Probab. Th. Rel. Fields 71, 19-35 (1986)

Probability Theory and Related Fields

On the Variation Distance for Probability Measures Defined on a Filtered Space

Yu.M. Kabanov[1], R.Sh. Liptser[2], and A.N. Shiryaev[3]

[1] Central Economical-Mathematical Institute, Krasikova 32, Moscow, USSR
[2] Institute of Control Science, Profsojuznaja 65, Moscow, USSR
[3] Mathematical Institute of Academy of Science, Vavilova 42, Moscow, USSR

Summary. We study the distance in variation between probability measures defined on a measurable space $(\Omega, \mathscr{F})$ with right-continuous filtration $(\mathscr{F}_t)_{t \leqq 0}$. To every pair of probability measures P and $\tilde{P}$ an increasing predictable process $h = h(P, \tilde{P})$ (called the Hellinger process) is associated. For the variation distance $\|P_T - \tilde{P}_T\|$ between the restrictions of P and $\tilde{P}$ to $\mathscr{F}_T$ (T is a stopping time), lower and upper bounds are obtained in terms of h. For example, in the case when $P_0 = \tilde{P}_0$,

$$2(1-(E\exp(-h_T))^{1/2}) \leqq \|P_T - \tilde{P}_T\| \leqq 4(Eh_T)^{1/2}.$$

In the cases where P and $\tilde{P}$ are distributions of multivariate point processes, diffusion-type processes or semimartingales h are expressed explicitly in terms of given predictable characteristics.

1. Introduction

In the present paper we study the proximity in variation of probability measures defined on a measurable space $(\Omega, \mathscr{F})$ with a right-continuous filtration $F = (\mathscr{F}_t)_{t \geqq 0}$. With every pair of probability measures P and $\tilde{P}$ we associate a predictable increasing process $h = h(P, \tilde{P})$ called the Hellinger process, [14], [17]. For the variation distance $\|P_T - \tilde{P}_T\|$ between the restrictions of P and $\tilde{P}$ to the σ-algebra $\mathscr{F}_T$ where T is a stopping time, lower and upper bounds are obtained in terms of h. For example, in the case when $P_0 = \tilde{P}_0$,

$$2(1-\sqrt{E\exp(-h_T)}) \leqq \|P_T - \tilde{P}_T\| \leqq 4\sqrt{Eh_T}$$

(see Corollary 2.1 and Theorems 2.1 and 2.2).

The criterion for strong convergence is a corollary of the above estimates. It asserts (see Theorem 2.3) that P_T^n converges to $P_T (T \in R_+)$ in variation if and only if the random variables $h_T(P^n, P)$ converge to zero in probability P.

In the case where P and $\tilde{P}$ are the distributions of multivariate point processes, diffusion-type processes, processes with independent increments or, more generally, semimartingales, the Hellinger process h can be expressed explicitly in terms of the given predictable characteristics, (see Sect. 6).

Thus for the cases mentioned above, the general results lead to efficient estimates and criteria. For example, if P and $\tilde{P}$ are the distributions of counting processes with continuous compensators A and $\tilde{A}$, then $h_t^{1/2}$ is the Hellinger distance between A and $\tilde{A}$ on $[0, t]$ i.e.

$$h_t = \tfrac{1}{2} \int_{[0,t]} (\sqrt{dA_s} - \sqrt{d\tilde{A}_s})^2 \overset{\text{def}}{=} \tfrac{1}{2} \int_{[0,t]} \left(\sqrt{\frac{dA_s}{dC_s}} - \sqrt{\frac{d\tilde{A}_s}{dC_s}}\right)^2 dC_s.$$

where $C = A + \tilde{A}$.

The results presented here have been announced in our note [6].

Our setting includes the case of "discrete time" when an increasing sequence of σ-algebras is given on $(\Omega, \mathscr{F})$. In this case a version of Theorem 2.3 has been established by L. Vostrikova [20]. It turns out that an estimate of the Hellinger integral of order α (Theorem 4.1) is useful to derive the lower bound on $\|P_T - \tilde{P}_T\|$. We adopted this idea from F. Liese who has obtained two-side inequalities for several particular cases, namely for counting processes with continuous compensators, Gaussian processes with independent increments and for diffusion-type processes, [9]–[11]. For arbitrary processes with independent increments two-side bounds in terms of the Hellinger process (which is deterministic in this case) are given in [17].

The inequalities for the variation distance between measures defined on a filtered space, different from ours but also based on the Hellinger process are obtained in [19].

It is possible to give upper bounds for $\|P_T - \tilde{P}_T\|$ in the general setting (but with some additional assumptions) in terms of another increasing process V (see our note [6]). These bounds imply the known estimates for multivariate point processes and counting processes [4], [7].

Note that the results of the present paper are closely connected with the problems of absolute continuity and contiguity of probability measures where the Hellinger process plays the fundamental role [5], [12], [14].

2. Main Results

1. At first, we describe the general setting for the considered problem. All necessary preliminaries can be found in [2], [5].

Let $(\Omega, \mathscr{F})$ be a measurable space with a right-continuous filtration $F = (\mathscr{F}_t)_{t \geq 0}$, $\mathscr{F} = \bigvee_{t \geq 0} \mathscr{F}_t$, and probability measures P, $\tilde{P}$ and $Q = (P + \tilde{P})/2$. For convenience of formulations we assume that the space $(\Omega, \mathscr{F}, Q)$ is complete and $\mathscr{F}_0$ contains all Q-null sets from $\mathscr{F}$ (i.e. the usual conditions with respect to Q are satisfied). We identify Q-indistinguishable processes. The relations between random variables are understood Q-almost surely. Expectations with respect to the probability measures P, $\tilde{P}$, Q, ... are denoted by E, $\tilde{E}$, E_Q, ...; $\mathscr{T}$

(correspondingly $\mathscr{T}_p$) is the set of all stopping times (correspondingly, the set of the predictable stopping times) with respect to F. We denote by P_T the restriction of P to the σ-algebra $\mathscr{F}_T$.

The variation distance between P_T and $\tilde{P}_T$ is defined by the following formula

$$\|P_T-\tilde{P}_T\|=2\sup_{A\in\mathscr{F}_T}|P(A)-\tilde{P}(A)|.$$

Let $\varkappa=(\varkappa_t)_{t\geqq 0}$ be the density process of the measure P with respect to Q. It means that $\varkappa$ is the process with sample paths from the space of càdlàg functions D, and for any $T\in\mathscr{T}$

$$\varkappa_T=dP_T/dQ_T.$$

The process $\varkappa$ is a nonnegative bounded Q-martingale: $0\leqq\varkappa\leqq 2$. Let κ be the jump measure of $\varkappa$, η be the Q-compensator (dual predictable projection) of κ. It is well known, [2], that $\varkappa$ can be represented in the following way:

$$\varkappa_t=\varkappa_0+\varkappa_t^c+x*(\kappa-\eta)_t$$

where $\varkappa^c$ is a continuous Q-martingale starting from zero[1].

Similarly we introduce the density process $\tilde{\varkappa}$ of the probability measure $\tilde{P}$ with respect to Q and associated objects $\tilde{\kappa}$, $\tilde{\eta}$, $\tilde{\varkappa}^c$.

Note that $\varkappa+\tilde{\varkappa}=2$ and consequently $\Delta\varkappa=-\Delta\tilde{\varkappa}$. Recall that for a random process X with sample paths from D the following notations are used: $X_-=(X_{t-})_{t\geqq 0}$ where $X_{0-}=X_0$, $\Delta X=X-X_-$; if $T\in\mathscr{T}$ then $X^T=(X_{t\wedge T})_{t\geqq 0}$.

2. If R is an arbitrary probability measure dominating P as well as $\tilde{P}$ (i.e. $P\ll R$ and $\tilde{P}\ll R$) then the quantity

$$H(P,\tilde{P})=E_R(dP/dR)^{1/2}(d\tilde{P}/dR)^{1/2}$$

does not depend on R. It is called the Hellinger integral for P and $\tilde{P}$.

In particular,

$$H(P_T,\tilde{P}_T)=E_Q\,\varkappa_T^{1/2}\,\tilde{\varkappa}_T^{1/2}.$$

We introduce an increasing process $B=B(P,\tilde{P})$ with

$$B_t=(1/2)\,((\varkappa_-\tilde{\varkappa}_-)^{-2}\circ\langle\varkappa^c\rangle_t+(\sqrt{1+x/\varkappa_-}-\sqrt{1-x/\tilde{\varkappa}_-})^2*\eta_t), \tag{2.1}$$

with $0/0=0$. Note that $\Delta B\leqq 1$.

Let

$$\sigma_n=\inf(t\colon \varkappa_t\wedge\tilde{\varkappa}_t\leqq 1/n),\qquad \sigma=\inf(t\colon\varkappa_t\wedge\tilde{\varkappa}_t=0), \tag{2.2}$$

$$\Gamma=\bigcup_{n\geqq 1}[\![0,\sigma_n]\!].$$

Definition. We call a predictable increasing right-continuous process h with $\Delta h\leqq 1$ the Hellinger process if

$$I_\Gamma\circ B=I_\Gamma\circ h$$

(i.e. the processes $I_\Gamma\circ B$ and $I_\Gamma\circ h$ are Q-indistinguishable).

---

[1] Throughout the sequel the symbols $\circ$, $\bullet$, $*$ denote, correspondingly, integrals with respect to a process with bounded variation, a semimartingale and a random measure

For the Hellinger process h the Doleàns-Dade exponential is defined by

$$\mathscr{E}_t(-h) = \exp(-h_t) \prod_{s \leqq t} (1 - \Delta h_s)\, e^{\Delta h_s}. \tag{2.3}$$

Evidently, $\mathscr{E}(-h)$ is a nonnegative decreasing process with $\mathscr{E}_0(-h) = 1$.

Put

$$d_0 = \|P_0 - \tilde{P}_0\|, \qquad H_T = H(P_T, \tilde{P}_T).$$

Theorem 2.1. *For any* $T \in \mathscr{T}$, $\varepsilon \in]0, 1[$ *we have*

$$\|P_T - \tilde{P}_T\| \geqq 2(1 - \sqrt{H_0 E \mathscr{E}_T(-h)}), \tag{2.4}$$

$$\|P_T - \tilde{P}_T\| \leqq 3\sqrt{2}\sqrt{1 - \varepsilon H_0} + 2P(\mathscr{E}_T(-h) \leqq \varepsilon). \tag{2.5}$$

Since $\mathscr{E}_T(-h) \leqq \exp(-h_T)$ and $H_0 \leqq 1 - d_0^2/8$ (see (3.1)), (2.4) implies

Corollary 2.1. *For any* $T \in \mathscr{T}$

$$\|P_T - \tilde{P}_T\| \geqq 2(1 - \sqrt{(1 - d_0^2/8)\, E \exp(-h_T)}).$$

We give two other bounds for the variation distance between P_T and $\tilde{P}_T$ which, similarly to the latter estimate, are based only on the value of h in the moment T.

Theorem 2.2. *Let* $T \in \mathscr{T}$, $\varepsilon > 0$. *Then*

$$\|P_T - \tilde{P}_T\| \leqq d_0 + 4\sqrt{E h_T}, \tag{2.6}$$

$$\|P_T - \tilde{P}_T\| \leqq (3/2)\, d_0 + 3\sqrt{2\varepsilon} + 2P(h_T \geqq \varepsilon). \tag{2.7}$$

Remark 2.1. Since $B(P, \tilde{P}) = B(\tilde{P}, P)$, Theorems 2.1 and 2.2 and Corollary 2.1 are valid after replacing P and E by $\tilde{P}$ and $\tilde{E}$.

3. Let P^n, $\tilde{P}^n$ be probability measures on a filtered space $(\Omega^n, \mathscr{F}^n, F^n = (\mathscr{F}_t^n)_{t \geqq 0})$ satisfying the usual conditions for $Q^n = (P^n + \tilde{P}^n)/2$. Put $h^n = h(P^n, \tilde{P}^n)$. Then Corollary 2.1 and (2.7) imply the following criterion for strong convergence of probability measures:

Theorem 2.3. *Let* $T \in R_+$. *Then the following conditions are equivalent:*

a) $\lim_n \|P_T^n - \tilde{P}_T^n\| = 0$,

b) $\lim_n \|P_0^n - \tilde{P}_0^n\| = 0$, $\lim_n P^n(h_T^n \geqq \varepsilon) = 0$, $\forall \varepsilon > 0$.

3. Hellinger Integrals and Hellinger Processes

1. For the proof of the lower bound for the variation distance, we need the concept of the Hellinger integral and the Hellinger process of order α. Their properties are studied in this section.

Let R be a probability measure dominating P and $\tilde{P}$. The Hellinger integral of order α with $\alpha\in]0,1[$ is defined by

$$H(\alpha, P, \tilde{P}) = E_R (dP/dR)^\alpha (d\tilde{P}/dR)^{1-\alpha}.$$

The value of $H(\alpha, P, \tilde{P})$ does not depend on the choice of R. Clearly,

$$0 \leqq H(\alpha, P, \tilde{P}) \leqq 1.$$

In agreement with the previous notation we omit the symbol $\alpha = 1/2$ in the sequel.

Let us recall some elementary inequalities connecting the Hellinger integral with the variation distance.

Lemma 3.1 (see [8], [16], [18]). *For any* $\alpha\in]0,1[$

$$2(1 - H(\alpha, P, \tilde{P})) \leqq \|P - \tilde{P}\| \leqq \sqrt{C_\alpha (1 - H(\alpha, P, \tilde{P}))} \tag{3.1}$$

where C_α *is a constant; for* $\alpha = 1/2$ *it is possible to choose* $C_\alpha = 8$.

2. Lemma 3.1 shows that estimates for $\|P_T - \tilde{P}_T\|$ can be derived from estimates for the Hellinger integrals

$$H_T(\alpha) \stackrel{\text{def}}{=} H(\alpha, P_T, \tilde{P}_T) = E_Q\, \varkappa_T^\alpha\, \tilde{\varkappa}_T^{1-\alpha}.$$

Thus, we must study the structure of the process

$$X(\alpha) = \varkappa^\alpha\, \tilde{\varkappa}^{1-\alpha}$$

once its additive and then its multiplicative decompositions have been constructed. The latter plays a crucial role in the proof of the inequalities of Theorem 2.1.

The function $u^\alpha v^{1-\alpha}$ (known in mathematical economics as the Cobb-Douglas function) is concave. So the nonnegative bounded process $X(\alpha) (\leqq 2)$ is a Q-supermartingale (recall that $\varkappa$ and $\tilde{\varkappa}$ are Q-martingales). Its Doob-Meyer decomposition has the form

$$X_t(\alpha) = X_0(\alpha) - A_t(\alpha) + M_t(\alpha) \tag{3.2}$$

where $A(\alpha)$ is a predictable increasing process, $M(\alpha)$ is a Q-martingale, $A_0(\alpha) = M_0(\alpha) = 0$.

It is easy to see that $E_Q A_\infty(\alpha) \leqq E_Q X_0(\alpha) \leqq 2$ and so

$$E_Q \sup_t |M_t(\alpha)| \leqq E_Q \sup_t X_t(\alpha) + E_Q A_\infty(\alpha) \leqq 4.$$

Let σ_n and σ be defined by (2.2). By a well-known property of nonnegative supermartingales (see, e.g. [2, (6.20)]) we have $\sigma = \lim \sigma_n$ and $X(\alpha) = X^\sigma(\alpha)$ on the set $(\sigma < \infty)$. Therefore the uniqueness of the Doob-Meyer decomposition implies, in particular, that $A^\sigma(\alpha) = A(\alpha)$ (and $M^\sigma(\alpha) = M(\alpha)$).

We note also the following property of $A(\alpha)$:

$$\Delta A_\sigma(\alpha) = 0 \quad \text{on} \quad (X_{\sigma-}(\alpha) = 0). \tag{3.3}$$

Indeed, by [1, V-T-10] for any $T \in \mathcal{T}_P$ we have

$$E(I_{(T<\infty)} \Delta M_T(\alpha) \mid \mathcal{F}_{T-}) = 0. \tag{3.4}$$

The random variables $\Delta A_T(\alpha)$ and $X_{T-}(\alpha)$ are $\mathcal{F}_{T-}$-measurable. Hence (3.2) implies that for any $T \in \mathcal{T}_P$

$$I_{(T<\infty, X_{T-}(\alpha)=0)} \Delta A_T(\alpha) = E_Q(\Delta X_T(\alpha) I_{(T<\infty, X_{T-}(\alpha)=0)} \mid \mathcal{F}_{T-}). \tag{3.5}$$

But the set $(\Delta A(\alpha)=0)$ is contained in the most countable union of the graphs of predictable stopping times. Thus (3.3) follows from (3.5).

For $\alpha \in]0, 1[$ put

$$\varphi_\alpha(u, v) = \alpha u + (1-\alpha) v - u^\alpha v^{1-\alpha}. \tag{3.6}$$

Evidently, $\varphi_\alpha(u, v)$ is a nonnegative convex function. We introduce the predictable increasing process (compare with (2.1))

$$B(\alpha) = 2\alpha(1-\alpha)(\varkappa_- \tilde{\varkappa}_-)^{-2} \circ \langle \varkappa^c \rangle + \varphi_\alpha(1 + x/\varkappa_-, 1 - x/\tilde{\varkappa}_-) * \eta. \tag{3.7}$$

Note that $\Delta B(\alpha) \leqq 1$.

Definition. We say that a predictable right-continuous process $h(\alpha)$ with $\Delta h(\alpha) \leqq 1$ is the Hellinger process of order α if

$$I_\Gamma \circ h(\alpha) = I_\Gamma \circ B(\alpha) \tag{3.8}$$

where $\Gamma = \bigcup_{n \geqq 1} [\![0, \sigma_n]\!]$ (see (2.2)).

Following Sect. 2 we omit the symbol $\alpha = 1/2$ in the sequel.

Lemma 3.2. *The process $A(\alpha)$ from the decomposition* (3.2) *can be represented as follows:*

$$A(\alpha) = X_-(\alpha) \circ h(\alpha) \tag{3.9}$$

where $h(\alpha)$ is any Hellinger process of order α.

The proof is given in [17].

Corollary 3.1. *Let $\theta = \inf(t\colon B_t(\alpha) = \infty$ or $\Delta B_t(\alpha) = 1)$. Then $\theta \geqq \sigma$.*

Proof. The representation

$$X_t(\alpha) = X_0(\alpha) - X_-(\alpha) \circ B(\alpha)_t + M_t(\alpha) \tag{3.10}$$

implies that $\inf(t\colon B_t(\alpha) = \infty) \geqq \sigma$ and $X_\theta(\alpha) = \Delta M_\theta(\alpha)$ on $(\Delta B_\theta(\alpha) = 1)$. Clearly, $\theta \in \mathcal{T}_P$ and $(\Delta B_\theta(\alpha) = 1)$ is a $\mathcal{F}_{\theta-}$-measurable set. Hence

$$E_Q(X_\theta(\alpha) I_{(\Delta B_\theta(\alpha)=1)} \mid \mathcal{F}_{\theta-}) = E_Q(\Delta M_\theta(\alpha) I_{(\Delta B_\theta(\alpha)=1)} \mid \mathcal{F}_{\theta-}) = 0$$

and, therefore, $X_\theta(\alpha) = 0$ on $(\Delta B_\theta(\alpha) = 1)$, i.e. $\theta \geqq \sigma$.

3. **Lemma 3.3.** *Let $h(\alpha)$ be a Hellinger process. Then*

$$X(\alpha) = \mathcal{E}(-h(\alpha))\, \mathcal{S}(\alpha) \tag{3.11}$$

where $\mathscr{E}(-h(\alpha))$ *is the Doleans-Dade exponential* (*see* (2.3)), $\mathscr{S}(\alpha)$ *is a nonnegative Q-supermartingale with the following properties:*

(i) *for any* $\tau \in \mathscr{T}$ *such that* $\mathscr{E}_\tau(-h(\alpha))>0$, *the process* $\mathscr{S}^\tau(\alpha)$ *is a local Q-martingale;*

(ii) *for any* $\tau \in \mathscr{T}$ *such that* $\mathscr{E}_\tau(-h(\alpha)) \geqq c>0$, *the process* $\mathscr{S}^\tau(\alpha)$ *is a bounded Q-martingale;*

(iii) *if* $\tilde{P} \overset{\text{loc}}{\sim} P$ (*i.e.* $\tilde{P}_t \sim P_t$, $\forall t \in R_+$) *then* $\mathscr{S}(\alpha)$ *is a local Q-martingale.*

Proof. Let $\mathscr{E}(-h(\alpha))$ be the Dolèan-Dade exponential. Define

$$\mathscr{S}(\alpha)=X(\alpha)\mathscr{E}^{-1}(-h(\alpha))$$

setting $0/0=0$. The identity (3.11) holds by virtue of Corollary 3.1. Since $\Delta h(\alpha) \leqq 1$, $\mathscr{S}(\alpha)$ is a nonnegative process.

Let $\tau \in \mathscr{T}$ and $\mathscr{E}_\tau(-h(\alpha))>0$. Then the process $z(\alpha)=\mathscr{E}^{-1}(-h(\alpha))$ on the set $[\![0, \tau]\!]$ satisfies the equation

$$z(\alpha)=1+z_-(\alpha)(1-\Delta h(\alpha))^{-1} \circ h(\alpha). \tag{3.12}$$

Using the Itô formula for a product and relations (3.10), (3.12) we have

$$\mathscr{S}_{\tau \wedge t}(\alpha)=X_0(\alpha)+z(\alpha) \cdot M(\alpha)_{\tau \wedge t}.$$

This representation implies (i). Properties (i) and (iii) follow from (i).

Since the stopping time θ (see Corollary 3.1) is predictable, there is a sequence of $\theta_k \in \mathscr{T}$, $k \geqq 1$, announcing θ. The nonnegative local Q-martingale $\mathscr{S}^{\theta_k}(\alpha)$ is a Q-supermartingale. The Fatou lemma implies that the nonnegative process

$$\bar{\mathscr{S}}(\alpha)=\mathscr{S}(\alpha)\, I_{[\![0,\theta[\![}+\mathscr{S}_{\theta-}(\alpha)\, I_{[\![\theta,\infty[\![}$$

is a Q-supermartingale. But $\mathscr{S}(\alpha)=\bar{\mathscr{S}}(\alpha)\, I_{[\![0,\theta[\![}$. Thus $\mathscr{S}(\alpha)$ is also a Q-supermartingale. The lemma is proved.

Theorem 3.1. *Let* $0<\alpha<\beta<1$, $p=(1-\alpha)/(1-\beta)$, $q=(1-\alpha)/(\beta-\alpha)$, $T \in \mathscr{T}$ *and* $h(\alpha)$ *be a Hellinger process of order* α.

Then

$$H_T(\beta) \leqq H_0^{1/p}(\alpha)(E\mathscr{E}_T^{q/p}(-h(\alpha)))^{1/q}. \tag{3.13}$$

Proof. From the obvious identity $X_T(\beta)=X_T^{1/p}(\alpha)\, z_T^{1/q}$ and the multiplicative decomposition of $X(\alpha)$ (Lemma 3.3) we get the following representation

$$X_T(\beta)=\mathscr{S}_T^{1/p}(\alpha)(z_T\, \mathscr{E}_T^{q/p}(-h(\alpha)))^{1/q} \tag{3.14}$$

where $\mathscr{S}(\alpha)$ is a Q-supermartingale with $\mathscr{S}_0(\alpha)=X_0(\alpha)$. Taking the Q-expectation of both sides of (3.14) and applying the Hölder inequality, we obtain the estimate

$$H_T(\beta) \leqq (E_Q\, \mathscr{S}_T(\alpha))^{1/p}(E_Q\, z_T\, \mathscr{E}_T^{q/p}(-h(\alpha))^{1/q}.$$

This implies (3.2) because $E_Q\, S_T(\alpha) \leqq E_Q\, S_0(\alpha)=H_0(\alpha)$.

Corollary 3.2. *The following inequality is valid:*

$$H_T(3/4) \leqq \sqrt{H_0\, E\mathscr{E}_T(-h)}. \tag{3.15}$$

For the proof it is sufficient to put in (3.13) $\alpha=1/2$, $\beta=3/4$.

Remark 3.1. Similarly to (3.13), it can be shown that

$$H_T(\alpha) \leqq H_0^{1/r}(\beta)(\tilde{E}\mathscr{E}_T^{1/r}(-h(\beta)))^{1/l}$$

where $r=\beta/\alpha$, $l=\beta/(\beta-\alpha)$.

4. At the end of this section we give some results allowing us to use any probability measure R dominating P and $\tilde{P}$ instead of Q in order to calculate the Hellinger process.

Let $z^R(\tilde{z}^R)$ be the density process of the measure $P(\tilde{P})$ with respect to R. The nonnegative R-supermartingale $X^R(\alpha)=(z^R)^\alpha(\tilde{z}^R)^{1-\alpha}$ has the Doob-Meyer decomposition

$$X^R(\alpha)=X_0^R(\alpha)-A^R(\alpha)+M^R(\alpha)$$

where $A^R(\alpha)$ is a predictable increasing process, $M^R(\alpha)$ is a R-martingale.

Lemma 3.4. *The following representation holds:*

$$A^R(\alpha)=X_-^R(\alpha)\circ B(\alpha) \qquad (R\text{-a.s.})$$

Proof. We need to prove that $\tilde{M}(\alpha)=X^R(\alpha)-X_0^R(\alpha)-X_-^R(\alpha)\circ B(\alpha)$ is a local R-martingale. In accordance with [2, (1.20)] it is sufficient to show that for any $T\in\mathscr{T}$ and $t\in R_+$

$$E_R\tilde{M}_{t\wedge T}(\alpha)=0.$$

Let ρ be the local density of Q with respect to R. Then $X^R(\alpha)=\rho X(\alpha)$. Making use of Lemma 3.2 and [2, (1.47)], we have

$$\begin{aligned}E_R\tilde{M}_{t\wedge T}(\alpha)&=E_R(X_{t\wedge T}^R(\alpha)-X_0^R(\alpha)-X_-^R(\alpha)\circ B(\alpha)_{t\wedge T})\\&=E_R\,\rho(X_{t\wedge T}(\alpha)-X_0(\alpha)-X_-(\alpha)\circ B(\alpha)_{t\wedge T})=E_Q\,M_{t\wedge T}(\alpha)=0.\end{aligned}$$

The Lemma is proved.

Corollary 3.3. *Let*

$$h^R(\alpha)=(X_-^R(\alpha))^{-1}\circ A^R(\alpha) \qquad (Q\text{-a.s.}).$$

Then $h^R(\alpha)$ is a Hellinger process of order α.

Indeed, by Lemma 3.4 we have

$$I_\Gamma I_{(\rho_->0)}(X_-^R(\alpha))^{-1}\circ A^R(\alpha)=I_\Gamma I_{(\rho_->0)}\circ B(\alpha) \qquad (Q\text{-a.s.})$$

and one needs to note only that $Q(\inf_t \rho_t>0)=1$, [5, Lemma 3].

Let $z^{R,c}$ be the continuous martingale part of z^R, κ^R be the jump measure of z^R, η^R be the R-compensator of κ^R. The notations $\tilde{z}^{R,c}$, $\tilde{\kappa}^R$, $\tilde{\eta}^R$ are evident.

Let $\sigma_n^R=\inf(t\colon z_t^R\wedge\tilde{z}_t^R\leqq 1/n)$, $\sigma^R=\inf(t\colon z_t^R\wedge\tilde{z}_t^R=0)$. Put

$$M_t^{R,n}=(z_-^R)^{-1}\cdot z_{t\wedge\sigma_n^R}^{R,c}, \qquad \tilde{M}_t^{R,n}=(\tilde{z}_-^R)^{-1}\cdot\tilde{z}_{t\wedge\sigma_n^R}^{R,c},$$

$$g_t^{R,n}=\sum_{s\leqq t\wedge\sigma_n^R}\varphi_\alpha(z_s^R/z_{s-}^R,\ \tilde{z}_s^R/\tilde{z}_{s-}^R).$$

It is easy to check that

$$\varphi_\alpha(u, v) \leqq K(\alpha)(|u-v| \wedge |u-v|^2)$$

where $K(\alpha)$ is a constant depending on α only.

It follows that

$$\Delta g^{R,n}(\alpha) \leqq 2K(\alpha)\, n^2((|\Delta \varkappa^R| + |\Delta \tilde{\varkappa}^R|) \wedge (|\Delta \varkappa^R|^2 + |\Delta \tilde{\varkappa}^R|^2).$$

The processes $\varkappa^R$ and $\tilde{\varkappa}^R$ are uniformly integrable R-martingales. Thus, the last estimate implies that $g^{R,n}(\alpha)$ is a right-continuous locally R-integrable process (in fact, $E_R\, g^{R,n}_{t \wedge T_k}(\alpha) < \infty$ for $T_k = \inf(t\colon \varkappa^R_t \wedge \tilde{\varkappa}^R_t \wedge g^{R,n}_t \geqq k)$).

Denote the R-compensator of $g^{R,n}(\alpha)$ by $G^{R,n}(\alpha)$.

Theorem 3.2. *Let $h^R(\alpha)$ be a right-continuous predictable increasing process with $\Delta h^R(\alpha) \leqq 1$ such that for any n*

$$h^R(\alpha) = (1/2)\, \alpha(1-\alpha) \langle M^{R,n} - \tilde{M}^{R,n} \rangle + G^{R,n}(\alpha) \quad \text{on} \quad [\![0, \sigma^R_n]\!]$$

(P- and $\tilde{P}$-a.s.).

Then $h^R(\alpha)$ is the Hellinger process of order α for P and $\tilde{\mathrm{P}}$.

Proof. It is easy to see that we only need to establish the equality $B(\alpha) = h^R(\alpha)$ on the sets $[\![0, \sigma^R_n \wedge \sigma_n]\!]$, $n \geqq 1$. For this reason as well as for notational convenience we shall consider all the processes as stopped at $\sigma^R_n \wedge \sigma_n$ and omit the superscript $\sigma^R_n \wedge \sigma_n$.

At first we prove that

$$\langle M^R - \tilde{M}^R \rangle = \langle M - \tilde{M} \rangle. \tag{3.16}$$

(the symbol Q is omitted).

Since $Q \ll R$, a R-semimartingale Y is a Q-semimartingale and the "brackets" $\langle Y^c \rangle$ of the continuous martingale parts of Y with respect to R and Q are Q-indistinguishable (see [2, Chap. 7]). The same holds for the stochastic integrals $H \cdot Y$.

By the above remark, the process $\langle M^R - \tilde{M}^R \rangle$ which is the "bracket" of the continuous part of R-semimartingales $z = (\varkappa^R_-)^{-1} \cdot \varkappa^R - (\tilde{\varkappa}^R_-)^{-1} \cdot \tilde{\varkappa}^R$ can be calculated as the "bracket" of the continuous part of the Q-semimartingale z. The continuous martingale parts of $\varkappa^R = \varkappa\rho$ and $\tilde{\varkappa}^R = \tilde{\varkappa}\rho$ with respect to Q are $\varkappa_- \cdot \rho^c + \rho_- \cdot \varkappa^c$ and $\tilde{\varkappa}_- \cdot \rho^c + \rho_- \cdot \tilde{\varkappa}^c$, correspondingly, where ρ^c is the continuous Q-martingale part of ρ. It follows that the continuous Q-martingale part of z with respect to Q has the form

$$\begin{aligned} z^c &= (\varkappa^R_-)^{-1} \varkappa_- \cdot \rho^c + (\varkappa^R_-)^{-1} \rho \cdot \varkappa^c - (\tilde{\varkappa}^R_-)^{-1} \tilde{\varkappa}_- \cdot \rho^c - (\tilde{\varkappa}^R_-)^{-1} \rho_- \cdot \tilde{\varkappa}^c \\ &= \varkappa_-^{-1} \cdot \varkappa^c - \tilde{\varkappa}_-^{-1} \cdot \tilde{\varkappa}^c \end{aligned}$$

and (3.16) is established.

Now we prove the equality

$$G^R(\alpha) = G(\alpha). \tag{3.17}$$

Due to the uniqueness of the compensator we only need to show that

$$E_Q f \circ G^R(\alpha)_\infty = E_Q f \circ G(\alpha)_\infty$$

for any positive predictable process f.

Note that $g(\alpha) = \rho_- \rho^{-1} I_{(\rho>0)} \circ g^R(\alpha)$ (Q-a.s.) and $I_{(\rho=0)} \circ g^R(\alpha) = 0$. By [2, (1.47)], omitting the symbol α for brevity, we have

$$\begin{aligned} E_Q f \circ G_\infty &= E_Q f \circ g_\infty = E_Q f \rho_- \rho^{-1} I_{(\rho>0)} \circ g^R_\infty \\ &= E_R \rho_\infty (f \rho_- \rho^{-1} I_{(\rho>0)} \circ g^R_\infty) = E_R f \rho_- \rho \rho^{-1} I_{(\rho>0)} \circ g^R_\infty \\ &= E_R f \rho_- I_{(\rho>0)} \circ g^R_\infty = E_R f \rho_- \circ g^R_\infty \\ &= E_R f \rho_- \circ G^R_\infty = E_R \rho_\infty (f \circ G^R_\infty) = E_Q f \circ G^R_\infty . \end{aligned}$$

Thus, (3.17) holds.

The assertion follows from (3.16) and (3.17) because

$$(1/2)\, \alpha(1-\alpha) \langle M - \tilde{M} \rangle + G(\alpha) = B(\alpha).$$

4. Preliminary Lemmas

This section contains some technical results.

1. **Lemma 4.1.** *For any* $T \in \mathscr{T}$

$$\|P_T - \tilde{P}_T\| \leqq d_0 + 2\sqrt{E_Q \langle \varkappa \rangle_T}$$

where $d_0 = \|P_0 - \tilde{P}_0\|$.

Proof. Using the identity $\varkappa + \tilde{\varkappa} = 2$ we have

$$|\varkappa_T - \tilde{\varkappa}_T| \leqq |\varkappa_0 - \tilde{\varkappa}_0| + |(\varkappa_T - \varkappa_0) - (\tilde{\varkappa}_T - \tilde{\varkappa}_0)| = |\varkappa_0 - \tilde{\varkappa}_0| + 2|\varkappa_T - \varkappa_0|.$$

Thus,

$$\begin{aligned} \|P_T - \tilde{P}_T\| &= E_Q |\varkappa_T - \tilde{\varkappa}_T| \leqq d_0 + 2E_Q |\varkappa_T - \varkappa_0| \\ &\leqq d_0 + 2\sqrt{E_Q(\varkappa_T - \varkappa_0)^2} = d_0 + 2\sqrt{E_Q \langle \varkappa \rangle_T}. \end{aligned}$$

Remark 4.1. It is possible to derive the following lower bound for the variation distance between P_T and $\tilde{P}_T$:

$$\|P_T - \tilde{P}_T\| \geqq 2(1 - \sqrt{1 - d_0^2/4 - E_Q \langle \varkappa \rangle_T}).$$

Indeed,

$$\|P_T - \tilde{P}_T\| \geqq E_Q (\sqrt{\varkappa_T} - \sqrt{\tilde{\varkappa}_T})^2 = 2(1 - \sqrt{E_Q \varkappa_T \tilde{\varkappa}_T})$$

and the result follows from the relations

$$E_Q \varkappa_T \tilde{\varkappa}_T = 2 - E_Q \varkappa_0^2 - E_Q \langle \varkappa \rangle_T \quad \text{and} \quad d_0^2 \leqq 4(E_Q \varkappa_0^2 - 1).$$

2. Recall the following

Definition. An increasing process A^1 strictly dominates an increasing process A^2 if $A_0^1 \geqq A_0^2$ and the difference forms an increasing process.

Lemma 4.2. *Both processes $2B$ and $4\varkappa_- \circ B$ strictly dominate the process $\langle \varkappa \rangle$.*

Proof. For $\varkappa_- > 0$, $\tilde{\varkappa}_- > 0$ and $-\varkappa_- \leqq x \leqq \tilde{\varkappa}_-$ put

$$W(\varkappa_-, \tilde{\varkappa}_-, x) = \frac{4x^2}{\varkappa_- \tilde{\varkappa}_- (\sqrt{\tilde{\varkappa}_-(\varkappa_- + x)} + \sqrt{\varkappa_-(\tilde{\varkappa}_- - x)})^2}$$

and note that

$$W(\varkappa_-, \tilde{\varkappa}_-, x) = (\sqrt{1 + x\varkappa_-^{-1}} - \sqrt{1 - x\tilde{\varkappa}_-^{-1}})^2.$$

From this and the definition of B we have

$$2B = (\varkappa_- \tilde{\varkappa}_-)^{-2} \circ \langle \varkappa^c \rangle + W(\varkappa_-, \tilde{\varkappa}_-, x) * \eta,$$

$$4\varkappa_- \circ B = 2\varkappa_- \tilde{\varkappa}_-^{-2} \circ \langle \varkappa^c \rangle + 2\varkappa_- W(\varkappa_-, \tilde{\varkappa}_-, x) * \eta.$$

Comparing these expressions with $\langle \varkappa \rangle = \langle \varkappa^c \rangle + x^2 * \eta$ and making use the inequalities $\varkappa_- \tilde{\varkappa}_- \leqq 1$, $\tilde{\varkappa}_- \leqq 2$,

$$\sqrt{\tilde{\varkappa}_-(\varkappa_- + x)} + \sqrt{\varkappa_-(\tilde{\varkappa}_- + x)} \leqq 2$$

we obtain easily the desired assertions.

5. Proofs of the Main Results

1. Proof of Theorem 2.1. The estimate (2.4) follows from the left inequality (3.1) with $\alpha = 3/4$ and (3.15) (see Corollary 3.2).

It is sufficient to prove the inequality (2.5) for finite T only. We use the multiplicative decomposition of the process $X = X(1/2)$ (Lemma 3.3, $\alpha = 1/2$). Put

$$\tau = \inf(t\colon \mathscr{E}_t(-h) \leqq \varepsilon), \qquad \varepsilon \in]0, 1[.$$

Let (τ_k) be a sequence annoucing the predictable stopping time τ. By Lemma 3.3

$$X_{\tau_k} = \mathscr{E}_{\tau_k}(-h)\, \mathscr{S}_{\tau_k} \geqq \varepsilon \mathscr{S}_{\tau_k}$$

and the process $\mathscr{S}^{\tau_k}$ is a bounded Q-martingale.

Thus,

$$H_{T \wedge \tau_k} = E_Q X_{T \wedge \tau_k} \geqq \varepsilon E_Q \mathscr{S}_{T \wedge \tau_k} = \varepsilon E_Q X_0 = \varepsilon H_0$$

and by virtue of the right inequality (3.1) with $\alpha = 1/2$ we have

$$\|P_{T \wedge \tau_k} - \tilde{P}_{T \wedge \tau_k}\| \leqq 2\sqrt{2}(1 - \varepsilon H_0)^{1/2}.$$

The above inequality and Lemma 4.3 imply that

$$\|P_T - \tilde{P}_T\| \leqq 3\sqrt{2}(1 - \varepsilon H_0)^{1/2} + 2P(\tau_k < T).$$

To prove (2.5) we note that for finite T we have

$$\lim_k P(\tau_k < T) \leqq P(\tau \leqq T) = P(\mathscr{E}_T(-h) \leqq \varepsilon).$$

2. *Proof of Theorem 2.2.* According to Lemma 4.2 and [2, Proposition (1.47)]

$$E_Q \langle \varkappa \rangle_T \leqq 4 E_Q \varkappa_- \circ B_T = 4 E_Q \varkappa_T B_T = 4 E B_T.$$

Hence (2.6) follows from Lemma 4.1 (note that any Hellinger process h strictly dominates B).

It is sufficient to prove the inequality (2.7) only for the finite stopping time T.

Put $\theta = \inf(t : \langle \varkappa \rangle_t \geqq 2\varepsilon)$. Then $\theta \in \mathscr{T}_P$. Let (θ_k) be an announcing sequence of stopping times for θ. For the finite T we have $P(\theta \leqq T) = P(\langle \varkappa \rangle_T \geqq 2\varepsilon)$. Lemma 4.2 implies that $P(\langle \varkappa_T \rangle \geqq 2\varepsilon) \leqq P(B_T \geqq \varepsilon)$. By virtue of Lemmas 4.3 and 4.1 we have

$$\begin{aligned} \|P_T - \tilde{P}_T\| &\leqq (3/2) \|P_{T \wedge \theta_k} - \tilde{P}_{T \wedge \theta_k}\| + 2P(\theta_k < T) \\ &\leqq (3/2) d_0 + 3\sqrt{2\varepsilon} + 2P(\theta_k < T). \end{aligned}$$

Since $\lim_k P(\theta_k < T) \leqq P(\theta \leqq T)$, (2.7) follows from this relations because $h_T \geqq B_T$.

6. Examples

1. We give an expression for the Hellinger process $h = h(P, \tilde{P})$ in the case when P and $\tilde{P}$ are the distributions of the (weak) solutions of the stochastic differential equations

$$\begin{aligned} dY_t &= a(t, Y)\, dt + dW_t, \qquad Y_0 = 0, \\ d\tilde{Y}_t &= \tilde{a}(t, \tilde{Y})\, dt + dW_t, \qquad \tilde{Y}_0 = 0 \end{aligned}$$

where a and $\tilde{a}$ are nonanticipating functionals on the space $(C, \mathscr{C})$ of continuous functions $x = (x_t)_{t \geqq 0}$. In other words, P (or $\tilde{P}$) is a measure on $(C, \mathscr{C})$ such that

$$x_t - \int_{[0,t]} a(s, x)\, ds \qquad \left(\text{or } x_t - \int_{[0,t]} \tilde{a}(s, x)\, ds\right)$$

is the Wiener process with respect to P (or $\tilde{P}$).

In considering the case $\Omega = C$, $F = (\mathscr{F}_t)_{t \geqq 0}$ is the minimal filtration satisfying the usual conditions with respect to Q and such that $\mathscr{F}_t \supseteq \mathscr{C}_t = \sigma(x_s, s \leqq t)$ for any t.

Theorem 5.1. *If for any* $t \in R_+$

$$\int_{[0,t]} a^2(s, x)\, ds < \infty \quad P\text{-a.s.}, \qquad \int_{[0,t]} \tilde{a}^2(s, x)\, ds < \infty \quad \tilde{P}\text{-a.s.}$$

then the process

$$h_t = (1/8) \int_{[0,t]} (a(s, x) - \tilde{a}(s, x))^2\, ds$$

is the Hellinger process for P *and* $\tilde{P}$.

Proof. By [15], Theorem 7.5 $P \overset{\text{loc}}{\ll} R$, $\tilde{P} \overset{\text{loc}}{\ll} R$ where R is the measure on $(C, \mathscr{C})$ with respect to which of the coordinate process $x = (x_t)_{t \geqq 0}$ is the Wiener process. Let $\varkappa^R$ and $\tilde{\varkappa}^R$ be the density processes of P and $\tilde{P}$ with respect to R. It is known (see, e.g. [15, Chap. 7]) that

$$\varkappa^R = 1 + \varkappa^R a \cdot x, \qquad \tilde{\varkappa}^R = 1 + \tilde{\varkappa}^R \tilde{a} \cdot x \quad (R\text{-a.s.}).$$

Let $\sigma_n^R = \inf(t\colon \varkappa_t^R \wedge \tilde{\varkappa}_t^R \leqq 1/n)$. It follows easily from the Ito formula that

$$X^R_{t \wedge \sigma_n^R} = 1 - (1/8)\, X^R_- (a - \tilde{a})^2 \circ \langle x \rangle_{t \wedge \sigma_n^R} + (1/2)\, X^R_- (a + \tilde{a}) \cdot x_{t \wedge \sigma_n^R}.$$

Since zero is an absorbing state for X^R we have

$$A^R = (1/8)\, X^R_- (a - \tilde{a})^2 \circ \langle x \rangle$$

and the desired assertion holds by Lemma 3.4 (note that $\langle x \rangle_t \equiv t$ R-a.s.).

2. Now we calculate the Hellinger process for the distributions of multivariate point processes.

Let $(\Omega, \mathscr{F}, F)$ be a filtered space, $(E, \mathscr{E})$ be a Lusin space (i.e. a Borel subset of a compact metric space), Δ be an extra point, $\mathscr{P}$ be the σ-algebra of predictable sets on $\Omega \times R_+$, $\hat{\mathscr{P}} = \mathscr{P} \otimes \mathscr{E}$ be the σ-algebra on $\Omega \times R_+ \times E$, $E^\Delta = E \cup \Delta$, $\mathscr{E}^\Delta = \sigma(\mathscr{E}, \{\Delta\})$.

According to [3] and [2, Chap. 3], a multivariate point process is a sequence $\Pi = (T_n, X_n)_{n \geqq 1}$ such that

a) $T_n \in \mathscr{T}$, $0 < T_n(\omega) \leqq T_{n+1}(\omega)$ and $T_n(\omega) < T_{n+1}(\omega)$ if $T_n(\omega) < \infty$;

b) random variables X_n with values $(E^\Delta, \mathscr{E}^\Delta)$ is $\mathscr{F}_{T_n}$-measurable and $X_n(\omega) = \Delta$ if and only if $T_n(\omega) = \infty$.

The integer-valued measure $\mu = \mu(dt, dx)$ is associated with Π by the formula

$$\mu([0, t] \times \Gamma) = \sum_{n \geqq 1} I_{(T_n \leqq t)} I_{(X_n \in \Gamma)}.$$

Let $\mathscr{F}_t^\mu = \sigma(\mu([0, t] \times \Gamma),\ s \leqq t,\ \Gamma \in \mathscr{E})$.

Suppose that $F = F^\mu = (\mathscr{F}_t^\mu)_{t \geqq 0}$. Let P and $\tilde{P}$ be the probability measures on $(\Omega, \mathscr{F})$ such that the random measure μ has as its compensators the predictable random measures ν and $\tilde{\nu}$, respectively.

Put $\Lambda = \nu + \tilde{\nu}$, $a_t = \nu(\{t\}, E)$, $\tilde{a}_t = \tilde{\nu}(\{t\}, E)$,

$$\rho_t^2(\nu, \tilde{\nu}) = (1/2)(\sqrt{\lambda} - \sqrt{\tilde{\lambda}})^2 * \Lambda_t$$

where $\lambda = \lambda(t, x)$ and $\tilde{\lambda} = \tilde{\lambda}(t, x)$ are $\hat{\mathscr{P}}$-measurable functions such that

$$\nu = \lambda \Lambda, \qquad \tilde{\nu} = \tilde{\lambda} \Lambda.$$

It is clear that the value $\rho_t(\nu, \tilde{\nu})(\omega)$ is the Hellinger distance between the restrictions of $\nu(\omega, \cdot)$ and $\tilde{\nu}(\omega, \cdot)$ to $\mathscr{B}_{[0,t]} \otimes \mathscr{E}$ ($\mathscr{B}_{[0,t]}$ is the Borel σ-algebra on $[0, t]$).

Theorem 6.2. *The process h defined by the formula*

$$h_t = \rho_t^2(\nu, \tilde{\nu}) + \tfrac{1}{2} \sum_{s \leqq t} \sqrt{1 - a_s} - \sqrt{1 - \tilde{a}_s})^2$$

is the Hellinger process for P and $\tilde{P}$.

Proof. We need several simple auxiliary statements which we formulate as lemmas.

Lemma 6.1. *The compensator* ν^Q *of the random measure* μ *with respect to* $Q=(P+\tilde{P})/2$ *is given by*

$$\nu^Q=(\varkappa_-\nu+\tilde{\varkappa}_-\tilde{\nu})/2$$

where $\varkappa$ *and* $\tilde{\varkappa}$ *are the density processes of* P *and* $\tilde{P}$ *with respect to* Q.

Proof. Using [2, (1.47)], for any nonmegative $\hat{\mathscr{P}}$-measurable function f we have:

$$\begin{aligned}2E_Q f*\mu_\infty &= Ef*\mu_\infty+\tilde{E}f*\mu_\infty=Ef*\nu_\infty+\tilde{E}f*\tilde{\nu}_\infty\\ &=E_Q\varkappa_\infty(f*\nu_\infty)+E_Q\tilde{\varkappa}_\infty(f*\tilde{\nu}_\infty)\\ &=E_Q f\varkappa_-*\nu_\infty+E_Q f\tilde{\varkappa}_-*\tilde{\nu}_\infty=E_Q f*(\varkappa_-\nu+\tilde{\varkappa}_-\tilde{\nu})_\infty.\end{aligned}$$

The desired result follows now from the uniqueness of the compensator.

According to [3], there exist $\hat{\mathscr{P}}$-measurable functions Y and $\tilde{Y}$ such that

$$\nu=Y\nu^Q \quad P\text{-a.s.}, \qquad \tilde{\nu}=\tilde{Y}\nu^Q \quad \tilde{P}\text{-a.s.} \tag{6.3}$$

In general, it is not true that the equalities (6.3) hold Q-a.s. Nevertheless, we have the following

Lemma 6.2. *Let* $\Gamma=(\varkappa_->0)$, $\tilde{\Gamma}=(\tilde{\varkappa}_->0)$. *Then*

$$I_\Gamma\nu=I_\Gamma Y\nu^Q, \qquad I_{\tilde{\Gamma}}\tilde{\nu}=I_{\tilde{\Gamma}}\tilde{Y}\nu^Q \quad Q\text{-a.s.} \tag{6.4}$$

Proof. For any $\hat{\mathscr{P}}$-measurable function $f\geqq 0$ we have

$$EfI_\Gamma*\nu_\infty=E_Q\varkappa_\infty(fI_\Gamma*\nu_\infty)=E_Q\varkappa_-fI_\Gamma*\nu_\infty.$$

On the other hand,

$$EfI_\Gamma*\nu_\infty=EfI_\Gamma Y*\nu^Q_\infty=E_Q\varkappa_-fI_\Gamma Y*\nu^Q_\infty.$$

Thus,

$$E_Q fI_\Gamma*\nu_\infty=E_Q fI_\Gamma Y*\nu^Q_\infty$$

and the first equality in (6.4) holds by the uniqueness of the compensator. The proof of the second equality in (6.4) is similar.

Put $a^Q_t=\nu^Q(\{t\}\times E)$ and note that by the above lemmas $a^Q=(a\varkappa_-+\tilde{a}\tilde{\varkappa}_-)/2$ and

$$I_{\Gamma\cap\tilde{\Gamma}}I_{(a^Q>0)}=I_{\Gamma\cap\tilde{\Gamma}}I_{(a+\tilde{a}>0)}.$$

According to the known results on the structure of the density process for multivariate point processes (see, e.g. [3], [5, § 12]) we have

$$\varkappa=1+\varkappa_-\cdot L$$

where

$$L_t=(Y-1+(\hat{Y}-a^Q)/(1-a^Q))*(\mu-\nu^Q)_t, \qquad t<\sigma,$$

$$\hat{Y}_t=\int_E Y(t,x)\,\nu^Q(\{t\}\times dx)$$

and σ is defined by (2.2). A similar representation holds for $\tilde{z}$. It follows from (2.1) that B^{σ_n} (σ_n is defined by (2.2)) is a Q-compensator of the increasing process b^{σ_n} where

$$b_t^{\sigma_n}=(1/2)\sum_{s\leqq t\wedge\sigma_n}(\sqrt{1+\Delta L_s}-\sqrt{1+\Delta\tilde{L}_s})^2=(1/2)(\sqrt{Y}-\sqrt{\tilde{Y}})^2*\mu_{t\wedge\sigma_n}$$
$$+(1/2)\sum_{s\leqq t\wedge\sigma_n} I_{(a_s^Q>0)}\left(\sqrt{\frac{1-a_s}{1-a_s^Q}}-\sqrt{\frac{1-\tilde{a}_s}{1-a_s^Q}}\right)^2(1-\mu(\{s\}\times E)).$$

Thus,

$$B_t^{\sigma_n}=(1/2)((\sqrt{Y}-\sqrt{\tilde{Y}})^2*\nu^Q_{t\wedge\sigma_n}+\sum_{s\leqq t\wedge\sigma_n} I_{(a_s^Q>0)}(\sqrt{1-a_s}-\sqrt{1-\tilde{a}_s})^2).$$

With the help of Lemmas 6.1 and 6.2 we can represent B^{σ_n} in the following form:

$$B_t^{\sigma_n}=(1/2)((\sqrt{\lambda}-\sqrt{\tilde{\lambda}})^2*\Lambda_{t\wedge\sigma_n}+\sum_{s\leqq t\wedge\sigma_n} I_{(a_s+\tilde{a}_s>0)}(\sqrt{1-a_s}-\sqrt{1-\tilde{a}_s})^2)$$
$$=\rho^2_{t\wedge\sigma_n}(\nu,\tilde{\nu})+(1/2)\sum_{s\leqq t\wedge\sigma_n}(\sqrt{1-a_s}-\sqrt{1-\tilde{a}_s})^2.$$

The assertion of Theorem 6.2 follows from the definition of the Hellinger process.

Remark 6.1. Let δ be a point outside E, $E^\delta=E\cup\delta$, $\mathscr{E}^\delta=\sigma(\mathscr{E},\{\delta\})$. Consider the random measure ν^0 on $(R_+\times E^\delta, \mathscr{B}_{[0,\infty[}\otimes\mathscr{E}^\delta)$ with

$$\nu^0(dt,dx)=I_E\,\nu(dt,dx)+I_{(a_t>0)}\,\varepsilon_\delta(dx)$$

where $\varepsilon_\delta(dx)$ is the Dirac measure. The random measure $\tilde{\nu}^0$ is defined similarly. Then $h_t=\rho_t^2(\nu^0,\tilde{\nu}^0)$. This relation of h with the Hellinger distance justifies the name "Hellinger process".

Remark 6.2. The Hellinger process h defined by Theorem 6.2 is constructed directly from the predictable characteristics of multivariate point processes (namely, the compensators). In particular, if the compensators are deterministic then h is also a deterministic process while B is random (it is stopped in the random moment σ). Thus, in a certain sense, h is the simplest Hellinger process.

3. Let $\Omega=D$, $\mathscr{F}$ be the completion of $\sigma(x_s, s\geqq 0)$ with respect to $Q=(P+\tilde{P})/2$, F be the minimal filtration satisfying the usual conditions with respect to Q and majorizing the filtration generated by coordinate maps in D.

Assume that under P (corr. under $\tilde{P}$) the coordinate process $x=(x_t)_{t\geqq 0}$ is a semimartingale with the triplet of predictable characteristics $T=(B, C, \nu)$ (corr. $\tilde{T}=(\tilde{B},\tilde{C},\tilde{\nu})$), see [2, Chap. 3], [5, §2]. Thus, under P we have the representation

$$x_t=x_0+B_t+uI_{(|u|>1)}*\mu_t+x_t^c+uI_{(|u|\leqq 1)}*(\mu-\nu)_t$$

where μ is the jump measure of x, $x^c=(x_t^c)_{t\geqq 0}$ is the continuous P-martingale part of x with $\langle x^c\rangle=C$. The similar representation for x holds under $\tilde{P}$.

Put

$$V_t = \mathrm{Var}_t(B-\tilde{B}) + |u|\, I_{(|u|\leqq 1)} * \mathrm{Var}(\nu-\tilde{\nu})_t,$$
$$\tau = \inf(t\colon V_t = \infty).$$

Define the predictable increasing process A in the following way:

$$A_t = \begin{cases} 0 & \text{if } \tau = 0 \\ B_t - \tilde{B}_t - uI_{(|u|\leqq 1)} * (\nu-\tilde{\nu})_t & \text{if } t<\tau, \\ \overline{\lim_{s\uparrow\tau}}\, A_s & \text{if } t \geqq \tau > 0. \end{cases}$$

Let

$$h_t = (1/8)\,|dA/dC| \circ \mathrm{Var}_t(A) + \rho_t^2(\nu,\tilde{\nu}) + \tfrac{1}{2} \sum_{s\leqq t} (\sqrt{1-a_s} - \sqrt{1-\tilde{a}_s})^2$$

where dA/dC is the $\mathscr{P}$-measurable version of the Radon-Nikodym derivative of the absolute continuous part of A with respect to C.

$$a_t = \nu(\{t\}, R_0), \qquad \tilde{a}_t = \tilde{\nu}(\{t\}, R_0), \qquad R_0 = R \setminus \{0\}.$$

Following [2, Chap. 12], we shall say that P has the local uniqueness property if for any $\theta \in \mathscr{T}$ and any measure $\hat{P}$ on $(\Omega, \mathscr{F})$ for which x is the semimartingale with the triplet of the predictable characteristics $T^\theta = (B^\theta, C^\theta, \nu^\theta)$ (where $\nu^\theta = I_{[\![0,\theta]\!]}\, \nu$) the restrictions of P and $\hat{P}$ on $\mathscr{F}_{\theta-}$ coincide.

Theorem 6.3 ([13]). *Assume that $C = \tilde{C}$ and at least one of the measures P or $\tilde{P}$ has the local uniqueness property.*

Then h is a Hellinger process for P and $\tilde{P}$.

Remark 6.3. For the case where P and $\tilde{P}$ are the distributions of processes with independent increments (i.e. the triplets T and $\tilde{T}$ are deterministic), the local uniqueness property holds [2, (13.2)]. The Hellinger process h is deterministic here, $H_T = \mathscr{E}_T(-h)$ for $T \in R_+$ and estimates for $\|P_T - \tilde{P}_T\|$ follows directly from Lemma 2.1 (see [17]).

References

1. Dellacherie, C.: Capacités et processus stochastiques. Berlin-Heidelberg-New York: Springer 1972
2. Jacod, J.: Calcul stochastique et problèmes des martingales Lect. Notes in Math. **714**. Berlin-Heidelberg-New York: Springer 1979
3. Jacod, J.: Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. Z. Wahrscheinlichkeitstheor. Verw. Geb. **31**, 235–253 (1975)
4. Kabanov, Yu.M., Liptser, R.Sh.: On convergence in variation of the distributions of multivariate point processes. Z. Wahrscheinlichkeitstheor. Verw. Geb. **63**, 475–485 (1983)
5. Kabanov, Yu.M., Liptser, R.Sh., Shiryaev, A.N.: Absolute continuity and singularity of locally absolute continuous probability distributions, I, II. Math. Sbornik, **107**, 364–415 (1978), **108**, 30–61 (1979) (in Russian). English translation: Math. USSR Shornik, **35**, 631–680 (1979), **36**, 31–58 (1980)
6. Kabanov, Yu.M., Liptser, R.Sh., Shiryaev, A.N.: On the proximity in variation of probability measures. Dokl. Akad. Nauk SSSR. **278**, 265–268 (1984) (in Russian)

7. Kabanov, Yu.M., Liptser, R.Sh., Shiryaev, A.N.: Weak and strong convergence of distributions of counting processes. Theor. Probab. Appl. **28**, 288–319 (1983) (in Russian)
8. Kraft, C.: Some conditions for consistency and uniform consistency of statistical procedures. Univ. California Publ. Stat. **2**, 125–141 (1955)
9. Liese, F.: An estimation of Hellinger integrals of point processes. Preprint (1983)
10. Liese, F.: Hellinger integrals of Gaussian processes with independent increments. Stochastics, **6**, 81–96 (1982)
11. Liese, F.: Hellinger integrals of diffusion processes. Preprint (1983)
12. Liptser, R.Sh., Pukelsheim, F., Shiryaev, A.N.: Necessary and sufficient conditions for contiguity and entire asymptotic separation of probability measures. Uspekhi Mat. Nauk. **37**, 97–124 (1982) (in Russian). English translation: Russian Math. Surveys **37**, 107–136 (1982)
13. Liptser, R.Sh., Shiryaev, A.N.: On contiguity of probability measures corresponding to semimartingales. Analysis Mathematica **11**, 93–124 (1985)
14. Liptser, R.Sh., Shiryaev, A.N.: On the problem of "predictable" criteria of contiguity. Lecture Notes in Math. **1021**, 384–418. Berlin-Heidelberg-New York: Springer 1983
15. Liptser, R.Sh., Shiryaev, A.N.: Statistics of random processes, I, II. Berlin-Heidelberg-New York: Springer 1977, 1978
16. Matusita, K.: Decision rules, based on the distance, for problem of fit, two samples, and estimation. Ann. Math. Stat. **26**, 631–640 (1955)
17. Mémin, J., Shiryaev, A.N.: Distance de Hellinger-Kakutani des lois correspondant à deux processus à accroissement indépendants: critères d'absolue continuité et de singularité. Preprint (1983)
18. Vajda, I.: On the f-divergence and singularity of probability measures. Period. Math. Hung. **2**, 223–234 (1972)
19. Valkeila, E., Vostrikova, L.: An integral representation for the Hellinger distance. Stochastics. [To appear]
20. Vostrikova, L.: On the "predictable" criteria of the variation convergence of probability measures. Uspekhi Math. Nauk. **39**, 143–144 (1984) (in Russian)

Received April 20, 1984; in final form April 26, 1985