COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE

18,3 (1977)

ON LOGARITHMIC INFORMATION IN POINT PROCESSES

Petr MANDL, Praha

Abstract: Pairs P^o, P^1 of probability distributions of point processes are considered. The respective logarithmic information is expressed in terms of the intensity (hazard function) ratio. Whence a sufficient condition for absolute continuity of P^1 with respect to P^o is obtained. The proofs given require much simpler mathematical apparatus than the derivation of similar results using general theory of point processes.

Key words: Point processes, hazard function, information, absolute continuity.

AMS: 60K99 Ref. Ž.: 9.657

---

1. In this note a point process is a random sequence of points $\{\tau_n, n = 1,2,\dots\}$ on the time axis $(0,\infty)$. The points can be interpreted as times of occurrence of an event, e.g., the failure of a machine. The intensity (or hazard function) at time t is a number Q_t such that the probability of an event in the interval $(t,t+dt)$ conditioned by all the past equals $Q_t dt$. We adopt the general approach to intensities as presented in [2], Chapters 18 and 19, but without requiring from the reader the knowledge of that book. For two point processes we express the logarithmic information with aid of the ratio of intensities. Then we obtain sufficient conditions for the absolute continuity of their probability distributions. For the properties of information we refer to [1].

2. The probability distribution of a point process is a measure on the space $(S, \mathcal{S}_\infty)$. S is the set of all non-decreasing sequences $s = \{s_1, s_2, \dots\}$ of positive numbers inclusively ∞ with the following properties. i) $\lim_{n\to\infty} s_n = \infty$. ii) $s_n < s_{n+1}$, whenever $s_n < \infty$. We introduce

$$\tau_n(s) = s_n, \; n = 1,2,\dots, \quad \tau_0(s) = 0, \; s \in S,$$

$$N_t(s) = \sum_{n=1}^{\infty} \chi_{\{s_n \leq t\}}, \; t \in [0,\infty), \; s \in S.$$

The counting process N provides a complete description of the random point s. Thus, $N. \sim s$.

Next we define an increasing system of σ-algebras

(1) $$\mathcal{S}_t = \sigma a\{N_{u\wedge t}, u \in [0,\infty)\}, \; t \in [0,\infty].$$

$u \wedge t$ is $\min(u,t)$. Definition (1) can be generalized. For σ a stopping time with respect to $\mathcal{S}$ we denote

$$\mathcal{S}_\sigma = \sigma a\{N_{u\wedge\sigma}, u \in [0,\infty)\}.$$

Let us recall the Galmarino Lemma on stopping times.

Lemma 1 (A.R. Galmarino). A non-negative random variable σ on $(S, \mathcal{S}_\infty)$ is a stopping time if and only if for $t \in [0,\infty)$, $s, s' \in S$,

$$N_u(s) = N_u(s'), \; u \leq t, \; \sigma(s) \leq t \Rightarrow \sigma(s) = \sigma(s').$$

Two consequences of Lemma 1 will be used in the sequel.

Lemma 2. Let σ be a stopping time, $\sigma(s) \sim \sigma(N.)$. Then

(2) $$\sigma(N.) = \sigma(N._{\wedge\sigma}), \quad s \in S.$$

Proof. Take $s \in S$. Let s' be such that $N._{\wedge\sigma}(s) = N.(s')$. Then $N_u(s) = N_u(s')$, $u \leqq \sigma(s)$. Hence, $\sigma(s) = \sigma(s')$, which is the same as (2). □

Lemma 3 ([2]). Let σ be a stopping time. Then

$$(3) \qquad \sigma \wedge \tau_n = \sigma(N._{\wedge\tau_{n-1}}) \wedge \tau_n, \quad n = 1,2,\ldots$$

Proof. If $\sigma(N.) < \tau_n$, then from Lemma 1 follows $\sigma(N._{\wedge\tau_{n-1}}) = \sigma(N.)$. Consequently, (3) holds. If $\sigma(N.) \geqq \tau_n$, then $\sigma(N._{\wedge\tau_{n-1}}) < \tau_n$ is impossible, because this would imply $\sigma(N.) = \sigma(N._{\wedge\tau_{n-1}})$. Again, (3) holds. □

3. Let two probability measures P^0, P^1 be defined on $(S, \mathcal{S}_\infty)$ by means of conditional distribution functions

$$(4) \qquad F_1^i(t), F_2^i(t/t_1), \ldots, F_n^i(t/t_1, \ldots, t_{n-1}), \ldots, \quad i = 0,1.$$

That is,

$$F_n^i(t/\tau_1, \ldots, \tau_{n-1}) = P^i(\tau_n \leqq t/\tau_1, \ldots, \tau_{n-1}), \quad t \in [0,\infty],$$
$$n = 1,2,\ldots, i = 0,1.$$

We assume that functions (4) are continuous on $[0,\infty)$. We define on S cumulative intensities (or compensators)

$$(5) \quad A_t^i = A_{\tau_{n-1}}^i + \int_{\tau_{n-1}}^t \frac{dF_n^i(u/\tau_1, \ldots, \tau_{n-1})}{1 - F_n^i(u/\tau_1, \ldots, \tau_{n-1})}, \quad \tau_{n-1} \leqq t < \tau_n,$$
$$n = 1,2,\ldots, \quad A_0^i \equiv 0.$$

The integrand on the right-hand side is a generalization of the hazard function known from renewal theory.

Further we assume that

(6) $$\frac{dF_n^1(t/t_1,\dots,t_{n-1})}{1-F_n^1\ (t/t_1,\dots,t_{n-1})} = \ell_n(t/t_1,\dots,t_{n-1})\cdot$$

$$\cdot\frac{dF_n^o(t/t_1,\dots,t_{n-1})}{1-F_n^o(t/t_1,\dots,t_{n-1})}\ ,\ t\in[0,\infty)\ .$$

$0 \leqq \ell_n < \infty$ is the Radon-Nikodym density of the measures on $[0,\infty)$ specified by the differentials with the convention $0\cdot\infty = 0/0 = 0$. Thus, it is possible to define the intensity ratio

(7) $$L_t = \ell_n(t/\tau_1,\dots,\tau_{n-1}),\quad \tau_{n-1}\leqq t<\tau_n,\ n = 1,2,\dots$$

We have

$$A_t^1 = \int_0^t L_u dA_u^o,\quad t\in[0,\infty).$$

The mathematical expectation under P^i will be denoted by E^i.

Finally we introduce the information measure. Let σ be a stopping time. We denote by $I_\sigma(P^1,P^o)$ the information in P^1 with respect to P^o on $\mathcal{F}_\sigma$. I.e., if $P^1 \prec P^o$ on $\mathcal{F}_\sigma$, and Z_σ is the corresponding Radon-Nikodym density, then

$$I_\sigma(P^1,P^o) = E^o Z_\sigma \log Z_\sigma = E^1 \log Z_\sigma\ ,\ Z_\sigma = \frac{dP^1}{dP^o}\Big|_{\mathcal{F}_\sigma}\ .$$

If $P^1 \prec P^o$ does not hold, then $I_\sigma(P^1,P^o) = \infty$.

4. Next we give an auxiliary result. It concerns point processes with at most one event. Let two probability distributions on $[0,\infty]$ have distribution functions F^o,F^1, respectively. Let F^o,F^1 be continuous on $[0,\infty)$, $F^o(0) = 0 = F^1(0)$. Define measures on $[0,\infty)$ by the relation

(8) $$da^i(t) = \frac{dF^i(t)}{1-F^i(t)}, \quad t \in [0,\infty), \quad i = 0,1.$$

Further let

(9) $$da^1(t) = \ell(t)da^o(t), \quad t \in [0,\infty),$$

where $0 \leqq \ell(t) < \infty$, $t \in [0,\infty)$.

For the information we get the following formula.

Lemma 4.

(10) $$I(F^1,F^o) = \int_0^{\infty}\int_0^{t-} (1+\ell(u)\log \ell(u)-\ell(u))da^o(u)dF^1(t).$$

Proof. The integral in (10) exists, since

$1 + x \log x - x \geqq 0$, $x \in [0,\infty)$. Consider first the case $F^1(\infty-) < 1$, $F^o(\infty-) = 1$. Then obviously $I(F^1,F^o) = \infty$. Moreover,

$$\int_0^{\infty-}\ell(u)da^o(u) = \int_0^{\infty-}da^1(u) < \infty \quad , \quad \int_0^{\infty-}da^o(u) = \infty .$$

The right-hand side of (10) is not less than

$$(1 - F^1(\infty-))(\int_0^{\infty-}(1-e^{-1})da^o(u) - \int_0^{\infty-}da^1(u)) = \infty .$$

Hence, (10) holds.

For the rest of the proof we may assume

(11) $$F^1(\infty-) < 1 \implies F^o(\infty-) < 1.$$

Set $\ell(\infty) = 1$. From (8),(9),(11) follows that the density of F^1 with respect to F^o is

(12) $$\frac{dF^1}{dF^o}(t) = \frac{1 - F^1(t-)}{1 - F^o(t-)}\ell(t), \quad t \in [0,\infty]$$

Consequently,

(13) $I(F^1,F^o) = \int_0^\infty \log\left(\frac{1 - F^1(t-)}{1 - F^o(t-)}\, \ell(t)\right) dF^1(t).$

The subsequent transformations lead from the right-hand side of (10) to that of (13) and vice versa. Their feasibility will be discussed afterwards.

$$\int_0^\infty \int_0^{t-} (1+\ell(u)\log \ell(u) - \ell(u))da^o(u)dF^1(t) =$$

$$= \int_0^\infty \left[\int_0^{t-} \frac{dF^o(u)}{1-F^o(u)} - \int_0^{t-} \frac{dF^1(u)}{1-F^1(u)} \right] dF^1(t) +$$

(14)

$$+ \int_0^{\infty-} \log \ell(u)\, \frac{1-F^1(u)}{1-F^o(u)}\, \ell(u)dF^o(u) =$$

$$= \int_0^\infty (-\log(1-F^o(t-)) + \log(1-F^1(t-)) + \log \ell(t))dF^1(t) =$$

$$= \int_0^\infty \log\left(\frac{1-F^1(t-)}{1-F^o(t-)}\, \ell(t)\right)dF^1(t).$$

We have used (12) and Fubini's Theorem.

If $F^1(\infty-) < 1$, then from the finiteness of either the left or the right-hand side of (1) follows the finiteness of all integrals occurring in (14). Thus, for this case, (10) is demonstrated. If $F^1(\infty-) = 1$, denote

$$\bar{t} = \inf\{t : F^1(t) = 1\}.$$

Define for $n = 1,2,\ldots$

$${}^nF^i(t) = F^i(t \wedge (\bar{t}-n^{-1}) \wedge n), \quad t \in [0,\infty), \quad {}^nF^i(\infty) = 1,$$
$$i = 0,1.$$

Apply (10) to $I({}^nF^1, {}^nF^o)$, and let $n \to \infty$. From the continuity of information, (10) follows. □

5. Theorem 1. Let σ be a stopping time. Then

(15) $$I_{\sigma}(P^1,P^0) = E^1\int_0^{\sigma-}(1+L_t\log L_t-L_t)dA_t^0.$$

Proof. Denote

$$K[x] = 1 + x\log x - x, \quad x\in[0,\infty).$$

$I_{\sigma\wedge\tau_n}(P^1,P^0)$ expressed with aid of $I_{\sigma\wedge\tau_{n-1}}(P^1,P^0)$ and of the conditional information contained in the event at time $\tau_n \leqq \sigma$ equals

(16) $$\begin{aligned} I_{\sigma\wedge\tau_n}(P^1,P^0) &= I_{\sigma\wedge\tau_{n-1}}(P^1,P^0) + \\ &+ E^1\chi_{\{\sigma\geqq\tau_{n-1},\tau_{n-1}<\infty\}}I_{\sigma\wedge\tau_n}(P^1(.|\mathcal{G}_{\sigma\wedge\tau_{n-1}}),P^0(.|\mathcal{G}_{\sigma\wedge\tau_{n-1}}))+ \\ &+ E^1(\chi_{\{\sigma<\tau_{n-1}\}}+\chi_{\{\sigma=\tau_{n-1}=\infty\}})I_{\sigma\wedge\tau_n}(P^1(.|\mathcal{G}_{\sigma\wedge\tau_{n-1}}), \\ &\qquad P^0(.|\mathcal{G}_{\sigma\wedge\tau_{n-1}})),\quad n = 1,2,\dots \end{aligned}$$

The last term is zero, since the conditional information vanishes.

To deal with the before last term, we note that by Lemma 3

$$\sigma\wedge\tau_n = \sigma(N_{.\wedge\tau_{n-1}})\wedge\tau_n = z_n(\tau_1,\dots,\tau_{n-1})\wedge\tau_n,$$

where $z_n(t_1,\dots,t_{n-1})$ is a Borel function of $t_1,\dots,t_{n-1}$. Thus, given $\sigma\geqq\tau_{n-1}$, $\tau_1,\dots,\tau_{n-1}<\infty$, the conditional distribution is

$$\bar{F}_n^i(t/\tau_1,\dots,\tau_{n-1}) = F_n^i(t\wedge z_n(\tau_1,\dots,\tau_{n-1})/\tau_1,\dots$$
$$\dots,\tau_{n-1}),\ t\in[0,\infty),$$
$$\bar{F}_n^i(\infty/\tau_1,\dots,\tau_{n-1}) = 1,\quad i = 0,1.$$

By Lemma 4, the conditional information equals

$$\int_0^\infty \int_0^{t\wedge\tau_n -} K[\ell_n(u/\tau_1,\dots,\tau_{n-1})]\,\frac{dF_n^o(u/\tau_1,\dots,\tau_{n-1})}{1-F_n^o(u/\tau_1,\dots,\tau_{n-1})}\,\cdot$$

$$\cdot\, dF_n^1(t/\tau_1,\dots,\tau_{n-1}) = E^1\{\int_{\tau_{n-1}}^{\sigma\wedge\tau_n -} K[L_u]dA_u^o \mid \tau_1,\dots,\tau_{n-1}\}.$$

Consequently, (16) implies

$$I_{\sigma\wedge\tau_n}(P^1,P^o) = I_{\sigma\wedge\tau_{n-1}}(P^1,P^o)+E^1\int_{\sigma\wedge\tau_{n-1}}^{\sigma\wedge\tau_n -} K[L_u]dA_u^o,\ n=1,2,\dots,$$

or

$$I_{\sigma\wedge\tau_n}(P^1,P^o) = E^1\int_0^{\sigma\wedge\tau_n -} K[L_u]dA_u^o,\ n=1,2,\dots$$

From here, (15) follows letting $n\longrightarrow\infty$, and using the continuity of information. □

Theorem 1 yields the following sufficient condition for $P^1 \prec P^o$.

<u>Theorem 2</u>. Let

(17) $$P^1\left(\int_0^{\infty -}(1+L_t\log L_t-L_t)dA_t^o<\infty\right) = 1.$$

Then $P^1\prec P^o$.

Proof. Let (17) hold. Define

$$\sigma_n = \inf\{t: \int_0^t K[L_u]dA_u^o \geq n\},\ n=1,2,\dots$$

By Theorem 1,

$$I_{\sigma_n}(P^1,P^o) = E^1\int_0^{\sigma_n -} K[L_u]dA_u^o \leq n.$$

Hence,

(18) $$P^1 \prec P^o \text{ on } \mathcal{F}_{\sigma_n},\ n=1,2,\dots$$

Let $B \in \mathcal{F}_\infty$, $P^0(B) = 0$. We have

$$P^1(B) \leq P^1(N_{\cdot \wedge \sigma_n} \in B,\ \sigma_n = \infty) + P^1(\sigma_n < \infty).$$

Further,

$$0 = P^0(B) \geq P^0(N_\cdot \in B,\ \sigma_n = \infty) = P^0(N_{\cdot \wedge \sigma_n} \in B,\ \sigma_n = \infty).$$

According to Lemma 2, $\sigma_n = \sigma_n(N_{\cdot \wedge \sigma_n})$, and hence

$$\{ N_{\cdot \wedge \sigma_n} \in B,\ \sigma_n = \infty \} \in \mathcal{F}_{\sigma_n}.$$

Thus, with regard to (18),

$$P^1(N_{\cdot \wedge \sigma_n} \in B,\ \sigma_n = \infty) = 0.$$

We conclude that $P^1(B) \leq P^1(\sigma_n < \infty)$. (17) implies

$$\lim_{n \to \infty} P^1(\sigma_n < \infty) = 0,$$

i.e., $P^1(B) = 0$. This establishes $P^1 \prec P^0$. □

6. Assume that in (6)

(19) $\qquad \ell_n(t/t_1,\ldots,t_n) > 0,\ t \in [0,\infty),\ n = 1,2,\ldots$

The hypotheses are then symmetrical with respect to P^0 and P^1, and

$$A_t^0 = \int_0^t L_u^{-1} dA_u^1, \quad t \in [0,\infty).$$

By Theorem 1,

$$I_\sigma(P^0,P^1) = E^0 \int_0^{\sigma-} K[L_t^{-1}] dA_t^1 = E^0 \int_0^{\sigma-} (L_t - \log L_t - 1) dA_t^0.$$

Further,

$$(1 + x \log x - x) + (x - \log x - 1) = (x-1)\log x,\ x \in [0,\infty),$$

where the expressions in the brackets on the left-hand side are non-negative. Theorem 2 has the following corollary.

Corollary 1. Let (19) hold together with

$$P^i\left(\int_0^{\infty-}(1-L_t)\log L_t\, dA_t^o<\infty\right)=1,\quad i=0,1,$$

then $P^1 \sim P^o$.

Example 1. Under P^1, let N be the pure birth Markov process with transition rates q_n from n-1 to n, n = 1,2,... Under P^o, let N be the Poisson process with intensity q_o. Condition (17) of Theorem 2 is

$$(20)\qquad P^1\left(\sum_{n=1}^{\infty} K[q_n/q_o]\, q_o(\tau_n-\tau_{n-1})<\infty\right)=1.$$

(20) holds if and only if

$$\infty > I_\infty(P^1,P^o)=\sum_{n=1}^{\infty}((q_o/q_n)-\log(q_o/q_n)-1).$$

Example 2. Let P^1 be the probability distribution of a doubly stochastic Poisson process $\bar{N}$ defined on a probability space $(\Omega,\mathcal{A},\bar{P})$. Let $\{Q_t, t\geq 0\}$ be the intensity of $\bar{N}$. Further, let P^o be the probability distribution of a Poisson process with variable intensity $q(t)>0$, $t\in[0,\infty)$. Then $dA_t^o = q(t)dt$, and

$$(21)\qquad L_t(\bar{N})=\bar{E}\{Q_t \mid \bar{N}_u, u\in[0,t)\}/q(t),\quad t\in[0,\infty).$$

(See [2] for the proof of (21)). From Jensen's inequality follows

$$I_\infty(P^1,P^o)=\bar{E}\int_0^{\infty}K[L_t(\bar{N})]\,q(t)dt \leq \int_0^{\infty}\bar{E}\bar{E}\{K[Q_t/q(t)] \mid \bar{N}_u,$$
$$u\in[0,t)\}\,q(t)dt \;=\; \int_0^{\infty}\bar{E}(q(t)+Q_t\log(Q_t/q(t))-Q_t)dt.$$

Consequently, the finiteness of the last integral is sufficient for $P^1 \prec P^0$.

References

[1] KULLBACK S.: Information theory and statistics. New York-London: J. Wiley and Chapman and Hall 1959

[2] LIPTSER R.S., SHIRYAEV A.N.: Statistics of random processes. Berlin-Heidelberg-New York: Springer-Verlag (in print)

Matematicko-fyzikální fakulta
Universita Karlova
Sokolovská 83, 18600 Praha 8
Československo

(Oblatum 13.7.1977)

Sčt 18