Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

E. Valkeila

A Prohorov bound for a Poisson process and a Bernoulli process

Department of Mathematical Statistics Report MS-R8701 January

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

A Prohorov Bound for a Poisson Process and a Bernoulli Process

Esko Valkeila
Computing Centre, University of Helsinki,
Tukholmankatu 2, SF-00250 Helsinki, FINLAND

Let M be a counting process with the compensator A and let N be the Poisson process with the compensator B. We give an upper bound for the Prohorov distance between the two processes in terms of the compensators, when the other process is a Bernoulli process with dependency.

1980 Mathematics Subject Classification: Primary, 60F14, 60G55, secondary 60K05.
Key Words & Phrases: Counting process, compensator, Prohorov distance, weak convergence, Bernoulli process.
Note: This research was done during the author's stay at the CWI. I want to thank the Ella and Georg Ehrnrooth Foudation, the Finnish Ministry of Education, the Netherlands Ministry of Education and CWI for making this stay possible.

1. Introduction

1.1. Let $(M^n, F^n)_{n \geq 1}$ be a sequence of counting processes with the compensators A^n. Let N be a non-homogeneous Poisson process with the compensator B. Brown (1978 , 1981) and Kabanov et al (1980, 1983) proved the following weak convergence result. If for all $t \geq 0$ we have

$$A_t^n \to^P B_t \text{ , as } n \to \infty, \tag{1.1}$$

where $\to^P$ means convergence in probability, then

$$M^n \to^D N, \text{ as } n \to \infty, \tag{1.2}$$

where $\to^D$ means weak convergence of the distributions of the sequence (N^n) to the distribution of M in the Skorohod space D.

1.2. Brown (1983) and Kabanov et al (1983) discuss also the proximity of the distributions of the sequence $(M^n)_{n \geq 1}$ and the distribution of the limit process N. To estimate the distance between M^n and N they use the variation distance between the finite dimensional distributions of the processes. These methods give also some results for the variation distance between the distributions on the interval $[0, T]$, where $T > 0$ is some constant. Later it turned out that in some cases it is more convenient to work with the Hellinger distance instead of the variation distance. For more details of the Hellinger distance and the associated Hellinger process in this context see Kabanov et al (1986) and Valkeila and Vostrikova (1986). Note, however, that the topology induced by the variation distance or the Hellinger distance is stronger than the topology of weak convergence in the space D.

The following simple example clarifies the discussion above. Let $(X_{k,n})_{1 \leq k \leq n}$ be independent Bernoulli variables with $P(X_{k,n} = 1) = 1/n$, where $k = 1,...,n$ and put $X_{0,n} = 0$. Define the Bernoulli process with jump probability 1/n by

Report MS-R8701
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

$$M_t^n = \sum_{k=0}^{[nt]} X_{k,n} ,$$

where $[x]$ is the integer part of x . Then the compensator A_t^n of M^n is $[nt]/n$. Let N be the standard Poisson process with unit intensity. If $n \to \infty$ then the condition (1.1) above is fulfilled, but the sequence (M^n) does not converge strongly (i.e. in the variation norm) in the space D. So the results of BROWN (1983) and KABANOV et al (1983) mentioned above can not be used to estimate the rate of convergence in (1.2) in the space D.

1.3. In NIKUNEN and VALKEILA (1985) we gave a rate for this convergence in terms of the Prohorov distance in the space D under the assumption that the compensator processes A^n are continuous functions of t. So the bound derived there can not be applied to the above case, either. Now we can modify the method of the proof in NIKUNEN and VALKEILA (1985) to get an upper bound for the Prohorov distance, even when the compensator processes can have jumps(and no continuous increasing part). In our proof we benefit a lot from the work of KABANOV et al (1983). We have from our general bound the following special result, which has been dealt with DUDLEY (1972) by another method. Let N be the standard Poisson process with unit intensity and M a Bernoulli process with jump probability $1/n$. Then

$$\rho(P_M, P_N) \leqslant 4/n , \tag{1.3}$$

where ρ is the Prohorov distance between the distributions of M and N in the unit interval [0,1]. We give more applications of our bound given in Theorem 2.1 in the last section of this paper.

1.4. The next section contains the formulation of our results after some notation and definitions. The third section contains the proofs. In the fourth section we prove (1.3) and apply our general bound in (2.1) to give estimates between the Prohorov distance between sums of N-valued random variables and a Poisson process.

1.5. I want to thank Martti Nikunen for reading the first draft of this note. His comments lead to some improvements in representation and to correct formulations of the results.

2. NOTATION, DEFINITIONS AND THE RESULTS

2.1. Denote by $D[0,T]$ the space of right continuous real functions x on $[0,T]$ with left hand limits. Let $x,y \in D$ and denote by $m_T(x,y)$ the uniform distance between x and y . We recall the definition of the Skorohod distance . Let Λ be the class of strictly increasing, continuous mappings of $[0,T]$ onto itself. If $\lambda \in \Lambda$, then $\lambda(0)=0$ and $\lambda(T)=T$. Define the Skorohod distance d_T by

$$d_T(x,y) = \inf_{\lambda \in \Lambda} \{ m_T(\lambda \circ x, y) + m_T(\lambda \circ I, I) \} ,$$

where $\lambda \circ x(t) = x(\lambda(t))$ and I is the function $I_t = t$. Then the space $(D[0,T], d_T)$ is a separable metric space (BILLINGSLEY (1968)). Note also that $d_T(x,y) \leqslant m_T(x,y)$.

2.2. If $x,y \in R$ then let $\min(x,y) = x \wedge y$ and $\max(x,y) = x \vee y$. For a nonnegative random variable X defined on a probability space (Ω , F , P) we put

$$\nu(X) = \inf \{ \epsilon \geqslant 0 : P(X \geqslant \epsilon) \leqslant \epsilon \}.$$

It is easy to see that $\nu(X)^2 \leqslant E(X)$ and for any $\epsilon > 0$ we have $\nu(X) \leqslant \epsilon \vee P(X \geqslant \epsilon)$. If Y is another nonnegative random variable, then $\nu(X+Y) \leqslant \nu(X) + \nu(Y)$, and if $X \leqslant Y$ $(P-a.s.)$, then $\nu(X) \leqslant \nu(Y)$.

Let (S,d) be a separable metric space and X,Y random elements from a probability space (Ω ,F, P) into S. Define the Ky-Fan metric $\alpha\ (X,Y;d)$ by

$$\alpha\ (X,Y;d) = \nu\ (d(X,Y)).$$

Recall that this metric metricizes convergence in probability.

Let $B(S)$ be the $\sigma-$algebra of Borel sets of S . For a Borel set C denote by C^ϵ the set

$$C^\epsilon = \{x \in S : d(x,C) < \epsilon \}.$$

If R and Q are two probability measures defined on $(S,\ B(S))$ then define the Prohorov distance $\rho(R,Q;d)$ by

$$\rho(R,Q;d) = \inf \{\epsilon > 0 : R(F) \leqslant \epsilon + Q(F^\epsilon) \text{ for all closed } F\}.$$

We use the following device to estimate the Prohorov distance. Let X and Y be S-valued random elements defined on a common probability space such that X induces the measure R and Y induces the measure Q on S. The pair (X,Y) is called a coupling of the measures R and Q. Let $C(R,Q)$ be the set of all possible couplings of R and Q (where also the space (Ω,F,P) may vary). Strassen and Dudley (see DUDLEY (1968)) proved that

$$\rho(R,Q;d) = \inf (\ \alpha(X,Y;d) : (X,Y) \in C(R,Q)\ ).$$

Hence $\rho(R,Q;d) \leqslant \alpha(X,Y;d)$ for any coupling (X,Y) .

2.3. Let (Ω,F,P) be a probability space with a filtration $(F_t)_{t\geqslant 0}$. We suppose that the filtration is right continuous and that F_0 contains all $P-$ null sets. A counting process N adapted to the filtration $(F_t)_{t\geqslant 0}$ is a process having right continuous increasing piece-wise constant paths with unit jumps and $N_0 = 0$. We suppose that $P(N_t < \infty) = 1$ for each $t \geqslant 0$. The compensator A with respect to the filtration $(F_t)_{t\geqslant 0}$ and the measure P is a predictable increasing process such that the process $N - A$ is a square-integrable local martingale. It can be shown that $\Delta A_t \leqslant 1$ (P -a.s.) , where $\Delta A_t = A_t - A_{t-}$. If A is a deterministic and continuous function then the process N is a nonhomogeneous Poisson process. We say that M is a Bernoulli process with dependency, if there exists fixed time points $\{s_1, \ldots, s_m\}$ and $s_i < s_{i+1}$ for $i=1,...,m-1$ and a sequence of 0,1-valued random variables X_i such that

$$M_{s_i} = M_{s_{i-1}} + X_i\ ,$$

and M is constant between the intervals $[s_{i-1}, s_i[$. Note that the distribution of M as a random element may depend of the order of the sequence (X_i). In below, when we speak about a Bernoulli process, we assume that we are given the fixed jump times $\{s_1, \ldots, s_m\}$. Note also that the compensator A of M has the form

$$A_t = \sum_{s_i \leqslant t} \Delta A_{s_i}\ .$$

For more details on counting processes we refer to LIPTSER and SHIRYAYEV (1978).

2.4. Now we formulate the results. The counting process induces a measure on the space $(D[0,T], d_T)$ and we denote it by P_N. For two counting processes N and M defined on the same probability space we write $\rho_T(P_M,P_N)$ instead of $\rho(P_M,P_N;d_T)$ and similarily $\alpha_T(A,B) = \alpha(A,B,m_T)$ for two increasing funtions A and B.

THEOREM 2.1. *Let $(M_t,F_t)_{t\geqslant 0}$ be a Bernoulli process on the space (Ω,F,P) with the compensator A and let N be a nonhomogeneous Poisson process on the same space with a strictly increasing compensator B.*

Then we have for $T > s_m$

$$\rho_T(P_M,P_N) \leq \alpha_T\,(B^{-1}\circ A,I) + E\,|A_T - B_T| + 2E\sum_{t\leq T}(\Delta A_t)^2\,. \tag{2.1}$$

REMARK 2.1. We prove the bound (1.3) in the section 4. We give other examples in the same section.

REMARK 2.2. Given any $\epsilon > 0$ we can replace the term

$$E\,|\,A_T - B_T\,|$$

in (2.1) by

$$E(\epsilon\wedge|A_T - B_T|) + P(|\,A_T - B_T\,| \geq \epsilon)\,.$$

REMARK 2.3. In NIKUNEN and VALKEILA (1985) we showed that for continuous compensator A the bound (2.1) involves only the first two terms. We were not able to combine the two different bounds to a single bound.

3. PROOFS

3.1. We start the proof and give some notation, which will be used later. If X is a process and $\{0=t_0,\ldots,t_n=T\}$ is a partition of the interval $[0,T]$, then by $f^n(X)$ we mean the discreticized process

$$f_t^n(X)=X_{t_k} \text{ if } t \in [t_k,t_{k+1}[$$

for $k = 0,\ldots,n-1$. If X and Y are two stochastic processes and ρ is a distance between two probability measures, then we write $\rho(P_X^n,P_Y^n)$ for the distance between the finite dimensional distributions of the processes X and Y at time points $t_0,\ldots,t_n$ in the space R^{n+1}.

The next Lemma is almost obvious.

LEMMA 3.1. *Let M be a Bernoulli process with jump times $\{s_1,\ldots,s_m\}$ and let $\{0=t_0,t_1,\ldots,t_n=T\}$ be a partition of the interval* $[0,T]$ *with $s_i \neq t_j$ for $i=1,\ldots,m$ and for $j=0,\ldots,n$. Then we have*

$$\rho_T(P_M,P_{f^n(M)}) \leq \max_{i\leq n}|t_i - t_{i-1}|. \tag{3.1}$$

PROOF Define a scaling function λ in such a way that $\lambda(t_i)=t_i$, if there is no jump point s_j in the interval $]t_{i-1},t_i[$ and $\lambda(s_j)=t_i$, if $s_j \in]t_{i-1},t_i[$ (we may assume that $s_1 > t_2$ and that $s_m < t_{n-1}$) and continue λ to be linear between these points. Then

$$M_t = f_{\lambda(t)} \text{ and } |\lambda(t)-t| \leq \max_{i\leq n}|t_i - t_{i-1}|.$$

From these observations we have (3.1). □

3.2. Our next step is to use the following Lemma due to KABANOV et al (1983). Before we formulate it, we have to make some assumptions. Denote by X^T the process $X_t^T = X_{t\wedge T}$. Assume that $B^T = B$, $A^T = A$ and $A^T \leq c$, where c is a constant. Let $\{t_0,\ldots,t_n\}$ be a partition of the interval

$[0,T]$. We can also assume that $P(\Delta A_{t_i}>0) = 0$ for $i=1,...,n$. Finally we assume that on the probability space (Ω,F,P) there is a sequence of independent Poisson processes not depending on F_∞ . We shall use the following notation in below. If f is a bounded predictable process , then by $f{*}N$ we mean the process

$$(f*N)_t = \int_0^t f_s \, dN_s ,$$

where the integral is a Lebesque-Stieltjes integral . Also, if A is a set, then by 1_A we shall denote the indicator function of the set A. In the next Lemma we suppose that a fixed partition of the interval $[0,T]$ is given. We denote by $\mathrm{Var}(P^n(N),P^n(M))$ the total variation distance between the finite dimensional distributions of M and N .

LEMMA 3.2. (KABANOV et al). *Let the compensator A of the counting process N satisfy the above assumptions . Let S be a predictable stopping time such that for some $j<n$,*

$$t_{j-1}< S <t_j \ (P-a.s.) \ \text{ on } \{ S<\infty\} ,$$

and $\Delta A_S>0$ $(P-a.s.)$ on $\{ S<\infty\}$. Then there exists a counting process $\tilde{N}$ with a compensator $\tilde{A}$ with respect to the filtration $(\tilde{F}_t)$ (which is finer than the original filtration (F_t)) with the following properties :

$$\mathrm{Var}\ (P^n(N) ,P^n(\tilde{N})) \leq E(\Delta A_S)^2 , \tag{a}$$

$$\tilde{A}_{t_i} = A_{t_i} ,i = 1,...,n , \tag{b}$$

$$\{ \Delta\tilde{A} >0 \} = \{ \Delta A >0 \} \ \setminus [S], \text{ where } [S] = \{ \omega \in\Omega , t >0 : S(\omega) = t \}. \tag{c}$$

Moreover, the compensator $\tilde{A}$ can be such that

$$\tilde{A}_t = A_t - (1_{[S]} *A)_t + 1_{[S,\infty[}(t)\Delta A_S \frac{t_j\wedge t-S}{t_j-S} . \tag{3.2}$$

For the proof we refer to KABANOV et al (1983).

Because we need the concrete form of the process $\tilde{M}$ below, we write it for our case. Let $(\pi_i)_{1\leq i\leq m}$ be m independent standard Poisson processes, which are also independent of the σ-field F_∞ . If $s_j \in]t_i,t_{i+1}[$ then define $\tilde{\pi}_j$ by the following way: $\tilde{\pi}_j(t)=0$, if $t<s_j$ and

$$\tilde{\pi}_j(t)=\pi(s_j+h(t))-\pi_j(s_j), \text{ where } h(t)=\Delta A_{s_j}\frac{t_{i+1}\wedge t-s_j}{t_{i+1}-s_j} , \tag{3.3}$$

and after applying Lemma 3.2. m times

$$\tilde{M}=\sum_{j=1}^{m}\tilde{\pi}_j .$$

3.3. The next Lemma shows how the Prohorov distance can be approximated by the corresponding distance between the finite dimensional distributions of the process.

LEMMA 3.3. (KUBILIUS and MIKULEVICIUS). *Let X and Y be two processes with paths in D and let $\{ t_0, \cdots ,t_n \}$ be a partition of the interval $[0,T]$. Then*

$$\rho_T(P_{f^*(X)}, P_{f^*(Y)}) \leq \rho(P^n_X, P^n_Y) . \tag{3.4}$$

PROOF We sketch the proof of Kubilius and Mikulevicius. The space of piece-wise constant functions $D^n = (x \in D : f^n(x) = x)$ is a closed subspace of D. Note also that the topologies induced by the metrics d_T and m_T coincide on D^n. It remains to note that the space D^n with metric m_T is isomorphic to the space R^{n+1} with the metric $r(x,y) = \max_{0 \leq i \leq n}(|x_i - y_i|)$, where $x = (x_0, \ldots, x_n) \in R^{n+1}$. □

REMARK 3.2. Note that we have not used the ordinary metric in the space R^{n+1} in the proof above. Because we apply this Lemma to estimate the distance between finite dimensional distributins in terms of the variation metric, we need not take care about this fact in below.

3.4. Now we proceed to the proof of Theorem 2.1. Take first a partition $\{0 = t_0, \ldots, t_n = T\}$ in such a way that $s_j \neq t_i$ and also with the property that $\max_{i \leq n} |t_i - t_{i-1}| < \epsilon$ for a given $\epsilon > 0$. Then we have Lemma 3.1 that

$$\rho_T(P_{f^*(M)}, P_M) \leq \epsilon. \tag{3.5}$$

Apply now Lemma 3.2 m times and we obtain a counting process $\tilde{M}$ with a countinuous compensator $\tilde{A}$ such that

$$\mathrm{Var}\,(P^n(M), P^n(\tilde{M})) \leq E \sum_{j \leq m} (\Delta A_{s_j})^2 . \tag{3.6}$$

So we have

$$\rho_T(P_M, P_N) \leq \rho_T(P_N, P_{\tilde{M}}) + \rho_T(P_{\tilde{M}}, P_{f^*(\tilde{M})}) + \rho_T(P_{f^*(\tilde{M})}, P_{f^*(M)}) + \rho_T(P_{f^*(M)}, P_M) . \tag{3.7}$$

According to Theorem 1 in NIKUNEN and VALKEILA (1985) we have for the first term on the right hand side of (3.7) that given any $\epsilon > 0$

$$\rho_T(P_N, P_{\tilde{M}}) \leq \alpha_T(B^{-1} \circ \tilde{A}, I) + E(\epsilon \wedge |\tilde{A}_T - B_T|) + P(|\tilde{A}_T - B_T| \geq \epsilon) . \tag{3.8}$$

Note that $\tilde{A}_{t_i} = A_{t_i}$ and that $\tilde{A}_T = A_T$ in (3.8).

Next we show that

$$\alpha_T(B^{-1} \circ \tilde{A}, I) \leq \alpha_T(B^{-1} \circ A, I) + \max_{i \leq n} |t_i - t_{i-1}| . \tag{3.9}$$

Indeed, if $t \in [t_{i-1}, t_i[$, then

$$0 < B^{-1} \circ \tilde{A}_t - t \leq B^{-1} \circ \tilde{A}_{t_i} - t_i + t_i - t_{i-1} = \alpha_T(B^{-1} \circ A, I) + \max_{i \leq n} |t_i - t_{i-1}| .$$

and symmetrically, if $0 < t - B^{-1} \circ \tilde{A}_t$, using t_{i-1} in place of t_i, we have (3.9).

To get an estimate for the third term on the right hand side of (3.7) we use (3.6), (3.4) and the fact that the Prohorov distance is less than the corresponding variation distance. So we have

$$\rho_T(P_{f^*(M)}, P_{f^*(\tilde{M})}) \leq E \sum_{t \leq T} (\Delta A_t)^2 . \tag{3.10}$$

Our final step is to prove that

$$\rho_T(P_{f^*(\tilde{M})}, P_{\tilde{M}}) \leq E \sum_{s \leq T} (\Delta A_s)^2 + \max_{i \leq n} |t_i - t_{i-1}| . \tag{3.11}$$

To do this, we modify the argument in DUDLEY (1972) a little. Put

$$B = \bigcup_{j \leq m} \{\tilde{\pi}_i(T) \geq 2\}.$$

On the complement of the set B the process $\tilde{M}$ increases only by jumps of size 1 in some interval $]t_{i-1}, t_i[$. Define $\lambda_{t_i}(\omega) = \inf(s \mid \tilde{\pi}_j(s) = \tilde{\pi}_j(t_i))$ if there is a jump and $\lambda_{t_i}(\omega) = t_i$ otherwise . Extend $\lambda(\omega)$ to be linear between these points. Then we have

$$d_T(\tilde{M}, f^n(\tilde{M})) \leq \max_{i \leq n} |t_i - t_{i-1}| + 1_B \sum_{j \leq m} \tilde{\pi}_j(T). \tag{3.12}$$

But $\pi_j(T)$ is conditionally on F_∞ Poisson distributed with parameter ΔA_{s_j}. Hence we have for any δ

$$P(1_B \sum_{j \leq m} \tilde{\pi}_j(T) \geq \delta) \leq P(B) \leq \sum_{j \leq m} E(\Delta A_{s_j})^2 .$$

From this we have (3.11) using the properties of the function ν mentioned in Section 2. Because $\max_{i \leq n} |t_j - t_{j-1}| < \epsilon$, we have , through the steps in (3.7) - (3.12) , proved the inequality (2.1).

4. EXAMPLES AND APPLICATIONS

EXAMPLE 4.1. First we give an example, where we get the bound in (1.3) as a special case. Let $(X_k)_{1 \leq k \leq n}$ be a sequence of Bernoulli- variables with

$$P(X_k = 1 | X_1, \ldots, X_{k-1}) = p_k(X_1, \ldots, X_{k-1}) = p_k . \tag{4.1}$$

Let $X_0 = 0$ and define the process M by

$$M_t = \sum_{k=0}^{[nt]} X_k . \tag{4.2}$$

Note that the distribution of the process M may depend of the ordering of the sequence (X_k) . If $F_t = F_t^M = \sigma(X_k : k \leq [nt])$ then the compensator A of the process M is

$$A_t = \sum_{k=0}^{[nt]} p_k , \tag{4.3}$$

where $p_0 = 0$. Let M be a non-homogeneous Poisson process with the compensator B . Then , from (2.1) we have that

$$\rho_1(P_M, P_N) \leq \alpha_1(B^{-1} \circ A, I) + E|A_T - B_T| + 2E \sum_{k=1}^{n} p_k^2 , \tag{4.4}$$

provided that $p_n = 0$. If this is not the case, then we can instead of process M consider the process M^ϵ, which jumps at the time points $i/n - \epsilon$, where $0 < \epsilon < 1/n$ and $i = 1,...,n$. It is clear that we have

$$\rho_1(P_M, P_{M^\epsilon}) \leq \epsilon + Ep_n \leq \epsilon + E \max_{i \leq n} p_k .$$

Hence we have instead of (4.4) the following bound

$$\rho_1(P_M, P_N) \leq \alpha_1(B^{-1} \circ A, I) + E|A_1 - B_1| + 2E \sum_{s \leq 1} (\Delta A_s)^2 + E(\max_{i \leq n} p_k) . \tag{4.5}$$

Now suppose that $p_k = 1/n$ for $k = 1,...,n$ and M that is a standard Poisson process. Then we have $A_1 = B_1 = 1$, $\alpha_1(A, I) \leq 1/n$, $\sum_{k=1}^{n} p_k^2 = 1/n$ and finally $\max_{1 \leq k \leq n} p_k = 1/n$. So we have

$$\rho_1(P_M,P_N) \leq 4/n ,$$

and we have (1.3). WHITT (1973) gives also this upper bound as a special case of his multivariate results. He considers independent Bernoulli processes. To finish this example we note that DUDLEY (1972) showed also that

$$\rho_1(P_M,P_N) > o(1/n)$$

in the case, when N is a Bernoulli process with jump probabilities $1/n$ and M is a standard Poisson process .

EXAMPLE 4.2. SERFLING (1978) discusses the upper bounds in the case, where the jumps of the process M can take values in N_+ . Next we indicate, how our bound (2.1) can be applied to this case. Let $(X_k)_{k\leq n}$ be a sequence of integer valued random variables. Put $X_0=0$ and define the process M by (4.2). Let p_k be as in (4.1). Let μ be the jump measure of M and let $\tilde{\mu}$ be its compensator with respect to the measure P and (F_t^M). For details about these conceps we refer to JACOD (1979). Define two counting processes M^1 and M^2 by

$$M_t^1 = \int_{\{x=1\}} \mu(]0,t],dx) = \sum_{i\leq[nt]} 1_{\{X_i=1\}}$$

and

$$M_t^2 = \int_{\{x>1\}} \mu(]0,t],dx) = \sum_{i\leq[nt]} 1_{\{X_i\geq2\}}.$$

Then there compensators A^1 and A^2 are

$$A_t^1 = \tilde{\mu}(]0,t],\{1\}) = \sum_{i\leq[nt]} p_i$$

and

$$A_t^2 = \tilde{\mu}(]0,t],\{x\geq2\}) = \sum_{i\leq[nt]} d_i,$$

where $d_i = P(X_i\geq2 \mid X_1,\ldots,X_{i-1})$. Put $B = \bigcup_{i\leq m} \{X_i\geq2\}$ and then we have that

$$d_1(M,M^1) \leq 1_B(M-M^1)_1. \tag{4.6}$$

Note that for any $0<\delta<1$ we have $P(1_B(M-M^1)_1\geq\delta) = P(M_1^2\geq1)$ and we have from (4.6) that

$$\rho_1(P_M,P_{M^1}) \leq P(M_1^2\geq1) \leq EA_1^2. \tag{4.7}$$

Then we have, if N is again a Poisson process with a strictly increasing compansator B , from (4.5) and (4.6) the following bound

$$\rho_1(P_M,P_N) \leq \alpha_1(B^{-1}\circ A^1,I) + E\,|A_1^1 - B_1| + 2E\sum_{t\leq1}(\Delta A_t^1)^2 + E\sup_{t\leq1}|\Delta A_t^1| + EA_1^2. \tag{4.8}$$

EXAMPLE 4.3. Our last example shows how we can approximate the Prohorov distance between a renewal process and a Poisson process in terms of the interrenewal distributions. First we introduce some notation. Let G be the distribution function and suppose that $(U_k)_{k\geq1}$ are independent and identically distributed according to G. We suppose that $G(0)=0$ and that G is either continuous or discrete with span d (by this we mean that $P(U\in\{d, 2d, 3d,\ldots\})=1$) . Define the integrated hazard H of G by

$$H(t) = \int_0^t \frac{G(ds)}{1-G(s-)}$$

and let R be the renewal function. Define the counting process M by

$$M_t = \sum_{k \geq 1} 1_{\{T_k \leq t\}} \ ,$$

where $T_k = U_1 + \cdots + U_k$ and put $T_0 = 0$. Suppose that T is a fixed time such that $T \neq d[dT]$. Then we have $R(T) = E(M_T + 1)$. If N is a Poisson process with strictly increasing compensator $B_t = \mu t$, then

$$\rho_T(P_M, P_N) \leq (\frac{R(T)}{\mu} m_T(H, \mu I)^{1/2} + \frac{R(T)}{\mu} m_T(H, \mu I) + 2R(T) \sum_{s \leq T} (\Delta H_s)^2. \tag{4.9}$$

Note that if G is continuous, then the bound involves only the first two terms on the right hand side of (4.8).

We show how to get (4.8) from (2.1). According to Bremaud and Jacod (1977) the compensator A of M with respect the filtration F_t^M has the form

$$A_T = A_{T_n} + H(t - T_n) \ , \ \text{ if } \ T_n < t \leq T_{n+1}. \tag{4.10}$$

Note also that we have $B_t^{-1} = \frac{t}{\mu}$ and because $t = T_n - t - T_n$ we have that

$$|A_T - B_T| \leq (M_T + 1) m_T(H, \mu I) \text{ and } m_T(B^{-1} \circ A, I) \leq \frac{(M_T + 1)}{\mu} m_T(A, \mu I). \tag{4.11}$$

We have also, if the distribution G is discrete, that

$$\sum_{s \leq T} (\Delta A_s)^2 \leq (M_T + 1) \sum_{s \leq T} (\Delta H_s)^2. \tag{4.12}$$

From (4.10) and (4.11) we have (4.8) using the properties of the metric α mentioned at the point 2.2.

To finish this example suppose that F is an exponential distribution with parameter μ and $G^n \to^d F$, as $n \to \infty$, where $\to^d$ means the weak convergence of distribution functions. Denote by M^n the corresponding renewal counting process sequence. We can use then (4.8) to get some information about the rate of convergence of M^n to a Poisson limit.

References

1. P. Billingsley (1968). *Convergence of probability measures,* Wiley, New York.
2. P. Bremaud and J. Jacod (1977). *Processus ponctueles et martingales: resultants recentes sur la modelisation et le filtrage,* Adv. Appl. Probab., 9, 362-416.
3. T.C. Brown (1978). *A martingale approach to the Poisson convergence of simple point processes,* Ann. Probab., 6, 615-618.
4. T.C. Brown (1981). *Compensators and Cox convergence,* Math. Proc. Cambridge Phil. Soc., 90, 305-319.
5. T.C. Brown (1983). *Some Poisson approximations using compensators,* Ann. Probab., 11, 726-744.
6. R.M. Dudley (1968). *Distances of probability metrics and random variables,* Ann. Math. Statist., 39, 1563-1572.
7. R.M. Dudley (1972). *Speeds of metric probability convergence,* Z. Wahrsch. Theor., 22, 323-332.

8. J. Jacod (1979). *Calcul stochastique et problemes de martingales,* Lect. Notes Math. 714, Springer-Verlag, New York.

9. Yu.M. Kabanov, R.Sh. Liptser and A.N. Shiryayev (1980). *Some limit theorems for simple point processes (a martingale approach),* Stochastics, 3, 203-218.

10. Yu.M. Kabanov, R. Sh. Liptser and A.N. Shiryayev (1983). *Weak and strong convergence of distributions of counting processes,* Theor. Probab. Appl., 28, 303-335.

11. Yu.M. Kabanov, R.Sh. Liptser and A.N. Shiryayev (1986). *On the variation distance of probability measures defined on a filtered space,* Prob. Theor. Rel. Fields, 71, 19-36.

12. K. Kubilius and R. Mikulevicius (1986). *On the rate of convergence of distributions of semimartingales,* in: *Proceedings of the IV Vilnius Conference of Probability Theory and Mathematical Statistics* (to appear).

13. R.Sh. Liptser and A.N. Shiryayev (1978). *Statistics of Random Processes,* Vol. II, Springer-Verlag, New York.

14. M. Nikunen and E. Valkeila (1985). *On metric distances between counting processes, 377-388,* in: N.V. Krylov, R.Sh. Liptser and A.A. Novikov (ed). *Statistics and control of random processes,* Optimization Software Inc., Springer-Verlag, New York.

15. R.J. Serfling (1978). *Some elementary results on Poisson approximation in a sequence of Bernoulli trials,* Siam Review, 20, 567-579.

16. E. Valkeila and L. Vostrikova (1986). *An integral representation for the Hellinger distance,* Math. Scand., to appear.

17. W. Whitt (1973). *On the quality of Poisson approximations,* Z. Wahrsch. Theor., 28, 23-36.