Math. Proc. Camb. Phil. Soc. (1981), **90**, 305
Printed in Great Britain

Compensators and Cox convergence

BY T. C. BROWN
University of Bath

(*Received* 1 *August* 1980, *revised* 5 *January* 1981)

1. *Introduction, notation and definitions*

In a previous paper (4), the author showed that the martingale approach to point processes may be used to derive Poisson distributional limit theorems. Here this approach is extended to Cox convergence, general increasing processes are considered, and a regularity condition needed in (4) is removed (that of calculability). In Section 3 we give examples to show the need for the various hypotheses used in the main theorem of Section 2. This theorem is applied to various (dependent) thinnings and compoundings of point processes in Section 4.

In (16), Kallenberg has showed that Papangelou's conditional intensity measures (18) may be used to obtain Cox convergence. Apart from the difference that these involve conditioning on two sides instead of one, the contrast with the present work is that conditional intensity measures have only been defined for the σ-fields generated by the process; most of the applications given here use larger σ-fields. In (1), Aldous has defined a stronger mode of weak convergence and shown that the Poisson case of Theorem 1 is true in this mode for point processes. Kabanov, Lipster and Shiryayev (14) have extended the Poisson case to point processes whose compensators are deterministic (but not necessarily continuous). Finally, Jacod and Memin (12) have considered the more general question of convergence of semimartingales to processes of independent increments.

Let $\mathbb{R}$ denote the reals, $\mathbb{R}^+ = [0, \infty)$ and $\mathbb{N}$ the natural numbers $\{1, 2, \ldots\}$. Let $\mathscr{D}$ denote the set of right continuous left limit functions from $\mathbb{R}^+$ to $\mathbb{R}$. Let $\mathscr{D}\uparrow$ denote the subset of $\mathscr{D}$ whose members, f, are increasing and satisfy $f(0) = 0$. Let $\mathscr{C}(\mathscr{C}\uparrow)$ denote the continuous functions in $\mathscr{D}(\mathscr{D}\uparrow)$. Let $\mathscr{Y}$ be the subset of $\mathscr{D}\uparrow$ whose members have integer values and $\mathscr{X}$ the subset of $\mathscr{Y}$ consisting of those x for which

$$\Delta x(t) = 0 \quad \text{or} \quad 1$$

where

$$\Delta x(t) = x(t) - x(t-).$$

If ∂_j is the function $\partial_j : \mathscr{D}\uparrow \to \mathbb{R}^+$ where

$$\partial_j(x) = \inf\{t : x(t) \geqslant j\},$$

then $x \in \mathscr{X}$ has the property

$$x(\partial_j(x)) = j, \quad \text{if} \quad \partial_j(x) < \infty.$$

Let $(\Omega, \mathscr{A}, P)$ be a probability space. A *stochastic process*, $(N, \mathscr{F})$, is a mapping $N : \mathbb{R}^+ \times \Omega \to \mathbb{R}$ together with an increasing family $\mathscr{F} = \{\mathscr{F}(t)\}_{t \geqslant 0}$, of sub σ-fields of $\mathscr{A}$

0305-0041/81/0000-8540 $03.50 © 1981 Cambridge Philosophical Society

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

to which N is adapted (we do not require the 'conditions habituelles' for the reasons given by Jacod (11)). $\mathscr{F}$ is called a *history for* N and it is called *intrinsic* if

$$\mathscr{F}(t) = \mathscr{F}(0) \vee \sigma(N(z), z \leqslant t).$$

If N has paths in $\mathscr{D}\uparrow$, it induces a measure, $\tilde{N}$, on the $\mathscr{F}$-predictable σ-algebra, $\mathscr{P}$, of $\mathbb{R}^+ \times \Omega$ via

$$\tilde{N}(C) = E\left(\int IC\, dN\right) \quad (C \in \mathscr{P}). \tag{1·1}$$

We call $(N, \mathscr{F})$ an *increasing process* if N has paths in $\mathscr{D}\uparrow$ and the measure $\tilde{N}$ is σ-finite. The latter restriction amounts to a mild integrability condition, which is certainly satisfied if $EN(t) < \infty$ for all t. Unexplained terminology for stochastic processes is in (7).

As in (11), we define the *compensator* (or *dual predictable projection*) of an increasing process, $(N, \mathscr{F})$, to be the unique predictable process, A, for which

$$\tilde{A} = \tilde{N}.$$

It follows from (1·1) that

$$E\left(\int Y\, dN\right) = E\left(\int Y\, dA\right) \tag{1·2}$$

for all non-negative predictable Y, an equation which is the basis of most of the calculations here.

We shall call an increasing process $(N, \mathscr{F})$ a *jump process* if it has paths in $\mathscr{Y}$ and a *point process* if it has paths in $\mathscr{X}$. The *jump times* of a jump process $(N, \mathscr{F})$ are the sequence of stopping times $\tau_1, \tau_2, \ldots$, defined by $\tau_0 = 0$ and

$$\tau_i = \inf\{t > \tau_{i-1} : \Delta N(t) > 0\} \quad (i = 1, 2, \ldots).$$

A process $(N, \mathscr{F})$ with paths in $\mathscr{X}$ is automatically a point process since the sets $\{0\} \times \Omega$,

$$\{(t, \omega) : \tau_{i-1}(\omega) < t \leqslant \tau_i(\omega)\} \quad (i = 1, 2, \ldots)$$

constitute a predictable decomposition of $\mathbb{R}^+ \times \Omega$ whose elements have $\tilde{N}$ measure $\leqslant 1$. A jump process, $(N, \mathscr{F})$, is called a *Poisson process with mean function* m if for $s \leqslant t$, $N(t) - N(s)$ has a Poisson distribution with parameter $m(t) - m(s)$, independently of $\sigma(N(z), z \leqslant s)$. A jump process $(N, \mathscr{F})$ is called a *Cox process directed by* A, if N has the same distribution as $\hat{N} \circ A$, where $\hat{N}$ is an independent *unit rate* Poisson process (i.e. $\hat{N}$ has the identity as mean function).

Topologize $\mathscr{D}\uparrow$ with the *vague* topology in which

$$f_n \to f$$

if, for each continuity point, t, of f,

$$f_n(t) \to f(t).$$

This topology is Polish; for one complete separable metric ρ_M see (5). It is equivalent to the vague topology on measure space (cf. (15)) and convergence to an element of $\mathscr{X}$ is equivalent to Skorohod J_1 convergence (cf. (20)). *Convergence in distribution* of random elements of $\mathscr{D}\uparrow$, denoted $\to_d$, hereafter always means convergence in this

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

topology. From the last statement convergence in distribution to a point process is equivalent to Skorohod J_1 convergence in distribution.

Let *Proj* (t, k, N) (N an increasing process, $t \geqslant 0, k \in \mathbb{N}$) denote the set of mappings $\pi: \mathscr{D}\uparrow \to (\mathbb{R}^+)^k$ where

$$\pi(f) = (f(t_1), \ldots, f(t_k)),$$

for some

$$0 \leqslant t_1 < t_2 \ldots < t_k \leqslant t,$$

such that

$$P(\Delta N(t_i) > 0, \quad \text{for some } i) = 0.$$

By omitting some (or all) of N, t, k we enlarge *Proj* by taking a union over the possible values of that (or those) argument(s). We note that

$$N_n \to_d N$$

if, and only if,

$$\pi(N_n) \to_d \pi(N)$$

for all $\pi \in \text{Proj}\,(N)$ ((15), theorem 4·2).

Suppose $(N, \mathscr{F})$ is an increasing process and that $0 < \epsilon < 1$. Form the *ϵ point process* $(N_\epsilon, \mathscr{F})$ by

$$N_\epsilon(t) = \text{number of } s \leqslant t \text{ such that } \Delta N(s) \in (1-\epsilon, 1+\epsilon).$$

Define the *ϵ remainder process* $(N^\epsilon, \mathscr{F})$ by

$$N^\epsilon(t) = N(t) - \Sigma \Delta N(s),$$

where the sum is over $s \leqslant t$ such that $\Delta N(s) \in (1-\epsilon, 1+\epsilon)$. If $(N_n, \mathscr{F}_n)$ is a sequence of increasing processes we call it an *asymptotic point process sequence*, if, for each ϵ in $(0, 1)$,the compensator, B_n^ϵ, of $(N_n^\epsilon, \mathscr{F}_n)$ satisfies

$$B_n^\epsilon(t) \to_d 0,$$

for all $t > 0$. This concept is weaker than the one of the same name in (4), since

$$E(B_n^\epsilon(t)) = E(N_n^\epsilon(t)),$$

which in turn is dominated by the right side of equation (1·1) in (4). Also, it allows one to consider a sequence in which N_n only has jumps of size $1 + 1/n$.

2. *Results*

The first convergence result we prove is

THEOREM 1. *Let $(N_n, \mathscr{F}_n)$, $n \geqslant 1$, be an asymptotic point process sequence with corresponding compensators A_n. Suppose that, for each $t \geqslant 0$,*

$$A_n(t) \to_P A(t)$$

where A is an increasing $\mathscr{F}_n(0)$ measurable process $(\forall n)$ whose paths are in $\mathscr{C}\uparrow$. Then N_n converges in distribution to a Cox process directed by A.

If A is deterministic then it is certainly $\mathscr{F}_n(0)$ measurable. Moreover, in that case, one may replace convergence in probability by convergence in distribution. Thus,

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

Poisson convergence is obtained from the one dimensional distributional convergence of the compensators to a continuous function. Indeed, in this case, inspection of the proof here shows that $(\Omega, \mathscr{A}, P)$ may be allowed to vary with n.

The method of proof of Theorem 1 is a common one in distributional limit theorems; first, it is shown that the hypotheses of the theorem imply tightness and then it is demonstrated that all limit points of the sequence $(N_n, \mathscr{F}_n)$ have the correct limit. At one point in the second part of the proof we shall need the notion of stability; a sequence of random elements (X_n), $n \geqslant 1$, of a space converges *stably in distribution* if it converges in distribution to some X and

$$P(X_n \in B, F) \to$$

for each X – continuity set B and each $F \in \mathscr{A}$. The benefit of stable convergence for us is that N_n converges stably to some N implies

$$(N_n, A_n) \to_d (N, A)$$

under the hypotheses of the theorem. We remark that a proof of Theorem 1 which reduced it to the case where A is deterministic (as, for example, in (8)) is complicated by the possibility that a regular conditional distribution might not exist for some $\mathscr{F}_n(t)$. It is certainly true that the theorem could be established by the methods of (14). However the present proof, based on stability, is completely different and seems somewhat simpler.

Of central importance in the proof of Theorem 1 is the following characterization of a Cox process, which follows from (10) and also from the uniqueness result of (11) (cf. Remark below). The Theorem contains Watanabe's characterization of a Poisson process.

THEOREM 2. *Suppose $(N, \mathscr{F})$ is a point process whose compensator, A, is continuous and $\mathscr{F}(0)$ measurable. Then N is a Cox process directed by A.*

Remark. If the hypotheses of the theorem are satisfied then A is the compensator of the intrinsic point process $(N, \mathscr{G})$ where $\mathscr{G}(0) = \sigma(A)$. Any intrinsic Cox process, $(N', \mathscr{G}')$, directed by A, in which $\mathscr{G}'(0) = \sigma(A)$, has compensator A. Thus we may apply Theorem 3·4 of (11) to deduce that

$$N =_d N'.$$

However, we give an elementary proof based on (1·2), as it seems attractively simple.

Proof. Suppose first that $A(\infty) = \infty$. Let $\hat{A}(t)$, $t \geqslant 0$, be the stopping time defined by

$$\hat{A}(t) = \inf\{z : A(z) > t\}$$

and $(\hat{N}, \hat{\mathscr{F}})$ be the increasing process given by

$$\hat{N}(t) = N \circ \hat{A}(t)$$

$$\hat{\mathscr{F}}(t) = \mathscr{F}(\hat{A}(t)).$$

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

Now $(\hat{N}, \hat{\mathscr{F}})$ has the identity as compensator and if $Y(z,\omega)$ is $I(t, \hat{A} \circ A(t,\omega))(z)$ then Y is adapted and left continuous. Thus

$$\begin{aligned} E|\hat{N} \circ A(t) - N(t)| &= E(\int Y\, dN) \\ &= E(\int Y\, dA) \\ &= E(A \circ \hat{A} \circ A(t) - A(t)) \\ &= 0. \end{aligned}$$

Thus, $\hat{N} \circ A$ equals N except possibly on an evanescent set.

Let $\{\hat{\tau}_m\}$ be the jump times of $\hat{N}$ and $\hat{\tau}_0 = 0$. Fix $\lambda > 0$, $m = 0, 1, \ldots$, and $B \in \hat{\mathscr{F}}(\hat{\tau}_m)$. Define,

$$X(t,\omega) = \exp(-\lambda\{t - \hat{\tau}_m(\omega)\})\, IB(\omega)\, I(\hat{\tau}_m(\omega), \hat{\tau}_{m+1}(\omega))(t)$$

so that X is predictable with respect to $\hat{\mathscr{F}}$. Since N is a point process, so also is $\hat{N}$ and thus

$$\begin{aligned} E(\exp(-\lambda(\hat{\tau}_{m+1} - \hat{\tau}_m)) : B) &= E(\int X\, d\hat{N}) \\ &= E(\int X\, dt) \\ &= \{P(B) - E(\exp(-\lambda(\hat{\tau}_{m+1} - \hat{\tau}_m)) : B)\}/\lambda, \end{aligned}$$

by the first paragraph of the proof. From this equation we deduce $\hat{\tau}_{m+1} - \hat{\tau}_m$ is an exponential (1) random variable, conditional on $\hat{\mathscr{F}}(\hat{\tau}_m)$. Additionally, since $\hat{\mathscr{F}}(\hat{\tau}_m) \supset \hat{\mathscr{F}}(0) \supset \sigma(A)$, it also follows that A is independent of $\sigma(\hat{\tau}_1, \ldots) = \sigma(\hat{N})$. But this means that $\hat{N} \circ A$ is Cox directed by A, and by the first paragraph of the proof so is N.

In the general case, note that we require only that $\pi \circ N$ has the appropriate distribution, for each π in $\mathrm{Proj}(N)$. Let Z be a Poisson process independent of $\mathscr{F}(\infty)$, t be in $\mathbb{R}^+$ and $N'(s)$ be $N(s)$, if $s \leqslant t$ and $N(t) + Z(s-t)$, for $s < t$. If $\pi \in \mathrm{Proj}(t, N)$, then $\pi \circ N$ and $\pi \circ N'$ have the same distribution. Further, the compensator of N' is $A(s)$, for $s \leqslant t$, and $A(t) + s - t$, for $s > t$. This reduces the general case to the case already treated.

To prove Theorem 1 we need the following lemma, which may be of independent interest.

LEMMA 1. *Let A be the compensator of an increasing process $(N, \mathscr{F})$ with $EN(t)$ finite for all t. Let B be the compensator of the ϵ point process of $(N, \mathscr{F})$. Then*

$$E|A(t) - B(t)| \leqslant \epsilon E\{N(t)\} + E\{N^\epsilon(t)\}$$

where N^ϵ is the ϵ remainder process of $(N, \mathscr{F})$.

Proof. For each $0 \leqslant z < z'$,

$$|N(z') - N(z) - \{N_\epsilon(z') - N_\epsilon(z)\}| \leqslant \{N^\epsilon(z') - N^\epsilon(z)\} + \epsilon\{N(z') - N(z)\}.$$

Now $A(t)$ and $B(t)$ may be obtained as weak L_1 limits of discrete approximations $A_m(t)$ and $B_m(t)$ (for example, (4)). From the previous equation

$$E|A_m(t) - B_m(t)| \leqslant \epsilon E\{N(t)\} + E\{N^\epsilon(t)\}$$

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

Thus

$$\begin{aligned} E|A(t)-B(t)| &= E(\{A(t)-B(t)\}\operatorname{sign}\{A(t)-B(t)\}) \\ &= \lim_{m\to\infty} E(\{A_m(t)-B_m(t)\}\operatorname{sign}\{A(t)-B(t)\}) \\ &\leqslant \limsup_{m\to\infty} E|A_m(t)-B_m(t)| \\ &\leqslant \epsilon E\{N(t)\}+E\{N^\epsilon(t)\}, \end{aligned}$$

by (2·7).

Proof of Theorem 1. By using a diagonal argument, we may, for any subsequence of integers, find a further subsequence of integers, $\{n_k\}$, and a P null set, E, for which

$$A_{n_k}(q,\omega)\to A(q,\omega), \quad \omega\notin E,$$

and q a rational in $[0,t]$. But increasing functions, defined on $[0,t]$, which converge at rational points of their domain to a continuous function on $[0,t]$, actually converge uniformly. Thus

$$\sup_{0\leqslant s\leqslant t}|A_{n_k}(s)-A(s)|\to_{\text{a.s.}} 0,$$

and hence

$$\sup_{0\leqslant s\leqslant t}|A_n(s)-A(s)|\to_P 0. \tag{2·1}$$

Let Q_A be the Cox measure on $\mathscr{Y}$, which is directed by the distribution of A. We are required to show that, for each $k\in\mathbb{N}$, $\alpha\in\mathbb{N}^k$, $s>0$, and $\Pi\in\operatorname{Proj}(s,k)$, we have

$$P(\Pi\circ N_n\leqslant\alpha)\to Q_A(\Pi\leqslant\alpha). \tag{2·2}$$

To show this, we may assume that $N_n(t)=N_n(s)$, $A(t)=A(s)$ and $A_n(t)=A_n(s)$, for all $t\geqslant s$, since this does not alter the hypotheses of the Theorem and does not change any of the quantities in (2·2). Since $A(s)$ is finite,

$$[A(s)\leqslant L]\uparrow\Omega, \quad L\to\infty.$$

Now, if G and H are events,

$$|P(G\,|\,H)-P(G)|\leqslant\{1-P^2(H)\}/P(H).$$

Hence, as $P(H)\to 1$, $P(G\,|\,H)\to P(G)$, uniformly in G. Thus to verify (2·2), it suffices to show that, for each $L\in\mathbb{N}$,

$$P(\Pi\circ N_n\leqslant\alpha\,|\,A(s)\leqslant L)\to Q_{\mathbf{L}}(\Pi\leqslant\alpha),$$

where $\mathbf{L}$ is the distribution of A conditional on $[A(s)\leqslant L]$. Since $A\in\mathscr{F}_n(0)$, it is straightforward to show that the compensator of $(N_n,\mathscr{F}_n)$ with respect to P conditional on $[A(s)\leqslant L]$ is A_n (with all processes restricted to $[A(s)\leqslant L]$ ((4), lemma 5)). Combining all this we see that we may assume that $A(s)\leqslant L$, for all s.

Next we reduce to the case where both $\{A_n(t)\}_n$, $\{N_n(t)\}_n$ are uniformly integrable, for each t. Define the stopping times, S_n, by

$$S_n=\inf\{t:A_n(t)\geqslant A(t)+1\}.$$

Since $A\in\mathscr{F}_n(0)$, S_n is predictable. To see this, note that S_n is the debut of the predictable set $\{(t,\omega):A_n(t,\omega)-A(t,\omega)\geqslant 1\}$, which contains the graph of S_n and use (7), IV, T16.

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

Thus

$$Z_n(t,\omega) = I(0, S_n(\omega))(t)$$

is a predictable process. Hence

$$N_n'(t) = \int_0^t Z_n dN_n$$

is an increasing process, which has compensator

$$A_n'(t) = \int_0^t Z_n dA_n \tag{2·3}$$

since the latter is easily seen to be $\mathscr{F}_n$ predictable and satisfy equation (1.2). M it is easy to see that $\{N_n'\}$ is asymptotically a point process sequence and that, for

$$A_n'(t) \sim_P A_n(t), N_n'(t) \sim_P N_n(t).$$

But

$$A_n'(t) \leqslant L+1, \tag{2·4}$$

by (2·3), so that $\{A_n'(t)\}$ is uniformly integrable. Thus, we may assume $\{A_n(t)\}$ uniformly bounded.

We may also assume that each N_n is a point process. To see this let $t > 0$ be arbitrary and choose $\epsilon_n \to 0$ so that if C_n is the compensator of the ϵ_n remainder process

$$C_n(t) \to_P 0.$$

If B_n is the compensator of $(M_n, \mathscr{F}_n)$, the ϵ_n point process of $(N_n, \mathscr{F}_n)$ stopped at t, then for all s, Lemma 1 gives

$$E|B_n(s) - A_n(s \wedge t)| \leqslant \epsilon_n E\{A_n(t)\} + E\{C_n(t)\} \to 0, \quad \text{uniformly in } s,$$

because $C_n \leqslant A_n \leqslant L+1$. Thus, for all s,

$$B_n(s) \to_P A(s \wedge t).$$

Further $E|M_n(s) - N_n(s \wedge t)|$ is also bounded by the right side of the last inequality. Thus, if Theorem 1 is true for point processes then $N_n(\cdot \wedge t)$ converges to a Cox process directed by $A(\cdot \wedge t)$. Because t was arbitrary, this is enough.

By Lemma 18·12 of (17) we have for all t

$$\begin{aligned} E\{N_n^2(t)\} &\leqslant 2E\{N_n(t) A_n(t)\} + E(A_n(t)) \\ &\leqslant 2(L+1)^2 + (L+1), \end{aligned}$$

by (2·4). Hence $\{N_n(t)\}$ is also uniformly integrable.

But this last statement also gives tightness in the vague topology ((15), lemma 4·5). Hence, it remains only to show that all limits point of $\{N_n\}$ are Cox processes directed by A. To do this, let us suppose that $\{N_n\}$ satisfies the hypotheses of the Theorem and that $N_n \to_d Q$, a measure on $\mathscr{D}\uparrow$. Since we only wish to identify Q, we may assume that $\{N_n\}$ shares any desirable properties that subsequences of it possess.

We first show $Q(\mathscr{X}) = 1$. We need only show that for each $j = 1, 2, \ldots$

$$Q\{y: \delta_{j+1}(y) - \delta_j(y) < \epsilon \quad \text{and} \quad \delta_j(y) < \infty\} \to 0, \quad \epsilon \to 0, \tag{2·5}$$

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

where $\delta_0 = 0$. It is straightforward to see that $\delta_j : (\mathscr{Y}, \rho_M) \to [0, \infty]$ is continuous, for each j. Fix j and write U_n for $\delta_j \circ N_n$, V_n for $\delta_{j+1} \circ N_n$. For large enough t

$$\begin{aligned} Q(\delta_{j+1} - \delta_j < \epsilon, \delta_j < \infty) &\leqslant Q(\delta_{j+1} - \delta_j < \epsilon, \delta_j < t - \epsilon) + \epsilon \\ &\leqslant \liminf_n P(V_n - U_n < \epsilon, U_n < t - \epsilon) + \epsilon. \end{aligned}$$

Let J be the predictable process $J(z, \omega) = I(U_n(\omega), U_n(\omega) + \epsilon](z)$ for $z \leqslant t$ and 0 otherwise. Then, as before, we obtain

$$\begin{aligned} P(V_n - U_n < \epsilon, U_n < t - \epsilon) &\leqslant E(\int J \, dA_n) \\ &\leqslant E(\sup_{0 \leqslant s \leqslant t} \{A_n(s + \epsilon) - A_n(s)\}) \\ &\to E(\sup_{0 \leqslant s \leqslant t} \{A(s + \epsilon) - A(s)\}) \end{aligned}$$

in view of (2·1) and the fact that the supremum in brackets is dominated by $A_n(t + \epsilon)$, which forms a uniformly integrable sequence. Equation (2·5) now follows by the continuity of A and the fact that $EA(t + \epsilon) < \infty$. Thus $Q(\mathscr{X}) = 1$.

Let $\mathscr{I}$ be a countable field which generates $\mathscr{B}$, the Borel sets of $(\mathscr{D}\uparrow, \rho_M)$. For each D in $\mathscr{I}$, $\{I[N_n \in D]\}_n$ is uniformly integrable and thus there is a (possibly D dependent) subsequence along which $I[N_n \in D]$ converges in the weak L_1 topology. By the diagonal method, we remove any dependence on D in the subsequence. Further, since $\{(N_n, IE)\}_n$ is tight, for each E in $\mathscr{A}$, we see that $N_n \to_d Q$ (stably) along the subsequence. Passing to this subsequence, apply the argument of (2) to obtain the existence of a point process, N, for which $N_n \to_d N$ (stably); although the argument given by Aldous and Eagleson is for random variables, it extends to the present case in a straightforward way, since $(\mathscr{D}\uparrow, \rho_M)$ is Polish. Certainly, (2·1) implies that $\rho_M(A_n, A) \to_P 0$, so that $(N_n, A_n) \to_d (N, A)$ (stably), by the remark after Theorem 1′ in (2).

Let $\mathscr{F}$ be the intrinsic history of N with $\mathscr{F}(0) = \sigma(A)$. Let $0 \leqslant s \leqslant t$. Note that $\mathscr{F}(s)$ is generated by the class of sets, F, of the form $[A \in B] \cap [\pi \circ N = m]$, where B is in $\mathscr{B}$, $k \in N, \pi \in \mathrm{Proj}(s, k, N)$, and $m \in \mathbb{N}^k$. This class is closed under finite intersections. Let $F_n = [A \in B] \cap [\pi \circ N_n = m]$. Since $F_n \in \mathscr{F}_n(s)$, the martingale property of $N_n - A_n$ gives

$$E(N_n(t) - A_n(t) : F_n) = E(N_n(s) - A_n(s) : F_n).$$

Moreover, the random variables on the left and right sides of this equation are uniformly integrable and converge in distribution to their counterparts with all n's removed $(n \to \infty)$, by the stable convergence of (N_n, A_n). Thus

$$E(N(t) - A(t) : F) = E(N(s) - A(s) : F),$$

which suffices to show that $N - A$ is an $\mathscr{F}$ martingale, from the first four sentences of the paragraph. The proof is finished by applying Theorem 2.

Straightforward generalizations of Theorem 1 exist in two directions. It is very simple to extend the Theorem to point processes on $\mathbb{R}$. Additionally, one may obtain a result for random measures on $\mathbb{R}^+ \times E$, where E is some other space; the limit will then be a marked point process. Further, one may obtain rates of convergence using

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

an embedding technique. Finally, some compensator statements may be made about the convergence of distributions on $\mathscr{X}$ to non Cox limits; however these are less applicable. We shall only describe the second generalization (details of the others are in (5)).

We use the setup of (11) with some minor modifications. The space E is assumed to be locally compact, second countable and Hausdorff (LCCH). Note that since such a space is Polish, it is also Lusin, as required by Jacod. An *equipped random measure*, $(N, \mathscr{G})$, on $\mathbb{R}^+ \times E$ will be a random measure, N, on $\mathbb{R}^+ \times E$, together with a history, $\mathscr{G}$, of N (that is a filtration, $\mathscr{G}$, such that for all Borel sets B of E, $N((0,s] \times B)) \in \mathscr{G}(s)$) such that M_N is σ-finite (cf. (11) for the definition of M_N). The *compensator* of $(N, \mathscr{G})$ will then be the 'dual predictable projection' of (11), Lemma 2·2. For each B, let $(N^B, \mathscr{G})$ be the increasing process given by

$$N^B(s) = N((0,s] \times B).$$

The compensator, A^B, of $(N^B, \mathscr{G})$ then satisfies

$$A^B(s) = A((0,s] \times B),$$

where A is the compensator of $(N, \mathscr{G})$. Define an *asymptotic marked point process sequence* to be an equipped sequence of random measures, N_n, on $\mathbb{R}^+ \times E$ such that N_n^E is an asymptotic point process sequence. We then have

PROPOSITION 1. *Suppose $(N_n, \mathscr{G}_n)$ is an asymptotic marked point process sequence with compensators, A_n. Let A be a random measure on $\mathbb{R}^+ \times E$ such that $A \in \mathscr{G}_n(0)$, for all n, and that*

$$A(\{s\} \times E) = 0,$$

for all $s \leqslant 0$. If for each $s \geqslant 0$ and B, a Borel set of E,

$$A_n((0,s] \times B) \to_P A((0,s] \times B),$$

then N_n converges in distribution to a Cox process, N, directed by A (i.e. *a random measure which, conditional on A, is Poisson with mean measure A*).

Proof. Since N_n^E satisfies the conditions of Theorem 1, $N_n((0,s] \times E)$ converges in distribution for all s. From (15), Lemma 4·5, this is enough for the tightness of $\{N_n\}$. Since A is certainly diffuse, a Cox process, N, directed by A is a marked point process. Thus, from Theorem 4·7 of (15), it is sufficient to show that $N_n(T)$ converges in distribution to $N(T)$, for all T in a suitable DC-ring. Choose the ring, T, consisting of finite unions of sets of the form $I \times B$, where I is a finite interval of $\mathbb{R}^+$ and B is a Borel set of E.

Each T in $\mathscr{T}$ may be written as a union of sets of the form $I \times B$, where the I's are disjoint. Let

$$N_{nT}(s) = \Sigma N_n((0,s] \cap I \times B) \tag{2·6}$$

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

where the sum is over the $I \times B$ in T with the I's disjoint. From the definition of A_n, it is straightforward to show that if A_{nT} is the compensator of $(N_{nT}, \mathscr{G}_n)$ then (2·6) holds with A replacing N throughout. Thus, for each s,

$$A_{nT}(s) \to A_T(s) = \Sigma A((0, s] \cap I \times B)$$

with the same summation as in (2·6). Further A_T is continuous and $\mathscr{G}_n(0)$ measurable by assumption. Also the compensator of the ϵ remainder process of $(N_{nT}, \mathscr{G}_n)$ is dominated by the compensator of the ϵ remainder process of $(N_n^E, \mathscr{G}_n)$. Choosing s to be the supremum of the projection of T onto $\mathbb{R}^+$, Theorem 1 implies $N_{nT}(s) = N_n(T)$ converges in distribution to $N(T)$.

3. *Examples*

In this section some examples are presented, which show that various plausible extensions of Theorem 1 are false.

(1) *Convergence in probability.* It would be desirable to replace the hypothesis of convergence in probability of the compensators by convergence in distribution. However, let $\hat{N}$ be a unit rate Poisson process and A a continuous increasing process, which is independent of $\hat{N}$. Suppose that the distribution of A has two atoms of probability $1/2$; the first, f_1, is the function $t \mapsto t \ln 2$, while the second function, f_2, agrees with f_1 on $[0, 1]$, but is constant (at $\ln 2$) on $(1, \infty)$. Let $\mathscr{F}$ be the intrinsic filtration of N with $\mathscr{F}(0) = \sigma(A)$, where $N = \hat{N} \circ A$, so that $(N, \mathscr{F})$ is a Cox process with compensator A. Let ξ be the same as N on $[0, 1]$, while $\xi(t) = \hat{N}(t \ln 2)$, if $\xi(1) > 0$, while $\xi(t) = 0$, if $\xi(1) = 0$, for all $t \geqslant 1$. It is easy to see that ξ is adapted to $\mathscr{F}$ and that its compensator has the same distribution as A. However,

$$P(\xi(2) = 0) = 1/2,$$

whereas

$$P(N(2) = 0) = 3/8.$$

Letting $N_n = \xi$ for all n, N_n does not converge in distribution to N, but all compensators have the same distribution.

(2) *Asymptotic point process sequence.* It would seem more natural to define an asymptotic point process sequence by requiring that, for all $t \geqslant 0$,

$$P(N_n^t \notin \mathscr{X}) \to 0.$$

However, let N_{ni} $(i, n = 1, 2, \ldots)$ be a counting process with at most one jump, which occurs at i/n. Suppose the jump has size n and probability $1/n^2$ of occurrence. Suppose that $\{N_{ni}\}_{i=1}^{\infty}$ is an independent sequence and that $\mathscr{F}_n(t) = \sigma(N_{ni}(z), z \leqslant t, i = 1, 2, \ldots)$. It is easily seen that the compensator of $(\Sigma_{i=1}^{\infty} N_{ni}, \mathscr{F}_n)$ is $[nt]/n$ at t. This converges pointwise to the identity in $\mathscr{C}\uparrow$, as $n \to \infty$. On the other hand

$$P(N_n(1) > 0) \leqslant 1/n \to 0,$$

while

$$P(N_n^t \notin \mathscr{X}) \leqslant [nt]/n^2 \to 0.$$

(3) *Measurability of A.* It would be nice to dispense with the assumption that $A \in \mathscr{F}_n(0)$, for all n. A slight modification of example (1) shows that this is not possible.

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

Construct N' and ξ', with ξ' independent of N' and $d(\xi') = d(\xi), d(N') = d(\hat{N})$. Let A' be f_1 on $[\xi'(1) > 0], f_2$ on $[\xi'(1) = 0]$, $\mathscr{F}'$ be the intrinsic history of N' with $\mathscr{F}'(0) = \sigma(\xi')$, $\mathscr{G}'$ be the purely intrinsic history of ξ'. It is easy to check that N^1, with $N^1 = N' \circ A'$, is a Cox process with compensator A', and that $(\xi', \mathscr{G}')$ also has compensator A'.

(4) *Continuity of* A. It is, of course, necessary to have A continuous to obtain convergence to a Poisson process. The trivial example of a deterministic point process shows this. However, Kabanov, Lipster and Shiryayev (**14**) have obtained convergence to a non Poisson limit if A is deterministic and noncontinuous.

(5) *Converse.* A trivial counter-example to the converse of Theorem 1 may be obtained by taking a Poisson process with an intrinsic history, $\mathscr{F}$, containing *all* the history of the process in $\mathscr{F}(0)$. More interesting examples will be given in the next section.

4. *Applications*

Many applications of Theorem 1 concern thinnings of point processes. In much of the literature on thinnings of point processes, particular requirements are made for the dependence of the thinning operation on the underlying process. The compensator approach makes it possible to state a general criterion for Cox convergence, which is quite easy to verify in many different cases. In fact, we consider the more general concept of compounding; a jump process, $(N, \mathscr{F})$, is said to be a *compound* of a point process $(M, \mathscr{G})$ if, for each ω, the jump times of $N(\omega)$ are contained in those of $M(\omega)$. The process N is a *thinning* of M if N is also a point process.

PROPOSITION 2. *If $(N, \mathscr{F})$ is a compound of a predictable point process $(M, \mathscr{F})$ then the compensator, A, of $(N, \mathscr{F})$ is given by*

$$A(t) = \int_0^t p(u)\, dM(u) \tag{4·1}$$

where, if $\tau_1, \tau_2, \ldots$ are the jump times of M, then

$$p(\tau_i) = E(\Delta N(\tau_i)\, 1_{\mathscr{F}}(\tau_i -)) \quad (i = 1, 2, \ldots) \tag{4·2}$$

and $\Delta N(\infty)$ is defined to be zero.

Proof. By assumption

$$N(t) = \int_0^t \Delta N(u)\, dM(u).$$

Thus we may apply (**7**), V, T31 to conclude that A is given by (4·1), with p the predictable projection of ΔN. The result T16 of (**7**), *V* gives equation (4·2) for p.

An immediate corollary to the above Proposition and Theorem 1 is:

COROLLARY 1. *Suppose $(N_n, \mathscr{F}_n)$ is a compound of a predictable point process $(M_n, \mathscr{F}_n)$ and that p_n is defined by equation (4·2) with n subscripts. Further, suppose that $(N_n, \mathscr{F}_n)$ $(n = 1, 2, \ldots)$ is an asymptotic point process sequence and that for all t,*

$$\int_0^t p_n(u)\, dM_n(u) \to_P A(t), \tag{4·3}$$

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

where A has continuous paths and is $\mathscr{F}_n(0)$ measurable. Then N_n converges in distribution to a Cox process directed by A.

Examples to illustrate the use of this corollary follow.

(1) *Renyi thinning.* Suppose that M_n is a continuous random time change of a fixed point process M, so that

$$M_n(t) = M(\beta_n(t))$$

for some $\sigma(M)$, strictly increasing, continuous process β_n. Setting $\mathscr{F}_n(0) = \sigma(M)$, a simple case of the corollary occurs, if $(N_n, \mathscr{F}_n)$ is an intrinsic point process and p_n is $\sigma(M)$ measurable. Here, (4·3) and the continuity of A suffices for Cox convergence, as A is automatically $\mathscr{F}_n(0)$ measurable. This includes Renyi's theorem on thinning where M was supposed a renewal process, p_n a constant (in both variables) and $\beta_n(t) = t/p_n$. Of course, in many examples the probabilistic structure is simple and other proofs show that the distributional form of (4·3) is necessary and sufficient (cf. (6)), provided $p_n \to_d 0$ and that conditional on $\sigma(M)$ the thinnings are independent. If the latter requirement is dropped, then interesting counter-examples to the necessity of (4·3) can be found. A simple one is obtained from a renewal process, M, with inter-renewal times $X_1, X_2, \ldots$ distributed according to F. Let $\epsilon_n \to 0$, $q_n = F(\epsilon_n)$ and $\beta_n(t) = t/q_n$. Set $\Delta N_n(t)$ equal to the indicator of the event that, for some i, $\beta_n(t)$ is the ith renewal of M_n and $X_i \leqslant \epsilon_n$. Then $p_n(t)$ is the indicator of the same event, since N_n is $\sigma(M)$ measurable. However, it is clear that N_n converges in distribution to a stationary Poisson process (cf. (13)).

(2) *Success runs in Bernoulli trials.* Suppose that $Z_1, Z_2, \ldots$ is a sequence of Bernoulli trials with $P(Z_1 = 1) = \rho$ and that the trials occur at the epochs of the jumps of an independent point process M. We say a *success run of length r* occurs at t if, for some i, M has its ith jump at t and the last r Z's give a new non-overlapping success run i.e. if the event

$$E_i = [Z_i \ldots Z_{i-r+1} = 1] \cap \{[Z_{i-r} = 0] \cup [Z_{i-r} \ldots Z_{i-2r+1} = 1]\}$$

occurs (cf. (9), p. 305). The point process whose jumps occur at the epochs of r-success runs is called a *point process of r-success runs for M*. We have a special case of example (1) for a fixed M, but we prefer to allow M to vary with n.

COROLLARY 2. *If N_n is a point process of r_n-success runs for M_n, then $r_n \to \infty$ and for all t,*

$$M_n(t)\,\rho^{r_n}(1-\rho) \to_d t$$

implies N_n converges in distribution to a unit rate Poisson process.

Proof. We omit most n subscripts to avoid complicated expressions. Set $\mathscr{F}(0) = \sigma(M)$ and let $\{Z_i\}, \{E_i\}$ be defined as above. Let

$$\mathscr{F}(t) = \sigma(E_1, \ldots, E_{M(t)}, Z_1, \ldots, Z_{M(t)-r}).$$

Let $\{\tau_i\}$ be the jump times of M. It is easy to see that

$$p(\tau_i) = \rho^r\{I[Z_{i-r} = 0] + IE_{i-r}\}. \tag{4·4}$$

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

Let $m(t)$ be the integer part of

$$t/\{\rho^r(1-\rho)\},$$

and let

$$X(t) = \sum_{i=r}^{m(t)} p(\tau_i)$$

$$\Phi(t) = \rho^r(1-\rho)\,M(t).$$

Using (4·4), we see that the random variable $X(t)$ has two terms the second of which has expectation bounded by

$$\rho^{2r}m(t) \to 0 \quad (n \to \infty),$$

since

$$P(E_{i-r}) \leqslant \rho^r.$$

The first term, on the other hand, converges in distribution to t, since the Z's are independent and identically distributed. Thus

$$(X_n, \Phi_n) \to_d (i, i)$$

where i is the identity function. But

$$\int_0^t p_n(u)\,dM_n(u) = X_n \circ \Phi_n(t)$$

and the argument for convergence of this to t is given by Billingsley ((3), p. 145).

We have looked at a particularly simple case here to minimize notational complexity, but only these difficulties are encountered when considering other Poissonian limits or allowing ρ to vary with n.

(3) *U-statistics.* Poisson limit theorems for U-statistics have been proved in (19), but here we indicate how use of Corollary 1 makes their proofs simpler.

Let g_n be a symmetric 0–1 valued function of 2 arguments in some space, $E (n = 1, 2, \ldots)$. Let $X_1, X_2, \ldots$ be independent and identically distributed E-valued random variables. Set

$$M_n(t) = [nt]$$

$$\mathscr{F}_n(t) = \sigma(X_1, X_2, \ldots, X_{[nt]})$$

and, for $j = 2, 3, \ldots$,

$$\Delta N_n(j/n) = \sum_{i=1}^{j-1} g_n(X_i, X_j).$$

It is easy to see that

$$p_n(j/n) = \sum_{i=1}^{j-1} E(g_n(X_i, X_j) \mid X_i)$$

so that after some manipulation

$$\int_0^t p_n(u)\,dM_n(u) = \sum_{i=2}^{[nt]} (n-i+1)\,E(g_n(X_i, X_1) \mid X_i).$$

The latter has mean

$$\binom{[nt]}{2} E\{g_n(X_1, X_2)\}$$

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

and variance dominated by

$$[nt]\, n^2 E(g_n(X_1, X_2)\, g_n(X_1, X_3)).$$

We impose the conditions

$$\binom{n}{2} E\{g_n(X_1, X_2)\} \to \lambda > 0 \tag{4·5}$$

and

$$n^3 E\{g_n(X_1, X_2)\, g_n(X_1, X_3)\} \to 0. \tag{4·6}$$

In this case it is then clear that

$$\int_0^t p_n(u)\, dM_n(u) \to_d \lambda t^2.$$

The compensator of the ϵ-remainder process of $(N_n, \mathscr{F}_n)$ is B_n where

$$B_n(t) = \sum_{j=2}^{[nt]} E(\Delta N_n(j/n) : \Delta N_n(j/n) > 1 \mid X_1, \ldots, X_{j-1}), \quad (0 < \epsilon < 1).$$

Hence

$$\begin{aligned} EB_n(t) &= \sum_{j=2}^{[nt]} \sum_{1 \leqslant i < k < j} E\{g_n(X_i, X_j)\, g_n(X_k, X_j)\} \\ &\to 0. \end{aligned}$$

Hence, under (4·5), $(N_n, \mathscr{F}_n)$ is an asymptotic point process sequence. Thus, for example, under (4.5) and (4.6), $N_n(1)$ converges to a Poisson (λ) random variable, giving Theorem A of (19) for $m = 2$.

In many applications (cf. (19)) it is convenient to mark a point of N_n with, for example, the corresponding value of X. Then Proposition 1 is easily applied. For definiteness, suppose $\{h_n\}$ is another sequence of functions from E^2 into a (LCCH) space E'. Let T_n be the random measure on $\mathbb{R}^+ \times E'$ defined by

$$T_n((0, t] \times B) = \sum_{1 \leqslant i < j \leqslant [nt]} g_n^B(X_i, X_j)$$

where, for each Borel set B of E',

$$g_n^B(X_i, X_j) = g_n(X_i, X_j)\, I[h_n(X_i, X_j) \in B].$$

If now (4·6) holds and

$$\binom{n}{2} E\{g_n^B(X_i, X_j)\} \to \mu(B)$$

for some measure μ on E', then the argument above and Proposition 1 shows that T_n converges in distribution to a Poisson process with mean measure, μ', given by

$$\mu'((0, t] \times B) = \mu(B)\, t^2.$$

We may use this result to see, for example, that the point process consisting of the positions of each of the close pairs in a random sample, $X_1, \ldots, X_n$, of points on the plane is asymptotically Poisson (cf. (5), for details).

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press

It is a pleasure to thank G. K. Eagleson, J. Jacod, F. Papangelou, and J. Grandell for helpful comments.

REFERENCES

(1) Aldous, D. Weak convergence of stochastic processes, for processes viewed in the Strasbourg manner. Preprint (1979).
(2) Aldous, D. and Eagleson, G. K. On mixing and stability of limit theorems. *Ann. Probability* **6** (1978), 325–331.
(3) Billingsley, P. *Convergence of probability measures* (John Wiley, New York, 1968).
(4) Brown, T. C. A martingale approach to the Poisson convergence of simple point processes. *Ann. Probability* **6** (1978), 615–628.
(5) Brown, T. C. Some distributional approximations for random measures. PhD thesis, University of Cambridge, 1979.
(6) Brown, T. C. Position dependent and stochastic thinning of point processes. *Stochastic Processes and Appl.* **9** (1979), 189–193.
(7) Dellacherie, C. *Capacités et processus stochastiques* (Springer-Verlag, Berlin, 1972).
(8) Eagleson, G. K. Martingale convergence to mixtures of infinitely divisible laws. *Ann. Probability* **3** (1975), 557–562.
(9) Feller, W. *An introduction to probability theory and its applications*, vol. 1, 3rd edition (John Wiley, New York, 1968).
(10) Grigelionis, B. The characterization of stochastic processes with conditionally independent increments. *Liet. Matem. Rink.* **18** (1977), 75–86.
(11) Jacod, J. Multivariate point processes: predictable projection, Radon–Nikodym derivatives, representation of martingales. *Zeit. Wahrscheinlichkeitstheorie* **31** (1975), 235–253.
(12) Jacod, J. and Memin, J. Sur la convergence des semimartingales vers un processus à accroissements indépendents. Séminaire Probabilités. XIV. *Springer Lecture Notes in Maths*, **784** (1980).
(13) Jagers, P. and Lindvall, T. Thinning and rare events in point processes. *Zeit. Wahrscheinlichkeitstheorie* **28** (1974), 89–98.
(14) Kabanov, Y. Lipster, R. and Shiryayev, A. Some limit theorems for simple point processes (martingale approach). *Stochastics* **3** (1980), 203–216.
(15) Kallenberg, O. *Random measures* (Academic Press, London, 1976).
(16) Kallenberg, O. On the asymptotic behaviour of line processes and systems of non-interacting particles. *Zeit. Wahrscheinlichkeitstheorie* **43** (1978), 65–95.
(17) Lipster, R. S. and Shiryayev, A. N. *Statistics of random processes.* II (Springer-Verlag, New York, 1978).
(18) Papangelou, F. The conditional intensity of general point processes and an application to line processes. *Zeit. Wahrscheinlichkeitstheorie* **28** (1974), 207–226.
(19) Silverman, B. and Brown, T. C. Short distances, flat triangles and Poisson limits. *J. Appl. Probab.* **15** (1978), 815–825.
(20) Whitt, W. Representation and convergence of point processes on the line. Preprint (1973).

https://doi.org/10.1017/S030500410005876X Published online by Cambridge University Press