POISSON LIMITS OF MULTIVARIATE RUN DISTRIBUTIONS[1]

By Carol E. Fuchs and H. T. David

Iowa State University

0. Summary. n balls on a circle are colored white or black according to n mutually independent binomial trials. It is shown here that, when their expectations converge with n, (a) counts of runs of various lengths are asymptotically independent Poisson; (b) counts of certain configurations other than runs yield asymptotic correlated Poisson distributions; (c) counts of configurations with structure independent of n can be partitioned into equivalence classes, with asymptotic equivalence (equality with probability one) and asymptotic independence respectively within and among classes. It is also shown that (d) there cannot, essentially, exist configurations whose counts, asymptotically, are marginally, but not multivariate, Poisson.

1. Introduction. In 1921, von Mises [14] showed that the count of binomially generated runs on a circle is asymptotically Poisson, when the expectation of this count converges with n. This paper extends von Mises' work in two directions: to the multivariate case and to configurations other than runs; the type of result obtained is indicated in the preceding section.

The asymptotic normal theory of runs, multivariate as well as univariate, has been treated by Mood [15]; related work is to be found in [6], [18] and [19].

Certain of the auxiliary results in the Appendix, specifically Theorem A.6, may be of some independent interest.

2. Definitions. The m-dimensional distribution designated here as *multivariate Poisson* has characteristic function (c.f.)

$$c(t_1, t_2, \cdots, t_m) = \exp\left(\textstyle\sum_{i=1}^m a_i z_i + \sum_{i<j} a_{ij} z_i z_j + \cdots + a_{12\cdots m} z_1 z_2 \cdots z_m - A_m\right), \tag{1}$$

where

$$A_m = \textstyle\sum_{i=1}^m a_i + \sum_{i<j} a_{ij} + \cdots + a_{12\cdots m}, \qquad a_{j_1 j_2 \ldots j_k} \geqq 0, \qquad z_j = \exp(it_j).$$

This form is that used by Teicher ([17], pp. 5–6) and Dwass and Teicher ([7], p. 467). This distribution is less general than the class of linear transformations of it which Loève ([13], p. 84) calls multivariate Poisson. Any multivariate Poisson with c.f. as in (1) can be interpreted as the joint distribution of possibly overlapping sums of independent Poisson random variables. A more general result, applicable to a larger class of multivariate Poisson distributions, is given by Dwass and Teicher ([7], pp. 463–466).

The bivariate Poisson will be of special interest. It has means $a_1 + a_{12}$, $a_2 + a_{12}$,

Received 23 April 1963; revised 14 July 1964.

[1] Research supported in part by the National Institutes of Health.

and covariance a_{12}. Its correlation coefficient, given by Campbell ([2], p. 20), is $r = [a_{12}^2/(a_1 + a_{12})(a_2 + a_{12})]^{\frac{1}{2}}$. Note that, when $a_{12} = 0$, the bivariate Poisson distribution is that of two independent Poissons. When a_1 (or a_2) $= 0$, the bivariate Poisson will be called semi-Poisson because it assigns non-zero probability only to that half of the positive quadrant where $X_1 \leqq X_2$. When $a_1 = a_2 = 0$, the two Poisson variables (each with parameter a_{12}) are equal with probability one, so that non-zero probability occurs only on the line $X_1 = X_2$.

The m-dimensional distribution designated here as multivariate multiple Poisson is defined as the joint distribution of arbitrary sub-sums of random variables whose joint distribution is multivariate Poisson. It can also be interpreted as the joint distribution of arbitrary sub-sums of non-negative integer multiples of independent Poisson random variables.

The symbols $\to$ and $\to_d$ both will denote convergence in distribution, respectively when the general term of a sequence of distributions and the general term of a sequence of random variables appears on the left ([4], p. 82–83).

3. Poisson limits of run distributions. This section is devoted to the results designated by (a) in Section 0.

Consider a circle with n positions, each of which can be filled by a white ball, with probability p, or by a black ball, with probability $q = 1 - p$. A "run" of white balls of length k is a succession of k white balls, preceded and succeeded by at least one black ball; similarly for a run of black balls of length l. The dependence of p, k and l on n will be brought out by writing $p(n)$, $k(n)$ and $l(n)$.

Let r and s be arbitrarily selected non-negative integers. Let $k_\alpha(n)$, $\alpha: 1, 2, \cdots, r$, and $l_\alpha(n)$, $\alpha: 1, 2, \cdots, s$, be $r + s$ arbitrary functions from the positive integers to the positive integers, with $k_\alpha(n) \neq k_{\alpha'}(n)$ and $l_\alpha(n) \neq l_{\alpha'}(n)$. Then $I_\alpha(n)$, $\alpha: 1, 2, \cdots, r$, is defined to be the number of positions on the circle at which a run of $k_\alpha(n)$ white balls begins, and $I_\alpha(n)$, $\alpha: r + 1, r + 2, \cdots, r + s$, is defined to be the number of positions on the circle at which a run of $l_{\alpha-r}(n)$ black balls begins.

A further matter of notation is the translation of Lemma A.6, which is to be used in the proof of Theorem 3.1, into the present context. A glossary is as follows: $m = r + s$; $N_\alpha = n$. Also, let the event $\xi(\alpha, t)$, $\alpha: 1, 2, \cdots, r$, be the beginning of a run of white balls of length $k_\alpha(n)$ at position t on the circle, and let $\xi(\alpha, t)$, $\alpha: r + 1, r + 2, \cdots, r + s$, be the beginning of a run of black balls of length $l_{\alpha-r}(n)$ at position t on the circle; then Ω_α is the set of n events $\xi(\alpha, t)$, $t: 1, 2, \cdots, n$, ω_α is a subset of Ω_α with t restricted to a subset τ_α of $(1, 2, \cdots, n)$, $n(\omega_\alpha)$ is the number of elements of ω_α (i.e., the number of integers in τ_α), and ν_α is a particular value for $n(\omega_\alpha)$. Further, $P(\omega)$ is the probability that, at least for the $n(\omega_1)$ positions t specified by ω_1, the events $\xi(1, t)$ materialize, and so on for $\alpha = 2, 3, \cdots, r + s$. Finally, exactly as in Lemma A.6, $V(\nu)$ is the class of $\prod_\alpha \binom{n}{\nu_\alpha}$ distinct vectors $(\omega_1, \omega_2, \cdots, \omega_{r+s})$—in effect, starting point combinations—that can be formed under the restrictions $n(\omega_\alpha) = \nu_\alpha$.

In order to relate the asymptotic behavior of run counts purely to the convergence of the expectations of these counts, it will be useful, as a last preliminary, to recognize two consequences of this convergence. The first of these, discussed in Lemma 3.1, involves the order of run length; the second, discussed in Lemma 3.2, involves the asymptotic behavior of the probabilities $p(n)$ and $q(n)$ of materializations of white and black balls respectively.

LEMMA 3.1. *Suppose* $\lim_{n\to\infty} E[I_\alpha(n)] = C_\alpha > 0$. *Then all run lengths* ($k_\alpha(n)$ *in the case of* $\alpha \leqq r$, *and* $l_{\alpha-r}(n)$ *in the case of* $\alpha > r$) *are of order at most* $n^{\frac{1}{2}} \log n$.

PROOF. Consider any $\alpha \leqq r$. Then, by definition and hypothesis $np(n)^{k_\alpha(n)}q(n)^2 = E[I_\alpha(n)] = C_\alpha \phi_\alpha(n)$, where $C_\alpha > 0$ and $\phi_\alpha(n)$ tends to 1 for large n. The equality of the left-hand and right-hand sides is now exploited in two ways. First, the equality implies that $nq(n)^2 \geqq C_\alpha \phi_\alpha(n)$, which, for n large, implies in turn that $-\log p(n) \geqq [C_\alpha \phi_\alpha(n)/n]^{\frac{1}{2}}$, so that $1/-\log p(n)$ is of order at most $n^{\frac{1}{2}}$. Second, the equality is solved for $k_\alpha(n)$, which yields an expression whose denominator is $-\log p(n)$, whose numerator is of order at most $\log n$, and which, therefore, is of order at most $n^{\frac{1}{2}} \log n$. The argument for $\alpha > r$ is similar.

LEMMA 3.2. (a) *If* $\lim_{n\to\infty} E[I_\alpha(n)] = C_\alpha > 0$ *for some* $\alpha \leqq r$, *then* $\lim_{n\to\infty} nq(n) = \infty$;

(b) *if* $\lim_{n\to\infty} E[I_{\alpha'}(n)] = C_{\alpha'} > 0$ *for some* $\alpha' > r$, *then* $\lim_{n\to\infty} np(n) = \infty$;

(c) *if both limits pertain for some* (α, α'), *then* $\lim_{n\to\infty} np(n)q(n) = \infty$.

PROOF. (a) Referring to the proof of Lemma 3.1, $nq(n)^2 \geqq C_\alpha \phi_\alpha(n)$, so that $[nq(n)]^2 \geqq C_\alpha n \phi_\alpha(n)$; similarly for (b). (c) Since it is assumed throughout that $k_\alpha(n) \geqq 1$, comparing the left-hand and right-hand sides in the proof of Lemma 3.1 yields $np(n)q(n)^2 \geqq C_\alpha \phi_\alpha(n)$; similarly, $nq(n)p(n)^2 \geqq C_{\alpha'}\phi_{\alpha'}(n)$, and multiplication of these two inequalities, followed by multiplication by n, yields $[np(n)q(n)]^3 \geqq C_\alpha C_{\alpha'} n \phi_\alpha(n)\phi_{\alpha'}(n)$.

Letting $I(n)$ be the vector of $(r + s)$ counts $I_\alpha(n)$, and C the vector of $(r + s)$ positive constants C_α, we are led to

THEOREM 3.1. *If* $\lim_{n\to\infty} E[I(n)] = C$, *then* $I(n) \to_d$ *the distribution of* $r + s$ *mutually independent Poisson variables with parameters the respective components of* C.

PROOF. We first compute the joint factorial moments of the variables $I_\alpha(n)$, using Lemma A.6. This will involve computing $S(\nu) = \sum_{V(\nu)} P(\omega)$. The dependence of $S(\nu)$ on n will be brought out by writing $S_n(\nu)$. Also, setting $\sum_{\alpha=1}^{r+s} \nu_\alpha$ equal to K, we note that, for some set vectors $\omega = (\omega_1, \omega_2, \cdots, \omega_{r+s})$, $P(\omega) = 0$, because of the impossibility of beginning K runs of the required types at the K positions specified by ω. $P(\omega) \neq 0$ either when the vector ω specifies an arrangement of K runs with no overlaps, or when it specifies an arrangement of K runs with one or more overlaps of the following types: (1) overlap on exactly one white ball; (2) overlap on exactly one black ball; (3) overlap on two adjacent balls of different colors. With every vector ω such that $P(\omega) \neq 0$ we now associate the vector (ρ_1, ρ_2, ρ_3), where ρ_i equals the number of overlaps of type (i), i: 1, 2, 3, occurring in the arrangement specified by ω. We note that

$$\rho_i < K. \tag{2}$$

It is further convenient to define $N_n(\rho_1, \rho_2, \rho_3)$ as the number of distinct vectors such that $P(\omega) \neq 0$ and such that the arrangement of K runs specified by ω shows ρ_i overlaps of type (i). It is easy to verify that

$$N_n(\rho_1, \rho_2, \rho_3) \leqq O(n^{K-\rho_1-\rho_2-\rho_3}) \tag{3}$$

and, in particular,

$$N_n(0, 0, 0) = n^K/\textstyle\prod_\alpha \nu_\alpha! + o(n^K), \tag{4}$$

where (4) follows from the upper bound for run length order given in Lemma 3.1. Recalling the definition of $S(\nu)$ in Lemma A.6, we also have $S_n(\nu) = \sum_{(\rho_1,\rho_2,\rho_3)} N_n(\rho_1, \rho_2, \rho_3)\cdot$

$$\frac{\prod_{\alpha=1}^{r} [p(n)^{k_\alpha(n)}q(n)^2]^{\nu_\alpha} \prod_{\alpha=r+1}^{r+s} [p(n)^2q(n)^{l_\alpha(n)}]^{\nu_\alpha}}{p(n)^{\rho_1+\rho_3}q(n)^{\rho_2+\rho_3}}. \tag{5}$$

But by the given of the theorem, $p(n)^{k_\alpha(n)}q(n)^2 = E[I_\alpha(n)]/n = C_\alpha\phi_\alpha(n)/n$ for α: 1, 2, $\cdots$, r, and $p(n)^2q(n)^{l_\alpha(n)} = E[I_\alpha(n)]/n = C_\alpha\phi_\alpha(n)/n$ for a: $r+1, \cdots, r+s$, where

$$\lim_{n\to\infty} \phi_\alpha(n) = 1. \tag{6}$$

Hence, substituting these expressions, as well as (4), in (5), we obtain $S_n(\nu) = AB(n)/\prod_{\alpha=1}^{r+s} \nu_\alpha ! + o(1) + \sum_{(\rho_1,\rho_2,\rho_3)\neq(0,0,0)} N\ (\rho_1, \rho_2, \rho_3)$ $AB(n)/n^Kp(n)^{\rho_1+\rho_3}q(n)^{\rho_2+\rho_3}$, where $A = \prod_{\alpha=1}^{r+s} C_\alpha^{\nu_\alpha}$ and $B(n) = \prod_{\alpha=1}^{r+s} \phi_\alpha(n)^{\nu_\alpha}$ Hence, in view of (3) and (6)

$$\begin{aligned} |S_n(\nu) - A/\textstyle\prod_{\alpha=1}^{r+s} \nu_\alpha !| &\leqq o(1) \\ &+ \sum_{(\rho_1,\rho_2,\rho_3)\neq(0,0,0)} O(1/[np(n)]^{\rho_1}[nq(n)]^{\rho_2}[np(n)q(n)]^{\rho_3}). \end{aligned} \tag{7}$$

But $\rho_1 > 0$ implies $s \geqq 1$, $\rho_2 > 0$ implies $r \geqq 1$, and $\rho_3 > 0$ implies $r, s \geqq 1$. Hence Lemma 3.2 implies that each of the terms of the last summation tends to zero. Hence, in view of (2) and (7) $\lim_{n\to\infty} S_n(\nu) = \prod_{\alpha=1}^{r+s} C_\alpha^{\nu_\alpha}/\prod_{\alpha=1}^{r+s} \nu_\alpha !$, or, in view of Lemma A.6, $\lim_{n\to\infty} \mu_{\nu,n}^* = \prod_{\alpha=1}^{r+s} C_\alpha^{\nu_\alpha}$, which is the ν'th factorial moment of the joint distribution of $r+s$ mutually independent univariate Poissons. It is clear that conditions (a) and (b) of Lemma A.4 are satisfied by the sequence $\{\mu_{\nu,n}\}$ of zero-moments corresponding to the sequence of factorial moments $\{\mu_{\nu,n}^*\}$. Hence Lemma A.5 yields the desired result.

Corollary 3.1. *If* $\lim_{n\to\infty} E[I(n)] = C > 0$, *then the distribution of* $\sum_{\alpha=1}^{r+s} I_\alpha(n)$ *converges to the univariate Poisson with parameter* $\sum_{\alpha=1}^{r+s} C_\alpha$.

Proof. Apply Theorem 3.1 and Corollary A.1.

Note that, when $r = s = 1$ and $k_1(n) = l_1(n)$, we have essentially the result of von Mises [14], namely that the asymptotic distribution of the total number of runs of given length (i.e., number of runs of both types) is Poisson.

Corollary 3.2. *If* $\lim_{n\to\infty} E[I(n)] = C > 0$, *the bivariate distribution of the number of white ball runs of length* $k(n)$ *and the number of runs of length* $k(n)$ *is asymptotically semi-Poisson with parameters* $a_2 = C_2$, $a_{12} = C_1$.

PROOF. Apply Theorem 3.1, Corollary A.2 and the definition of a semi-Poisson given in Section 2.

COROLLARY 3.3. *If* $\lim_{n\to\infty} E[I(n)] = C > 0$, *the asymptotic multivariate distribution of sums of numbers of runs of arbitrary type and length is multivariate Poisson.*

PROOF. Apply Theorem 3.1 and Corollary A.3.

4. Poisson limits of configuration distributions. It has been shown above that the asymptotic distribution of numbers of "runs" in the usual sense, under the restriction that their expectations converge with n, is multivariate independent Poisson. Runs can be generalized to configurations that are not asymptotically independent; this is illustrated now, in Theorem 4.1 below.

DEFINITION 4.1. Let $U_\alpha(n)$ equal the number of configurations $G_\alpha(n)$, $\alpha = 1, 2$, which materialize on the circle, where $G_\alpha(n)$ is a succession of $k_\alpha(n) \geqq 1$ white balls immediately followed by a succession of $l_\alpha(n) \geqq 1$ black balls, $[k_1(n), l_1(n)] \neq [k_2(n), l_2(n)]$.

DEFINITION 4.2. Let $\mathrm{K}(n) = \max_\alpha [k_\alpha(n)]$, $\kappa(n) = \min_\alpha [k_\alpha(n)]$, $\Lambda(n) = \max_\alpha [l_\alpha(n)]$, $\lambda(n) = \min_\alpha [l_\alpha(n)]$.

Lemma A.6 is also used in proving Theorem 4.1. The translation of Lemma A.6 is entirely analogous to that in Section 3. Thus, $m = 2$, $N_\alpha = n$, and, if the $\xi(\alpha, t)$, α: 1, 2, is the beginning of a configuration $G_\alpha(n)$ at position t on the circle, then Ω_α, ω_α, τ_α, $n(\omega_\alpha)$, ν_α, $P(\omega)$ and $V(\nu)$ are defined in a manner analogous to their definitions for the $r + s$ run types of Section 3.

Lemmas 4.1 and 4.2, below, are the analogues of Lemmas 3.1 and 3.2; as before, these lemmas deal, respectively, with the orders of configuration length and of $[p(n), q(n)]$.

LEMMA 4.1. *If* $\lim_{n\to\infty} E[U_\alpha(n)] = Q_\alpha > 0$, $\alpha = 1, 2$, *and if* $[k_1(n), k_2(n)] \neq (1, 1)$ *and* $[l_1(n), l_2(n)] \neq (1, 1)$ *for large* n, *then* $k_1(n)$, $k_2(n)$, $l_1(n)$, *and* $l_2(n)$ *all are of order at most* $n^{\frac{1}{2}} \log n$.

PROOF. Consider for example the subsequence of $\{n\}$ for which $k_1(n), l_1(n) \geqq 2$ and $k_2(n), l_2(n) = 1$. Then $np(n)^{k_1(n)} q(n)^{l_1(n)} = E[U_1(n)] = Q_1\phi_1(n)$, which implies, as in the proof of Lemma 3.1, that both $1/-\log p(n)$ and $1/-\log q(n)$ are of order at most $n^{\frac{1}{2}}$. Further, as in the proof of Lemma 3.1, we find that solving for $k_1(n)$ or $l_1(n)$ yields an expression whose denominator is either $-\log p(n)$ or $-\log q(n)$, whose numerator is of order at most $n^{\frac{1}{2}}$, and which, therefore, is of order at most $n^{\frac{1}{2}} \log n$; this completes the proof of the lemma, since a similar argument holds for the other eight subsequence types possible under the hypothesis of the lemma.

LEMMA 4.2. *If* $\lim_{n\to\infty} E[U_\alpha(n)] = Q_\alpha > 0$, $\alpha = 1, 2$, *and if*

$$\lim_{n\to\infty} p(n)^{\mathrm{K}(n)-\kappa(n)} q(n)^{\Lambda(n)-\lambda(n)} = D, \quad 0 \leqq D \leqq 1,$$

then $\lim_{n\to\infty} np(n)^{\mathrm{K}(n)} q(n)^{\Lambda(n)} = (Q_1 Q_2 D)^{\frac{1}{2}}$.

PROOF. Consider the subsequence of $\{n\}$ such that

$$[\mathrm{K}(n), \Lambda(n)] = [k_1(n), l_2(n)].$$

For this subsequence

$$np(n)^{K(n)}q(n)^{\Lambda(n)} = Q_1\phi_1(n)q(n)^{\Lambda(n)-\lambda(n)} = Q_2\phi_2(n)p(n)^{K(n)-\kappa(n)}.$$

Since $\phi(n) > 0$ for large n, division is possible and $p(n)^{K(n)-\kappa(n)}q(n)^{\lambda(n)-\Lambda(n)} = [Q_1\phi_1(n)/Q_2\phi_2(n)]$. However, by hypothesis, $p(n)^{K(n)-\kappa(n)}q(n)^{\Lambda(n)-\lambda(n)} = D\psi(n)$, where $\psi(n)$ tends to 1 for large n. Hence

$$q(n)^{\Lambda(n)-\lambda(n)} = \{[Q_2D\phi_2(n)\psi(n)/Q_1\phi_1(n)]\}^{\frac{1}{2}}.$$

Finally, substitute this into the above expression for $np(n)^{K(n)}q(n)^{\Lambda(n)}$. Similar demonstrations hold for the subsequences of $\{n\}$ where $[K(n), \Lambda(n)] = [k_2(n), l_1(n)]$, $[k_1(n), l_1(n)]$, or $[k_2(n), l_2(n)]$, which demonstrates the lemma.

THEOREM 4.1. *Consider the vector* $U(n)$ *defined in Definition* 4.1, *and assume that* $[k_1(n), k_2(n)] \neq (1, 1)$, $[l_1(n), l_2(n)] \neq (1, 1)$. *If* $\lim_{n\to\infty} E[U(n)] = Q > 0$ *and if* $\lim_{n\to\infty} p(n)^{K(n)-\kappa(n)}q(n)^{\Lambda(n)-\lambda(n)} = D$, $0 \leqq D \leqq 1$, *then* $U(n) \to_d$ *a bivariate Poisson distribution with correlation coefficient* $D^{\frac{1}{2}}$.

PROOF. As in the case of Theorem 3.1, we begin by computing $S(\nu)$, whose dependence on n is brought out by writing $S_n(\nu)$. Also, setting $\nu_1 + \nu_2$ equal to K, we note that $P(\omega) \neq 0$ either when the vector ω specifies an arrangement of K configurations with no overlaps, or when it specifies an arrangement of K configurations with one or more overlaps. Such overlaps can, in this case, be of only one type: overlap on $\kappa(n)$ white balls and $\lambda(n)$ black balls. The number of such overlaps corresponding to a particular ω with $P(\omega) \neq 0$ is denoted by ρ, and we note that $\rho \leqq \nu \equiv \min(\nu_1, \nu_2)$.

In addition, we define $N_n(\rho)$ as the number of distinct vectors ω such that $P(\omega) \neq 0$ and such that the arrangement of K runs specified by ω shows precisely ρ overlaps. It is now easy to verify that

$$N_n(\rho) = n^{K-\rho}/(\nu_1 - \rho)!(\nu_2 - \rho)!\rho! + o(n^{K-\rho}), \tag{8}$$

where, analogously to Section 3, the upper bound for run length order given in Lemma 4.1 is required for (8). Further,

$$S_n(\nu_1, \nu_2) = \textstyle\sum_{\rho=0}^{\nu} N_n(\rho)[p(n)^{k_1(n)}q(n)^{l_1(n)}]^{\nu_1-\rho} \cdot [p(n)^{k_2(n)}q(n)^{l_2(n)}]^{\nu_2-\rho}[p(n)^{K(n)}q(n)^{\Lambda(n)}]^{\rho}. \tag{9}$$

But, by the given of the theorem,

$$p(n)^{k_\alpha(n)}q(n)^{l_\alpha(n)} = E[U_\alpha(n)]/n = Q_\alpha\phi_\alpha(n)/n, \qquad \alpha = 1, 2, \tag{10}$$

and Lemma 4.2, in addition to the given, implies that $p(n)^{K(n)}q(n)^{\Lambda(n)} = (Q_1Q_2D)^{\frac{1}{2}}\phi_3(n)/n$, where all three functions ϕ satisfy

$$\lim_{n\to\infty} \phi(n) = 1. \tag{11}$$

Now substitute (8) and (10) in (9). Then, using (11),

$$S_n(\nu_1, \nu_2) = \sum_{\rho=0}^{\nu} \frac{Q_1^{\nu_1-\rho}Q_2^{\nu_2-\rho}(Q_1Q_2D)^{\rho/2}}{(\nu_1 - \rho)!(\nu_2 - \rho)!\rho!} + o(1),$$

so that the limit of $\nu_1 ! \nu_2 ! S_n(\nu_1, \nu_2)$ is the factorial moment of order (ν_1, ν_2) of a correlated bivariate Poisson distribution with marginal means Q_1 and Q_2 and correlation coefficient $D^{\frac{1}{2}}$([2], pp. 20–21). Arguments analogous to those used at the end of the proof of Theorem 3.1 then show that $U(n)$ tends in distribution to this correlated bivariate Poisson.

Theorem 4.1 leads to several corollaries, stated below with the given of Theorem 4.1 implied.

COROLLARY 4.1. *When $D = 0$, $0 < D < 1$, or $D = 1$, $U(n)$ is respectively asymptotically bivariate independent Poisson, bivariate correlated Poisson, or bivariate singular Poisson.*

COROLLARY 4.2. *When $Q_1 < Q_2$ and $D = Q_1/Q_2$ the asymptotic distribution of $U(n)$ is semi-Poisson with parameters $a_2 = Q_2 - Q_1$ and $a_{12} = Q_1$.*

COROLLARY 4.3. *$U_1(n) + U_2(n)$ is asymptotically univariate multiple Poisson, with parameters $Q_1 + Q_2 - 2(Q_1Q_2D)^{\frac{1}{2}}$ and $(Q_1Q_2D)^{\frac{1}{2}}$.*

PROOF. By Theorem 4.1 and Equation (1) in Section 2, the asymptotic distribution of $U_1(n) + U_2(n)$ has characteristic function

$$c(t, t) = \exp [A(z - 1) + B(z^2 - 1)],$$

with $A = Q_1 + Q_2 - 2(Q_1Q_2D)^{\frac{1}{2}}$ and $B = (Q_1Q_2D)^{\frac{1}{2}}$; this is the characteristic function of $P_1 + 2P_2$, where P_1 and P_2 are mutually independent univariate Poissons, with parameters respectively A and B.

COROLLARY 4.4. *When $D = 1$ and $Q_1 = Q_2 = Q$, $U_1(n) + U_2(n)$ asymptotically is distributed as $2P$, where P is a Poisson with parameter Q.*

PROOF. As for Corollary 4.3, but with the given particular parametric values.

5. Limit distributions of arbitrary configurations of fixed structure. This section discusses briefly the asymptotic multivariate distribution of numbers of configurations on the circle, under the conditions (a) that the structures of the various configurations involved remain fixed as n becomes large, and (b) that the expected number(s) of the configurations whose pattern has (have) the fewest white balls is (are) convergent with n. This asymptotic distribution is a special case of the multivariate Poisson in the sense that it can only involve mutually independent, equivalent (random variables which are equal with probability one), or degenerate random variables (random variables which are zero with probability one.)

Consider white and black balls from a binomial population arranged on a circle, as in Section 0. Let $p(n)$ be the probability of a white ball. Let W_{ij} be the number of configurations of type (ij) appearing on the circle where $i = 1, 2, \cdots, m$ enumerates distinct white balls patterns (e.g. *oxo* and *oxoxx* have the same i), while $j = 1, 2, \cdots, r_i$ enumerates different black ball patterns superimposed on the ith pattern of white balls. Let k_i be the number of white balls in configuration (ij). Let $k = \min_i k_i$. Details of derivation will be omitted, but it is clear that, under the condition $np(n)^k \to \lambda$, the asymptotic character of the joint distribution of the $\sum r_i$ random variables W_{ij} follows from the following asymptotic considerations: (1) configurations with more than k white balls occur with probability zero; (2) black balls materialize with probability one, so that, for the configurations with k white balls, only the white ball pattern matters; (3) indeed, any two such configurations that have the same white ball pattern occur simultaneously with probability one, with the respective counts equal, while two such configurations with differing white ball patterns occur simultaneously with probability zero, with the respective counts independent. These considerations lead to

CONCLUSION 5.1. Configuration counts for configurations with $k_i > k$ tend to zero in probability; configuration counts for configurations with $k_i = k$ tend in distribution to that of several sets of Poisson (parameter λ) variables, with mutual independence among sets, and equivalence within.

CONCLUSION 5.2. By Lemma A.1, $\sum_{i=1}^{m} \sum_{j=1}^{r_i} W_{ij}$ tends in distribution to the distribution of $\sum_{k_i=k} r_i P_i$, where the P_i are mutually independent Poissons with parameter λ.

6. The prevalence of multivariate Poissons as limit distributions of configurations. The theorem presented in this section states sufficient conditions for the asymptotic joint distribution of arbitrary configurations on the circle to be multivariate Poisson. It is based on the fact that the only infinitely divisible distribution with Poisson marginals is the multivariate Poisson ([7], p. 467).

This theorem essentially eliminates the possibility that configurations on the circle may exist whose numbers are asymptotically marginally Poisson but whose joint distribution is not multivariate Poisson.

THEOREM 6.1. *Let $Y_\alpha(n)$, $\alpha = 1, 2, \cdots, m$, be the number of arbitrary configurations $H_\alpha(n)$ of $k_\alpha(n) \geqq 1$ white balls and $l_\alpha(n) \geqq 1$ black balls, materializing on n positions of a circle. Let*

(a) $[k_1(n), \cdots, k_m(n)] \neq (1, \cdots, 1)$ *and* $[l_1(n), \cdots, l_m(n)] \neq (1, \cdots, 1)$ *for large n*, and $E[Y_\alpha(n)] \rightarrow \lambda_\alpha$;

(b) *on any arc of $\gamma n + o(n)$ successive positions,*

$$Y(\gamma, n) = [Y_1(\gamma, n), \cdots, Y_m(\gamma, n)] \rightarrow_d$$

a multivariate distribution whose marginals are Poissons with parameters $\gamma\lambda_1$, $\cdots$, $\gamma\lambda_m$, where $Y_\alpha(\gamma, n)$ is the number of configurations $H_\alpha(n)$ materializing on this arc.

Then $Y(n) = [Y_1(n), \cdots, Y_m(n)] \rightarrow_d$ a multivariate Poisson distribution.

PROOF. By assumption (b) with $\gamma = 1$, $Y(n) \rightarrow_d$ a multivariate distribution with Poisson marginals. To prove that this multivariate distribution is Poisson, we show that it is infinitely divisible, since, by [7] p. 467, the only infinitely divisible distribution with Poisson marginals is the multivariate Poisson.

Consider an arbitrary positive integer M. Divide the circle into $2M$ arcs, M of them large and M of them small, with the small arcs alternating with the large ones. The type of argument given in Lemmas 3.1 and 4.1 will hold also here under assumption (a) so that there will exist a constant L such that, for large n, configuration length will be less than $Ln^{\frac{1}{2}} \log n$. Let a large arc have length $\doteq n/M -$

$Ln^{\frac{1}{2}} \log n$, a small arc have length $\geqq Ln^{\frac{1}{2}} \log n$, and let all large arcs have exactly the same length. Let the random m-vector $\epsilon_{1,M,n}$ count the configurations $H_1, \cdots, H_m$ beginning on the first small arc, and similarly for $\epsilon_{2,M,n}, \cdots, \epsilon_{M,M,n}$. Let the random m-vector $Y_{1,M,n}$ count the configurations $H_1, \cdots, H_m$ that begin on the first large arc, and similarly for $Y_{2,M,n}, \cdots, Y_{M,M,n}$. Then

$$Y(n) = \sum_{i=1}^{M} Y_{i,M,n} + \sum_{i=1}^{M} \epsilon_{i,M,n} \equiv A(M, n) + B(M, n).$$

But each of the components of a particular $\epsilon_{i,M,n}$ converges to zero in probability, by assumption (b). Hence $B(M, n)$ converges in probability to $(0, \cdots, 0)$, so that $Y(n)$ and $A(M, n)$ converge in distribution to the same distribution ([4], p. 300). Hence, by the multivariate continuity theorem ([4], pp. 102–103), the c.f.'s of $Y(n)$ and $A(M, n)$, say $c_n(t)$ and $C_{M,n}(t)$, converge to the same limit c.f., say $c(t)$.

However, $c_{M,n}(t) = [\phi_n(t)]^M$, where $\phi_n(t)$ is the c.f. of $Y_{i,M,n}$, since the $Y_{i,M,n}$ are independent in view of assumption (a) and the choice of L, and are identically distributed in view of the assumption that all large arcs have exactly the same length. Hence $c(t) = \lim_{n\to\infty} [\phi_n(t)]^M = [\lim_{n\to\infty} \phi_n(t)]^M$, where $\lim_{n\to\infty} \phi_n(t)$ is a c.f. by assumption (b).

APPENDIX

The first auxiliary result is a special case of a theorem by Chernoff ([3], p. 8, [1], pp. 146–147).

LEMMA A.1. *Let $\{\Phi_n\}$ be a sequence of probability measures on R_m. Let $f = (f_1, f_2, \cdots, f_q)$, where f_α is a sub-sum of components of $x \varepsilon R_m$, α: 1, 2, $\cdots$, q. Let Ψ_n be the probability measure induced on R_q by f and Φ_n. Then, if $\Phi_n \to \Phi$, then $\Psi_n \to \Psi$, where Ψ is the probability measure induced on R_q by f and Φ.*

Lemma A.1 yields

COROLLARY A.1. *If $X_1^{(n)}, X_2^{(n)}, \cdots, X_m^{(n)} \to_d$ a multivariate independent Poisson with parameters a_j, then $\sum_{j=1}^m X_j^{(n)} \to_d$ a univariate Poisson with parameter $\sum_{j=1}^m a_j$.*

COROLLARY A.2. *If $X_1^{(n)}, X_2^{(n)}, \cdots, X_m^{(n)} \to_d$ a multivariate independent Poisson, then subset sums of the $X_j^{(n)} \to_d$ a multivariate Poisson.*

COROLLARY A.3. *If $X_1^{(n)}, X_2^{(n)}, \cdots, X_m^{(n)} \to_d$ a multivariate Poisson, then subset sums of the $X_j^{(n)} \to_d$ a multivariate multiple Poisson.*

Another useful auxiliary result, due to Cramér and Wold ([5], pp. 291–292), generalizes a sufficient condition proposed by Carleman for determinacy of the Hamburger moment problem ([16], pp. xi, 1, 4, 11).

LEMMA A.2. *Suppose there exists at least one m-dimensional probability distribution Φ with zero-moments $\mu_{\nu_1,\cdots,\nu_m}$; $\nu_1, \cdots, \nu_m = 0, 1, 2, \cdots$. Let $\lambda_{2k} = \mu_{2k,0,\cdots,0} + \cdots + \mu_{0,\cdots,0,2k}$. Then, if $\sum_{k=1}^{\infty} \lambda_{2k}^{-1/2k}$ diverges, Φ is substantially unique.*

Lemma A.2 now yields

LEMMA A.3. *The moments of the multivariate Poisson distribution satisfy the condition of Lemma A.2.*

PROOF. For k sufficiently large, all kth order marginal univariate Poisson moments do not exceed $2(k^k)$, so that $\lambda_{2k} \leqq 2m(2k)^{2k}$, and the series $\sum_{k=1}^{\infty} \lambda_{2k}^{-1/2k}$ diverges.

In the body of this paper a sequence of moments is considered which corresponds to a sequence of distributions $\{\Phi_n\}$. When these moments converge to those of the multivariate Poisson, the question arises whether or not the latter distribution is indeed the limit of the sequence of distributions $\{\Phi_n\}$. The following theorem by Haviland ([11], p. 632) gives sufficient conditions under which convergence of moments implies convergence in distribution.

LEMMA A.4. *Let* $\{\Phi_n\}$ *be a sequence of distributions on* R_m *such that*

(a) *the zero moments* $\mu_{\nu_1,\cdots,\nu_m}(\Phi_n)$, $\nu_\alpha : 0, 1, \cdots$, *exist for all* n;

(b) $|\mu_{\nu_1,\cdots,\nu_m}(\Phi_n)| \leqq K(\nu_1, \cdots, \nu_m)$ *for all* n, $\nu_\alpha : 0, 1, \cdots$;

(c) $\lim_{n\to\infty} \mu_{\nu_1,\cdots,\nu_m}(\Phi_n) = \mu_{\nu_1,\cdots,\nu_m}$ *exists for all* $\nu_\alpha : 0, 1, \cdots$.

Then there exists at least one distribution, say Φ, *such that* $\mu_{\nu_1\cdots\nu_m}(\Phi) = \mu_{\nu_1,\cdots,\nu_m}$, *and a subsequence* $\{\Phi_{n_i}\}$ *can be extracted from* $\{\Phi_n\}$ *such that* $\Phi_{n_i} \to \Phi$.

If, in addition, the sequence $\{\mu_{\nu_1,\nu_2,\cdots,\nu_m}\}$ *is such that* Φ *is substantially unique, then the sequence* $\{\Phi_n\}$ *itself* $\to \Phi$.

Lemmas A.2, A.3 and A.4 now lead to

LEMMA A.5. *Let* $\lambda_{\nu_1,\cdots,\nu_m}$; $\nu_1, \cdots, \nu_m = 0, 1, 2, \cdots$, *be the moments of a multivariate Poisson distribution* Φ. *Consider a sequence of distribution* $\{\Phi_n\}$ *defined on* R_m *with moments* $\mu_{\nu_1,\nu_2,\cdots,\nu_m}(\Phi_n)$. *If these moments satisfy conditions* (a) *and* (b) *of Lemma* A.4, *and if* $\lim_{n\to\infty} \mu_{\nu_1,\nu_2,\cdots,\nu_m}(\Phi_n) = \lambda_{\nu_1,\nu_2,\cdots,\nu_m}$, *then* $\Phi_n \to \Phi$.

The final auxiliary result, stated below as Lemma A.6, will be used to obtain the factorial moments for the distribution of numbers of runs and numbers of other configurations. In the univariate case, this result has been derived and used by von Mises [14] and by Fréchet [9], and has recently been rediscovered by Iyer [12]. The multivariate result is easily derived either by extending the arguments of Fréchet and von Mises, or that of Iyer.

LEMMA A.6. *Consider a finite set* Ω *of events, divided in some fashion into* m *subsets* Ω_α *containing respectively* N_α *events. Let* ω_α *be a particular subset of* Ω_α *containing* $n(\omega_\alpha)$ *events. Let* $P(\omega) \equiv P(\omega_1, \cdots, \omega_m)$ *be the probability that all* $\sum n(\omega_\alpha)$ *events in* ω_α *materialize, and let* $V(\nu) = V(\nu_1, \cdots, \nu_m)$ *be the class of all* $\binom{N_1}{\nu_1} \cdots \binom{N_m}{\nu_m}$ *set vectors* $(\omega_1, \cdots, \omega_m)$ *that can be formed under the restriction* $n(\omega_\alpha) = \nu_\alpha$, $\alpha: 1, \cdots, m$. *Finally, let* $I = (I_1, \cdots, I_m)$ *be the m-dimensional chance variable whose* αth *component counts the number of events of* Ω_α *that materialize. Then, if* $\mu_\nu^* = \mu_{\nu_1,\cdots,\nu_m}^*$ *is the factorial moment of order* ν *of* I, $\mu_\nu^* = S(\nu)\prod_{\alpha=1}^m (\nu_\alpha !)$, *where* $S(\nu) = \sum_{V(\nu)} P(\omega)$.

The content of Lemma A.6 is illustrated by considering five events E_1, E_2, E_3, E_4, E_5 with $\Pr\{E_1, E_2, E_3, E_4, E_5\} = \frac{1}{2}$ and $\Pr\{\bar{E}_1, \bar{E}_2, \bar{E}_3, \bar{E}_4, E_5\} = \Pr\{\bar{E}_1, E_2, E_3, E_4, \bar{E}_5\} = \Pr\{E_1, \bar{E}_2, E_3, E_4, E_5\} = \frac{1}{6}$. Then if $\Omega_1 : E_1, E_2$ and $\Omega_2 : E_3, E_4, E_5$, $\mu_{1,2}^* = E[(I_1)(I_2)(I_2 - 1)] = (2)(3)(2)(\frac{1}{2}) + \cdots + (1)(3)(2)(\frac{1}{6}) = 7\frac{1}{3}$, while $S(1, 2) = \Pr\{E_1; E_3, E_4\} + \cdots + \Pr\{E_2; E_4, E_5\} = (\frac{1}{2} + \frac{1}{6}) + \cdots + (\frac{1}{2}) = 3\frac{2}{3}$.

REFERENCES

[1] ANDERSON, T. W. (1963). Asymptotic theory for principal component analysis. *Ann. Math. Stat.* **34** 122–148.

[2] CAMPBELL, J. T. (1934). The Poisson correlation function. *Proc. Edinburgh Math. Soc. Ser.* 2 **4** 18–76.

[3] CHERNOFF, HERMAN (1957). Lectures on order relations and convergence in distribution. Mimeograph, transcribed by John W. Pratt at Stanford University.

[4] CRAMÉR, HARALD (1957). *Mathematical methods of statistics.* Princeton Univ. Press.

[5] CRAMÉR, HARALD and WOLD, H. (1936). Some theorems on distribution functions. *London Math. Soc. J.* **11** 290–294.

[6] DAVID, HERBERT T. (1962). The sample mean among the moderate order statistics. *Ann. Math. Stat.* **33** 1160–1166.

[7] DWASS, MEYER and TEICHER, HENRY (1957). On infinitely divisible random vectors. *Ann. Math. Stat.* **28** 461–470.

[8] FELLER, WILLIAM (1957). *An introduction to probability theory and its applications,* **1**. Wiley, New York.

[9] FRÉCHET, MAURICE (1940). Les probabilités associées à un système d'evènements compatibles et dependants. *Actualités Scientifiques et Industrielles.* No. 859.

[10] GURLAND, JOHN (1955). Lectures in advanced probability. Mimeograph, Department of Statistics, Iowa State Univ.

[11] HAVILAND, E. K. (1934). On the theory of absolutely additive distribution functions. *Amer. J. Math.* **56** 625–658.

[12] IYER, P. V. KRISHNA (1958). A theorem on factorial moments and its application. *Ann. Math. Stat.* **29** 254–261.

[13] LOÈVE, MICHEL (1950). On sets of probability laws and their limit elements. *Univ. Calif. Publ. Statist.* **1** 53–88.

[14] MISES, R. VON (1921). Das problem der iterationen. *Zeitchrift für ang. Math. und Mech.* **1** 298–307.

[15] MOOD, A. M. (1940). The distribution theory of runs. *Ann. Math. Stat.* **11** 367–392.

[16] SHOHAT, J. A. and TAMARKIN, J. D. (1943). *The problem of moments.* American Mathematical Society, New York.

[17] TEICHER, HENRY (1954). On the multivariate Poisson distribution. *Skand. Aktuarietidskr.* **37** 1–9.

[18] WALD, A. and WOLFOWITZ, J. (1940). On a test whether two samples are from the same population. *Ann. Math. Stat.* **11** 147–162.

[19] WISHART, J. and HIRSCHFELD, H. O. (1936). A theorem concerning the distribution of joins between line segments. *London Math. Soc. J.* **11** 227–235.