

## COMPETING RISKS AND INDEPENDENT MINIMA: A MARKED POINT PROCESS APPROACH

ELJA ARJAS,\* *University of Oulu*

PRISCILLA GREENWOOD,\*\* *University of British Columbia*

### Abstract

Given an indexed family of rate functions, we construct a family of independent random variables such that their minimum and the corresponding index form a *marked point* having the same rates. If the rates have common discontinuities, the random variables are conditionally independent, given a random allocation of these jumps. Between any two adjacent points of a marked point process there is a similar structure. Coherent functions in system lifetime theory provide examples.

Non-identifiability of multivariate lifetime distributions from mortality data is interpreted in this context.

COMPETING RISKS; INDEPENDENT MINIMA; MARKED POINT PROCESSES; MARTINGALE METHODS; NON-IDENTIFIABILITY

### 1. Introduction and summary

In many areas of practical interest one wants to study events associated with a random discrete set of points. Experimental results are then of the form (*k*th time when something happens), (identity of *k*th event))<sub>*k*≧1</sub>, a sequence of times and 'marks'. Examples arise in system reliability studies, and in biomedical survival studies. In this paper we look at problems from these areas in a point-process setting. Sequences of the above kind are then understood to be sample paths of a marked point process. The result is, we hope, a clarification of some essential ideas about competing risks and independent minima (cf. Evans (1980)).

The conversion of risk models involving dependent random variables into models involving only independent ones has been accomplished under various hypotheses by Esary and Marshall (1974), Tsiatis (1975), Miller (1977), Langberg, Proschan and Quinzi (1977), (1978). A general theorem of Langberg et al. (1978), stated in Section 2, gives necessary and sufficient conditions for the existence of a family of independent random variables such that their

---

Received 22 December 1980.

\* Department of Applied Mathematics and Statistics, University of Oulu, SF-90570 Oulu 57, Finland.

\*\* Department of Mathematics, University of British Columbia, Vancouver, B.C. Canada.

minimum and the corresponding index have the same joint distribution as the lifetime and the failure pattern of a given system. The theorem specifies distributions which could be used in simulation. However, the condition imposed, that failure pattern distributions have no common jumps, is a severe restriction in applications which may be entirely discrete.

In Section 3 we prove a similar theorem, thinking of the time of occurrence of an event of interest (e.g. system failure) together with the exact pattern of its happening as a *marked point*, i.e., a marked point process with only one point. Starting with given rate functions we construct independent random variables such that their minimum and its index form a marked point with the given rates (Theorem 3.1). If the rates have common jumps, the construction is made using conditionally independent random variables, and simulation is still possible. Because the rates of the marked point are equivalent to the probability ratios which appear in Langberg et al. (1977), (1978) and other treatments of competing risks, our construction is an extension of theirs. In Corollaries 3.1 and 3.2 we formulate the result starting with probability ratios as in Langberg et al. (1977), (1978) and starting with a marked point.

In Section 4 we restate the results of Section 3 as they apply between successive points of a marked point process. We focus on the following example. A multicomponent system with its component failure times  $(T_1, \dots, T_n)$  was treated by Arjas (1981) as a marked point process, and the definition of hazard function was extended to this context as the compensator of a counting process with respect to the  $\sigma$ -fields generated by the system's past. Here we write such a hazard function, on certain stochastic intervals, in terms of competing risks.

The non-identifiability of a multivariate lifetime distribution from mortality data, as established e.g. by Tsiatis (1975), is given an interpretation in terms of an associated marked point process in Section 5. We comment in a similar vein on the article of Prentice et al. (1978) on competing risks in a medical setting.

## 2. Preliminaries

We introduce notation and structures needed in the sequel. Additional definitions and information about point processes, marked point processes, and other applications are given by Brémaud and Jacod (1977) and Brémaud (1981).

A marked point process will be denoted by  $(p_i, \xi(p_i))$ , where  $p_i$  is the  $i$ th time point,  $\xi(p_i) \in \mathcal{S}$  is the mark at that point and  $\mathcal{S}$  is the (countable) set of all marks. A marked point process consisting of only one point will be called simply a *marked point*, and will be denoted by  $(p, \xi_p)$ .

Each point process can be identified with a counting process. In the case of a

marked point  $(p, \xi_p)$ , let

$$N(t) = 1_{\{p \leq t\}}, \quad N_I(t) = 1_{\{p \leq t, \xi_p = I\}}, \quad t \geq 0, \quad I \in \mathcal{I}.$$

Then

$$(2.1) \quad N(t) = \sum_I N_I(t).$$

If  $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$  is an increasing, right-continuous, completed family of  $\sigma$ -fields, let  $A^{\mathcal{F}}(t)$  denote the  $\mathcal{F}$ -compensator of  $N(t)$  and  $A_I^{\mathcal{F}}(t)$  the  $\mathcal{F}$ -compensator of  $N_I(t)$ . If  $N(t)$  (or  $N_I(t)$ ) is not  $\mathcal{F}_t$ -measurable, replace  $N(t)$ , here, by a cadlag version of  $E[N(t) | \mathcal{F}_t]$  and let  $E[N(t) | \mathcal{F}_t] = m_t + A^{\mathcal{F}}(t)$  be its Doob–Meyer decomposition. Then

$$(2.2) \quad A^{\mathcal{F}}(t) = \sum_I A_I^{\mathcal{F}}(t).$$

Each  $\mathcal{F}$ -compensator can be understood as a risk or hazard function, cf. Arjas (1981). Equation (2.2) tells us, in these terms, that the risk associated with  $p$  is the sum over  $I \in \mathcal{I}$  of the risks that  $p$  happens due to ‘cause’  $\mathcal{I}$ . The precise meaning of (2.2) depends, of course, on the history  $\mathcal{F}$  which is under consideration.

Consider in particular the histories

$$\mathcal{G}_t = \sigma(p; p \leq t) = \sigma(N(s); s \leq t)$$

and

$$\mathcal{H}_t = \sigma((p, \xi_p); p \leq t) = \sigma(N_I(s); s \leq t, I \in \mathcal{I}).$$

Notice that although  $N_I(t)$  is not  $\mathcal{G}_t$ -measurable we still have

$$N(t) = E(N(t) | \mathcal{G}_t) = \sum_I E(N_I(t) | \mathcal{G}_t).$$

Writing  $F_I(t) = P(p \leq t, \xi_p = I)$  and  $F(t) = \sum_I F_I(t) = P(p \leq t)$ , the following explicit integral representations hold almost surely for the compensator processes  $A^{\mathcal{G}}(t)$ ,  $A_I^{\mathcal{G}}(t)$  and  $A_I^{\mathcal{H}}(t)$ .

*Proposition 2.1.*

$$(2.3) \quad A^{\mathcal{G}}(t) = \int_0^{t \wedge p} \frac{dF(s)}{1 - F(s-)}$$

$$(2.4) \quad A_I^{\mathcal{G}}(t) = A_I^{\mathcal{H}}(t) = \int_0^{t \wedge p} \frac{dF_I(s)}{1 - F(s-)}.$$

*Proof.* (2.3) is well known and is a special case of, for example, Theorem 18.2 in Liptser and Shirayev (1978). Its proof consists of verifying that the

right-hand-side expression (as a deterministic function stopped at the  $\mathcal{G}$ -stopping time  $p$ ) is  $\mathcal{G}$ -predictable and that it satisfies the martingale condition: for  $s < t$

$$E\left(N(t) - \int_0^{t \wedge p} \frac{dF(u)}{1 - F(u-)} \mid \mathcal{G}_s\right) = N(s) - \int_0^{s \wedge p} \frac{dF(u)}{1 - F(u-)} \quad \text{a.s.}$$

This is checked through a straightforward integration. In (2.4), the integral representation of  $A_I^{\mathcal{H}}(t)$  is a special case of Jacod's (1975) Proposition 3.1, where we choose  $dx = \{I\}$  ( $\mathcal{I}$  is countable) and  $n = 1$ . The right-hand side integral is even  $\mathcal{G}$ -predictable and it satisfies the following  $(\mathcal{H}_t)$ -martingale condition. For  $s < t$ :

$$E\left(N_I(t) - \int_0^{t \wedge p} \frac{dF_I(u)}{1 - F(u-)} \mid \mathcal{H}_s\right) = N_I(s) - \int_0^{s \wedge p} \frac{dF_I(s)}{1 - F(s-)} \quad \text{a.s.}$$

That it satisfies also the corresponding  $(\mathcal{G}_t)$ -martingale condition follows by taking conditional expectations on both sides with respect to  $\mathcal{G}_s$ . (By the convention made above, since  $(N_I(t))_{t \geq 0}$  is not adapted to  $(\mathcal{G})_{t \geq 0}$ , we replace it by  $(E(N_I(t) \mid \mathcal{G}_t))_{t \geq 0}$ .)

*Remark.* The intuitive reason why the compensators  $A_I^{\mathcal{G}}(t)$  and  $A_I^{\mathcal{H}}(t)$  coincide is that on  $\{p > s\}$  there is no difference in the martingale conditioning, i.e.,  $\mathcal{G}_s \cap \{p > s\} = \mathcal{H}_s \cap \{p > s\}$ , and on  $\{p \leq s\}$  the function

$$N_I(t) - \int_0^{t \wedge p} \frac{dF_I(s)}{1 - F(s-)}$$

remains constant for all  $t \geq s$ .

The following lemma establishes the relationship between probability measures and their rate functions.

*Lemma 2.1* (Jacod (1975), p. 243). Let  $\mathcal{A}$  be the class of right-continuous non-decreasing functions  $R$  such that  $R(0) = 0$ ,  $\Delta R(t) = R(t) - R(t-) \leq 1$  for all  $t$  and  $R(s) = R(t)$  for all  $s \geq t$  whenever  $\Delta R(t) = 1$ . Then the formulas

$$(2.5) \quad F(t) = F(0, t] = 1 - e^{-R^c(t)} \prod_{s \leq t} (1 - \Delta R(s)),$$

where  $R^c$  is the continuous part of  $R$ , and

$$(2.6) \quad R(t) = \int_0^t \frac{F(ds)}{1 - F(s-)},$$

define a bijective correspondence between functions  $R \in \mathcal{A}$  and distribution functions (or probability measures)  $F$  on  $(0, \infty]$ .

*Remark.* If  $F$  is a defective distribution function, i.e.,  $\lim_{t \rightarrow \infty} F(t) < 1$ , then we make the convention  $F(\{+\infty\}) = 1 - \lim_{t \rightarrow \infty} F(t)$ .

Finally, let us see how marked point processes arise in reliability applications. Let  $\mathbf{T} = (T_1, \dots, T_n)$  denote the random lifetimes of a set of  $n$  components, labelled  $1, \dots, n$ , i.e.,  $\mathbf{T}$  is a random vector with values in the first orthant of  $\mathbb{R}^n$ . Let  $T_{(i)}$  denote the  $i$ th of these lifetimes taken in increasing order with coincident values counted once, and let  $X_{(i)}$  denote the set of indices of those components which fail at precisely time  $T_{(i)}$ . As an alternative,  $X_{(i)}$  can be defined to be the set of indices of those components which fail up to (and including) time  $T_{(i)}$ . In either case, the  $(X_{(i)})$  take values in  $\mathcal{J}$ , the collection of subsets of  $\{1, \dots, n\}$ . If  $q$  is the number of distinct values among  $T_1, \dots, T_n$ , let  $T_{(i)} = \infty$  and  $X_{(i)} = \phi$  for  $i \geq q + 1$ . The pair  $(T_{(i)}, X_{(i)})$  denotes the *marked point process corresponding to  $\mathbf{T}$* .

We say that  $\tau$  is a *coherent functional* of  $\mathbf{T}$  or of  $(T_{(i)}, X_{(i)})$ , if there is a collection of subsets  $A_1, \dots, A_m$  of  $\{1, \dots, n\}$  such that  $\tau = \min_k \max_{i \in A_k} T_i$ . With  $\xi_\tau = X_{(i)}$  on  $\{\tau = T_{(i)}\}$ ,  $(\tau, \xi_\tau)$  is a *marked point*. We call  $\xi_\tau$  the failure pattern.

The main result of Langberg et al. (1978) can be stated, including the content of their Remark 4.2, as follows.

*Theorem 2.1.* Let  $\tau$  be a coherent functional of a lifelength vector  $\mathbf{T} = (T_1, \dots, T_n)$ . Let  $F$  be the distribution of  $\tau$  and let  $F_I$  be the (sub)distribution of  $\tau$  restricted to  $\{\xi_\tau = I\}$ ,  $I \in \mathcal{J}$ . Let  $\alpha(F) = \sup\{x : F(x, \infty) > 0\}$ . Then a necessary and sufficient condition for the existence of a set of independent random variables  $\{H_I, I \in \mathcal{J}\}$  such that  $H = \min_I H_I$  and  $\xi_H$ , the index of the minimum, satisfy

$$(2.7) \quad (H, \xi_H) \stackrel{d}{=} (\tau, \xi_\tau),$$

is that the sets of discontinuities of  $F_I$  be pairwise disjoint on the interval  $[0, \alpha(F)]$ .

The distribution  $G_I$  of  $H_I$  is given by

$$(2.8) \quad G_I(t) = 1 - \exp\left(-\int_0^t \frac{dF_I^c(s)}{1 - F(s-)}\right) \times \prod_{s \leq t} \left(1 - \frac{\Delta F_I(s)}{1 - F(s-)}\right), \quad t \in [0, \alpha(F)],$$

where  $F_I^c$  is the continuous part of  $F_I$  and  $\Delta F_I(s) = F_I(s) - F_I(s-)$  is its jump at point  $s$ .

*Remark.* Apparently, in Formula (4.1) of Langberg et al. (1978),  $F_I$  should be  $F$  in the denominator.

**3. The competing risks theorem in terms of a marked point**

We now come to the main result which, together with its two corollaries, rephrases and extends Theorem 2.1. The method of proof differs from earlier approaches in that compensators, and therefore martingales, are involved.

*Theorem 3.1.* Let  $\{R_I, I \in \mathcal{I}\}$  be a countable family of functions in the class  $\mathcal{A}$  of Lemma 2.1 such that  $\sum R_I \in \mathcal{A}$ . Then there exists a family of random distribution functions  $\hat{G}_I$  and random variables  $S_I (I \in \mathcal{I})$  such that, given  $\mathcal{R} = \sigma(\hat{G}_I; I \in \mathcal{I})$ , the  $S_I$  are independent, have distributions  $\hat{G}_I$ , the  $\hat{G}_I$  have almost surely no pairwise common discontinuities and:

- (i) Each  $S_I$  has the given rate function  $R_I$ , i.e., if  $\mathcal{F}_t^I = \sigma(S_I; S_I \leq t)$ , the  $(\mathcal{F}_t^I)$ -compensator of  $S_I$  is  $R_I(t \wedge S_I)$ ;
- (ii) Let  $S = \min_I S_I$  and  $\xi_S = I$  on  $\{S = S_I\}$  and denote  $\mathcal{F}_t = \sigma(S_I; S_I \leq t, I \in \mathcal{I}) = \bigvee_{I \in \mathcal{I}} \mathcal{F}_t^I$ . Then the (multivariate)  $(\mathcal{F}_t)$ -compensator of the marked point  $(S, \xi_S)$  is  $(R_I(t \wedge S); t \geq 0, I \in \mathcal{I})$ ;
- (iii) The functions  $K_I(t) = P(S \leq t, \xi_S = I)$ ,  $K = \sum K_I$ , satisfy

$$R_I(t) = \int_0^t \frac{dK_I(s)}{1 - K(s-)}.$$

*Proof.* Let  $G_I$  be the probability distribution on  $(0, \infty]$  which arises from  $R_I$  through Lemma 2.1., i.e.,

$$(3.1) \quad G_I(t) = G_I(0, t] = 1 - e^{-R_I^*(t)} \prod_{s \leq t} (1 - \Delta R_I(s)), \quad t \geq 0.$$

Let then  $\hat{G}_I$  be a random distribution function, which has the same increments as  $G_I$  except at those points  $s$  where two or more of the distributions have a discontinuity; at such points  $s$  the total amount of jumps  $\sum_I \Delta G_I(s)$ , is randomly allocated to one of the distributions  $\hat{G}_I$  by using the weights

$$\frac{\Delta G_I(s)}{\sum_J \Delta G_J(s)} = \frac{\Delta R_I(s)}{\sum_J \Delta R_J(s)}$$

as probabilities. Formally, if  $D$  is the set of all discontinuity points in  $\{R_I; I \in \mathcal{I}\}$ , we let

$$(3.2) \quad \Pr \left( \Delta \hat{R}_K(s) = \sum_J \Delta R_J(s) \cdot \delta_{JK}; K \in \mathcal{I} \right) = \frac{\Delta R_I(s)}{\sum_J \Delta R_J(s)}, \quad s \in D, \quad I \in \mathcal{I},$$

taking the vectors  $(\Delta \hat{R}_I(s); I \in \mathcal{I})$  to be independent for different  $s \in D$ , and define

$$(3.3) \quad \hat{G}_I(0, t] = 1 - e^{-R_I^*(t)} \prod_{s \leq t} (1 - \Delta \hat{R}_I(s)), \quad t \geq 0, \quad I \in \mathcal{I}.$$

Notice that  $E(\Delta \hat{R}_I(s)) = \Delta R_I(s)$ . Let  $\mathcal{R} = \sigma(\Delta \hat{R}_I(s); s \in D, I \in \mathcal{I}) = \sigma(\hat{G}_I; I \in \mathcal{I})$

be the  $\sigma$ -field generated by the above randomization. Given  $\mathcal{R}$ , we let  $\{U_I; I \in \mathcal{I}\}$  be a family of independent  $[0, 1]$ -uniformly distributed random variables, and define

$$(3.4) \quad S_I = \inf \{s : \hat{G}_I(s) \geq U_I\}, \quad I \in \mathcal{I}.$$

Given  $\mathcal{R}$ ,  $S_I$  has the distribution  $\hat{G}_I$ .

(i)  $S_I$  has the distribution  $G_I$ , since for all  $t \geq 0$

$$\begin{aligned} \Pr(S_I \leq t) &= E(\Pr(S_I \leq t \mid \mathcal{R})) \\ &= E(\hat{G}_I(t)) \\ &= E\left(1 - e^{-R_I^c(t)} \prod_{s \leq t} (1 - \Delta \hat{R}_I(s))\right) \\ &= 1 - e^{-R_I^c(t)} \prod_{s \leq t} (1 - \Delta R_I(s)) \\ &= G_I(t). \end{aligned}$$

On the other hand, by applying Proposition 2.1 to  $S_I$  in place of  $p$  we see that the  $(\mathcal{F}_t^I)$ -compensator of  $S_I$  almost surely has the expression

$$\int_0^{t \wedge S_I} \frac{P(S_I \in ds)}{P(S_I \geq s)} = \int_0^{t \wedge S_I} \frac{dG_I(s)}{1 - G_I(s-)}.$$

By the converse part of Lemma 2.1. the right-hand side is the same as  $R_I(t \wedge S_I)$ .

The function  $R_I(t \wedge S_I)$  is also the  $(\mathcal{F}_t)$ -compensator of  $S_I$ . This fact, which is used in proving (ii) below, is seen as follows: The  $(\mathcal{F}_t^I)$ -martingale condition for  $R_I(t \wedge S_I)$  is

$$(3.5) \quad \text{for } s < t: E(N^I(t) - R_I(t \wedge S_I) \mid \mathcal{F}_s^I) = N^I(s) - R_I(s \wedge S_I) \quad \text{a.s.}$$

where  $N^I(t) = 1_{\{S_I \leq t\}}$ ,  $t \geq 0$ , is the counting process corresponding to  $S_I$ . Since  $R_I(t \wedge S_I)$  is obviously  $(\mathcal{F}_t)$ -predictable we need only check that (3.5) holds also when  $\mathcal{F}_s^I$  is replaced by  $\mathcal{F}_s = \bigvee_{I \in \mathcal{I}} \mathcal{F}_s^I$ . It is clearly enough to consider the case  $\{S_I > s\}$ . But from the fact that the randomizations on  $(0, s]$  and  $(s, \infty]$  were assumed independent and the conditional independence of  $S_I$  given  $\mathcal{R}$  it follows that  $\mathcal{F}_\infty^I$  (i.e.,  $S_I$ ) and  $\bigvee_{J \neq I} \mathcal{F}_s^J$  are conditionally independent given  $\{S_I > s\}$ , and consequently on that set, for any  $\mathcal{F}_\infty^I$ -measurable random variable  $Z$ ,

$$E(Z \mid \mathcal{F}_s^I) = E\left(Z \mid \mathcal{F}_s^I \vee \bigvee_{J \neq I} \mathcal{F}_s^J\right) = E(Z \mid \mathcal{F}_s) \quad \text{a.s.}$$

(ii) First observe that  $\xi_S$  is almost surely uniquely defined. Let  $N^I(t) = 1_{\{S_I \leq t\}}$ ,  $I \in \mathcal{I}$ , be the univariate counting processes as defined in (i) and let

$(N_I(t); I \in \mathcal{I}) = (1_{\{S \leq t, \xi_S = I\}}; I \in \mathcal{I})$  be the multivariate counting process corresponding to the marked point  $(S, \xi_S)$ . By inspection we find that almost surely

$$(3.6) \quad N^I(t \wedge S) = N_I(t \wedge S) = N_I(t) \quad \text{for all } t \geq 0 \quad \text{and } I \in \mathcal{I},$$

and hence their  $(\mathcal{F}_t)$ -compensators are almost surely equal. Our claim is that the  $(\mathcal{F}_t)$ -compensator of  $N_I(t)$  is  $R_I(t \wedge S)$ .

We have seen in (i) above that  $R_I(t \wedge S_I)$  is the  $(\mathcal{F}_t)$ -compensator of  $N^I(t)$  (or of  $S_I$ ). But then  $R_I(t \wedge S_I \wedge S) = R_I(t \wedge S)$  is the  $(\mathcal{F}_t)$ -compensator of the stopped process  $N^I(t \wedge S)$  (see e.g. Corollary 1.2.7 of Stroock and Varadhan (1979)), and by (3.6) it is also the  $(\mathcal{F}_t)$ -compensator of  $N_I(t)$ .

(iii) By Proposition 2.1 the  $(\mathcal{F}_t)$ -compensator of  $N_I(t) = 1_{\{S \leq t, \xi_S = I\}}$  is

$$\int_0^{t \wedge S} \frac{P(S \in ds, \xi_S = I)}{P(S \geq s)} = \int_0^{t \wedge S} \frac{dK_I(s)}{1 - K(s-)}.$$

On the other hand we have just seen that it is also  $R_I(t \wedge S)$  so that (iii) holds.

*Remark.* It follows from our construction that if  $\sum R_I(t)$ , which was assumed to be in the class  $\mathcal{A}$ , either approaches  $\infty$  with  $t$  or has a 'final' jump of size 1 at some point, then at least one of the random functions  $\hat{R}_I(t) = R_I^c(t) + \sum_{s \leq t} \Delta \hat{R}_I(s)$  must have that same property and then the corresponding distribution  $\hat{G}_I$  is proper (cf. Theorem 1 of Miller (1977)).

*Remark.* If there are no common discontinuities of the  $R_I$ 's, the functions  $G_I$  as defined in (3.1) are the distributions of  $S_I$ , independent, and it is evident that  $R_I(t \wedge S_I)$  is the  $(\mathcal{F}_t)$ -compensator of  $S_I$ .

The following corollaries make explicit how Theorem 3.1 applies if we start with given distributions or with a marked point.

*Corollary 3.1.* Given distributions  $F$  and  $F_I, I \in \mathcal{I}$ , such that  $F = \sum_{I \in \mathcal{I}} F_I$ , Theorem 3.1 holds with

$$R_I(t) = \int_0^t \frac{dF_I(s)}{1 - F(s-)}$$

and  $F_I(t) = P(S \leq t, \xi_S = I)$ .

*Proof.* From the given  $R_I$ , Theorem 3.1 gives us  $(S, \xi_S)$  and  $K_I, I \in \mathcal{I}$ , and  $K$  such that

$$(3.7) \quad \int_0^t \frac{dF_I(s)}{1 - F(s-)} = \int_0^t \frac{dK_I(s)}{1 - K(s-)}.$$

Summing on  $I$  we obtain

$$\int_0^t \frac{dF(s)}{1 - F(s-)} = \int_0^t \frac{dK(s)}{1 - K(s-)}.$$

It follows from Lemma 2.1 that  $F = K$ , and from (3.7) we have  $F_I = K_I$ .



*Corollary 3.2.* Let  $(p, x)$  be a marked point with  $x$  taking values  $I \in \mathcal{I}$ . There exist  $S_I, I \in \mathcal{I}$ , which are conditionally independent as in Theorem 3.1, and for  $S = \min_I S_I$  we have

$$(3.8) \quad (S, \xi_S) \stackrel{d}{=} (p, x).$$

*Proof.* Let  $F_I(ds) = P(p \in ds, x = I)$ , and  $F(ds) = P(p \in ds)$ . Let

$$R_I(t) = \int_0^t \frac{dF_I(s)}{1 - F(s-)}.$$

Theorem 3.1 gives us  $(S, \xi_S)$  and  $K_I$ , and Corollary 3.1 says that  $F_I = K_I$ , i.e. (3.8) holds.

*Application.* Let  $\tau$  be a coherent function of a random lifetime vector  $(T_1, \dots, T_n)$ . Then  $(\tau, \xi_\tau)$  is a marked point with rates

$$R_I(t) = \int_0^t \frac{dF_I(s)}{1 - F(s-)},$$

where  $F(t) = P(\tau \leq t)$ ,  $F_I(t) = P(\tau \leq t, \xi_\tau = I)$ . The mark set  $\mathcal{I}$  is the set of possible failure patterns. Corollaries 3.1 and 3.2 are alternate formulations of an extension of the theorem of Langberg et al. (1978) (restated here as Theorem 2.1).

#### 4. Conditioning on the past history

Now we consider a marked point process  $(p_i, x_i)_{i \geq 1}$ . In particular this could be the process  $(T_{(i)}, X_{(i)})$  corresponding to a lifetime vector  $\mathbf{T} = (T_1, \dots, T_n)$ , defined in Section 2. Suppose that we have been watching the process up to time  $t$ , learning exactly those  $(p_k, x_k)$  for which  $p_k \leq t$ , and we are interested in the future behaviour of the process, conditionally on our knowledge of the past.

We let  $\mathcal{F}_t = \sigma((p_i, x_i), p_i \leq t)$  be the history generated by the marked point process up to time  $t$  and look for a result similar to Theorem 3.1.

The  $\mathcal{F}$ -compensator of  $(p_i, x_i)_{i \geq 1}$ , which we denote simply by  $A(t, I)$ , can be written as (Jacod (1975))

$$(4.1) \quad A(t, I) = \sum_k A^{(k)}(t, I),$$

$$(4.2) \quad A^{(k)}(t, I) = \int_{p_k}^{p_{k+1} \wedge t} dR_I^{(k)}(s)$$

$$R_I^{(k)}(ds) = \frac{P(p_{k+1} \in ds, x_{k+1} = I \mid \mathcal{F}_{p_k})}{P(p_{k+1} \geq s \mid \mathcal{F}_{p_k})}.$$

Expression (4.2) looks similar to the integral representation (2.4) of the compensator  $\mathcal{A}_I^{\mathcal{K}}$  of the marked point  $(p, x)$ , and it can be proved by the same argument as Proposition 2.1.

Conditionally on  $\mathcal{F}_{p_k}$ , the functions

$$B^{(k)}(t, I) = \int_{p_k}^{p_k+t} dR_I^{(k)}(s), \quad t \geq 0, \quad I \in \mathcal{I}$$

are almost surely in class  $\mathcal{A}$  and there exists corresponding distributions  $G_I$  and  $\hat{G}_I$  as in the proof of Theorem 3.1. We suppress  $\omega$  in the notation and continue ‘conditionally on  $\mathcal{F}_{p_k}$ ’. Now we can write

$$(4.3) \quad \int_0^{t \wedge (p_{k+1} - p_k)} dR_I^{(k)}(p_k + s) \stackrel{d}{=} \int_0^{t \wedge \Psi} \frac{dG_I(s)}{1 - G_I(s-)}$$

where  $\Psi$  is the minimum of conditionally independent random variables with distributions  $G_I$ . This corresponds to the identity in distribution of the marked points, conditionally on  $\mathcal{F}_{p_k}$ ,

$$(4.4) \quad (p_{k+1} - p_k, x_{k+1}) \stackrel{d}{=} (\Psi, \xi_\Psi).$$

We can interpret this identity as saying that the result of Langberg et al., and others, holds conditionally between successive points of a marked point process, i.e., there exist, conditionally on  $\mathcal{F}_{p_k}$  and a possible randomization, independent random variables such that their minimum and its index are the same in ‘law and patterns’ as the interpoint interval and the next mark of the process. Since we have such a statement for each  $k$ , (4.4) says that the entire process can be viewed in terms of conditionally independent competing risks.

### 5. Non-identifiability

An additional question about competing risks was studied by Tsiatis (1975). Suppose that  $F$  is the distribution of a failure time  $\tau$ , where failure is due to, or associated with, one of a set of  $n$  risks numbered by  $i = 1, \dots, n$ . For example,  $\tau$  is the time of death owing to one of a set of  $n$  diseases. In the present notation  $F_i(dt) = P(\tau \in dt, \xi_\tau = i)$ ,  $F = \sum_i F_i$  are regarded as given. Tsiatis proved a version of Theorem 2.1 in this setting. He asked additionally: suppose that  $Y_1, \dots, Y_n$  are random variables such that  $\Psi = \min_i Y_i$  satisfies (in our notation)

$$(5.1) \quad (\Psi, \xi_\Psi) \stackrel{d}{=} (\tau, \xi_\tau).$$

Then what can be said about the joint distribution of  $(Y_1, \dots, Y_n)$ ? His answer is that the class of possible distributions is large. In terms of a marked point process the class can be described as follows. Let  $\{F_i\}$  and  $F$  be given and

consider an arbitrary random vector  $(Y_1, \dots, Y_n)$ ,  $Y_i \geq 0$ . The joint distribution of  $(Y_1, \dots, Y_n)$  determines, and is determined by, a marked point process  $(T_{(i)}, X_{(i)})$ ,  $i = 1, \dots, n$ . Condition (5.1) specifies that  $(\tau, \xi_\tau) \stackrel{d}{=} (T_{(1)}, X_{(1)})$ . This condition does not determine the joint distribution of  $(Y_1, \dots, Y_n)$ , but specifies the class of  $n$ -variate distributions yielding the given  $(T_{(1)}, X_{(1)})$ . That this class contains at most one product measure is a version of Theorem 2.1. The construction which gives the product measure is

$$(5.2) \quad P(Y_i \in dt \mid Y_i \geq t) = P(\tau \in dt, \xi_\tau = i \mid \tau \geq t).$$

To resolve the 'non-identifiability' of  $(Y_1, \dots, Y_n)$  one would need to observe (and therefore have given in distribution) the entire associated process  $(T_{(i)}, X_{(i)})$ . However, when  $T_{(1)}$  corresponds to 'death', as Tsiatis suggests, a practical solution must involve the observation of medical records before or at the time of death.

The problem of finding a practical meaning of the joint distribution of  $(Y_1, \dots, Y_n)$  or even of its marginal distributions in the medical context is discussed by Prentice et al. (1978) and Prentice and Kalbfleisch (1979), in connection with competing risks data arising in the treatment of leukemia. They prove that the (marginal) distributions of the  $Y_i$  are non-identifiable using the fact that an identifiable random variable must be uniquely defined in terms of the right-hand side of (5.2). They define the density function  $f_i(t, Z^*)$  of time to failure and cause of failure, by writing (we combine our notation and theirs)

$$(5.3) \quad P(\tau \in dt, \xi_\tau = i \mid \tau \geq t, Z(t)) = \lambda_i(t, Z) dt$$

where  $Z(t)$  denotes the value of a covariate or regression vector at time  $t$ ,  $Z^* = Z^*(t) = \{Z(u); u \leq t\}$  and

$$(5.4) \quad \begin{aligned} \bar{F}(t, Z^*) &= \exp \left( - \sum_i \int_0^t \lambda_i(u, Z) du \right), \\ f_i(t, Z^*) &= \lambda_i(t, Z) \bar{F}(t, Z^*), \end{aligned}$$

and they study a corresponding likelihood function. They note the implicit assumption that the  $Z(t)$  are deterministic or generated by an 'external' stochastic mechanism. One might replace  $Z(t)$  in (5.3) by  $\mathcal{F}_{t-}$ , generated by a random past, and view  $\lambda_i(t, Z) dt$  as the measure compensating for the marked point  $(\tau, \xi_\tau)$ . However, in the case where  $Z$  is not deterministic (or  $\mathcal{F}_0$ -measurable),  $\lambda_i(t, Z)$  is not directly convertible into a probability distribution of  $(\tau, \xi_\tau)$  in the way (5.4) indicates.

## References

- AALLEN, O. (1978) Nonparametric inference for a family of counting processes. *Ann. Statist.* **6**, 701-726.

- ARJAS, E. (1981) The failure and hazard processes in multivariate reliability systems. *Math. Operat. Res.* To appear.
- BRÉMAUD, P. (1981) *Point Processes and Queues: Martingale Dynamics*. Springer-Verlag, Berlin.
- BRÉMAUD, P. AND JACOD, J. (1977) Processus ponctuels et martingales: résultats récents sur la modélisation et le filtrage. *Adv. Appl. Prob.* **9**, 362–416.
- ESARY, J. D. AND MARSHALL, A. (1974) Multivariate distributions with exponential minimums. *Ann. Statist.* **2**, 84–98.
- EVANS, R. A. (1980) Independent or mutually exclusive. *IEE Trans. Reliability* **29**, 289 (editorial).
- GILL, R. D. (1980) *Censoring and Stochastic Integrals*. Mathematical Centre Tracts **124**, Amsterdam.
- JACOD, J. (1975) Multivariate point processes: Predictable projection, Radon–Nikodym derivatives, representation of martingales. *Z. Wahrscheinlichkeitsth.* **31**, 235–253.
- LANGBERG, N., PROSCHAN, F. AND QUINZI, A. J. (1977) Converting dependent models into independent ones, with applications in reliability. In *The Theory and Applications of Reliability, with Emphasis on Bayesian and Nonparametric Methods*, **1**, ed I. Shimi and C. P. Tsokos, Academic Press, New York, 259–276.
- LANGBERG, N., PROSCHAN, F. AND QUINZI, A. J. (1978) Converting dependent models into independent ones, preserving essential features. *Ann. Prob.* **6**, 174–181.
- LIPTSER, R. N. AND SHIRYAYEV, A. N. (1978) *Statistics of Random Processes, II. Applications*. Springer-Verlag, Berlin.
- MEYER, P. A. (1972) *Martingales and Stochastic Integrals*. Lecture Notes in Mathematics **284**, Springer-Verlag, Berlin.
- MILLER, D. R. (1977) A note on independence of multivariate lifetimes in competing risks models. *Ann. Statist.* **5**, 576–579.
- PRENTICE, R. L., KALBFLEISCH, J. D., PETERSON, A. V., JR., FLOURNOY, N., AND FAREWELL, V. T. (1978) The analysis of failure times in presence of competing risks. *Biometrics* **34**, 541–554.
- PRENTICE, R. L. AND KALBFLEISCH, J. D. (1979) Hazard rate models with covariates. *Biometrics* **35**, 25–39.
- STROOCK, D. W. AND VARADHAN, S. R. S. (1979) *Multidimensional Diffusion Processes*. Springer-Verlag, New York.
- TSIATIS, A. (1975) A nonidentifiability aspect in the problem of competing risks. *Proc. Nat. Acad. Sci. USA* **72**, 20–22.