

TR 85-42

ESTIMATION OF THE RATE OF A
DISCRETE-TIME MULTIVARIATE POINT PROCESS

John Gubner and Prakash Narayan¹

Electrical Engineering Department

University of Maryland

College Park, Maryland 20742

Abstract

We introduce the notion of a discrete-time, multivariate point process which can arise in the modeling of an optical communication system. We wish to estimate the rate of this process at time t given the past of the process up to time $t-1$. This requires the computation of a certain conditional expectation; we perform this computation by introducing an absolutely continuous change of measure and then applying the generalized Bayes' rule.

¹This research was sponsored by the Office of Naval Research under grant no.N0001485-G-0102 and by the Minta Martin Fund for Aerospace Research from the University of Maryland at College Park.

I. Introduction

Suppose that a laser beam, whose intensity is modulated by an information source, strikes a photodetector which is a part of the receiver of an optical communication system. Suppose that the beam is also influenced by external random phenomena; this occurs, for example, if the beam passes through a turbulent atmosphere [1]. The occurrence of photoelectrons at the detector can then be modeled as a doubly-stochastic, time-space Poisson process [1]. It is often necessary to compute the conditional expectation of the rate of such a process, for example, in performing a likelihood ratio test (Snyder [2], Chapter 2 and Chapter 6). In general, the computation of this conditional expectation is quite difficult, although results are available if the rate process has a Gaussian form, and the photodetector surface is very large; in addition, results for *linear* estimates are available when the photodetector surface is arbitrary, but the rate process still has a Gaussian form [3]. (see also [4, 5] for related filtering results).

Let D represent the photosensitive surface of the photodetector. It is well known that the probability is negligible that more than one photoelectron will occur in the entire region D during a time interval of length Δt , if Δt is sufficiently small. Let D_1, \dots, D_K , be a partition of D into disjoint subregions. Let Δt be "sufficiently small," and let $n_t(k)$ denote the number of photoelectrons occurring in the region D_k , during an interval $(t, t + \Delta t]$. We should expect that for each $1 \leq k \leq K$, $n_t(k)$ takes only the values 0 or 1. In addition, we can have $n_t(k) = 1$ for at most one k ; for $i \neq k$, we must have $n_t(i) = 0$. In this paper we formulate the discrete-time version of the rate-estimation problem by introducing the notion of a discrete-time, multivariate point process. By adapting the continuous-time procedures found in (Bremaud [6], Chapter IV; see also pp. 69-70, 80 of Chapter III), we solve the problem of estimating the rate of a discrete-time, multivariate point process. (For more general discrete-time procedures, see [7, 8]).

II. Problem Statement

Let $K \geq 1$ be a fixed integer. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space on which all the random variables in this section are defined. We call $\{n_t(k); t \geq 1, 1 \leq k \leq K\}$ a discrete-time, multivariate point process if each $n_t(k)$ takes only the values 0 and 1 and if the events $\{n_t(k) = 1\}, 1 \leq k \leq K$, are disjoint, so that simultaneous events do not occur. Next, let \mathcal{G}_0 denote the trivial σ -field on Ω , and set

$$\mathcal{G}_t \triangleq \sigma\{n_s(k); 1 \leq s \leq t, 1 \leq k \leq K\}.$$

Now, let \mathbf{X} be a sub- σ -field of \mathcal{F} , and let

$$\mathcal{F}_t \triangleq \mathcal{G}_t \vee \mathbf{X}; \quad t \geq 0,$$

denote the smallest σ -field containing $\mathcal{G}_t \cup \mathbf{X}$. We assume that the random variables

$$\lambda_t(k) \triangleq \mathbf{E}[n_t(k) \mid \mathcal{F}_{t-1}]; \quad t \geq 1,$$

are actually \mathbf{X} -measurable (note that $\mathbf{X} = \mathcal{F}_0$) with *known* joint distributions. For example, we might take $\mathbf{X} = \sigma(X)$ for some random variable X with known distribution function $F(x)$. Then each $\lambda_t(k)$ would be some Borel function of X . Clearly, the joint distributions of the $\{\lambda_t(k); t \geq 1, 1 \leq k \leq K\}$ would be known, at least in principle. Now, observe that $\lambda_t(k) = \mathbf{P}(n_t(k) = 1 \mid \mathcal{F}_{t-1})$; since we assume that simultaneous events do not occur,

$$\sum_{k=1}^K \lambda_t(k) = \mathbf{P}\left[\bigcup_{k=1}^K \{n_t(k) = 1\} \mid \mathcal{F}_{t-1}\right].$$

This implies that $p_t \triangleq 1 - \sum_{k=1}^K \lambda_t(k)$ is also a conditional probability, and hence, nonnegative and bounded above by 1.

Our objective is to compute

$$\hat{\lambda}_t(k) \triangleq \mathbf{E}[n_t(k) \mid \mathcal{G}_{t-1}]$$

in terms of the known joint distribution of $\{\lambda_s(k); 1 \leq s \leq t, 1 \leq k \leq K\}$.

III. A Reformulation of the Problem

We shall solve our problem by reformulating the probabilistic setting above on a different probability space, $(\bar{\Omega}, \bar{F}, \bar{Q})$. On $(\bar{\Omega}, \bar{F}, \bar{Q})$ let $\{\nu_t(k); t \geq 1, 1 \leq k \leq K\}$ be a discrete-time, multivariate point process. In a manner analogous to that outlined in the previous section, we take \bar{G}_0 to be the trivial σ -field on $\bar{\Omega}$, and set

$$\bar{G}_t \triangleq \sigma\{\nu_s(k); 1 \leq s \leq t, 1 \leq k \leq K\}; \quad t \geq 1.$$

Next, let \bar{X} be a sub- σ -field of \bar{F} , and set

$$\bar{F}_t \triangleq \bar{G}_t \vee \bar{X}; \quad t \geq 0.$$

We now assume that

$$E_Q[\nu_t(k) \mid \bar{F}_{t-1}] = \mu_t(k), \quad (1)$$

where the $\{\mu_t(k)\}$ are arbitrary constants satisfying $\mu_t(k) > 0$ and

$q_t \triangleq 1 - \sum_{k=1}^K \mu_t(k) > 0$. (The symbol E_Q denotes expectation with respect to the measure

Q). As a consequence of the above assumption, under the measure Q , each $\nu_t(k)$ is independent of \bar{G}_{t-1} . In addition, the σ -field \bar{G}_t is independent of the σ -field \bar{X} (see Appendix). Next, let the random variables $\{\bar{\lambda}_t(k); t \geq 1, 1 \leq k \leq K\}$ defined on $\bar{\Omega}$ be \bar{X} -measurable and have the same joint distributions under Q as $\{\lambda_t(k); t \geq 1, 1 \leq k \leq K\}$ (defined on Ω) under P .

Given the preceding probabilistic setting on $(\bar{\Omega}, \bar{F}, \bar{Q})$, we make the following definitions. Let

$$\begin{aligned} \bar{p}_t &\triangleq 1 - \sum_{k=1}^K \bar{\lambda}_t(k); \quad t \geq 1, \\ l_t &\triangleq \sum_{k=1}^K \left(\frac{\bar{\lambda}_t(k)}{\mu_t(k)} - \frac{\bar{p}_t}{q_t} \right) \nu_t(k) + \frac{\bar{p}_t}{q_t}; \quad t \geq 1, \end{aligned} \quad (2)$$

and

$$L_t \triangleq \prod_{s=1}^t l_s : t \geq 1,$$

with $L_0 \equiv 1$. Observe that the denominators in (2) are nonzero and deterministic. Also, $0 \leq \tilde{\lambda}_t(k) \leq 1$ \mathbf{Q} -a.s., and hence, \tilde{p}_t is clearly bounded and \mathbf{Q} -integrable. (Recall that the joint distributions of the $\{\tilde{\lambda}_t(k)\}$ under \mathbf{Q} are the same as those of the $\{\lambda_t(k)\}$ under \mathbf{P}). Consequently, l_t and L_t are \mathbf{Q} -integrable. Since $\sum_{k=1}^K \tilde{\lambda}_t(k) = 1 - \tilde{p}_t$, and $\sum_{k=1}^K \mu_t(k) = 1 - q_t$, it is easy to see that $\mathbf{E}_{\mathbf{Q}}[l_t | \tilde{\mathbf{F}}_{t-1}] = 1$. Since $L_t = L_{t-1}l_t$, it is clear that

$$\mathbf{E}_{\mathbf{Q}}[L_t | \tilde{\mathbf{F}}_{t-1}] = L_{t-1}; \quad (3)$$

i.e., L_t is an $\tilde{\mathbf{F}}_t$ -martingale under \mathbf{Q} . Since, $\mathbf{E}_{\mathbf{Q}}[L_t] = \mathbf{E}_{\mathbf{Q}}[L_1] = \mathbf{E}_{\mathbf{Q}}[l_1] = 1$, we can define a new measure $\tilde{\mathbf{P}}$ on $\tilde{\mathbf{F}}_t$ by

$$\tilde{\mathbf{P}}(F) \triangleq \int_F L_t d\mathbf{Q}; \quad F \in \tilde{\mathbf{F}}_t.$$

(Technically, we should show that the family $\{L_t\}$ is uniformly integrable. However, for our purposes this is not necessary since if we wish to compute $\mathbf{E}[\nu_t(k) | \mathcal{G}_{\tau-1}]$ for some τ , we can select any finite $T \geq \tau$ and then restrict our attention to $1 \leq t \leq T$). Observe that since $L_0 = 1$, $\tilde{\mathbf{P}} = \mathbf{Q}$ on $\tilde{\mathbf{X}}$. If $\tilde{\mathbf{E}}$ denotes expectation with respect to the measure $\tilde{\mathbf{P}}$, it is not hard to verify (since simultaneous events do not occur) that

$$\begin{aligned} \tilde{\mathbf{E}}[\nu_t(k) | \tilde{\mathbf{F}}_{t-1}] &= \mathbf{E}_{\mathbf{Q}}[\nu_t(k) \frac{L_t}{L_{t-1}} | \tilde{\mathbf{F}}_{t-1}] \\ &= \mathbf{E}_{\mathbf{Q}}[\nu_t(k) l_t | \tilde{\mathbf{F}}_{t-1}] \\ &= \tilde{\lambda}_t(k), \end{aligned} \quad (4)$$

where the first equality in (4) follows from the generalized Bayes' rule. Now, since $\tilde{\mathbf{P}} = \mathbf{Q}$ on $\tilde{\mathbf{X}}$,

$$\tilde{\lambda}_t(k) = \tilde{\mathbf{E}}[\nu_t(k) | \tilde{\mathbf{F}}_{t-1}]$$

under $\tilde{\mathbf{P}}$ is probabilistically equivalent to $\lambda_t(k)$ under \mathbf{P} . In fact, we can make the following

statement. The probabilistic relationships among

$$\{ n_t(k) \}, \{ \lambda_t(k) \}, \{ F_t \}, \{ G_t \}, \text{ and } \mathbf{X}$$

under \mathbf{P} are the same as those among

$$\{ \nu_t(k) \}, \{ \bar{\lambda}_t(k) \}, \{ \bar{F}_t \}, \{ \bar{G}_t \}, \text{ and } \bar{\mathbf{X}}$$

under $\bar{\mathbf{P}}$. In the next section we shall compute $\bar{\mathbf{E}} [\nu_t(k) \mid \bar{G}_{t-1}]$ explicitly as a Borel function of $\{ \nu_s(i); 1 \leq s \leq t-1, 1 \leq i \leq K \}$. From the statement above, it follows that $\mathbf{E} [n_t(k) \mid G_{t-1}]$ will be equal to the same Borel function applied to $\{ n_s(i); 1 \leq s \leq t-1, 1 \leq i \leq K \}$.

IV. Calculations

In this section we compute $\bar{\mathbf{E}} [\nu_t(k) \mid \bar{G}_{t-1}]$. By Bayes' rule,

$$\bar{\mathbf{E}} [\nu_t(k) \mid \bar{G}_{t-1}] = \frac{\mathbf{E}_Q [\nu_t(k) L_t \mid \bar{G}_{t-1}]}{\mathbf{E}_Q [L_t \mid \bar{G}_{t-1}]}. \quad (5)$$

Next, observe that

$$\mathbf{E}_Q [L_t \mid \bar{G}_{t-1}] = \mathbf{E}_Q [\mathbf{E}_Q [L_t \mid \bar{F}_{t-1}] \mid \bar{G}_{t-1}] = \mathbf{E}_Q [L_{t-1} \mid \bar{G}_{t-1}]$$

by equation (3). In addition,

$$\begin{aligned} \mathbf{E}_Q [\nu_t(k) L_t \mid \bar{G}_{t-1}] &= \mathbf{E}_Q [\mathbf{E}_Q [\nu_t(k) L_t \mid \bar{G}_t] \mid \bar{G}_{t-1}] \\ &= \mathbf{E}_Q [\nu_t(k) \mathbf{E}_Q [L_t \mid \bar{G}_t] \mid \bar{G}_{t-1}]. \end{aligned}$$

Hence, (5) can be rewritten as

$$\bar{\mathbf{E}} [\nu_t(k) \mid \bar{G}_{t-1}] = \frac{\mathbf{E}_Q [\nu_t(k) \mathbf{E}_Q [L_t \mid \bar{G}_t] \mid \bar{G}_{t-1}]}{\mathbf{E}_Q [L_{t-1} \mid \bar{G}_{t-1}]}. \quad (6)$$

Before proceeding further, it will be convenient to introduce the following notation. For each $t \geq 1$, let

$$\nu_t \triangleq [\nu_t(1), \dots, \nu_t(K)]' ,$$

where ' denotes transpose. Then ν_t is a K -dimensional random vector. Now, let

$$\bar{\nu}_t \triangleq [\nu_1, \dots, \nu_t]'$$

Clearly, $\bar{\nu}_t$ is a tK -dimensional random vector. Next, let

$$z_t \triangleq [z_t(1), \dots, z_t(K)]' \text{ and } \bar{z}_t \triangleq [z_1, \dots, z_t]'$$

denote "dummy" variables in \mathbb{R}^K and \mathbb{R}^{tK} , respectively. Define

$$h_t(z_t) = \sum_{k=1}^K \left\{ \frac{\lambda_t(k)}{\mu_t(k)} - \frac{p_t}{q_t} \right\} z_t(k) + \frac{p_t}{q_t}, \quad (7)$$

and

$$\bar{h}_t(z_t) = \sum_{k=1}^K \left\{ \frac{\bar{\lambda}_t(k)}{\bar{\mu}_t(k)} - \frac{\bar{p}_t}{\bar{q}_t} \right\} z_t(k) + \frac{\bar{p}_t}{\bar{q}_t}.$$

Clearly, $h_t(z_t)$ is an \mathbf{X} -measurable random variable and $\bar{h}_t(z_t)$ is an $\bar{\mathbf{X}}$ -measurable random variable. In fact, $\bar{h}_t(z_t)$ under \mathbf{Q} is probabilistically equivalent to $h_t(z_t)$ under \mathbf{P} . It is then clear that the same can be said of

$$H_t(\bar{z}_t) \triangleq \prod_{s=1}^t h_s(z_s), \quad (8)$$

and

$$\bar{H}_t(\bar{z}_t) \triangleq \prod_{s=1}^t \bar{h}_s(z_s).$$

Next observe that $l_t = \bar{h}_t(\nu_t)$, and that $L_t = \bar{H}_t(\bar{\nu}_t)$. We are now ready to compute $\mathbf{E}_{\mathbf{Q}}[L_t | \bar{\mathcal{G}}_t]$ in (6) for each $t \geq 1$. (Note that $\mathbf{E}_{\mathbf{Q}}[L_0 | \bar{\mathcal{G}}_0] = \mathbf{E}_{\mathbf{Q}}[L_0] = 1$). Observe that $L_t = \bar{H}_t(\bar{\nu}_t)$ is an $\bar{\mathbf{X}}$ -measurable function of a $\bar{\mathcal{G}}_t$ -measurable random vector. Since $\bar{\mathbf{X}}$ and $\bar{\mathcal{G}}_t$ are independent under \mathbf{Q} , it follows that

$$\mathbf{E}_{\mathbf{Q}}[\bar{H}_t(\bar{\nu}_t) | \bar{\mathcal{G}}_t] = \mathbf{E}_{\mathbf{Q}}[\bar{H}_t(\bar{z}_t)] | \bar{z}_t = \bar{\nu}_t.$$

From the remarks preceding equation (8), we have

$$\mathbf{E}_Q [L_t \mid \bar{\mathcal{G}}_t] = \mathbf{E} [H_t(\bar{z}_t) \mid \bar{z}_t = \bar{v}_t].$$

We now set

$$f_t(\bar{z}_t) \triangleq \mathbf{E} [H_t(\bar{z}_t)]. \quad (9)$$

Note that $f_t(\bar{z}_t)$ is a deterministic function of \bar{z}_t . Equation (6) becomes

$$\bar{\mathbf{E}} [\nu_t(k) \mid \bar{\mathcal{G}}_{t-1}] = \frac{\mathbf{E}_Q [\nu_t(k) f_t(\bar{v}_t) \mid \bar{\mathcal{G}}_{t-1}]}{f_{t-1}(\bar{v}_{t-1})}.$$

We can write

$$\mathbf{E}_Q [\nu_t(k) f_t(\bar{v}_t) \mid \bar{\mathcal{G}}_{t-1}] = \mathbf{E}_Q [\nu_t(k) f_t(\bar{v}_{t-1}, \nu_t) \mid \bar{\mathcal{G}}_{t-1}].$$

Now, equation (1) implies that ν_t is independent of $\bar{\mathcal{G}}_{t-1}$ under \mathbf{Q} . Therefore,

$$\mathbf{E}_Q [\nu_t(k) f_t(\bar{v}_t) \mid \bar{\mathcal{G}}_{t-1}] = \mathbf{E}_Q [\nu_t(k) f_t(\bar{z}_{t-1}, \nu_t) \mid \bar{z}_{t-1} = \bar{v}_{t-1}].$$

Since simultaneous events do not occur, it is trivial to compute

$$\begin{aligned} \mathbf{E}_Q [\nu_t(k) f_t(\bar{z}_{t-1}, \nu_t)] &= f_t(\bar{z}_{t-1}, e_k) \mathbf{Q}(\nu_t(k) = 1) \\ &= \mu_t(k) f_t(\bar{z}_{t-1}, e_k), \end{aligned}$$

where e_k is the standard unit vector in \mathbf{R}^K with a 1 in the k th position and a 0 in the other $K-1$ positions. We conclude that

$$\bar{\mathbf{E}} [\nu_t(k) \mid \bar{\mathcal{G}}_{t-1}] = \frac{\mu_t(k) f_t(\bar{v}_{t-1}, e_k)}{f_{t-1}(\bar{v}_{t-1})},$$

and hence,

$$\mathbf{E} [n_t(k) \mid \mathcal{G}_{t-1}] = \frac{\mu_t(k) f_t(\bar{n}_{t-1}, e_k)}{f_{t-1}(\bar{n}_{t-1})},$$

with the obvious meaning of the symbol \bar{n}_{t-1} .

V. Summary

We have shown that if $\{ n_t(k); t \geq 1, 1 \leq k \leq K \}$ is a discrete-time, multivariate point process residing in the probabilistic setting outlined in Section II, then

$$\mathbf{E} [n_t(k) \mid \mathcal{G}_{t-1}] = \frac{\mu_t(k) f_t(\bar{n}_{t-1}, e_k)}{f_{t-1}(\bar{n}_{t-1})},$$

where $f_t(\bar{z}_t)$ is given by equations (7), (8), and (9).

The $\mu_t(k)$ introduced in Section III were arbitrary; if we set $\mu_t(k) = \frac{1}{K+1}$ for all t and all k in the preceding equations, we find that

$$\mathbf{E} [n_t(k) \mid \mathcal{G}_{t-1}] = \frac{a_t(\bar{n}_{t-1}, e_k)}{a_{t-1}(\bar{n}_{t-1})},$$

where

$$a_t(\bar{z}_t) \triangleq \mathbf{E} \left[\prod_{s=1}^t \left(\sum_{k=1}^K (\lambda_s(k) - p_s) n_s(k) + p_s \right) \right].$$

Appendix

For $1 \leq s \leq t$, let z_s denote either the zero vector in \mathbb{R}^K or any standard unit vector in \mathbb{R}^K . In this appendix we prove that if $E \in \bar{\mathbf{X}}$, then (using the notation of Sections III and IV)

$$\mathbf{Q}(\nu_t = z_t, \dots, \nu_1 = z_1, E) = \left(\prod_{s=1}^t \mathbf{Q}(\nu_s = z_s) \right) \mathbf{Q}(E).$$

Since $\bar{\mathcal{G}}_t = \sigma\{\nu_1, \dots, \nu_t\}$, this will prove that $\bar{\mathcal{G}}_t$ and $\bar{\mathbf{X}}$ are independent under \mathbf{Q} .

Proof. Using the definition of conditional probability,

$$\mathbf{Q}(\nu_t = z_t, \dots, \nu_1 = z_1, E) = \int \mathbf{Q}(\nu_t = z_t \mid \bar{\mathbf{F}}_{t-1}) d\mathbf{Q}. \quad (10)$$

$\{\nu_{t-1} = z_{t-1}, \dots, \nu_1 = z_1, E\}$

Since simultaneous events do not occur, if $z_t = e_k$, $Q(\nu_t = z_t \mid \bar{F}_{t-1}) = \mu_t(k)$. If $z_t = 0$,

$$Q(\nu_t = z_t \mid \bar{F}_{t-1}) = 1 - \sum_{k=1}^K \mu_t(k). \text{ Since } Q(\nu_t = z_t \mid \bar{F}_{t-1}) \text{ is deterministic,}$$

$$Q(\nu_t = z_t \mid \bar{F}_{t-1}) = Q(\nu_t = z_t).$$

Hence (10) becomes

$$Q(\nu_t = z_t, \dots, \nu_1 = z_1, E) = Q(\nu_t = z_t)Q(\nu_{t-1} = z_{t-1}, \dots, \nu_1 = z_1, E).$$

The remainder of the proof by induction is clear.

QED

Acknowledgement

It is a pleasure to thank Professor A. M. Makowski for several illuminating discussions on the applications of an absolutely continuous change of measure to filtering problems.

References

- [1] D. L. Snyder, "Applications of Stochastic Calculus for Point Process Models Arising in Optical Communication," pp. 789-804 in *Communication Systems and Random Process Theory*, ed. J. K. Skwirzynski, Sijthoff and Noordhoff, Alphen aan der Rijn, The Netherlands (1978).
- [2] D. L. Snyder, *Random Point Processes*, John Wiley & Sons, New York (1975).
- [3] J. Gubner and P. Narayan, "Estimation of the Rate of a Doubly-Stochastic Time-Space Poisson Process," *Systems & Control Letters*, Submitted, (1985).
- [4] D. L. Snyder and P. M. Fishman, "How to Track a Swarm of Fireflies by Observing Their Flashes," *IEEE Transactions on Information Theory*, vol. IT-21, pp. 692-695 (Nov. 1975).
- [5] I. B. Rhodes and D. L. Snyder, "Estimation and Control Performance for Space-Time Point-Process Observations," *IEEE Transactions on Automatic Control*, vol. AC-22, No. 3, pp. 338-346 (June 1977).
- [6] P. Bremaud, *Point Process and Queues, Martingale Dynamics*, Springer-Verlag, New York (1981).
- [7] P. M. Bremaud and J. H. van Schuppen, *Discrete Time Stochastic Systems: Part I: Stochastic Calculus and Representations*, Report No. 7603a, Department of Systems Science and Mathematics, Washington University, St. Louis (1976).
- [8] P. M. Bremaud and J. H. van Schuppen, *Discrete Time Processes: Part II: Estimation Theory*, Report No. 7603b, Department of Systems Science and Mathematics, Washington University, St. Louis (1976).