Adv. Appl. Prob. **25**, 255–260 (1993)
Printed in N. Ireland

LETTER TO THE EDITOR

ON CONDITIONAL INTENSITIES AND ON INTERPARTICLE CORRELATION IN NON-LINEAR DEATH PROCESSES

TIMOTHY C. BROWN,* *University of Western Australia*
PETER DONNELLY,** *Queen Mary and Westfield College, London*

Abstract

Ball and Donnelly (1987) announced a result giving circumstances in which there is positive or negative correlation between the death times in a non-linear, Markovian death process. A proof is provided here, based on results concerning the distribution of optional random variables in terms of their conditional intensities.

CORRELATION; MARTINGALE

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 60J05

1. Introduction

In this note we establish a general formula for the joint distribution of a set of non-negative random variables, in terms of their conditional intensities with respect to a history for which all the random variables are optional (i.e. stopping) times. The tool for this is a simple result, whose proof is based on Fubini's theorem, which gives the tail of the distribution of a single optional time, conditional on an event known at a previous time. The general formula leads to a conditional intensity criterion for when there is positive or negative correlation between a set of optional times.

In Section 3, the criterion is applied to establish some circumstances in which there is positive or negative correlation between the death times in a non-linear Markovian death process. Positive correlation occurs when the death rate for an individual decreases with increasing population size and negative correlation when it decreases. This result was announced in Ball and Donnelly (1987), but the proof in their paper has an error. Apart from the conditional intensity criterion, the key tool in the proof is *multivariate total positivity of order 2* and its consequences, as given by Karlin and Rinott (1980).

2. General result

Let T be an optional time relative to a history $\{\mathscr{F}_t\}$, i.e. T is a non-negative random variable such that $[T \leqq t] \in \mathscr{F}_t$ (the information known at time t) for all $t \geqq 0$. We suppose throughout that t has a *conditional intensity* λ which is defined to be an adapted, non-negative stochastic process such that, if

$$N_t = I[T \leqq t], \tag{2.1}$$

Received 16 March 1992; revision received 25 August 1992.

* Present address: Department of Statistics, University of Melbourne, Parkville, VIC 3052, Australia.

** Postal address: School of Mathematical Sciences, Queen Mary and Westfield College, Mile End Road, London E1 4NS, UK.

then

$$Nt - \int_0^t \lambda_s \, ds$$

is a martingale (see Brémaud (1981), Chapter 1, for definitions and discussion).

The following theorem appears not to have been stated elsewhere, but the method of proof has much in common with Brémaud (1981), p. 77 and Melamed and Walrand (1986).

Theorem 1. Suppose T is an optional time with conditional intensity λ. Let $0 \leqq s \leqq t$ and A be an event in $\mathcal{F}_s$ such that $A \subseteq [T > s]$. Suppose further that, for each $z \geqq 0$, A_z, is an event such that $A_z \supseteq [\lambda_z \neq 0]$ and $\boldsymbol{P}(A \cap Z_z) > 0$. Then

$$\boldsymbol{P}(T > t \mid A) = 1 - \int_s^t \boldsymbol{E}(\lambda_z \mid A \cap A_z)\boldsymbol{P}(A_z \mid A) \, dz. \tag{2.2}$$

In particular,

$$\boldsymbol{P}(T > t \mid A) = \exp\left(-\int_s^t \boldsymbol{E}(\lambda_z \mid A \cap [T > z]) \, dz\right), \tag{2.3}$$

where the answer is interpreted as 0 if the integral is infinite.

Proof. Let N be as in (2.1). Then, because $[T < s]$ and A are disjoint,

$$\boldsymbol{P}(T > t \mid A) = 1 - \frac{\boldsymbol{E}\left\{\int_s^t IA \, dN_z\right\}}{\boldsymbol{P}(A)}$$

$$= 1 - \frac{\boldsymbol{E}\left\{\int_s^t IAIA_z\lambda_z \, dz\right\}}{\boldsymbol{P}(A)}$$

upon using the martingale property (2.1) and the fact that $IA_z\lambda_z = \lambda_z$. But the order of integration of dz and dP may be reversed by the Fubini–Tonelli theorem, and this leads to (2.2). For (2.3), take $A_z = [T > z]$ and note that $\boldsymbol{P}(A \cap [T > s]) = \boldsymbol{P}(A) > 0$, by assumption. Let u be the infimum of z such that $\boldsymbol{P}(A \cap [T > z]) = 0$ so that $u > s$. Moreover, $z < u$ implies $\boldsymbol{P}(A \cap [T > z]) > 0$. Equation (2.2) gives for $v < u$

$$\boldsymbol{P}(T > v \mid A) = 1 - \int_s^v \boldsymbol{E}(\lambda_z \mid A \cap [T > z])\boldsymbol{P}(T > z \mid A) \, dz.$$

The solution to the previous integral equation is the exponential formula (2.2) with t replaced by v for $v < u$ (see, for example, Brémaud (1981), p. 338). But $\boldsymbol{P}(T = u \cap A)$ is the expected value of the integral with respect to Lebesgue measure of λIA over the singleton $\{u\}$, and hence it is zero. Thus $\boldsymbol{P}(T > v \mid A) \to 0$ as $v \uparrow u$. Thus, since we have shown (2.3) is valid for $v < u$, the integral in (2.3) is infinite for $v \geqq u$, and the formula (2.3) is valid for all $v \geqq s$ with the interpretation given in the theorem.

Theorem 1 can be used to write down expressions for the joint distributions of optional times in terms of their conditional intensities.

Theorem 2. Suppose that relative to a common history $T_1, T_2, \cdots, T_n$ (for some $n \geqq 1$) are optional times with conditional intensities $\lambda_1, \lambda_2, \cdots, \lambda_n$. Let $\wedge$ denote the minimum, $S_0 = 1$, $S_1 = T_1 \wedge T_2 \wedge \cdots \wedge T_n, \cdots, S_i = T_i \wedge T_{i+1} \wedge \cdots \wedge T_n, \cdots, S_n = T_n$. Then, for $0 = t_0 \leqq t_1 \leqq t_2 \leqq \cdots \leqq t_n$,

$$\boldsymbol{P}(T_1 > t_1 \cap T_2 > t_2 \cap \cdots \cap T_n > t_n) = \exp\left(-\sum_{j=1}^n \int_0^{t_j} f_j(z) \, dz\right), \tag{2.4}$$

where, for $1 \leqq i \leqq j$ and $t_{i-1} < z \leqq t_i$,

$$f_j(z) = \boldsymbol{E}(\lambda_j(z) \mid S_0 > t_0 \cap \cdots \cap S_{i-1} > t_{i-1} \cap S_i > z). \tag{2.5}$$

(The event $[S_0 > t_0]$ is certain: it has been included merely to make the formula the same for $i = 1$.)

Proof. Because of the assumed ordering of the t's, the probability on the left of (2.4) is unchanged by replacing each occurrence of T with S. We may therefore write the probability as a product over $i = 1, \cdots, n$ of the conditional probabilities $\boldsymbol{P}(S_i > t_i \mid A_i)$ where $A_i = [S_0 > t_0 \cap \cdots \cap S_{i-1} > t_{i-1}]$. From the definition, the conditional intensity for S_i is easily seen to be $(\lambda_i + \cdots + \lambda_n)$ up to S_i, because of the assumption of a common history. Theorem 1 thus gives

$$\boldsymbol{P}(S_i > t_i \mid A_i) = \exp\left(-\int_{t_{i-1}}^{t_i} E(\lambda_i(z) + \cdots + \lambda_n(z) \mid A_i \cap S_i > z)\, dz\right). \tag{2.6}$$

Summing the arguments of these exponentials that are multiplied to give the probability on the left side of (2.4), expanding the conditional intensity on the right side of (2.6) as a sum, and collecting together the integrals in which $\lambda_j(z)$ appears, we get (2.4) and (2.5).

Theorem 2 can be used to compare the probability of the intersection of events in (2.4) with the product of the probabilities of the individual events.

Corollary 1. Suppose the conditions of Theorem 2 hold. Then

$$\boldsymbol{P}(T_1 > t_1 \cap T_2 > t_2 \cap \cdots \cap T_n > t_n) \overset{(\geqq)}{\leqq} \prod_{i=1}^{n} \boldsymbol{P}(T_i > t_i), \tag{2.7}$$

if, for each $i = 1, \cdots, n$ and $j \geqq i$

$$\boldsymbol{E}(\lambda_j(z) \mid S_0 > t_0 \cap S_1 > t_1 \cap \cdots \cap S_{i-1} > t_{i-1} \cap S_i > z) \overset{(\leqq)}{\geqq} \boldsymbol{E}(\lambda_j(z) \mid T_j > z). \tag{2.8}$$

Proof. The right-hand side of (2.8) is the function that arises as $f_j(z)$ when applying (2.4) to the single time T_j (that is, taking $n = 1$). If the resulting exponentials in (2.4) are multiplied together for $j = 1$ to n, the integrals in the exponents add to produce an exponential of exactly the same form as if (2.4) is applied to calculate the probability on the left of (2.7). Condition (2.8) specifies that, in each integral, one integrand dominates the other pointwise and the corollary then comes from the fact that the negative exponential is decreasing.

3. Non-linear death processes

In this section we use Corollary 1 to prove a result about interparticle correlation in death processes. Suppose that $\{X(T), t \geqq 0\}$ is a Markov death process with $X(0) = n$ and death rates $\mu_n, \mu_{n-1}, \cdots, \mu_1$. That is, $\{X(t), t \geqq 0\}$ is a continuous-time Markov chain with infinitesimal transition rates

$$\lim_{h \downarrow 0} h^{-1} \boldsymbol{P}(X(t+h) = j \mid X(t) = k) = \begin{cases} \mu_k & \text{if } j = k-1 \\ 0 & \text{if } j \neq k, k-1. \end{cases}$$

Augment the process to describe the fates of n individuals labelled from $\{1, 2, \ldots, n\}$ in the natural way: initially all individuals are alive and at each (downward) jump of the death process $X(\cdot)$, one of the existing living individuals is chosen (randomly and uniformly) to die, with each of these choices being independent of the past history of the process. Write T_i, $i = 1, 2, \cdots, n$ for the time at which individual i dies. For an equivalent definition define $\mathscr{F}_t = \sigma\{T_i I[T_i \leqq t], \; i = 1, 2, \cdots, n\}$ and stipulate that relative to this history, T_i, $i =$

1, 2, $\cdots$, n, has conditional intensity λ_i given by

$$\lambda_i(t) = X(t)^{-1}\mu_{X(t)}I[T_i > t].$$

Following a conjecture of Faddy (1985), Ball and Donnelly (1987) announce a result relating the correlation of the death times $T_1, \ldots, T_n$ to features of the death rates $\mu_n, \ldots, \mu_1$. This correlation structure is related to variability properties of the death process, see Ball and Donnelly (1987) for a discussion. Unfortunately, there is an error in their proof. Specifically, the second displayed equation on p. 760 is incorrect. (To see this, take $n = 2$, $t_1 \leqq t_2 = t$.) It seems that one cannot write such conditional probabilities in terms of conditional hazard rates. The result is nonetheless true. We show here how to apply the techniques of Section 2; an alternative proof, based on correlation inequalities for Markov processes, follows as in Donnelly (1993). Lefèvre and Michaletzky (1990), Barbour (personal communication) and Liggett (personal communication) have given alternative proofs of the positive correlation part of the result, in fact of a stronger correlation result. The approach here deals equally with both forms of correlation.

The error in Ball and Donnelly (1987) is effectively duplicated in Lefèvre and Michaletzky (1990) in their study of a linear death process in a random environment. In that context also, an alternative approach could be based on Corollary 1 here. In fact, it seems easier to proceed directly and exploit the conditional independence structure. This is done in Lefèvre and Milhaud (1990) to establish a much more general result.

Theorem 3. For the death process described above, with death times $T_1, T_2, \cdots, T_n$, if $0 \leqq t_1, t_2, \cdots, t_n$, then

$$\boldsymbol{P}(T_1 > t_1 \cap T_2 > t_2 \cap \cdots \cap T_n > t_n) \geqq \boldsymbol{P}(T_1 > t_1)\boldsymbol{P}(T_2 > t_2) \cdots \boldsymbol{P}(T_n > t_n) \tag{3.1}$$

if the average death rates $\mu_1, \mu_2/2, \cdots, \mu_n/n$ form a decreasing sequence, while

$$\boldsymbol{P}(T_1 > t_1 \cap T_2 > t_2 \cap \cdots \cap T_n > t_n) \subset \boldsymbol{P}(T_1 > t_1)\boldsymbol{P}(T_2 > t_2) \cdots \boldsymbol{P}(T_n > t_n) \tag{3.2}$$

if the average death rates $\mu_1, \mu_2/2, \cdots, \mu_n/n$ form an increasing sequence.

Proof. Because $T_1, T_2, \cdots, T_n$ are exchangeable, we may assume, without loss of generality, that $t_1 \leqq t_2 \leqq \cdots \leqq t_n$. We will establish (2.8) for these times. The left-hand side of (2.8) is $\boldsymbol{E}(X(z)^{-1}\mu_{X(z)} \mid A_i \cap [S_i > z])$ (with A_i as in the proof of Theorem 2) while the right-hand side is $\boldsymbol{E}(X(z)^{-1}\mu_{X(z)} \mid T_j > z)$. In view of the assumptions on the sequence μ_k/k the inequalities required for (3.1) and (3.2) will both follow from the fact that the distribution of $X(z)$ given A_i and $[S_i > z]$ is stochastically greater than (or equal to) the distribution of $X(z)$ given $T_j > z$. To show this fact, writing $B_j = [T_j > z]$, we need only show that for an increasing function $\phi : \mathbb{R} \to \mathbb{R}$

$$\boldsymbol{E}\{\phi(X(z))I(A_i \cap [S_i > z])\}\boldsymbol{P}(B_j) \geqq \boldsymbol{E}(\phi(X(z))IB_j)\boldsymbol{P}(A_i \cap [S_i > Z]). \tag{3.3}$$

Conditioning on $\boldsymbol{X} \equiv (X_1, \cdots, X_{i-1}, X_i) \equiv (X(t_1), \cdots, X(t_{i-1}), X(z))$ produces

$$\begin{aligned}\boldsymbol{P}(B_j \mid \boldsymbol{X}) &= \frac{X_i}{n}\\ &\equiv g(X_1, \cdots, X_i),\end{aligned}$$

say, and, from elementary but long-winded combinatorial arguments,

$$\begin{aligned}\boldsymbol{P}(A_i \cap [S_i > z] \mid \boldsymbol{X}) &= \binom{X_i}{n} I[X_j \geqq n - j + 1, j = 1, \cdots, i]\\ &\quad \times \frac{(X_i - 1)!}{(n-1)!\,(X_i - n + i - 1)!} \prod_{j=1}^{i-1} (X_j - n + j)\\ &\equiv g(X_1, \cdots, X_i)\psi(X_1, \cdots, X_i),\end{aligned}$$

https://doi.org/10.2307/1427505 Published online by Cambridge University Press

say. Let $\mathscr{X}=\mathbb{N}^i$ have the obvious partial order and endow it with the i-fold product, σ, of counting measure for the purpose of computing the density, f, of $(X_1, \cdots, X_i)$. We extend ϕ to $\mathscr{X}$ by defining $\phi(x_1, \cdots, x_i)=\phi(x_i)$ and then ϕ and ψ are increasing functions on $\mathscr{X}$.

A function $h:\mathscr{X}\to\mathbb{R}$ is said to be multivariate totally positive of order 2 (MTP_2) if, for $\boldsymbol{x}, \boldsymbol{y}\in\mathscr{X}$,

$$h(\max(x_1, y_1), \cdots, \max(x_k, y_k))h(\min(x_1, y_1), \cdots, \min(x_k, y_k)) \geqslant h(\boldsymbol{x})h(\boldsymbol{y})$$

(see Karlin and Rinott (1980)). Clearly g defined above is MTP_2 and, by Proposition 3.10 of Karlin and Rinott (1980) and the main result of Karlin and Macgregor (1959) the density f is also MTP_2. But Equation (1.17a) of Karlin and Rinott (1980) states that, for f and g MTP_2, and ϕ and ψ both increasing,

$$\left(\int_{\mathscr{X}} f(\boldsymbol{x})g(\boldsymbol{x})\sigma(d\boldsymbol{x})\right)\left(\int_{\mathscr{X}} f(\boldsymbol{x})g(\boldsymbol{x})\phi(\boldsymbol{x})\psi(\boldsymbol{x})\sigma(d\boldsymbol{x})\right) \geqq \left(\int_{\mathscr{X}} f(\boldsymbol{x})g(\boldsymbol{x})\phi(\boldsymbol{x})\sigma(d\boldsymbol{x})\right)\left(\int_{\mathscr{X}} f(\boldsymbol{x})g(\boldsymbol{x})\psi(\boldsymbol{x})\sigma(d\boldsymbol{x})\right)$$

which is (3.3) on using the definitions.

A key part of the argument for Theorem 3 is that the density f of the death process is MTP_2, as stated in the paper of Karlin and Rinott (1980). This is proved in Karlin and McGregor (1959), who also give a coupling argument which is there described as heuristic. In the current context the coupling proof is rigorous and it is so elegant that it seems worth repeating here.

It is easy to see that the density f is MTP_2 if for any $t>0$, natural numbers $x_1<x_2$ and $y_1<y_2$ we have

$$\text{(3.4)} \quad \boldsymbol{P}(X_t=y_1 \mid X_0=x_1)\boldsymbol{P}(X_t=y_2 \mid X_0=x_2) \geqq \boldsymbol{P}(X_t=y_2 \mid X_0=x_1)\boldsymbol{P}(X_t=y_1 \mid X_0=x_2).$$

To demonstrate (3.4), let X^1 and X^2 be two independent copies of the death process X with X^i starting in x_i, $(i=1,2)$. Consider the optional time (relative to the history generated by both processes) τ which is the first time that X^1 and X^2 meet, or ∞ if they do not meet. Let Y^i, $(i=1,2)$ be the same as X^i up to the time τ, but then swap the paths of X^1 and X^2 to obtain Y^1 and Y^2 after τ. In view of the facts that $x_1<x_2$, $y_1<y_2$, $Y_0^1=x_1$ and $Y_0^2=x_2$, to have $Y_t^1=y_2$ and $Y_t^2=y_1$ the paths of Y^1 and Y^2, and therefore also those of X^1 and X^2, must have crossed before t. From this and the definitions,

$$\text{(3.5)} \quad \begin{aligned} [X_t^1=y_1\cap X_t^2=y_2] &\supseteq [X_t^1=y_1\cap X_t^2=y_2\cap \tau\leqq t] \\ &=[Y_t^1=y_2\cap Y_t^2=y_1=y_1\cap\tau\leqq t] \\ &=[Y_t^1=y_2\cap Y_t^2=y_1]. \end{aligned}$$

By the strong Markov property, Y^i $(i=1,2)$ are also independent copies of X^i, so that the event on the right of (3.5) has the same probability as $[X_t^1=y_2\cap X_t^2=y_1]$, and (3.4) follows

Acknowledgement

The support is gratefully acknowledged of the Australia Research Council for a visit of the second author to WA. He was also supported in part by the UK SERC Advanced Fellowship B/AF/1255. The first author completed part of the work during a visit to Universität Zürich supported by Schweiz Nationalfonds Grant 21–25579.88.

References

Ball, F. and Donnelly, P. (1987) Interparticle correlation in death processes with application to variability in compartmental models. *Adv. Appl. Prob.* **18,** 755–766.

Brémaud, P. (1981) *Point Processes and Queues: Martingale Dynamics.* Springer-Verlag, Berlin.

DONNELLY, P. (1993) The correlation structure of epidemic models. To appear.

FADDY, M. J. (1985) Non-linear stochastic compartmental models. *IMA J. Math. Appl. Med. Biol.* **2,** 287–297.

KARLIN, S. AND MCGREGOR, J. L. (1959) Coincidence probabilities. *Pacific J. Math.* **9,** 1141–1164.

KARLIN, S. AND RINOTT, Y. (1980) Classes of orderings of measures and related correlation inequalities. 1. Multivariate totally positive distributions. *J. Multivariate Anal.* **10,** 467–498.

LEFÈVRE, C. AND MICHALETZKY, G. (1990) Interparticle dependence in a linear death process subjected to a random environment. *J. Appl. Prob.* **27,** 491–498.

LEFÈVRE, C. AND MILHAUD, X. (1990) On the association of the lifelengths of components subjected to a stochastic environment. *Adv. Appl. Prob.* **22**, 961–964.

MELAMED, B. AND WALRAND, J. (1986) On the one-dimensional distribution of counting processes with stochastic intensities. *Stochastics* **19,** 1–9.