$

where $\sigma^2 = a^*b^* + d^2 + d$. The decisive step now is to be able to represent $\tilde{Z}_N$ in terms of Z_N^*. It is easy to see that

$$\begin{aligned} &N^{-\frac{1}{2}}\sum_{n=1}^{N}(X_n - \bar{X})(Y_n - \bar{Y}) \\ &= N^{-\frac{1}{2}}\sum_{n=1}^{N}\{(X_n - a^*) - (\bar{X} - a^*)\}\{(Y_n - b^*) - (\bar{Y} - b^*)\} \\ &= Z_N^* + dN^{\frac{1}{2}} - (\bar{Y} - b^*)S_N^* - (\bar{X} - a^*)T_N^* + N^{-\frac{1}{2}}S_N^*T_N^* \end{aligned} \tag{22}$$

where

$$S_N^* = N^{-\frac{1}{2}}\sum_{n=1}^{N}(X_n - a^*) = S_N - dN^{\frac{1}{2}} = s - dN^{\frac{1}{2}}$$

and similarly,

$$T_N^* = N^{-\frac{1}{2}}\sum_{n=1}^{N}(Y_n - b^*) = T_N - dN^{\frac{1}{2}} = t - dN^{\frac{1}{2}}.$$

Under the alternative (a, b, c) when $N \to \infty$ we have,

$$\bar{X} - a^* \to 0, \qquad \bar{Y} - b^* \to 0$$

in probability. Furthermore, since s and t are finite constants we have from (22) that

$$N^{-\frac{1}{2}}\sum_{n=1}^{N}(X_n - \bar{X})(Y_n - \bar{Y}) \sim Z_N^* + dN^{\frac{1}{2}} + h$$

where,

$$h = (\bar{Y} - b^*)\,dN^{\frac{1}{2}} + (\bar{X} - a^*)\,dN^{\frac{1}{2}} - sd - td + d^2N.$$

Also

$$(\bar{X}\bar{Y})^{\frac{1}{2}} = \{(a + sN^{-\frac{1}{2}})(b + tN^{-\frac{1}{2}})\}^{\frac{1}{2}} \sim (ab)^{\frac{1}{2}}.$$

Finally we obtain

$$\tilde{Z}_N \sim (ab)^{-\frac{1}{2}}(Z_N^* + dN^{\frac{1}{2}} + h),$$

and therefore, asymptotically,

$$\beta(a, b, c) \sim \int P[Z_N^* + dN^{\frac{1}{2}} + h > k(ab)^{\frac{1}{2}} \mid s, t, a^*, b^*, d] \, dP_N(s, t, a^*, b^*, d)$$

$$= P[Z_N^* + dN^{\frac{1}{2}} + h > k(ab)^{\frac{1}{2}} \mid a^*, b^*, d] = (2\pi)^{-\frac{1}{2}} \int_{\{k(ab)^{\frac{1}{2}} - dN^{\frac{1}{2}} - h\}\sigma^{-1}}^{\infty} e^{-\frac{1}{2}u^2} \, du.$$

It is easily seen that $\beta(a, b, c) \to 1$ as $N \to \infty$ for fixed alternative (a, b, c). When $N \to \infty$, our interest is only in the values of c which are in the neighborhood of $c = 0$. If the convergence in (21) should hold also when $c \to 0$, it is necessary to have uniform convergence of the law in c for given fixed values of a and b. A sufficient condition for this uniformity of the convergence is (see Parzen [9], p. 38)

$$E\{(X_n - a^*)(Y_n - b^*) - d\}^4 < M$$

where M is independent of c. The existence of M can be easily verified. If in particular

$$cN^{\frac{1}{2}} \to \gamma(ab)^{-1} \quad \text{as} \quad N \to \infty$$

we obtain that

$$\{k(ab)^{-1} - dN^{\frac{1}{2}} - h\}\sigma^{-1} \to \{k(ab)^{\frac{1}{2}} - \gamma\}(ab)^{-\frac{1}{2}} = k - \gamma(ab)^{-\frac{1}{2}}.$$

We summarized the results proved in this section in

THEOREM 6. *Let $(X_n, Y_n) n = 1, 2, \cdots, N$ be independent observations on a bivariate Poisson vector (X, Y). Then if the test*

$$(23) \qquad \tilde{\varphi} = \tilde{\varphi}(\tilde{Z}_N \mid s, t) = \begin{cases} 1 \\ 0 \end{cases} \text{according as} \quad \tilde{Z}_N \gtrless k$$

is used for the independence of the Poisson variables X and Y; where k is chosen such that

$$(2\pi)^{-\frac{1}{2}} \int_k^{\infty} e^{-u^2/2} \, du = \alpha$$

and α is the asymptotic size of the test, then the asymptotic power of the test for the alternative (a, b, c) is given by

$$\beta(a, b, c) = (2\pi)^{-\frac{1}{2}} \int_{\{k(ab)^{\frac{1}{2}} - dN^{\frac{1}{2}} - h\}\sigma^{-1}}^{\infty} e^{-\frac{1}{2}u^2} \, du.$$

Furthermore, if $cN^{\frac{1}{2}} \to \gamma(ab)^{-1}$ as $N \to \infty$, the limiting asymptotic power of the test is given by

$$(2\pi)^{-\frac{1}{2}} \int_{k - \gamma(ab)^{-\frac{1}{2}}}^{\infty} e^{-\frac{1}{2}u^2} \, du.$$

3.5. Asymptotically LMP BRS test. Since the small sample distribution theory of the test statistic is not known, in practice we shall use, as suggested earlier, the asymptotic theory. If such is the case, we are in fact approximating a LMP BRS test. It is of interest to know whether this approximation of the LMP BRS test is also asymptotically LMP BRS test in the sense of definition 2.

The LMP BRS test of preassigned level of significance α was given in (4) in terms of Z. In terms of the random variable $\tilde{Z}_N$ the test is:

$$\varphi^* = \varphi^*(\tilde{Z}_N \mid \bar{X}, \bar{Y}) = \begin{cases} 1 \\ \gamma^* \\ 0 \end{cases} \text{according as} \quad \tilde{Z}_N \gtreqless k^*(N, \bar{X}, \bar{Y}) = k^*$$

and whatever be a and b, the constants γ^* and k^* are chosen to satisfy

$$\alpha = \sum_z \varphi^*(z) p_{\tilde{Z}_N}(z; a, b, 0),$$

where $p_{\tilde{Z}_N}(z; a, b, c)$ is the probability function of $\tilde{Z}_N$ for the system (a, b, c). The power of this test is

$$\beta_N^*\{\varphi^*; a, b, c\} = \sum_z \varphi^*(z) p_{\tilde{Z}_N}(z; a, b, c).$$

The corresponding asymptotic test is given in (23) and it has the size

$$\tilde{\alpha}_N = \sum_z \tilde{\varphi}_0(z) p_{\tilde{Z}_N}(z; a, b, 0)$$

and for the alternative (a, b, c) it has the power

$$\tilde{\beta}_N\{\tilde{\varphi}; a, b, c\} = \sum_z \tilde{\varphi}(z) p_{\tilde{Z}_N}(z; a, b, c).$$

Definition 2. The test $\tilde{\varphi}$ is an asymptotically LMP BRS test of size α if

(i) $\tilde{\alpha}_N \to \alpha$

and

(ii) $$\Delta_N(a, b) = N^{-\frac{1}{2}} \left| \left\{ \frac{\partial}{\partial c} \tilde{\beta}_N(\tilde{\varphi}; a, b, c) \right\} \Big|_{c=0} - \left\{ \frac{\partial}{\partial c} \beta_N^*(\varphi^*; a, b, c) \right\} \Big|_{c=0} \right| \to 0,$$

when $N \to \infty$, whatever be a and b.

Theorem 7. *The test* $\tilde{\varphi}$ *is asymptotically* LMP BRS *test of size* α.

Proof. To prove (i) consider

$$|\tilde{\alpha}_N - \alpha| = \left|\sum_z \{\tilde{\varphi}(z) - \varphi^*(z)\} p_{\tilde{Z}_N}(z; a, b, 0)\right|$$
$$\leq \sum_z I(k, k_N) p_{\tilde{Z}_N}(z; a, b, 0),$$

where $I(k, k_N)$ is the indicator of the closed interval with end points k and k_N. Let the probability distribution function of $\tilde{Z}_N$ and $\tilde{Z}_\infty$ be denoted by F_N and F_∞ respectively where we know already that F_∞ is $N(0, 1)$ and hence is continuous and differentiable. Therefore, given ϵ there exists by Theorem 5 an $N(\epsilon)$ such that

$$|F_N - F_\infty| < \epsilon,$$

for all $N \geqq N(\epsilon)$. This implies that there is an η depending on ϵ and k such that η approaches zero as ϵ approaches zero and

$$|k_N - k| < \eta(\epsilon, k).$$

for $N \geqq N(\epsilon)$. Also it is clear that

$$k - \eta \leqq k_N \leqq k + \eta.$$

Therefore,

$$\begin{aligned} |\tilde{\alpha}_N - \alpha| &\leqq \sum_z I(k - \eta, k + \eta) p_{\tilde{z}_N}(z; a, b, 0) \\ &= F_N(k + \eta) - F_N(k - \eta). \\ &\leqq F_\infty(k + \eta) + \epsilon - F_\infty(k - \eta) + \epsilon \\ &= \eta F'_\infty(k') + \eta F'_\infty(k'') + 2\epsilon, \end{aligned}$$

where k' is a suitable point between k and $k + \eta$ and similarly k'' is between k and $k - \eta$. Finally let $\epsilon \to 0$, then it follows also that $\eta \to 0$, hence using the finiteness of $F'_\infty(k')$ and $F'_\infty(k'')$, (i) is proved.

To prove (ii) let us consider

$$\begin{aligned} \Delta_N(a, b) &\leqq N^{-\frac{1}{2}} \left\{ \sum_z | \tilde{\varphi}_0(z) - \varphi_0^*(z) | \frac{\partial}{\partial c} \log p_{\tilde{z}_N}(z; a, b, c) |_{c=0} p_{\tilde{z}_N}(z; a, b, 0) \right\} \\ &= N^{-\frac{1}{2}} \left\{ \sum_z | \tilde{\varphi}_0(z) - \varphi_0^*(z) | \left[\sum_{n=1}^{N} (X_n Y_n - ab) \right] p_{\tilde{z}_N}(z; a, b, 0) \right\}. \end{aligned}$$

Therefore, using Schwartz' inequality

$$\Delta_N^2(a, b) \leqq (a^2b + ab^2 + ab) \left\{ \sum_z |\tilde{\varphi}_0(z) - \varphi_0^*(z)| p_{\tilde{z}_N}(z; a, b, 0) \right\} \to 0$$

as $N \to \infty$ because of (i) and the boundedness of $(a^2b + ab^2 + ab)$. Hence (ii) is also satisfied. This completes the proof of the theorem.

3.6. Applications. This testing problem arose in connection with the work on beetles which is being conducted by Professor Park, etc. (see Neyman, Park, Scott [8], p. 75) at The University of Chicago.

Suppose that a square tray, in which a thin layer of flour is spread, is divided into subsquares of the same size by drawing hypothetical lines parallel to the sides of the square tray. A very large number of beetles were kept in it and after sufficient time had elapsed, the number of male and female beetles in each subsquare was noted. The problem was to study if in the tray, the distribution of male and female beetles was independent of one another.

If we assume that each pair of beetles distributes itself independently of the other pair and uniformly so in the tray, then the stochastic derivation of the bivariate Poisson vector (see Loève [5], p. 84) shows that it is applicable to the present situation. Several similar situations could be easily conceived where the bivariate Poisson vector is applicable.

4. A class of bivariate random vectors and test of independence. In this section we describe a class of bivariate probability functions such that for any of these bivariate probability functions regarded as the joint probability function of the random variables X and Y, the LMP BRS test for the independence of X and Y is the same as the one obtained in the case of bivariate Poisson vector. Also, an example is given for which the LMP BRS test of independence is not of that form.

4.1. A class of bivariate random vectors. In Definition 1 the random variable Z is a Poisson Variable with expectation d. Let $f(t)$ be a continuous probability function and define a random variable W taking nonnegative integral values by the probability law

$$
\begin{aligned}
P\{W = w \mid \sigma \geqq 0\} &= \int_0^\infty e^{-d} \frac{d^w}{w!} f\left(\frac{d}{\sigma}\right) d\left(\frac{d}{\sigma}\right) \\
&= \int_0^\infty e^{-\sigma t} \frac{(\sigma t)^w}{w!} f(t)\, dt,
\end{aligned} \tag{24}
$$

where σ plays the role of a scale parameter. In particular, if $\sigma = 0$, the random variable W is degenerate at zero. This leads to

DEFINITION 3. The bivariate random vector (X, Y) is a bivariate Poisson* vector with respect to the continuous probability function $f(t)$ if $(X, Y) = (X^* + W, Y^* + W)$ where X^* and Y^* are two independent Poisson variables and W is a random variable independent of X^* and Y^* and having the probability function (24).

Using the above definition, we can easily verify that the random variables X and Y are independent if and only if $\sigma = 0$ and in case they are independent each is a Poisson variable. If $\sigma > 0$, the random variables X and Y are dependent and moreover the marginals X and Y are no longer necessarily distributed as Poisson variables.

4.2. Test of independence. Under the hypothesis of independence $\sigma = 0$ and the random variables X and Y are independent Poisson variables with respective expectations a and b. Furthermore, if $(X_n, Y_n)n = 1, 2, \cdots, N$ are independent observations on (X, Y), then (S, T), is a vector boundedly complete sufficient statistic for (a, b) as shown in Theorem 2(i). Therefore, all similar tests for this problem have Neyman structure. The conditional LMP BRS test for independence is given by

$$
\varphi(\chi \mid s, t) = \begin{cases} 1 \\ \gamma \\ 0 \end{cases} \text{according as}
$$

$$\frac{\partial}{\partial\sigma}\left\{\log\prod_{n=1}^{N} e^{-a-b}a^{X_n}b^{Y_n}\sum_{w=0}^{W_n^*}\frac{(ab)^{-s}}{(X_n-w)!\,w!(Y_n-w)!}\cdot\int_0^\infty e^{-\sigma t}(\sigma t)^w f(t)\,dt\,\Big|_{\sigma=0}\right\} \gtreqless k(s,t,N), \tag{25}$$

where $S = s$, $T = t$ and $W_n^* = \min(X_n, Y_n)$. If we assume that $f(t)$ is such that

$$\frac{\partial}{\partial\sigma}\int e^{-\sigma t}\frac{(\sigma t)^w}{w!}f(t)\,dt\,\Big|_{\sigma=0} = \int\frac{\partial}{\partial\sigma}\left\{e^{-\sigma t}\frac{(\sigma t)^w}{w!}\Big|_{\sigma=0}\right\}f(t)\,dt \tag{26}$$

for $\sigma \geqq 0$, $n = 0, 1, 2, \cdots$ and

$$\int tf(t)\,dt < \infty \tag{27}$$

then the condition (25) can be reduced to the equivalent form

$$Z = \sum_{n=1}^{N} X_nY_n \gtreqless k = k_N(s, t).$$

THEOREM 9. *Let* $(X_n, Y_n)n = 1, 2, \cdots, N$ *be independent observations on the bivariate Poisson* vector* (X, Y) *with respect to the probability function* $f(t)$ *satisfying the conditions* (26) *and* (27); *then the* LMP BRS *test for the independence of* X *and* Y *is given by*

$$\varphi_0(Z \mid s, t) = \begin{cases}1\\ \gamma_0\\ 0\end{cases} \textit{according as} \quad Z \gtreqless k_0 = k(N, s, t).$$

Therefore, the theory developed in Section 3.3 is applicable in these situations also.

4.3. Test of independence which is not a function of $\sum_{n=1}^{N} X_nY_n$. In Definition 3 instead of the probability function $f(t)$ let us assume that $t = 0$ and $1/\sigma$ with respective probabilities $1 - p$ and p. Then the probability function of (X, Y) is given by

$$P\{X = x, Y = y \mid a, b, p\} = (1-p)e^{-a-b}\frac{a^xb^x}{x!\,y!} + p\sum_{w=0}^{w^*} e^{-a-b-1}\frac{a^xb^y(ab)^{-w}}{(x-w)!\,w!(y-w)!}.$$

Also X and Y are independent if and only if $p = 0$.

In this case the LMP BRS test for $p = 0$ against $p > 0$ is given by the rule:

$$\varphi(\chi \mid s, t) = \begin{cases}1\\ \gamma\\ 0\end{cases} \text{according as} \quad \sum_{n=1}^{N}\sum_{w=0}^{W_n^*}\frac{X_n!\,Y_n!}{(X_n-w)!\,w!(Y_n-w)!} \gtreqless k(N, s, t).$$

Clearly this test of independence is not a function of $\sum_{n=1}^{N} X_nY_n$.

Acknowledgment. The author is deeply indebted to Professor Lucien M. LeCam for the patient guidance and generous assistance in the preparation of this work, and to Professors Jerzy Neyman, David Blackwell and Elizabeth L. Scott for their interest and encouragement. The author is also grateful to the referee for pointing out an error in the proof of Lemma 2 and for other constructive remarks which have clarified the exposition.

REFERENCES

[1] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, 1946.

[2] L. M. Graves, *The Theory of Functions of Real Variable*, McGraw-Hill Book Co., New York and London, 1946.

[3] L. M. LeCam, *Mimeographed Notes on Estimation*, Statistical Laboratory, University of California, Berkeley, 1954.

[4] E. Lehmann and H. Scheffé, "Completeness, similar regions and unbiased estimation, Part I," *Sankhyā*, Vol. 10 (1950), pp. 305–340.

[5] M. Loève, "On sets of probability laws and their limit elements," *University of California Publications in Statistics*, Vol. 1, University of California Press, Berkeley, 1950, pp. 53–88.

[6] M. Loève, *Probability Theory*, D. Van Nostrand, New York, 1955.

[7] J. Neyman and E. S. Pearson, "Contribution to the theory of testing statistical hypothesis," *Statistical Research Memoirs*, Vol. 1, Department of Statistics, University of London, University College, London, 1936, pp. 1–35.

[8] J. Neyman, T. Park, and E. L. Scott, "Struggle for existence. The Tribolium model: biological and statistical aspects," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 4, University of California Press, Berkeley, 1956, pp. 41–79.

[9] E. Parzen, "On uniform convergence of families of sequences of random variables," *University of California Publications in Statistics*, Vol. 2, University of California Press, Berkeley, 1954 pp. 23–45.

[10] H. Teicher, "On the multivariate Poisson distribution," *Skandinavisk Aktuarietidskrift*, Nos. 1–2 (1954), pp. 1–9.