Z. Wahrscheinlichkeitstheorie verw. Geb. 28, 23–36 (1973)

On the Quality of Poisson Approximations

Ward Whitt*

0. Summary

Poisson processes (possibly nonhomogeneous) are constructed in the function spaces $D_q \equiv D([0,1]^q, R)$ and $D_q \times \cdots \times D_q$ in order to approximate superpositions of uniformly sparse point processes and partial sums of infinitesimal integer-valued nonnegative random variables. Bounds for the Prohorov distance are computed, where the Prohorov distance is defined on the space of all probability measures on D_q, with the Skorohod metric being used on D_q. These bounds yield functional central limit theorems (invariance principles) and rates of convergence for functional central limit theorems involving convergence to the Poisson process. In this regard, this paper is an extension of Section 6 of Dudley [4].

For related references and general background, the reader is referred to: Billingsley [2] for convergence of probability measures; Çinlar [3] for superpositions of point processes; Dudley [4] for rates of convergence; Bickel and Wichura [1], Neuhaus [6], and Straf [7] for the space D_q; and Whitt [8] for methods to apply the rates of convergence results here to obtain rates of convergence for related functionals and processes.

This paper is organized as follows. Partial sums of integer-valued random variables are treated in Section 1; superpositions of point processes are treated in Section 2; and the extension of the results in Section 2 to the superposition of p-dimensional $(1 \leqq p \leqq \infty)$ point processes is outlined in Section 3. An intuitive understanding of the results can perhaps be achieved more quickly by examining the special cases in Examples 1.1 and 2.1.

1. Partial Sums from an Array of Random Variables

For each $n \geqq 1$, let $\{X_n(j_1, \ldots, j_q),\ 1 \leqq j_i \leqq n,\ 1 \leqq i \leqq q\}$ be a q-dimensional array of n^q independent nonnegative integer-valued random variables. Let the associated partial sums to $S_n(k_1, \ldots, k_q) = \sum_{j_1=1}^{k_1} \cdots \sum_{j_q=1}^{k_q} X_n(j_1, \ldots, j_q)$ for $1 \leqq k_i \leqq n$ and $1 \leqq i \leqq q$, with $S_n(k_1, \ldots, k_q) = 0$ if $0 \leqq k_i \leqq n$ and $1 \leqq i \leqq q$ with $k_j = 0$ for at least one j. Let $\mathscr{L}(X)$ denote the probability law of X, i.e., the image measure induced by the random variable X on its range. We thus write $\mathscr{L}(X_n) \to \mathscr{L}(X)$ for weak convergence of random elements X_n to X, cf. [2]. Let P_λ denote the Poisson probability

* Partially supported by National Science Foundation Grant GK-27866 and by a Yale University Junior Faculty Fellowship in the Social Sciences.

distribution with parameter λ on R^1. Let

$$a_n = \min\{P\{X_n(j_1, \dots, j_q)=0\}: 1 \leqq j_1 \leqq n, \dots, 1 \leqq j_q \leqq n\},$$

$$b_n = \sum_{j_1=1}^{k_1} \cdots \sum_{j_q=1}^{k_q} P\{X_n(j_1, \dots, j_q) > 1\},$$

and (1.1)

$$c_n = \sum_{j_1=1}^{k_1} \cdots \sum_{j_q=1}^{k_q} P\{X_n(j_1, \dots, j_q) = 1\}.$$

The classical Poisson approximation theorem in this setting is ([5], p. 132)

Theorem 1.1. *If* $a_n \to 1$, *then necessary and sufficient conditions for*

$$\mathscr{L}(S_n(n, \dots, n)) \to P_\lambda \quad \text{on } R^1 \quad \text{as } n \to \infty$$

are (i) $b_n \to 0$ *and* (ii) $c_n \to \lambda$.

Rates of convergence for Theorem 1.1 have also been obtained by several authors; a survey of these results appears in Section 6 of [3]. Let $\mathscr{P}(S)$ denote the set of all probability measures on the measurable space $(S, \mathscr{S})$. The rates of convergence reported in [3] involve the metrics σ_i on $\mathscr{P}(R^1)$, where

$$\sigma_i(P, Q) = \sup\{|P(E) - Q(E)|, E \in \mathscr{E}_i \subseteq \mathscr{S}\}, \tag{1.2}$$

$\mathscr{E}_1 = \mathscr{S}$, $\mathscr{E}_2 = \{\{n\}, n \geqq 0\}$, and $\mathscr{E}_3 = \{(-\infty, n], n \geqq 0\}$. (Note that for the distance d in (3.1) of [3], $d = 2\sigma_1$.)

Our object is to investigate the convergence as $n \to \infty$ of entire arrays $\{S_n(k_1, \dots, k_q), 1 \leqq k_i \leqq n\}$ to a Poisson random field, i.e., to a Poisson probability measure on the function space $D_q \equiv D([0,1]^q, R)$. What we shall do is establish rates of convergence for the functional generalization of Theorem 1.1, that is, the Poisson analogue of Donsker's invariance principle for a lattice of random variables, cf. [1] and references there. For $q=1$ and $\{X_n(k_1)\}$ i.i.d. Bernoulli for each n, such a functional limit theorem is outlined in Problem 3 on p. 143 of [2]. A rate of convergence result in the same setting appears in Section 6 of [4]. For $q > 1$, see Theorem 6.2 of [7].

As usual, our first step is to represent the array $\{S_n(k_1, \dots, k_q)\}$ as a random element of D_q. For this purpose, let

$$S_n \equiv S_n(t_1, \dots, t_q) = S_n([n t_1], \dots, [n t_q]), \tag{1.3}$$

for $0 \leqq t_i \leqq 1$, $1 \leqq i \leqq q$. Next, for n given, we construct a Poisson approximation to S_n in D_q. First, let $\{Y_n(j_1, \dots, j_q), 1 \leqq j_i \leqq n, 1 \leqq i \leqq q\}$ be an array of independent real-valued Poisson random variables with parameters

$$\lambda_n(j_1, \dots, j_q) = P\{X_n(j_1, \dots, j_q) = 1\} / P\{X_n(j_1, \dots, j_q) = 0\}. \tag{1.4}$$

Remark 1.1. If $X_n(j_1, \dots, j_q)$ has a Poisson distribution with parameter $\lambda_n(j_1, \dots, j_q)$, then the approximating Poisson variable $Y_n(j_1, \dots, j_q)$ is also Poisson with the same parameter. Consequently, if the original array is Poisson, then the approximation is just a replication. □

Just as for S_n, let $U_n(k_1, \dots, k_q) = \sum_{j_1=1}^{k_1} \cdots \sum_{j_q=1}^{k_q} Y_n(j_1, \dots, j_q)$ for $1 \leqq k_i \leqq n$ and $1 \leqq i \leqq q$, with $U_n(k_1, \dots, k_q) = 0$ if $0 \leqq k_i \leqq n$ and $1 \leqq i \leqq q$ with $k_j = 0$ for at least one j.

Let U_n also represent an element of D_q, defined by

$$U_n \equiv U_n(t_1, \ldots, t_q) = U_n([n\,t_1], \ldots, [n\,t_q]), \tag{1.5}$$

for $0 \leqq t_i \leqq 1$, $1 \leqq i \leqq q$. Let V_n be the associated Poisson process obtained from U_n by letting the $Y_n(j_1, \ldots, j_q)$ points of the jump at $(j_1\,n^{-1}, \ldots, j_q\,n^{-1})$ be independent and uniformly distributed over the q-dimensional open interval

$$\big((j_1-1)\,n^{-1}, j_1\,n^{-1}\big) \times \cdots \times \big((j_q-1)\,n^{-1}, j_q\,n^{-1}\big).$$

This makes V_n a Poisson process in D_q which is stationary in each of the n^q open intervals. (We call U_n the discretization of the possibly nonhomogeneous Poisson Process V_n.) Obviously the entire process is stationary whenever all the random variables $X_n(j_1, \ldots, j_q)$ in the original array have a common distribution. Note that our Poisson process V_n has no jumps at a boundary of any interval.

Let A_n and B_n be the events

$$A_n = \bigcap_{j_1=1}^{n} \cdots \bigcap_{j_q=1}^{n} \{X_n(j_1, \ldots, j_q) \leqq 1\}$$

and (1.6)

$$B_n = \bigcap_{j_1=1}^{n} \cdots \bigcap_{j_q=1}^{n} \{Y_n(j_1, \ldots, j_q) \leqq 1\}.$$

On A_n and B_n the arrays $\{X_n\}$ and $\{Y_n\}$ consist entirely of 0-1 variables. It is easy to see that we have constructed S_n and U_n so that

Lemma 1.1. *For any measurable set E in D_q,*

$$P(S_n \in E \mid A_n) = P(U_n \in E \mid B_n).$$

To obtain results without conditioning, we apply the elementary

Lemma 1.2. *Let B be an event in a probability space $(S, \mathscr{S}, P)$ with $P(B) > 0$. Then*

$$\sup\{|P(A) - P(A \mid B)|, A \in \mathscr{S}\} = P\{B^c\}.$$

Theorem 1.2. *If σ_1 is the distance on $\mathscr{P}(D_q)$ defined in* (1.2), *then*

$$\begin{aligned}\sigma_1(\mathscr{L}(S_n), \mathscr{L}(U_n)) &\leqq P(A_n^c) + P(B_n^c)\\ &\leqq b_n + c_n(1-a_n)\,2^{-1}\,a_n^{-2}\end{aligned}$$

where a_n, b_n, and c_n are as in (1.1).

Proof. First, $b_n = P(A_n^c)$. Next

$$\begin{aligned}P(B_n^c) &\leqq \sum_{j_1=1}^{n} \cdots \sum_{j_q=1}^{n} P\{Y_n(j_1, \ldots, j_q) > 1\}\\ &= \sum_{j_1=1}^{n} \cdots \sum_{j_q=1}^{n} \big(1 - e^{-\lambda_n(j_1, \ldots, j_q)} - \lambda_n(j_1, \ldots, j_q)\, e^{-\lambda_n(j_1, \ldots, j_q)}\big)\\ &\leqq \sum_{j_1=1}^{n} \cdots \sum_{j_q=1}^{n} [\lambda_n(j_1, \ldots, j_q)]^2/2, \quad \text{from the Taylor expansion,}\end{aligned}$$

$$\leqq 2^{-1} \sum_{j_1=1}^{n} \cdots \sum_{j_q=1}^{n} (P\{X_n(j_1,\ldots,j_q)=1\}/P\{X_n(j_1,\ldots,j_q)=0\})^2$$

$$\leqq 2^{-1} a_n^{-2} \sum_{j_1=1}^{n} \cdots \sum_{j_q=1}^{n} (P\{X_n(j_1,\ldots,j_q)=1\})^2$$

$$\leqq 2^{-1} a_n^{-2}(1-a_n)\, c_n. \quad \square$$

Corollary 1.1. *Let* $\lambda_n = \sum_{j_1=1}^{n} \cdots \sum_{j_q=1}^{n} \lambda_n(j_1,\ldots,j_q)$ *and* σ_1 *be the metric on* $\mathscr{P}(R^1)$ *in* (1.2). *Then*

$$\sigma_1(\mathscr{L}[S_n(n,\ldots,n)], P_{\lambda_n}) \leqq P(A_n^c) + P(B_n^c) \leqq b_n + c_n(1-a_n)\, 2^{-1} a_n^{-2}.$$

Proof. The proof of Theorem 1.2 applies. Alternatively, we can apply Theorem 3.4 of [8].

Remark 1.2. Corollary 1.1 gives a rate of convergence for Theorem 1.1. Note that

$$c_n \leqq \lambda_n = \sum_{j_1=1}^{n} \cdots \sum_{j_q=1}^{n} P\{X_n(j_1,\ldots,j_q)=1\}/P\{X_n(j_1,\ldots,j_q)=0\} \leqq c_n/a_n,$$

so that $\lambda_n \to \lambda$. Corollary 1.1 can be stated in terms of λ instead of λ_n using the fact that $\sigma_1(P_{\lambda_n}, P_\lambda) \leqq |\lambda_n - \lambda|$, cf. Example 3.5 of [3]. In the same way, rates of convergence are available for Corollary 1.1 and Theorem 1.2 when the limit is a Poisson process (or its discretization) other than the one constructed by the approximation above. Let U_n' and V_n' be such alternate processes constructed from an array of Poisson variables $\{Y_n'(j_1,\ldots,j_q)\}$ with parameters $\lambda_n'(j_1,\ldots,j_q)$. Then

$$\sigma_1[\mathscr{L}(U_n), \mathscr{L}(U_n')] \leqq \sum_{j_1=1}^{n} \cdots \sum_{j_q=1}^{n} |\lambda_n(j_1,\ldots,j_q) - \lambda_n'(j_1,\ldots,j_q)|,$$

cf. Example 3.6 of [3].

Example 1.1. If the variables $X_n(j_1,\ldots,j_q)$ only take values 0 and 1 with

$$P\{X_n(j_1,\ldots,j_q)=1\} = n^{-q},$$

then $a_n = (1-n^{-q})$, $b_n = 0$, and $c_n = 1$, cf. (1.1). The bound in Theorem 1.2 and Corollary 1.1 thus becomes $2^{-1} a_n^{-2} n^{-q}$ which is less than $2n^{-q}$ for $n \geqq 2$ and $q \geqq 1$. The approximating Poisson measure P_{λ_n} in Corollary 1.1 has parameter $(1-n^{-q})^{-1}$. Hence, by Remark 1.2, the rate of convergence to P_λ with $\lambda = 1$ in Theorem 1.1 is of order n^{-q}. In particular,

$$\sigma_1([S_n(n,\ldots,n)], P_1) \leqq 4n^{-q}$$

for $n \geqq 2$ and $q \geqq 1$. For further results concerning this special case, see Corollaries 1.3 and 1.4 plus Remarks 1.3, 1.4, and 1.5 at the end of this section. $\square$

It now remains to show how close U_n is to V_n. For this purpose, we use the metric d on D_q defined on p. 1289 of [6] or p. 1662 of [1]. Let Φ be the group of all transformations ϕ: $[0,1]^q \to [0,1]^q$ of the form $\phi(t) = \phi(t_1,\ldots,t_q) =$

$[\phi_1(t_1), \ldots, \phi_q(t_q)]$, where each ϕ_i: $[0,1] \to [0,1]$ is continuous, strictly increasing and fixes zero and one. Define d for any $x, y \in D_q$ by

$$d(x, y) = \inf\{\max(\|x - y(\phi)\|, g(\phi)) : \phi \in \Phi\}, \tag{1.7}$$

where $\|x\| = \sup\{|x(t)| : t \in [0,1]^q\}$, $g(\phi) = \sup\{|\phi(t) - t| : t \in [0,1]^q\}$, and $|t| = \max\{|t_i|, 1 \leqq i \leqq q\}$.

Lemma 1.3. *For* $q = 1$, $d(U_n, V_n) \leqq n^{-1}$ *on* $B_n \cap \{Y_n(n) = 0\}$, *where* B_n *is defined in* (1.6).

Proof. Following Theorem 6.1 of [4], we construct a time transformation ϕ_n so that $\|V_n - U_n(\phi_n)\| = 0$ and $g(\phi_n) \leqq n^{-1}$, as required for (1.7). Whenever V_n has a jump at t in $(n^{-1}(j-1), n^{-1}j)$, just let ϕ_n be constructed so that $\phi_n(t) = n^{-1}j$. Complete the definition of ϕ_n by linear interpolation. □

After some modification (some modification is necessary!), the argument outlined above for $q = 1$ extends to $q \geqq 2$. Let C_n be the event that $Y_n(j_1, \ldots, j_q) = 0$ whenever at least one $j_i = n$. Let D_n be the subset of B_n in (1.6) in which $Y_n(j_1, \ldots, j_q) = 0$ for all $(j_1, \ldots, j_q)$ with $(j_1, \ldots, j_q) \neq (k_1, \ldots, k_q)$ and at least one $j_i = k_i$ whenever $Y_n(k_1, \ldots, k_q) = 1$. For $q = 2$, D_n can be described by saying that given a jump in any square, there is no other jump in the same row or column. Having selected D_n appropriately, the same argument used for $q = 1$ yields

Lemma 1.4. *For* $q > 1$,

$$d(U_n, V_n) \leqq n^{-1} \quad \text{on } C_n \cap D_n.$$

Proof. It suffices to construct an appropriate time transformation $\phi_n = (\phi_{n1}, \ldots, \phi_{nq})$, cf. (1.7). On $C_n \cap D_n$ each ϕ_{ni} can be constructed on $[0,1]$ just as in the proof of Lemma 1.3. The bound would be $q^{\frac{1}{2}} n^{-1}$ instead of n^{-1} if we used the Euclidean norm instead of the supremum norm on R^k, cf. (1.7). □

Following [4], our main result will be stated in terms of Prohorov's metric. For any measurable subset A of a separable metric space (S, m), let $A^\varepsilon = \{y \in S : \exists x \in A, m(x, y) < \varepsilon\}$. Then *Prohorov's metric* $\rho \equiv \rho(m)$, which induces the topology of weak convergence on $\mathscr{P}(S)$, is defined by

$$\rho(P, Q) = \inf\{\varepsilon \geqq 0 : P(F) \leqq \varepsilon + Q(F^\varepsilon), F \text{ closed}\}. \tag{1.8}$$

It is significant that $\rho[\mathscr{L}(X), \mathscr{L}(Y)] \leqq \alpha(X, Y)$ where $\alpha \equiv \alpha(m)$ is the usual metric associated with convergence in probability of S-valued random elements, i.e.,

$$\alpha(X, Y) = \inf\{\varepsilon \geqq 0 : P[m(X, Y) \geqq \varepsilon] \leqq \varepsilon\}, \tag{1.9}$$

cf. [4] and references there.

As an immediate consequence of Lemmas 1.3 and 1.4, we obtain

Theorem 1.3. *Let* A_n, B_n, C_n, *and* D_n *be as in Lemmas 1.1 and 1.4. Let* a_n, b_n, *and* c_n *be as in* (1.1).

(i) *If* $q = 1$, *then*

$$\begin{aligned}\alpha(U_n, V_n) &\leqq \max\{n^{-1}, P(B_n^c \cup \{Y_n(n) = 0\})\}\\ &\leqq \max\{n^{-1}, c_n(1 - a_n)2^{-1}a_n^{-2} + (1 - a_n)\}.\end{aligned}$$

(ii) *If* $q>1$, *then*

$$\alpha(U_n, V_n) \leqq \max\{n^{-1}, P(C_n^c)+P(D_n^c)\}$$
$$\leqq \max\{n^{-1}, b_n+q\, n^{q-1}(1-a_n)(c_n+1)\}.$$

Proof. Lemmas 1.3 and 1.4 apply directly. It is easy to see that $P\{Y_n(n)>0\} \leqq 1-a_n$ for $q=1$, $P(C_n^c) \leqq q\, n^{q-1}(1-a_n)$ for $q>1$, and

$$P(D_n^c) \leqq q\, n^{q-1}(1-a_n)\, c_n + b_n. \quad \square$$

Corollary 1.2. *In the setting of Theorem 1.2,*

(i) *if* $q=1$, *then*

$$\rho[\mathscr{L}(S_n), \mathscr{L}(V_n)] \leqq P(A_n^c)+P(B_n^c)+\max\{n^{-1}, P(B_n^c)+P(C_n^c)\}$$
$$\leqq b_n + c_n(1-a_n)\, 2^{-1} a_n^{-2} + \max\{n^{-1}, c_n(1-a_n)\, 2^{-1} a_n^{-2}+(1-a_n)\};$$

(ii) *if* $q>1$, *then*

$$\rho[\mathscr{L}(S_n), \mathscr{L}(V_n)] \leqq P(A_n^c)+P(B_n^c)+\max\{n^{-1}, P(C_n^c)+P(D_n^c)\}$$
$$\leqq b_n + c_n(1-a_n)\, 2^{-1} a_n^{-2} + \max\{n^{-1}, b_n + q\, n^{q-1}(1-a_n)(c_n+1)\}.$$

Proof. Just apply the triangle inequality with Theorems 1.2 and 1.3. The metric ρ is appropriate because $\rho \leqq \alpha$ and $\rho \leqq \sigma_1$. $\square$

Corollary 1.3. *In the setting of Example 1.1 with* $n \geqq 2$,

$$\alpha(U_n, V_n) \leqq \begin{cases} 3n^{-1}, & q=1 \\ 2q\, n^{-1}, & q>1 \end{cases}$$

and

$$\rho[\mathscr{L}(S_n), \mathscr{L}(V_n)] \leqq \begin{cases} 5n^{-1}, & q=1 \\ 2q\, n^{-1} + 2n^{-q}, & q>1. \end{cases}$$

We can also express bounds involving a homogeneous Poisson process V with intensity $\lambda=1$ instead of V_n in Corollary 1.3.

Lemma 1.5. *If* V_1 *and* V_2 *are homogeneous Poisson processes in* D_q *with intensities* λ_1 *and* λ_2, *then*

$$\rho[\mathscr{L}(V_1), \mathscr{L}(V_2)] \leqq |\lambda_1 - \lambda_2|.$$

Proof. By the triangle inequality,

$$\rho[\mathscr{L}(V_1), \mathscr{L}(V_2)] \leqq \rho[\mathscr{L}(V_1), \mathscr{L}(U_n^1)] + \rho[\mathscr{L}(U_n^1), \mathscr{L}(U_n^2)] + \rho[\mathscr{L}(U_n^2), \mathscr{L}(V_2)],$$
$$\leqq \alpha(V_1, U_n^1) + \sigma[\mathscr{L}(U_n^1), \mathscr{L}(U_n^2)] + \alpha(U_n^2, V_2),$$

where U_n^1 and U_n^2 are discretizations of V_1 and V_2 depending on n, that is,

$$U_n^i(t_1, \ldots, t_q) = V_i([n\, t_1]/n, \ldots, [n\, t_q]/n)$$

for $0 \leqq t_i \leqq 1$, $1 \leqq i \leqq q$. In other words, U_n^i is constructed from V_i in the reverse of the way V_n was constructed from U_n in (1.5). Following Remark 1.2, we see that

$$\sigma_1[\mathscr{L}(U_n^1), \mathscr{L}(U_n^2)] \leqq \sum_{j=1}^{n^q} |\lambda_1 n^{-q} - \lambda_2 n^{-q}|$$
$$\leqq |\lambda_1 - \lambda_2|$$

for all $n \geqq 1$. On the other hand, $\alpha(U_n^i, V_i) \to 0$ as $n \to \infty$ by Theorem 1.3. Since n is arbitrary, the proof is complete. □

Corollary 1.4. *In the setting of Example 1.1 with $n \geqq 2$,*

$$\rho[\mathscr{L}(S_n), \mathscr{L}(V)] \leqq \begin{cases} 6n^{-1}, & q=1 \\ 2qn^{-1}+3n^{-q}, & q>1, \end{cases}$$

for any $n \geqq 2$ and $q \geqq 1$, where V is a homogeneous Poisson process in D_q with unit intensity.

Remark 1.3. Following Dudley ([4], p. 330), we see that the bounds in Corollaries 1.3 and 1.4 can not be improved to order $o(n^{-1})$. Dudley's negative result applies to D_q as well as D. The event that V_n (or V) has at least one jump at a point $t = t_1, \ldots, t_q$ such that $|t_i - j/n| \geqq (2^{-1} - \delta)/n$ for $0 < \delta < 2^{-1}$ has a positive probability bounded away from 0 for all n. Hence, $\rho[\mathscr{L}(U_n), \mathscr{L}(V_n)] \geqq K n^{-1}$ for some constant K.

Remark 1.4. Note that for $q > 1$ the larger contribution to the error in Corollary 1.4 comes from the difference between the Poisson process and its discretization. In summary,

(i) $\rho[\mathscr{L}(S_n), \mathscr{L}(U_n)] = 0(n^{-q})$,
(ii) $\rho[\mathscr{L}(U_n), \mathscr{L}(U)] = 0(n^{-q})$, and
(iii) $\rho[\mathscr{L}(U_n), \mathscr{L}(V_n)] = \rho[\mathscr{L}(U), \mathscr{L}(V)] = 0(n^{-1})$.

Remark 1.5. The metric d in (1.7) induces the Skorohod topology on D_q but is not complete. The Billingsley metric d_0 induces the same topology and is complete. It is defined for any $x, y \in D_q$ by

$$d_0(x, y) = \inf\{\max(\|x - y(\phi)\|, h(\phi)) : \phi \in \Phi\}, \tag{1.10}$$

cf. (1.7), where

$$h(\phi) = \sup\left\{\left|\log \frac{\phi_i(s) - \phi_i(t)}{s-t}\right| : \ s, t \in [0, 1], 1 \leqq i \leqq q\right\}, \tag{1.11}$$

cf. p. 113 of [2] and p. 1289 of [6]. Since changing from d to d_0 complicates matters, we shall use d throughout this paper. Here we only indicate briefly what the problem is and how it can be dealt with.

Within the subsets $B_n \cap \{Y_n = 0\}$ for $q = 1$ and $C_n \cap D_n$ for $q > 1$, $|\phi_i(s) - \phi_i(t)| \leqq n^{-1}$ for all $s, t \in [0, 1]$, cf. Lemmas 1.3 and 1.4, but it is possible for $|s - t|$ to be arbitrarily close to 0, thus allowing $h(\phi)$ in (1.11) to be unbounded.

The size of $h(\phi)$ can be controlled by working with bands of approximately $n^{\frac{1}{2}}$ intervals of length n^{-1}. (For simplicity, assume $n^{\frac{1}{2}}$ is an integer.) We permit no jump in a band next to any boundary. We permit no more than one jump in any interior band. We do not permit jumps in contiguous bands. Then

$$h(\phi) \leqq \left|\log \frac{(n^{\frac{1}{2}}+1)/n}{n^{\frac{1}{2}}/n}\right| = |\log(1 - n^{-\frac{1}{2}})| \leqq n^{-\frac{1}{2}}. \tag{1.12}$$

Thus on the smaller set above, $d_0(U_n, V_n) \leqq n^{-\frac{1}{4}}$. The probability of the complement of this set is of the order $n^{\frac{1}{4}}(1-a_n)(c_n 2^{-1} a_n^{-2}+1)$ for $q=1$ and

$$b_n + q\, n^{q-1}\, n^{\frac{1}{4}}(1-a_n)(c_n+1), \qquad q>1,$$

which in the setting of Example 1.1 is of order $n^{-\frac{1}{4}}$. Hence, in this setting we get

$$\rho[\mathscr{L}(S_n), \mathscr{L}(V_n)] \leqq 0(n^{-\frac{1}{4}})$$

where ρ is the metric in (1.8) on $\mathscr{P}(D_q, d_0)$. This leads us to conclude that the change from d to d_0 has a significant impact. In general it appears that bounds such as the ones computed in this paper will be sensitive to the choice of metrics on R^k, D_q, and $\mathscr{P}(D_q)$.

2. Superposition of Point Processes

Closely related to the problem of the previous section is the Poisson approximation for a superposition of uniformly sparse point processes, cf. [3]. Let $N_1, \ldots, N_n$ be independent point processes in D_q, i.e., independent random elements in D_q which are nonnegative, nondecreasing, integer-valued, and 0 at $(t_1, \ldots, t_q)$ if $t_j=0$ for some j. We interpret $N_k(t_1, \ldots, t_q)$ as the number of points in $[0, t_1] \times \cdots \times [0, t_q]$ for the k-th process. Let $N = N_1 + \cdots + N_n$ be the superposition process in D_q. Our object is to obtain a Poisson approximation for N with an estimate of the quality.

The general procedure is quite simple. We approximate $[0, 1]^q$ with a discrete lattice to obtain a discretization of N. This gives us an array of random variables which is almost, but not quite, independent. We then construct an approximating array of independent Poisson variables. The method of construction at this stage is not quite the same as in the last section due to the lack of independence. Finally, we construct a Poisson process in D_q from the array of independent Poisson random variables exactly as before. We are thus able to invoke Lemmas 1.3 and 1.4 for this part.

We have a degree of freedom in our discretization. We can partition the time domain $[0, 1]^q$ into as many parts as we wish, but there is the obvious tradeoff: as the partition gets finer the discrete approximation of the continuous improves but the size of the associated array increases. We shall partition each interval $[0, 1]$ into the m subintervals $[0, m^{-1}], (m^{-1}, 2m^{-1}], \ldots, ([m-1]\, m^{-1}, 1]$. Thus $[0, 1]^q$ is partitioned into m^q q-dimensional intervals each of Lebesgue measure m^{-q}.

Our discrete approximation for N in D_q is thus N^m, where

$$N^m(t_1, \ldots, t_q) = N([m\, t_1]/m, \ldots, [m\, t_q]/m) \tag{2.1}$$

for $0 \leqq t_i \leqq 1$ and $1 \leqq i \leqq q$. Let B_j denote a generic q-dimensional interval of the type described above. Our associated array of random variables then consists of $X_m(j) = N(B_j)$, $1 \leqq j \leqq m^q$, where we regard N as a set function here.

Let $F \equiv F_m$ denote the union of those B_j touching the upper boundary of $[0, 1]^q$. Let

$$A \equiv A^m = \{N(F)=0\}. \tag{2.2}$$

For $q=1$, let

$$B \equiv B^m = \bigcap_{j=1}^{m-1} \{N(B_j) \leqq 1\}. \tag{2.3}$$

For $q>1$, let $G_j \equiv G_{mj}$ denote the union of those B_k other than B_j with at least one side in common with B_j. (Here "side" means projection onto one of the coordinate axes of $[0,1]^q$.) Let $C \equiv C^m$ be the event

$$C \equiv C^m = \bigcap_j \left(\{N(G_j)=0, N(B_j)=1\} \cup \{N(B_j)=0\}\right). \tag{2.4}$$

Just as in Lemmas 1.3 and 1.4, we have

Lemma 2.1. *Let d be the metric on D_q defined in* (1.7).

(i) *For $q=1$,*

$$d(N, N^m) < m^{-1} \quad \text{on} \quad A \cap B,$$

and (ii) *for $q>1$,*

$$d(N, N^m) < m^{-1} \quad \text{on} \quad A \cap C,$$

where A, B, and C are as in (2.2)–(2.4).

To estimate the probabilities, let

$$a_m = \sup_{1 \leqq j \leqq m^q} \left\{ \sum_{k=1}^{n} P[N_k(B_j)=1] \right\}$$

and (2.5)

$$b = \sum_{k=1}^{n} P\{N_k([0,1]^q) > 1\},$$

cf. (3.9)–(3.12) of [3].

Lemma 2.2. *Let a_m and b be as in* (2.5). *For $q \geqq 1$,*

$$P\{d(N, N^m) > m^{-1}\} \leqq b + q\, m^{(q-1)}\, a_m + q\, m^{(2q-1)}\, a_m^2.$$

Proof. For $q=1$, $(A \cap B)^c = A^c \cup B^c \subseteq E_1 \cup E_2 \cup E_3$, where

$$E_1 = \bigcup_{k=1}^{n} \{N_k([0,1]^q) > 1\},$$

$$E_2 = \bigcup_{\substack{j \\ B_j \subset F}} \bigcup_{k=1}^{n} \{N_k(B_j)=1\},$$

and

$$E_3 = \bigcup_{j=1}^{m} \bigcup_{l=1}^{n} \bigcup_{\substack{k=1 \\ k \neq l}}^{m} \{N_k(B_j)=1, N_l(B_j)=1\}.$$

Hence, $P(A^c \cup B^c) \leqq b + q\, m^{(q-1)}\, a_m + m\, a_m^2$ with $q=1$.

For $q>1$, $(A \cap C)^c = A^c \cup C^c \subseteq E_1 \cup E_2 \cup E_4$, where

$$E_4 = \bigcup_{j=1}^{m^q} \bigcup_{l=1}^{n} \bigcup_{\substack{k=1 \\ k \neq l}}^{m} \bigcup_{\substack{p \\ B_p \subset G}} \{N_k(B_p)=1, N_l(B_j)=1\}$$

and $P(E_4) \leqq q\, m^{2q-1}\, a_m^2$. $\square$

Corollary 2.1. *If* $b \leqq K_1 m^{-1}$ *and* $a_m \leqq K_2 m^{-q}$, *where* a_m *and* b *are as in* (2.5), *then*

$$\alpha(N, N^m) \leqq K_3 m^{-1}$$

for any $q \geqq 1$, *where* α *is the metric defined in* (1.9) *and* $K_3 = \max\{1, K_1 + q(K_2 + K_2^2)\}$.

We now come to the second stage of our approximation. Recall that we have the array of real-valued random variables $\{X(j)\}$ associated with our discrete approximation N^m in D_q, where

$$X(j) = \sum_{k=1}^{n} X_k(j) = \sum_{k=1}^{n} N_k(B_j) = N(B_j). \tag{2.6}$$

If these random variables were independent, we could apply the results of the last section and be done, but of course this is not the case. Let $\{Y_k(j); 1 \leqq j \leqq m^q, 1 \leqq k \leqq n\}$ be independent Poisson random variables with the parameter of $Y_k(j)$ being

$$\lambda_k(j) = P\{N_k(B_j) = 1, N_k([0, 1]^q) = 1\}. \tag{2.7}$$

Then $\{Y(j)\}$ in the Poisson array corresponding to $\{X(j)\}$ has parameter $\lambda(j) = \lambda_1(j) + \cdots + \lambda_n(j)$. Then, by virtue of (3.7) or (6.29) of [3], we have

Lemma 2.3. *Let* σ_1 *be the metric in* (1.2) *on* $\mathscr{P}(R^{m^q})$. *Then*

$$\begin{aligned}\sigma_1[\mathscr{L}(\{X_k(j); 1 \leqq j \leqq m^q\}), \mathscr{L}(\{Y_k(j); 1 \leqq j \leqq m^q\})] \\ \leqq P\{N_k([0, 1]^q) > 1\} + P\{N_k([0, 1]^q) = 1\}^2.\end{aligned}$$

Proof. As we remarked above, this is (3.7) or (6.29) of [3], recalling that the metrics σ_1 here and d in [3] are related by $2\sigma_1 = d$. □

Remark 2.1. Our choice of parameters $\lambda_k(j)$ in (2.7) does not make the Poisson approximation of a Poisson process a replication, cf. Remark 1.1. To achieve this, we could use parameters

$$\lambda'_k(j) = \lambda_k(j)/P\{N_k([0, 1]^q) = 0\}.$$

We have not used $\lambda'_k(j)$ because the statement corresponding to Lemma 2.3 becomes more complicated.

In order to express related results, let

$$c = \sum_{k=1}^{n} P\{N_k([0, 1]^q) = 1\}^2. \tag{2.8}$$

Corollary 2.2. *If* σ_1 *is again the metric in* (1.2) *on* $\mathscr{P}(R^{m^q})$, *then*

$$\sigma_1[\mathscr{L}(\{X(j)\}), \mathscr{L}(\{Y(j)\})] \leqq b + c,$$

where b *and* c *are defined in* (2.5) *and* (2.8).

Proof. The arrays $\{X_k(j)\}$ are independent for different k. Hence

$$\begin{aligned}\sigma_1[\mathscr{L}(X_{k_1}(j) + X_{k_2}(j)), \mathscr{L}(Y_{k_1}(j) + Y_{k_2}(j))] \leqq \sigma_1[\mathscr{L}(X_{k_1}(j)), \mathscr{L}(Y_{k_1}(j))] \\ + \sigma_1[\mathscr{L}(X_{k_2}(j)), \mathscr{L}(Y_{k_2}(j))],\end{aligned}$$

cf. (3.4) of [3], and the desired conclusion holds. □

Now let U_{mk} and U_m be the random elements of D_q induced by the partial sums on j of $\{Y_k(j)\}$ and $\{Y(j)\}$ as in (1.5).

Corollary 2.3. *If σ_1 is the metric in* (1.2) *on $\mathscr{P}(D_q)$, then*

$$\sigma_1[\mathscr{L}(N^m), \mathscr{L}(U_m)] \leqq b+c,$$

where b and c are defined in (2.5) *and* (2.8).

We are now at the final stage of our approximation. Let V_{mk} and V_m be the Poisson processes in D_q corresponding to U_{mk} and U_m just as in Section 1. Just as in Lemma 2.1, we have

Lemma 2.4. *If A, B, and C be as in* (2.2)–(2.4) *with V_m in place of N, then*

(i) *for $q=1$,*

$$d(U_m, V_m) < m^{-1} \quad \text{on} \quad A \cap B,$$

and (ii) *for $q>1$,*

$$d(U_m, V_m) < m^{-1} \quad \text{on} \quad A \cap C.$$

Lemma 2.5. *Let a_m, b, and c be as in* (2.5) *and* (2.8). *For $q \geqq 1$,*

$$P\{d(V_m, U_m) > m^{-1}\} \leqq 2^{-1} c + q\, m^{(q-1)} a_m + q\, m^{(2q-1)} a_m^2.$$

Proof. Following the proof of Lemma 2.2 with V_{mk} instead of N_k, we have

$$\begin{aligned} P(E_1) &\leqq \sum_{k=1}^{n} (1 - e^{-\lambda_k} - \lambda_k e^{-\lambda_k}) \\ &\leqq 2^{-1} \sum_{k=1}^{n} \lambda_k^2 \\ &\leqq 2^{-1} c, \end{aligned}$$

where $\lambda_k = \sum_j \lambda_k(j) = P\{N_k([0,1]^q) = 1\}$;

$$\begin{aligned} P(E_2) &= \sum_{\substack{j \\ B_j \subseteq F}} \sum_{k=1}^{n} \lambda_k(j)\, e^{-\lambda_k(j)} \\ &\leqq \sum_{\substack{j \\ B_j \subseteq F}} \sum_{k=1}^{n} \lambda_k(j) \\ &\leqq q\, m^{(q-1)} a_m, \end{aligned}$$

where $\lambda_k(j) \leqq P\{N_k(B_j) = 1\} \leqq a_m$;

$$\begin{aligned} P(E_3) &\leqq \sum_{j=1}^{m} \sum_{l=1}^{n} \sum_{\substack{k=1 \\ k \neq l}}^{n} \lambda_k(j)\, \lambda_l(j) \\ &\leqq m\, a_m^2; \end{aligned}$$

$$\begin{aligned} P(E_4) &\leqq \sum_{j=1}^{m^q} \sum_{l=1}^{n} \sum_{\substack{k=1 \\ k \neq l}}^{n} \sum_{\substack{p \\ B_p \subseteq G}} \lambda_k(p)\, \lambda_l(j) \\ &\leqq q\, m^{(2q-1)} a_m^2. \end{aligned}$$

Theorem 2.1. *Let a_m, b, and c be as in* (2.5) *and* (2.8). *Let ρ be the Prohorov metric in* (1.8) *on $\mathscr{P}(D_q)$. Then*

$$\rho[\mathscr{L}(N), \mathscr{L}(V_m)] \leqq (b+c) + \max\{m^{-1}, b + q\, m^{(q-1)}\, a_m + q\, m^{(2q-1)}\, a_m^2\} + \max\{m^{-1}, 2^{-1} c + q\, m^{(q-1)}\, a_m + q\, m^{(2q-1)}\, a_m^2\}.$$

Proof. It is only necessary to apply the triangle inequality with Lemma 2.2, Corollary 2.3, and Lemma 2.5. In doing so, Lemmas 2.2 and 2.5 translate immediately into a statement involving the metric α in (1.9). Furthermore, $\rho \leqq \alpha$ and $\rho \leqq \sigma_1$. □

Example 2.1. Suppose that $N_1, \ldots, N_n$ are independent and identically distributed stationary point processes in D_q with

(i) $P\{N_1([0,1]^q) = 1\} = c_1 n^{-1}$,

(ii) $P\{N_1([0,1]^q) > 1\} = c_2 n^{-2}$,

(iii) $P\{N_1(B_1) = 1\} = c_3 n^{-(q+1)}$,

where $B_1 = [0, n^{-1}] \times \cdots \times [0, n^{-1}]$. These probabilities are chosen to be directly proportional to the Lebesgue measure of the set being counted and inversely proportional to the number of point processes being added, with multiple points being less likely.

Let $N = N_1 + \cdots + N_n$ and $m = n$. Then the quantities in (2.2) and (2.5) become: $a_n = c_3 n^{-q}$, $b = c_2 n^{-1}$, and $c = c_1^2 n^{-1}$. Theorem 2.1 yields $\rho[\mathscr{L}(N), \mathscr{L}(V_n)] \leqq q K n^{-1}$ in this setting. The approximating Poisson process in this case is stationary with intensity c_1 (independent of n). Consequently, if an array of point processes is defined with the component processes of the n-th row satisfying the assumptions above, then Theorem 2.1 determines a bound of $K n^{-1}$ for the rate of convergence to the stationary Poisson process with intensity c_1. We also note that if the metric d_0 were used instead of d on D_q, then we would have a bound of $K n^{-\frac{1}{2}}$, cf. Remark 1.5.

3. Superposition of p-Dimensional Point Processes

The results of the last section extend easily to the superposition of p-dimensional point processes, cf. (4.1) of [3]. Since the argument is similar, we only give the highlights.

Let D_{qp} be the p-dimensional product space $D_q \times \cdots \times D_q$. Let the metric on D_{qp} be defined for any $x = (x_1, \ldots, x_p)$ and $y = (y_1, \ldots, y_p)$ by

$$d(x, y) = \max_{1 \leqq i \leqq p} d_i(x_i, y_i), \tag{3.1}$$

where d_i is the metric on D_q in (1.7). Let $\underset{\sim}{N}_1, \ldots, \underset{\sim}{N}_n$ be independent point processes in D_{qp}, so that $\underset{\sim}{N}_i = (N_{i1}, \ldots, N_{ip})$ with N_{ij} being a point process in D_q of the type studied in the last section. It is important that the components of $\underset{\sim}{N}_i$ are *not* assumed to be independent. Let $\underset{\sim}{N} = \underset{\sim}{N}_1 + \cdots + \underset{\sim}{N}_n$ be the p-dimensional superposition process. We approximate $\underset{\sim}{N}$ by a p-dimensional Poisson process $\underset{\sim}{V}_m = (V_{m1}, \ldots, V_{mp})$ in D_{qp} where the component processes of $\underset{\sim}{V}_m$ *are* independent. As before, m indicates the size of the discrete partition we choose to use.

Remark 3.1. We could also have $p=\infty$. Then, instead of (3.1), a metric d can be defined for any $x=(x_1, \ldots)$ and $y=(y_1, \ldots)$ in D_q^∞ by

$$d(x, y)=\sum_{i=1}^{\infty} 2^{-i} d_i(x_i, y_i)/[1+d_i(x_i, y_i)]. \tag{3.2}$$

It is easy to check that the other arguments still apply.

The general procedure is as before. Let $\underset{\sim}{N}^m=(N_1^m, \ldots, N_p^m)$ be the discrete approximation for $\underset{\sim}{N}=(N_1, \ldots, N_p)$ as in (2.1). Then Lemma 2.1 carries over with d in (3.1) or (3.2) and

$$\begin{aligned} A&=\bigcap_{i=1}^{p} A_i=\bigcap_{i=1}^{p} \{N_i(F)=0\} \\ B&=\bigcap_{i=1}^{p} B_i=\bigcap_{i=1}^{p} \bigcap_{j=1}^{m} \{N_i(B_j)\leqq 1\} \\ C&=\bigcap_{i=1}^{p} C_i=\bigcap_{i=1}^{p} \bigcap_{j=1}^{m^q} (\{N_i(G_j)=0, N_i(B_j)=1\} \cup \{N_i(B_j)=0\}) \end{aligned} \tag{3.3}$$

instead of (2.2)–(2.4). To obtain the analog of Lemma 2.2, let

$$a_m=\sup_{1\leqq j\leqq m^q} \left\{\sum_{k=1}^{n} \sum_{i=1}^{p} P[N_{ik}(B_j)=1]\right\}$$

and

$$b_m=\sum_{k=1}^{n} P\left\{\sum_{i=1}^{p} N_{ki}([0,1]^q)>1\right\}, \tag{3.4}$$

cf. (4.2)–(4.4) of [3] and (2.5) here. The bound in Lemma 2.2 then applies for $P\{d(\underset{\sim}{N}, \underset{\sim}{N}^m)>m^{-1}\}$ with a_m and b as in (3.4) and d as in (3.1) or (3.2).

For the second stage of our approximation, we use the $n p m^q$-dimensional array $\{X_{ik}(j), 1\leqq i\leqq p, 1\leqq j\leqq m^q, 1\leqq k\leqq n\}$ defined by

$$X_{ik}(j)=N_{ik}(B_j), \tag{3.5}$$

cf. (2.6). Let $\{Y_{ik}(j);\ 1\leqq i\leqq p,\ 1\leqq j\leqq m^q,\ 1\leqq k\leqq n\}$ be the associated $n p m^q$-dimensional array of independent Poisson random variables with the parameter of $Y_{ik}(j)$ being

$$\lambda_{ik}(j)=P\left\{N_{ik}(B_j)=1,\ \sum_{i=1}^{p} N_{ik}([0,1]^q)=1\right\}, \tag{3.6}$$

cf. (2.7). Then Lemma 2.3 carries over with $\sum_{i=1}^{p} N_{ik}([0,1]^q)$ instead of $N_k([0,1]^q)$ in the two terms on the right. Corollaries 2.2 and 2.3 also apply if we redefine c in (2.8) to be

$$c=\sum_{k=1}^{n} P\left\{\sum_{i=1}^{p} N_{ik}([0,1]^q)=1\right\}^2. \tag{3.7}$$

Finally, arguments differing very little from those of Lemma 2.5 and Theorem 2.1 yield identical statements here. Recall that the constants a_m, b, and c, the metric d, and the processes have been altered, however.

Corollary 3.1. *If* $a_m \leqq K_1 m^{-q}$, $b \leqq K_2 m^{-1}$, *and* $c \leqq K_3 m^{-1}$ *for the constants in* (3.4) *and* (3.7), *then*

$$\rho[\mathscr{L}(N), \mathscr{L}(V_m)] \leqq K_4 m^{-1},$$

where ρ *is the Prohorov metric in* (1.9) *on* $\mathscr{P}(D_{qp})$.

References

1. Bickel, P. J., Wichura, M. J.: Convergence criteria for multi-parameter stochastic processes and some applications. Ann. Math. Statist. **42**, 1656–1670 (1971)
2. Billingsley, P.: Convergence of Probability Measures. New York: Wiley 1968
3. Çinlar, E.: Superposition of point processes. Stochastic Point Processes: Statistical Analysis, Theory and Applications, Ed. P. A. W. Lewis, New York: John Wiley and Sons 549–606, 1972
4. Dudley, R. M.: Speeds of metric probability convergence. Z. Wahrscheinlichkeitstheorie verw. Gebiete **22**, 323–332 (1972)
5. Gnedenko, B. V., Kolmogorov, A. N.: Limit Distributions for Sums of Independent Random Variables, Second Edition. Reading, Massachusetts: Addison-Wesley 1968
6. Neuhaus, G.: On weak convergence of stochastic processes with multidimensional time parameter. Ann. Math. Statist. **42**, 1285–1295 (1971)
7. Straf, M. L.: Weak convergence of stochastic processes with several parameters. Proc. Sixth Berk. Symp. Stat. Prob. **2**, 187–221 (1972)
8. Whitt, W.: Preservation of rates of convergence under mappings. Z. Wahrscheinlichkeitstheorie verw. Gebiete. To appear

Ward Whitt
Department of Administrative Sciences
Yale University
2 Hillhouse Avenue
New Haven, Connecticut 06520
USA

(Received October 10, 1972)