Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

On a posterior information process for parametric families of experiments

K. Dzhaparidze, P.J.C. Spreij, E. Valkeila

Probability, Networks and Algorithms (PNA)

PNA-R9818 December 1998

Report PNA-R9818
ISSN 1386-3711

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

On a Posterior Information Process for Parametric Families of Experiments

Kacha Dzhaparidze
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

Peter Spreij
Vrije Universiteit
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

Esko Valkeila
University of Helsinki
P.O. Box 4, 00014 Helsinki, Finland

ABSTRACT

In a filtered statistical experiment a priori and a posteriori probability measures are defined on an abstract parametric space. The information in the posterior, given the prior, is defined by the usual Kullback-Leibler formula. Certain properties of this quantity is investigated in the context of so-called arithmetic and geometric measures and arithmetic and geometric processes. Interesting multiplicative decompositions are presented that involve Hellinger processes indexed both by prior and by posterior distributions.

1991 Mathematics Subject Classification: 60G07, 60H30, 62B15, 94A17
Keywords and Phrases: filtered statistical experiment, prior and posterior distributions, Kullback-Leibler information, arithmetic and geometric measures, arithmetic and geometric processes, Hellinger integrals and processes.
Note: work carried out under project PNA3.3 Stochastic processes and applications, and supported in part by the Academy of Finland Project No. 37689 and the Research Grants Committee of the University of Helsinki.

1. Introduction

The setup of the present report is the same as in our previous paper [1], where certain aspects are studied of randomized filtered experiments indexed by an arbitrary parametric space. Some new aspects discussed in this report concern relationship between prior and posterior probability distributions on the parametric space. In section 3.3 we will introduce the notion of the Kullback-Leibler information in the posterior distribution given a prior, and in section 5.5 we will present its multiplicative representation. The predictable component of this representation (5.23) involves the Hellinger process $h(\alpha)$ indexed by the prior probability measure α. This notion is well-known in the particular case of binary experiments (see the book [5]) or in the case of a finite parametric space (see [3] and [4]). The present generalization to an arbitrary parametric space comes from [1]. The martingale component of the representation (5.23) is the density of the geometric (g-)mean measure with respect to the arithmetic (a-)mean measure whose definitions and properties are also taken over from [1]. Some additional aspects of the a-mean measure can be found in [9], section 65. The g-mean measure confined to a finite parametric space has been introduced in [2]. The expression (5.24) of the martingale involved in the representation (5.23) seems interesting because of a natural relationship between the prior and posterior expectations of the drift β (in the integrand of the continuous part) and the intensity density Y (in the integrand of the discrete part).

The information defined by (3.11) at any stopping time $T > 0$, is based on all past observations up to T. In section 7 a different, dynamical approach is pursued. First an instantaneous gain of information is defined by the usual Kullback-Leibler formula (cf (7.1) or (7.4)), that is provided by a single observation at each fixed time instant. By integrating then all the results up to certain stopping time $T > 0$, we get the so-called cumulative information contained in the posterior given a prior, cf definition (7.3) or (7.5). Parallel to the previous case (5.23), we again have an interesting multiplicative decomposition (7.8).

2. Randomized experiments

2.1 Statistical experiment

Consider a *statistical experiment* $(\Omega, \mathcal{F}, \{P_\theta\}_{\theta\in\Theta})$, where $\{P_\theta\}_{\theta\in\Theta}$ is a certain parametric family of probability measures defined on a measurable space $(\Omega, \mathcal{F})$ with a set of elementary events Ω and a σ-field $\mathcal{F}$. Suppose that each member of the family $\{P_\theta\}_{\theta\in\Theta}$ is equivalent to a certain probability measure Q, i.e.

$$\{P_\theta\}_{\theta\in\Theta} \quad \sim \quad Q. \tag{2.1}$$

For each fixed $\theta \in \Theta$ denote by p_θ the Radon-Nikodym derivative of P_θ with respect to Q:

$$p_\theta = \frac{dP_\theta}{dQ}. \tag{2.2}$$

So, for each $\theta \in \Theta$ and $B \in \mathcal{F}$

$$P_\theta(B) = \int_B p_\theta(\omega) Q(d\omega) = E_Q\{1_B p_\theta\}. \tag{2.3}$$

Here and elsewhere below we use the expectation sign E indexed by a probability measure.

2.2 Randomization

On the set of parameter values Θ define a σ-field $\mathcal{A}$ and consider a probability space $(\Theta, \mathcal{A}, \alpha)$ where α is a certain probability measure. In this way a statistical parameter ϑ is viewed as a random variable on the probability space $(\Theta, \mathcal{A}, \alpha)$ with the probability measure α determining *a priori* distribution of ϑ.

Consider now the direct product $(\mathbf{\Omega}, \boldsymbol{\mathcal{F}}, \mathbf{Q})$ of two probability spaces $(\Omega, \mathcal{F}, Q)$ and $(\Theta, \mathcal{A}, \alpha)$, where $\mathbf{\Omega} = \Omega\times\Theta$, $\boldsymbol{\mathcal{F}} = \mathcal{F}\otimes\mathcal{A}$ and $\mathbf{Q} = Q\times\alpha$. Along with $\mathbf{Q}$ define on $(\mathbf{\Omega}, \boldsymbol{\mathcal{F}})$ another probability measure $\mathbf{P}$ as follows: for each $B \in \boldsymbol{\mathcal{F}}$

$$\mathbf{P}(\mathbf{B}) = \int_{\mathbf{B}} p(\omega,\theta) Q(d\omega)\alpha(d\theta) \doteq E_{\mathbf{Q}}\{1_{\mathbf{B}} p\} \tag{2.4}$$

so that for each $\boldsymbol{\omega} = (\omega, \theta) \in \boldsymbol{\Omega}$ we have $p(\boldsymbol{\omega}) = \frac{d\mathbf{P}}{d\mathbf{Q}}(\boldsymbol{\omega})$. Obviously, by (2.2) we have the identity

$$p(\boldsymbol{\omega}) = \frac{d\mathbf{P}}{d\mathbf{Q}}(\boldsymbol{\omega}) = \frac{dP_\theta}{dQ}(\omega) = p_\theta(\omega); \tag{2.5}$$

Observe that in the present setting the probability measure P_θ defined for each $\theta \in \Theta$ by (2.3) (and satisfying $P_\theta(\Omega) = 1$), can be viewed as a regular conditional probability measure, under the condition that the statistical parameter ϑ takes on the particular value θ. In view of (2.3) we can rewrite (2.4) as follows: for each $\mathbf{B} = B \times A \in \boldsymbol{\mathcal{F}}$

$$\mathbf{P}(\mathbf{B}) = \int_A p_\theta(B)\alpha(d\theta) = E_\alpha\left\{1_A E_Q\{1_B p\}\right\} = E_Q\left\{1_B E_\alpha\{1_A p\}\right\},$$

since by Loève [8], theorem 8.2B, it is allowed to interchange the integration order.

The Kullback-Leibler information in this experiment $I(\mathbf{P}|\mathbf{Q}) = E_{\mathbf{Q}} \log \frac{d\mathbf{Q}}{d\mathbf{P}}$ is positive by assumption. Later (from section 3.3 onwards) we also assume that this information is finite, i.e.

$$0 < I(\mathbf{P}|\mathbf{Q}) < \infty. \tag{2.6}$$

Note the identity

$$I(\mathbf{P}|\mathbf{Q}) = E_\alpha I(P_\vartheta|Q) \tag{2.7}$$

where $I(P_\theta|Q) = E_Q \log \frac{dQ}{dP_\theta}$ is the Kullback-Leibler information in P_θ for each fixed $\theta \in \Theta$ with respect to a dominating measure Q. This is easily seen in view of (2.5), since

$$E_\alpha I(P_\vartheta|Q) = E_\alpha E_Q \log \frac{dQ}{dP_\vartheta} = E_{\mathbf{Q}} \log \frac{d\mathbf{Q}}{d\mathbf{P}}.$$

2.3 Arithmetic mean measure

It is often useful to make a concrete choice of a dominating measure Q. Like in [5], p 163, a new measure on the same measurable space $(\Omega, \mathcal{F})$, the so-called *arithmetic mean measure* $\bar{P} = \bar{P}_\alpha$, is defined as follows: for each $B \in \mathcal{F}$

$$\bar{P}(B) = \mathbf{P}(B \times \Theta) = E_\alpha P_\vartheta(B).$$

Lemma 3.2 in [1] tells us that under the assumption (2.1)

$$\bar{P}_\alpha \sim Q \quad \text{and} \quad \frac{d\bar{P}_\alpha}{dQ} = E_\alpha p_\vartheta. \tag{2.8}$$

We mention one specific usage of the arithmetic mean measure. In the Bayesian setup the measure α on $(\Theta, \mathcal{A})$ is viewed as *a priori* probability measure. Along with this, one may also define on the same space *a posteriori* probability measure α^1 as follows: for all $A \in \mathcal{A}$

$$\alpha^1(A) \doteq \int_A \frac{dP_\theta}{d\bar{P}_\alpha}\alpha(d\theta) \tag{2.9}$$

i.e. for each $\theta \in \Theta$

$$\frac{d\alpha^1}{d\alpha}(\theta) = \frac{dP_\theta}{d\bar{P}_\alpha}. \tag{2.10}$$

Note that for fixed $A \in \mathcal{A}$ the random variable $\alpha^1(A)$ is $\mathcal{F}$-measurable. Define now the Kullback-Leibler information in the posterior α^1 with respect to the prior α, given for each $\theta \in \Theta$ by

$$I(\alpha^1|\alpha) = E_\alpha \log \frac{d\alpha}{d\alpha^1}(\vartheta). \tag{2.11}$$

Since the interchange of the integration order is allowed (see the previous section) we have the identity $E_{\bar{P}} I(\alpha^1|\alpha) = E_\alpha I(P_\vartheta|\bar{P})$.

In section 5.4 another measure, called the *geometric mean measure*, will be introduced which, used as a dominating measure, yields the important equality (5.21); cf also section 6.1.

3. Randomized filtered experiment

3.1 Filtration

Let the measurable space $(\Omega, \mathcal{F})$ be equipped with a filtration $F = \{\mathcal{F}_t\}_{t\geq 0}$, an increasing and right continuous flow of sub-σ-fields of $\mathcal{F}$, so that $\bigvee_{t\geq 0} \mathcal{F}_t = \mathcal{F}_\infty = \mathcal{F}$. Assume that the filtered probability space $(\Omega, \mathcal{F}, F = \{\mathcal{F}_t\}_{t\geq 0}, Q)$ is a stochastic basis: $\mathcal{F}$ is Q-complete and each $\mathcal{F}_t$ contains the Q-null sets of $\mathcal{F}$. We also assume for simplicity that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ Q-a.s. The filtered probability space

$$(\Omega, \mathcal{F}, F, \{P_\theta\}_{\theta\in\Theta}, Q) \tag{3.1}$$

so defined is called a *filtered statistical experiment.* Consider the optional projections of the probability measures Q and P_θ with respect to F, and use the same symbols for resulting optional valued processes: for a F-stopping time T Q_T and $P_{\theta,T}$ are then the restrictions of the measures Q and P_θ to the sub-σ-field $\mathcal{F}_T$. Since $P_{\theta,T}$ is equivalent to Q_T for each $\theta \in \Theta$, we can define the Radon-Nikodym derivatives

$$z_T(\theta) \doteq \frac{dP_{\theta,T}}{dQ_T} = E_Q\{p_\theta|\mathcal{F}_T\}. \tag{3.2}$$

Thus according to [5], section III.3, for each fixed $\theta \in \Theta$ there is a unique (up to Q-indistinguishability) process $z(\theta) = z(\theta, Q)$ called the *density process.* We usually stress the dependence on a dominating measure Q. So $z_t(\theta, Q) = \frac{dP_{\theta,t}}{dQ_t}$ for all $t \geq 0$. The density process possesses the following properties (see [5], section III.3, proposition 3.5, for more details): for each $\theta \in \Theta$

(i) $\inf_t z_t(\theta, Q) > 0$ Q-a.s.

(ii) $\sup_t z_t(\theta, Q) < \infty$ Q-a.s.

(iii) $z(\theta, Q)$ is a (Q, F)-uniformly integrable martingale with $E_Q z_t(\theta, Q) = 1$, for all $t \in [0, \infty]$.

Consider now the situation of section 2.2. Let $(\mathbf{\Omega}, \boldsymbol{\mathcal{F}}, \mathbf{F}, \mathbf{P}, \mathbf{Q})$ be a binary experiment equipped with the filtration $\mathbf{F} = \{\mathcal{F}_t \otimes \mathcal{A}\}_{t\geq 0}$, which is call a *filtered randomized experiment.* Take again the optional projections of the probability measures $\mathbf{Q}$ and $\mathbf{P}$ with respect to $\mathbf{F}$. For a F-stopping time T (which is clearly $\mathbf{F}$-stopping time, as well) $\mathbf{Q}_T$ and $\mathbf{P}_T$ are then the restrictions of the measures $\mathbf{Q}$ and $\mathbf{P}$ to the sub-σ-field $\boldsymbol{\mathcal{F}}_T$, with the Radon-Nikodym derivative

$$\mathbf{z}_T \doteq \frac{d\mathbf{P}_T}{d\mathbf{Q}_T} = E_Q\{p|\boldsymbol{\mathcal{F}}_T\} \tag{3.3}$$

with p as in (2.5). We get then the identity $E_\alpha z_T(\vartheta, Q) = E_{\mathbf{Q}}\{p|\boldsymbol{\mathcal{F}}_T\}$.

All parametric families of processes $\{X(\theta)\}_{\theta\in\Theta}$ treated in this paper (such as the family of density processes $\{z(\theta)\}_{\theta\in\Theta}$ defined by (3.2)) are supposed to be *adapted* to the filtration $\mathbf{F}$, i.e. $\{\mathcal{F}_t\otimes\mathcal{A}\}$-measurable for each $t \geq 0$, and càdlàg for each $\theta \in \Theta$. A parametric family of processes $\{X(\theta)\}_{\theta\in\Theta}$ is called *predictable* if it is $\mathcal{P}\otimes\mathcal{A}$-measurable, where $\mathcal{P}$ is the predictable σ-field on $\Omega\times\mathbb{R}_+$. Let μ be a random measure defined on $\mathbb{R}_+\times E$ with an appropriate measurable space $(E, \mathcal{E})$. With a random measure μ and a probability measure Q we associate the Doléans measure M^Q_μ, defined on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ where $\tilde{\Omega} = \Omega\times\mathbb{R}_+\times E$ and $\tilde{\mathcal{F}} = \mathcal{F}\otimes B(\mathbb{R}_+)\otimes\mathcal{E}$. Recall that $M^Q_\mu(d\omega; dt, dx) = Q(d\omega)\mu(\omega; dt, dx)$. We will use the common notation $M^Q_\mu(\cdot|\tilde{\mathcal{P}})$ for the corresponding conditional expectation with respect to $\tilde{\mathcal{P}} = \mathcal{P}\otimes\mathcal{E}$ (for more details see [5], section III.3c, or [7], chapter 3). Define similarly the Doléans measure $M^{\mathbf{Q}}_\mu$ on $(\tilde{\Omega}\times\Theta, \tilde{\mathcal{F}}\otimes\mathcal{A})$, $M^{\mathbf{Q}}_\mu = M^Q_\mu\otimes\alpha$. Write $\tilde{\boldsymbol{\mathcal{P}}} = \tilde{\mathcal{P}}\otimes\mathcal{A}$. Let W be a nonnegative $\tilde{\mathcal{F}}\otimes\mathcal{A}$-measurable function. Then we define for each θ the function $W_\theta(\cdot,\cdot,\cdot) = W(\cdot,\theta,\cdot,\cdot)$, which is then $\tilde{\mathcal{F}}$-measurable. Likewise we also consider W_ϑ. Then we obtain from Fubini's theorem $M^{\mathbf{Q}}_\mu(W|\tilde{\boldsymbol{\mathcal{P}}}) = E_\alpha\{M^Q_\mu(W_\vartheta|\tilde{\mathcal{P}})\} = M^Q_\mu(E_\alpha W_\vartheta|\tilde{\mathcal{P}})$. Finally, let ν be the compensator of μ. Both μ and ν extend trivially to random measures –again denoted by μ and ν– on $\mathbb{R}_+\times E$ parameterized by ω, θ via $\mu(\omega, \theta; dt, dx) = \mu(\omega; dt, dx)$ and likewise for ν. Hence for a $\tilde{\mathcal{P}}\otimes\mathcal{A}$-measurable positive function W on $\tilde{\Omega}\times\Theta$ we can associate the process $\hat{W}$ in the usual way:

$$\hat{W}_t(\boldsymbol{\omega}) = \int_E W(\boldsymbol{\omega}; t, x)\nu(\omega; \{t\}\times dx). \tag{3.4}$$

In the sequel these results will be applied to the well known integer-valued random measure μ^X associated to (the jumps of) a càdlàg process X as defined in [5], section II.1, proposition 1.16.

3.2 Prior and posterior measures

Departing from the identity (2.8), we define a new process $a(\alpha, Q)$, called the *arithmetic mean process*, by restricting the density $\frac{d\bar{P}_\alpha}{dQ}$ to $\mathcal{F}_t$ for all $t \geq 0$ so that

$$a(\alpha, Q) = z(\bar{P}_\alpha, Q) \tag{3.5}$$

where $z_t(\bar{P}_\alpha, Q) = E_Q\{\frac{d\bar{P}_\alpha}{dQ}|\mathcal{F}_t\}$ for all $t \geq 0$. Note that with the choice $\bar{P}$ as the dominating measure it becomes particularly simple: identically $a(\alpha, \bar{P}) = 1$. We may also write

$$a(\alpha, Q) = E_\alpha z(\vartheta, Q) \tag{3.6}$$

with $a_t(\alpha, Q) = \mathbf{z}_t = E_{\mathbf{Q}}\{p|\boldsymbol{\mathcal{F}}_t\}$ for all $t \geq 0$, cf (3.3). The a-mean process possesses the following properties:

(i) $\inf_t a_t > 0$ Q-a.s.

(ii) $\sup_t a_t < \infty$ Q-a.s.

(iii) a is a (Q, F)-uniformly integrable martingale with $E_Q a_t = 1$ for all $t \geq 0$,

(iv) if X is a certain (Q, F)-semimartingale, then

$$\langle a, X^c\rangle = E_\alpha\langle z(\vartheta, Q), X^c\rangle \quad\text{and}\quad M^Q_{\mu^X}(a|\tilde{\mathcal{P}}) = E_\alpha M^Q_{\mu^X}(z(\vartheta, Q)|\tilde{\mathcal{P}}).$$

The first three statements are quite parallel to that of density processes in section 3.1. For the proof of property (iv), see [1], proposition 3.1.

Like in section 2.3 consider the Bayesian setup in which the measure α on $(\Theta, \mathcal{A})$ is interpreted *a priori* probability measure. Along with this define for each stopping time T on the same space the *a posteriori* probability measure α^T as follows: for $\forall A \in \mathcal{A}$

$$\alpha^T(A) \doteq \frac{\int_A z_T(\theta, Q)\alpha(d\theta)}{\int_\Theta z_T(\theta, Q)\alpha(d\theta)},$$

i.e.

$$\frac{d\alpha^T}{d\alpha}(\vartheta) = \frac{z_T(\vartheta, Q)}{\int_\Theta z_T(\theta, Q)\alpha(d\theta)}.$$

Compare these equations with (2.9) and (2.10) which we had initially, prior to the filtered setup. Obviously, the posterior α^T so defined is free of the choice of a dominating measure Q. Note that for fixed $A \in \mathcal{A}$ the random variable $\alpha^T(A)$ is $\mathcal{F}_T$-measurable. In view of the identity $a(\alpha, \bar{P}) \equiv 1$ mentioned above, we get with $\bar{P_\alpha}$ as the dominating measure that for each $\theta \in \Theta$

$$\frac{d\alpha^T}{d\alpha}(\theta) = \frac{z_T(\theta, \bar{P_\alpha})}{a_T(\alpha, \bar{P_\alpha})} = z_T(\theta, \bar{P_\alpha}). \tag{3.7}$$

3.3 Information in the posterior given a prior

Along with the a-mean process (3.6), we associate with the parametric family of density processes $\{z(\theta, Q)\}_{\theta\in\Theta}$ a so-called *geometric mean process*

$$g(\alpha, Q) = e^{E_\alpha \log z(\vartheta, Q)}. \tag{3.8}$$

By the Jensen inequality g-mean process is dominated by a-mean process identically, i.e.

$$g(\alpha, Q) \leq a(\alpha, Q) \tag{3.9}$$

so that the g-mean process shares property (ii) of the a-mean process, mentioned in section 3.2. As for the lower bound, we have assumed (2.6) in order to let the g-mean process share property (i) of the a-mean process, as well.

Proposition 3.1. *Assume (2.1) and (2.6). The geometric mean process $g = g(\alpha, Q)$ possesses the following properties:*
(i) $\inf_t g_t > 0$ *Q-a.s.*
(ii) $\sup_t g_t < \infty$ *Q-a.s.*
(iii) g is a (Q, F)-supermartingale of class (D) with $g_0 = 1$.

Property (i) is an immediate consequence of (2.6) and Jensen's inequality, while (ii) follows from equation (3.9). As for property (iii), we have that the g-mean process is indeed of class (D), since it is dominated by a process of class (D), a (Q, F)-uniformly integrable martingale a (see (3.9)). For the concluding inequality $E_Q\{g_t|\mathcal{F}_s\} \leq g_s$ as $s \leq t$, a consequence of Jensen's inequality, see [1], proposition 4.1. Observe that by the identity (3.5) and the inequality (3.9) we have

$$\frac{g(\alpha, Q)}{a(\alpha, Q)} = g(\alpha, \bar{P_\alpha}) \leq 1. \tag{3.10}$$

Surely, this fraction depends only on the prior α but not on the choice of a dominating measure Q.

Similarly to (2.11), we define at a stopping time $T > 0$ the Kullback-Leibler information in the posterior probability measure α^T with respect to the prior α by

$$I(\alpha^T|\alpha) = E_\alpha \log \frac{d\alpha}{d\alpha^T} \tag{3.11}$$

which is a non-negative quantity by the Jensen inequality. It is related to the arithmetic and geometric mean processes as follows:

Theorem 3.2. *Let $T > 0$ be a stopping time, let α and α^T be the prior and posterior probability measures on the parametric space $(\Theta, \mathcal{A})$ and let $I(\alpha^T|\alpha)$ be the Kullback-Leibler information in the posterior α^T with respect to the prior α, as defined in (3.11). Then*

$$e^{-I(\alpha^T|\alpha)} = \frac{g_T(\alpha, Q)}{a_T(\alpha, Q)} = g_T(\alpha, \bar{P}_\alpha). \tag{3.12}$$

Proof. The latter equality has already been presented in (3.10). By the definitions (3.7), (3.8) and (3.11) we have $e^{-I(\alpha^T|\alpha)} = e^{E_\alpha \log z_T(\vartheta, \bar{P}_\alpha)} = g_T(\alpha, \bar{P}_\alpha)$. □

Observe again that the information $I(\alpha^T|\alpha)$ depends only on the prior α but not on the choice of a dominating measure Q. In view of the properties of the arithmetic and geometric mean processes, presented above, we have

Proposition 3.3. *Assume (2.1) and (2.6). Let $I(\alpha^\cdot|\alpha)$ be the information process starting from zero, $I(\alpha^0|\alpha) = 0$, and at $t > 0$ defined by (3.11). Then it possesses the following properties:*
(i) $\inf_t I(\alpha^t|\alpha) > 0$ *Q-a.s.*
(ii) $\sup_t I(\alpha^t|\alpha) < \infty$ *Q-a.s.*
(iii) $I(\alpha^\cdot|\alpha)$ is a (P_α, F)-submartingale of class (D).

Proof. This is a direct consequence of theorem 3.2 and the corresponding property of the g-mean process mentioned above. □

4. Predictable characteristics

4.1 Triplet of characteristics

Suppose that we are given the observations that constitute a semimartingale X defined on $(\Omega, \mathcal{F}, F, Q)$, i.e. a (Q, F)-semimartingale, with the triplet of predictable characteristics $T = (B, C, \nu)$. This and all the triplets considered in the present paper are related to a fixed truncation function $\hbar : \mathbb{R} \to \mathbb{R}$, a bounded function with a compact support so that $\hbar(x) = x$ in a vicinity of the origin. By the Girsanov theorem for semimartingales (see [5], Theorem III.3.24, p 159, or [7], Theorem IV.5.3, p 232) X is also a (P_θ, F)-semimartingale for each $\theta \in \Theta$. Denote by $T(\theta) = (B(\theta), C(\theta), \nu(\theta))$ the corresponding triplet of predictable characteristics, which is related to the triplet T as follows:

$$\begin{cases} B(\theta) &= B + \beta(\theta) \cdot C + (Y(\theta) - 1)\, \hbar * \nu \\ C(\theta) &= C \\ \nu(\theta) &= Y(\theta) \cdot \nu \end{cases} \tag{4.1}$$

with certain processes $\beta(\theta) = \beta(\theta, Q)$ and $Y(\theta) = Y(\theta, Q)$ so that $|\beta(\theta)|^2 \cdot C_t < \infty$ and $(Y(\theta) - 1)\ \hbar * \nu_t < \infty$ Q-a.s. for all $t \geq 0$. According to [7], Lemma IV.5.6, p 231, these processes are described as follows. Let the density process be represented in as a Doléans-Dade exponential $z(\theta, Q) = \mathcal{E}(m(\theta, Q))$ with

$$m(\theta, Q) = z_{\_}(\theta, Q)^{-1} \cdot z(\theta, Q). \tag{4.2}$$

The continuous process $\beta(\theta, Q)$ satisfies $\beta(\theta, Q) \cdot C = \langle m(\theta, Q), X^c \rangle$. As for $Y(\theta, Q)$, a $\tilde{\mathcal{P}} \otimes \mathcal{A}$ -measurable positive function, it satisfies $Y(\theta, Q) - 1 = M^Q_{\mu^X}(\Delta m(\theta, Q) | \tilde{\mathcal{P}})$.

4.2 Characteristics w.r.t. the a-mean measure

In the situation of the previous section, the observations X constitute a semimartingale with respect to the a-mean measure $\bar{P}_\alpha$, as well. The following theorem, taken over from [1], section 3.3 (a generalization of a result by Kolomiets [6]; see also [5], Theorem III.3.40, p 163 or [7], Theorem IV.5.4, p 234), relates the triplet under $\bar{P}_\alpha$ to the triplets $T(\theta), \theta \in \Theta$:

Theorem 4.1. *Assume (2.1). Let X be a (P_θ, F)-semimartingale for each $\theta \in \Theta$ with the triplet $T(\theta)$ of predictable characteristics. Then it is a $(\bar{P}_\alpha, F)$-semimartingale as well, with the triplet $\bar{T} = (\bar{B}, \bar{C}, \bar{\nu})$ where*

$$\begin{cases} \bar{B} &= E_\alpha\{z_{\_}(\vartheta, \bar{P}_\alpha) \cdot B(\vartheta)\} \\ \bar{C} &= C \\ \bar{\nu} &= E_\alpha\{z_{\_}(\vartheta, \bar{P}_\alpha) \cdot \nu(\vartheta)\}. \end{cases} \tag{4.3}$$

Proof. See [1], theorem 3.3. □

Corollary 4.2. *Under the conditions of theorem 4.1 the local characteristics $\bar{\beta}$ and $\bar{Y}$ with respect to the arithmetic mean measure $\bar{P}_\alpha$ are the posterior expectations of $\beta(\vartheta)$ and $Y(\vartheta)$: for each $t > 0$*

$$\bar{\beta}_t = E_{\alpha^{t-}} \beta_t(\vartheta) \quad \textit{and} \quad \bar{Y}_t = E_{\alpha^{t-}} Y_t(\vartheta). \tag{4.4}$$

Proof. In view of the identity (3.7) the definitions (4.4) are equivalent to

$$\bar{\beta} = E_\alpha\{z_{\_}(\vartheta, \bar{P}_\alpha)\beta(\vartheta)\} \quad \text{and} \quad \bar{Y} = E_\alpha\{z_{\_}(\vartheta, \bar{P}_\alpha)Y(\vartheta)\}. \tag{4.5}$$

By (4.1) and (4.3)

$$\bar{B} = B + \bar{\beta} \cdot C + (\bar{Y} - 1)\ \hbar * \nu$$

with $\bar{\beta}$ and $\bar{Y}$ as in (4.5). This confirms the desired assertion. □

5. Explicit representations

5.1 Hellinger integrals and processes

Let T be a F-stopping time. We associate with the family of probability measures $\{P_{\theta,T}\}_{\theta \in \Theta}$, the so-called *Hellinger integral of order α* which is defined according to [5], section IV.1, and [1], section 4.2, as the Q-expectation of the g-mean process evaluated at T:

$$H(\alpha, T) = E_Q g_T(\alpha, Q). \tag{5.1}$$

Note that the Hellinger integral is independent of the choice of the dominating measure Q: if Q' is another dominating measure such that $Q \ll Q'$ and $Z = \frac{dQ}{dQ'}$, then $E_Q g(\alpha, Q) = E_{Q'} g(\alpha, Q')$, since $E_Q g(\alpha, Q) = E_{Q'}\{Z\, g(\alpha, Q)\}$ and by definition (3.8)

$$Z\, g(\alpha, Q) = e^{E_\alpha \log[Z\, z(\vartheta, Q)]} = e^{E_\alpha \log z(\vartheta, Q')} = g(\alpha, Q'), \tag{5.2}$$

cf [1], section 4.2.

Next, we define the Hellinger process of order α, denoted traditionally by $h(\alpha)$.

Theorem 5.1. *Assume (2.1) and (2.6). There exists a (unique up to Q-indistinguishability) predictable finite-valued increasing process $h(\alpha)$ starting from the origin $h_0(\alpha) = 0$, so that*

$$M(\alpha, Q) = g(\alpha, Q) + g_{\_}(\alpha, Q) \cdot h(\alpha) \tag{5.3}$$

is a (Q, F)-uniformly integrable martingale.

Proof. See [1], theorem 4.2. □

Like the Hellinger integrals, the Hellinger processes are independent of the choice of the dominating measure Q. Note also that up to a Q-evanescent set

$$\Delta h(\alpha) < 1 \tag{5.4}$$

so that the Doléans-Dade exponential of $-h(\alpha)$, a positive decreasing finite-valued process

$$\mathcal{E}(-h(\alpha)) = e^{-h(\alpha)} \prod_{s \le \cdot} (1 - \Delta h_s(\alpha)) e^{\Delta h_s(\alpha)}$$

is well defined and

$$\mathcal{E}(-h(\alpha))^{-1} = \mathcal{E}\left(\frac{1}{1 - \Delta h(\alpha)} \cdot h(\alpha)\right). \tag{5.5}$$

For these facts on the Hellinger processes we refer to [1], section 4.3; see also section 4.5, where the Hellinger process $h(\alpha)$ is characterized as the compensator of a certain (Q, F)-submartingale V of class (D) to be described next (see [1], sections 4.4 and 4.5 for more details).

The following notations will be used: if $\{X(\theta)\}_{\theta \in \Theta}$ is a certain parametric family of processes, then $a(X) = E_\alpha X(\vartheta)$ and (for a nonnegative family) $g(X) = e^{E_\alpha \log X(\vartheta)}$ denote its arithmetic and geometric mean processes, respectively (cf the special cases (3.6) and (3.8)). Denote by $\phi(X) = a(X) - g(X)$ the difference of the arithmetic and geometric process and note that this difference process is homogeneous in the sense that if C is a process independent of θ, then $\phi(CX) = C\phi(X)$. Note also that if the continuous part $X(\vartheta)^c$ possesses the *variance process*

$$v(X^c) \doteq var_\alpha X(\vartheta)^c = E_\alpha |X(\vartheta)^c|^2 - |E_\alpha X(\vartheta)^c|^2 \tag{5.6}$$

that is a (Q, F)-submartingale of class (D), then the compensator is given by

$$\tilde{v}(X^c) \doteq a(\langle X^c \rangle) - \langle a(X^c) \rangle. \tag{5.7}$$

Assume (2.1) and (2.6). Write m as a shorthand notation for $m(\vartheta, Q)$, a (Q,F)-uniformly integrable martingale given by (4.2). Let the process

$$V = \frac{1}{2}v(m^c) + \sum_{s\leq \cdot} \phi_s(1+\Delta m) \tag{5.8}$$

be a (Q,F)-submartingale. Then its compensator $\tilde{V}$ and the Hellinger process $h(\alpha)$ are Q-indistinguishable. This is exactly the assertion of theorem 4.7 in [1]. Upon further specification of the underlying model, we will be able in the next section to express this compensator in terms of the triplet of predictable characteristics of the observations.

5.2 Representation of Hellinger processes

In order to present the Hellinger processes explicitly, we need further specification of the randomized experiment in question. We turn therefore back to the setting of section 4.2 and suppose that a (Q,F)-semimartingale X is observed whose triplet of predictable characteristics is $T = (B, C, \nu)$. In addition to (2.1), assume that all local (Q,F)-martingales have the representation property relative to X, so that for each fixed $\theta \in \Theta$ the density process is represented as the Doléans-Dade exponential $z(\theta,Q) = \mathcal{E}(m(\theta,Q))$ of the (Q,F)-uniformly integrable martingale

$$m(\theta,Q) = \beta(\theta)\cdot X^c + \left(Y(\theta) - 1 + \frac{\hat{Y}(\theta) - \hat{1}}{1-\hat{1}}\right) * (\mu^X - \nu) \tag{5.9}$$

where $\beta(\theta) = \beta(\theta,Q)$ and $Y(\theta) = Y(\theta,Q)$ are the same as in section 4.2. According to the notation (3.4) the processes $\hat{1} = \hat{1}(Q)$ and $\hat{Y}(\theta) = \hat{Y}(\theta,Q)$ are associated with the third characteristics ν and $\nu(\theta)$ (cf (4.1)) so that

$$\hat{1}_t(\omega) = \nu(\omega;\{t\}\times\mathbb{R}) \quad \text{and} \quad \hat{Y}_t(\boldsymbol{\omega}) = \int_E Y_t(\omega,\theta,x)\nu(\omega,\{t\},dx) = \nu(\boldsymbol{\omega};\{t\}\times\mathbb{R}).$$

Add now the representation (5.9) to the conditions (2.1) and (2.6) of the previous section. It is needed to specify the compensator of V, i.e. the compensators of both terms in (5.8), that yields the Hellinger process $h(\alpha)$, as was noticed at the end of the preceding section. The result is asserted in the next theorem (cf [1], theorem 5.3; the proof is reproduced below, since the basic arguments are needed anew in section 7).

Theorem 5.2. *In the situation described in the previous section, assume (5.9). Then*

$$h(\alpha) = \frac{1}{2}v(\beta)\cdot C + \phi(Y)*\nu + \sum_{s\leq\cdot}\phi_s(1-\hat{Y}). \tag{5.10}$$

Proof. The first term in (5.8) is compensated as follows. The compensator $\tilde{v}(m^c)$ of the variance process $v(m^c)$ is $\tilde{v}(m^c) = v(\beta)\cdot C$. this is easily seen by applying (5.6) and (5.7) to $m(\theta,Q)^c = \beta(\theta)\cdot X^c$. Next, we have to show that the second term in (5.8) is compensated by the sum of the last two terms in (5.10), i.e. that

$$\sum_{s\leq\cdot}\phi_s(1+\Delta m) - \left\{\phi(Y)*\nu + \sum_{s\leq\cdot}\phi_s(1-\hat{Y})\right\} \tag{5.11}$$

is a (Q,F)-local martingale. But by the same considerations as in [5], Lemma IV.3.22, this claim holds true, provided

$$\phi(1+\Delta m)=\phi(Y(\cdot;\cdot,\Delta X))I_{\{\Delta X\neq 0\}}+\frac{\phi(1-\hat{Y})}{1-\hat{1}}I_{\{\Delta X=0\}}. \tag{5.12}$$

To prove (5.12), recall first the definition of the stochastic integral $W*(\mu^X-\nu)$: It is any purely discontinuous local martingale having the jumps $W(\cdot,\cdot,\Delta X)1_{\{\Delta X\neq 0\}}-\hat{W}$, cf [5], definition II.1.27 or [7], theorem 3.5.1. Apply this to the second term of $m(\theta,Q)$ in (5.9). We get

$$\begin{aligned}1+\Delta m(\theta,Q) &= 1+\{Y(\theta;\cdot,\Delta X)-1\}I_{\{\Delta X\neq 0\}}-\frac{\hat{Y}(\theta)-\hat{1}}{1-\hat{1}}I_{\{\Delta X=0\}}\\ &= Y(\theta;\cdot,\Delta X)I_{\{\Delta X\neq 0\}}+\frac{1-\hat{Y}(\theta)}{1-\hat{1}}I_{\{\Delta X=0\}}.\end{aligned} \tag{5.13}$$

From this we immediately obtain (5.12). The proof is complete. □

Remark 1. The explicit expression for the (Q,F)-local martingale (5.11) is given by the following decomposition:

$$\sum_{s\leq\cdot}\phi_s(1+\Delta m)=\left\{\phi(Y)-\frac{\phi(1-\hat{Y})}{1-\hat{1}}\right\}*(\mu^X-\nu)+\phi(Y)*\nu+\sum_{s\leq\cdot}\phi_s(1-\hat{Y}). \tag{5.14}$$

In section 5.4 we will make use of the following simple corollary.

Corollary 5.3. *Under the conditions of theorem 5.2, the positive valued process* $1-\Delta h(\alpha)$ *can be represented as follows:*

$$1-\Delta h(\alpha)=g(1-\hat{Y})+\hat{g}(Y). \tag{5.15}$$

Proof. We have already seen that the process on the left hand side of (5.15) is positive valued, cf (5.4). Here the notation (3.4) is used, so that for instance

$$\hat{g}_t(Y)(\omega)=\int_E e^{E_\alpha \log Y(\omega,\vartheta,x)}\nu(\omega;\{t\}\times dx).$$

Hence (5.10) yields $\Delta h(\alpha)=\phi(1-\hat{Y})+\hat{\phi}(Y)$ that is equivalent to (5.15), since $\phi(\cdot)=a(\cdot)-g(\cdot)$ and $a(1-\hat{Y})+\hat{a}(Y)=1$. The proof is complete. □

5.3 a-mean process as an exponential

In the setting of the previous section the a-mean and g-mean processes (defined in sections 3.2 and 3.3, respectively) have useful representations in terms of the Doléans-Dade exponentials. In this section we treat the a-mean process. The g-mean process will be treated in the next section. As was noticed in section 2.3 the a-mean process is in fact a certain density process, cf (3.5).

Theorem 5.4. *Under the conditions of theorem 5.2 the a-mean process defined by (3.6) (cf also (3.5)) may be represented as the Doléans-Dade exponential* $a(\alpha, Q) = \mathcal{E}(\bar{m}(\alpha, Q))$ *of the* (Q, F)*-uniformly integrable martingale*

$$\bar{m}(\alpha, Q) = \bar{\beta} \cdot X^c + \left(\bar{Y} - 1 + \frac{\hat{\bar{Y}} - \hat{1}}{1 - \hat{1}} \right) * (\mu^X - \nu) \tag{5.16}$$

with the posterior expectations $\bar{\beta}$ *and* $\bar{Y}$ *defined by (4.4).*

Proof. In virtue of corollary 4.2 it suffices to substitute β and Y in (5.9) by $\bar{\beta}$ and $\bar{Y}$ defined by (4.4). This yields (5.16). □

5.4 Multiplicative decomposition of g-mean process

In this section a multiplicative decomposition of the g-mean process will be presented (cf (5.20) or (5.22) below). In (5.21) a new dominating measure $G = G_\alpha$ occurs. This measure, called *geometric mean measure*, is defined as follows. Suppose once more that the observations constitute a semimartingale X that possesses the triplet of predictable characteristics $T = (B, C, \nu)$ with respect to the probability measure Q and the triplet $T(\theta) = (B(\theta), C(\theta), \nu(\theta))$ with respect to the probability measure $P_\theta, \theta \in \Theta$, cf (4.1). For any fixed α let $G = G_\alpha$ be a probability measure on the same space $(\Omega, \mathcal{F}, F)$, equivalent to Q, that prescribes to X the triplet of predictable characteristics $T^G = (B^G, C^G, \nu^G)$ where

$$\begin{cases} B^G &= a(B) + (Y^G - a(Y))\hbar * \nu \\ C^G &= C \\ \nu^G &= Y^G \cdot \nu \ \text{ with } \ Y^G = \frac{g(Y)}{g(1-\hat{Y}) + \hat{g}(Y)}. \end{cases} \tag{5.17}$$

Recall the notations of lemma 5.3 according to which $Y^G = \frac{g(Y)}{1-\Delta h(\alpha)}$. In the next theorem a characterization is given for the density process $z(G_\alpha, Q)$ which is defined at each $t \geq 0$ by $z_t(G_\alpha, Q) = E_Q\{\frac{dG_\alpha}{dQ}|\mathcal{F}_t\}$, cf [2].

Theorem 5.5. *Under the conditions of theorem 5.2 the density process* $z(G_\alpha, Q)$ *may be presented as a Doléans-Dade exponential*

$$z(G_\alpha, Q) = \mathcal{E}\left(\frac{1}{1 - \Delta h(\alpha)} \cdot N(\alpha, Q) \right) \tag{5.18}$$

of a (Q, F)*-uniformly integrable martingale*

$$N(\alpha, Q) = a(\beta) \cdot X^c + \left\{ g(Y) - \frac{g(1 - \hat{Y})}{1 - \hat{1}} \right\} * (\mu^X - \nu).$$

that is simply related to $M(\alpha, Q)$ *defined by (5.3):*

$$M(\alpha, Q) = g_{\_}(\alpha, Q) \cdot N(\alpha, Q). \tag{5.19}$$

Proof. The relation (5.19) follows from [1], lemma 5.4. From [7], theorem 2.5.1, and the decomposition (5.3) we get the multiplicative decomposition

$$g(\alpha, Q) = \mathcal{E}\left(\frac{1}{1 - \Delta h(\alpha)} \cdot N(\alpha, Q) \right) \mathcal{E}(-h(\alpha)) \tag{5.20}$$

with the process $N(\alpha, Q) = g_{\_}(\alpha, Q)^{-1} \cdot M(\alpha, Q)$ that clearly meets (5.19). From [1], theorem 5.6, we know that

$$g(\alpha, G) = \mathcal{E}(-h(\alpha)). \tag{5.21}$$

Substitute this in the expression on the right hand side of (5.20) and compare the result with the identity $g(\alpha, Q) = z(G, Q)\, g(\alpha, G)$, a consequence of (5.2). This confirms the desired equality (5.18). □

By (5.18), the multiplicative decomposition (5.20) of the g-mean process can be given an alternative form.

Corollary 5.6. *Under the conditions of theorem 5.2 the g-mean process possesses the multiplicative decomposition*

$$g(\alpha, Q) = z(G_\alpha, Q)\, \mathcal{E}(-h(\alpha)) \tag{5.22}$$

Proof. Apply theorem 5.5 to (5.20). □

Another important consequence is the following useful representation for the Hellinger integral.

Corollary 5.7. *Under the conditions of theorem 5.2*

$$H(\alpha, T) = E_G \mathcal{E}(-h(\alpha))_T.$$

Proof. Substitute Q in (5.1) by G and apply (5.21). □

5.5 Representation of a posterior information

At a stopping time $T > 0$, define by (3.11) the information $I(\alpha^T|\alpha)$ in the posterior α^T with respect to the prior α. It satisfies identity (3.12) and therefore we have

Corollary 5.8. *Under the conditions of theorem 5.2 the information $I(\alpha^T|\alpha)$ at a stopping time $T > 0$ can be presented as follows:*

$$e^{-I(\alpha^T|\alpha)} = z_T(G_\alpha, \bar{P}_\alpha)\, \mathcal{E}(-h(\alpha))_T \tag{5.23}$$

where the density process $z(G_\alpha, \bar{P}_\alpha)$ of the g-mean measure G_α with respect to the a-mean measure $\bar{P}_\alpha$ is the Doléans-Dade exponential

$$z(G_\alpha, \bar{P}_\alpha) = \mathcal{E}\left(\frac{1}{1 - \Delta h(\alpha)} \cdot N(\alpha, \bar{P}_\alpha)\right)$$

with

$$N(\alpha, \bar{P}_\alpha) = \big(a(\beta) - \bar{\beta}\big) \cdot X^c + \left\{ g\left(\frac{Y}{\bar{Y}}\right) - g\left(\frac{1 - \hat{Y}}{1 - \hat{\bar{Y}}}\right) \right\} * (\mu^X - \bar{\nu}) \tag{5.24}$$

where $\bar{\beta}$, $\bar{Y}$ and $\bar{\nu}$ are predictable characteristics of the observed process X with respect to the arithmetic mean measure $\bar{P}_\alpha$, as defined in section 4.2.

Proof. In view of identity (3.12) it suffices to substitute Q in (5.20) by $\bar{P}_\alpha$ and to verify that $N(\alpha, \bar{P}_\alpha)$ indeed has the asserted form. □

6. Examples

6.1 Processes with independent increments

Let the Hellinger process $h(\alpha)$ be deterministic. According to [5], remark VI.1.25, this is the case when the underlying process X has independent increments. It is directly seen from corollary 5.7 that in this case $H(\alpha, \cdot) = \mathcal{E}(-h(\alpha))$ at any instant t. Let us turn back to the setup of section 2.3 in which the a-mean measure $\bar{P} = \bar{P}_\alpha$ has been defined on a measurable space $(\Omega, \mathcal{F})$. Let us define on the same space another measure so that for each $B \in \mathcal{F}$

$$G(B) = \frac{\int_B g(\alpha, \omega) Q(d\omega)}{\int_\Omega g(\alpha, \omega) Q(d\omega)} \tag{6.1}$$

where the integrand is the geometric mean of p_ϑ given by (2.5), i.e. $g(\alpha, \omega) = e^{E_\alpha \log p_\vartheta(\omega)}$ (it depends on the dominating measure Q; as usual ω is suppressed). The denominator in (6.1) is the Hellinger integral $H(\alpha) = E_Q g(\alpha, Q)$, cf (5.1), so that we have the equality $g(\alpha, Q) = \frac{dG}{dQ} H(\alpha)$. The measure G just defined is in fact the geometric measure $G = G_\alpha$ of the preceding section. Take the logarithm of both sides in the latter equality and then the expectation with respect to Q. By (2.7) we get $I(\mathbf{P}|\mathbf{Q}) = I(G|Q) - \log H(\alpha)$. Note that by the Jensen inequality $H(\alpha) \le 1$, hence we have on the right the sum of two positive quantities. Moreover, with the special choice of the dominating measure $Q = G$ this equality reduces to $E_\alpha I(P_\vartheta|G) = -\log H(\alpha)$. We thus see that for any dominating measure Q we have

$$E_\alpha I(P_\vartheta|Q) \ge E_\alpha I(P_\vartheta|G), \tag{6.2}$$

which means that the g-mean measure G minimizes the average information. By the Jensen inequality the lower bound in (6.2) is estimated by $E_\alpha I(P_\vartheta|G) \ge I(\bar{P}|Q)$.

6.2 Discrete time

As confined to the special case of a discrete-time filtered space $(\Omega, \mathcal{F}, F = \{\mathcal{F}_n\}_{n\in\mathrm{N}})$, the present theory is quite straightforward. Let us therefore shortly review the results. Suppose that the present space is endowed with the family of probability measures $\{P_\theta\}_{\theta\in\Theta}$ that are all equivalent to a certain probability measure Q. Denote their restrictions to $\mathcal{F}_n$ by $\{P_{\theta,n}\}_{\theta\in\Theta}$ and Q_n. Often the n^{th} experiment is described by its outcomes, say vectors $(X_1, \ldots, X_n)$ that generate the σ-algebra $\mathcal{F}_n$, and the above restrictions are viewed as their distributions. For each n and $\theta \in \Theta$ denote by $z_n(\theta, Q)$ the density of $P_{\theta,n}$ with respect to Q_n. With the n^{th} experiment the posterior measure α^n is associated whose density with respect to the prior α is defined for each $\theta \in \Theta$ as follows:

$$\frac{d\alpha^n}{d\alpha}(\theta) = \frac{z_n(\theta, Q)}{\int_\Theta z_n(\theta, Q)\alpha(d\theta)} = z_n(\theta, \bar{P}) \tag{6.3}$$

where $\bar{P}_n$ is the restriction to $\mathcal{F}_n$ of the arithmetic mean measure. Its density with respect to Q is given by $a_n(\alpha, Q) = z_n(\bar{P}, Q) = E_\alpha z_n(\theta, Q)$, cf (3.5) and (3.6). This defines the a-mean sequence $a(\alpha, Q) = \{a_n(\alpha, Q)\}_{n\in\mathrm{N}}$. It is useful to express the g-mean sequence $g(\alpha, Q) = \{g_n(\alpha, Q)\}_{n\in\mathrm{N}}$ in terms of the geometric means $g_i = e^{E_\alpha \log r_i(\vartheta, Q)}$ of the ratios $r_i = z_i/z_{i-1}$, with convention $z_0 \equiv 1$. For we get $g_n(\alpha, Q) = g_1 \cdots g_n$. Obviously, this process $g(\alpha, Q)$ has the multiplicative decomposition (5.22) in discrete time, with

$$h_n(\alpha) = \sum_{i=1}^{n} E_Q \{1 - g_i|\mathcal{F}_{i-1}\}$$

and the density of the restriction to $\mathcal{F}_n$ of the g-mean measure G_α with respect to Q_n

$$z_n(G_\alpha, Q) = \prod_{i=1}^{n} \frac{g_i}{E_Q\{g_i|\mathcal{F}_{i-1}\}} = \mathcal{E}\left(\frac{1}{1-\Delta h(\alpha)} \cdot N(\alpha, Q)\right)_n$$

where $N_n(\alpha, Q) = \sum_{i=1}^n (g_i - E_Q\{g_i|\mathcal{F}_{i-1}\})$. In view of corollary 5.8, it is now easy to get the multiplicative decomposition (5.23) with

$$e^{-I(\alpha^n|\alpha)} = \frac{z_n(G_\alpha, Q)}{a_n(\bar{P}_\alpha, Q)} \prod_{i=1}^{n} E_Q\{g_i|\mathcal{F}_{i-1}\}. \tag{6.4}$$

6.3 Independent observations

Let $X_1, X_2, \ldots$ be a sequence of independent real-valued observations with X_i drawn according to a probability density (with respect to some σ-finite measure ρ) that belongs to a certain parametric family $\{f_i(\cdot, \theta)\}_{\theta\in\Theta}$. Suppose that for ρ-a.a. $x \in \mathbb{R}$

$$\gamma_i(x, \alpha) \doteq e^{E_\alpha \log f_i(x,\vartheta)} > 0,$$

so that by the Jensen inequality

$$0 < \Gamma_i(\alpha) \doteq \int_{-\infty}^{\infty} \gamma_i(x, \alpha)\rho(dx) < 1$$

(equality on the right hand side is excluded by the assumption that ϑ is nondegenerate under α) We can use the formulas of the preceding examples, taking into consideration the correspondence between the pairs g_i, $E_Q\{g_i|\mathcal{F}_{i-1}\}$ of section 6.2 and $\gamma_i(X_i, \alpha)$, $\Gamma_i(\alpha)$ of the present section. The Hellinger integral and the Hellinger sequence are then given by

$$H(\alpha, n) = \prod_{i=1}^{n} \Gamma_i(\alpha) \quad \text{and} \quad h_n(\alpha) = \sum_{i=1}^{n} (1 - \Gamma_i(\alpha))$$

with the relationship $H(\alpha, \cdot) = \mathcal{E}(-h(\alpha))$ as in section 6.1, since $h(\alpha)$ is deterministic. For a certain sample size n the posterior measure α^n on the parametric space is determined by its density with respect to the prior α

$$\frac{d\alpha^n}{d\alpha}(X_1, \ldots, X_n; \vartheta) = \frac{f_1(X_1, \vartheta) \cdots f_n(X_n, \vartheta)}{\int_\Theta f_1(X_1, \theta) \cdots f_n(X_n, \theta)\alpha(d\theta)}$$

where the denominator, denoted below by $a_n(X_1, \ldots, X_n; \alpha)$, is the density with respect to $\rho^{\otimes n}$ of the a-mean measure restricted to $\mathcal{F}_n$. The information in α^n given α has the representation

$$e^{-I(\alpha^n|\alpha)} = \frac{g_n(X_1, \ldots, X_n; \alpha)}{a_n(X_1, \ldots, X_n; \alpha)} H(\alpha, n)$$

where $g_n(X_1, \ldots, X_n; \alpha) = \prod_{i=1}^n \frac{\gamma_i(X_i, \alpha)}{\Gamma_i(\alpha)}$ is the density with respect to $\rho^{\otimes n}$ of the g-mean measure restricted to $\mathcal{F}_n$.

6.4 Diffusion

Let the observation process X be defined so that under each measure $P_\theta, \theta \in \Theta$,

$$X - \int_0^\cdot \beta_s(\theta)ds$$

is a Wiener process $W(\theta)$. Suppose that for each $s > 0$ the drift $\beta_s(\theta)$ has non-vanishing variance with respect to α, denoted as above by $v_s(\beta)$. Then the Hellinger processes

$$h(\alpha) = \frac{\sigma^2}{2} \int_0^\cdot v_s(\beta)ds$$

where σ^2 is the intensity of the Wiener processes $W(\theta), \theta \in \Theta$, are related to the Hellinger integrals evaluated at a certain stopping time T so that

$$H(\alpha, T) = E_G \mathcal{E}(-h(\alpha)) = E_G \left\{ e^{-\frac{\sigma^2}{2} \int_0^T v_s(\beta)ds} \right\}.$$

Under the g-mean measure $G = G_\alpha$

$$X - \int_0^\cdot a_s(\beta)ds$$

is a Wiener process. Hence, if under certain measure Q the observation process X itself is a Wiener process of intensity σ^2, then the density process of G_α with respect to Q is given by $z(G,Q) = e^{a(\beta)\cdot X - \frac{1}{2}\langle a(\beta)\cdot X \rangle}$ with $\langle a(\beta) \cdot X \rangle = \sigma^2 \int_0^\cdot a_s^2(\beta)ds$. Furthermore, the density process of $\bar{P}_\alpha$ with respect to Q is $z(\bar{P}, Q) = e^{\bar{\beta}\cdot X - \frac{1}{2}\langle \bar{\beta}\cdot X \rangle}$, cf (4.4). Hence, the density of the posterior $\alpha^\cdot$ with respect to the prior α is given by

$$\frac{d\alpha^\cdot}{d\alpha}(\theta) = e^{(\beta(\theta)-\bar{\beta})\cdot \bar{W} - \frac{1}{2}\langle (\beta(\theta)-\bar{\beta})\cdot \bar{W} \rangle}$$

and the information in $\alpha^\cdot$, given the prior α, satisfies

$$e^{-I(\alpha^\cdot|\alpha)} = e^{(a(\beta)-\bar{\beta})\cdot \bar{W} - \frac{1}{2}\langle (a(\beta)-\bar{\beta})\cdot \bar{W} \rangle} e^{-\frac{\sigma^2}{2} \int_0^\cdot v_s(\beta)ds}$$

where $\bar{W}$ is a Wiener process under the a-mean measure $\bar{P}_\alpha$.

6.5 Point processes

Consider a d-dimensional counting process $(N^1, \ldots, N^d)$ with the cumulative intensities $(\Lambda^1(\theta), \ldots, \Lambda^d(\theta))$ under the measure $P_\theta, \theta \in \Theta$. Suppose that the family $\{\Lambda^i(\theta)\}_{\theta\in\Theta}$ is equivalent to some positive increasing process A and that the densities $(Y^1(\theta), \ldots, Y^d(\theta))$ satisfy

$$E_\alpha \log \frac{Y_s^i(\vartheta)}{1 - \Delta\Lambda_s(\vartheta)} > -\infty \quad \text{with} \quad \Lambda = \Lambda^1 + \cdots + \Lambda^d$$

for all $s > 0$ and $i = 1, \ldots, d$. The Hellinger process of order α is given by

$$h(\alpha) = \int_0^\cdot \phi_s(\mathbf{Y})dA_s + \sum_{s\le\cdot} \phi_s(1 - \Delta\Lambda) \quad \text{with} \quad \phi(\mathbf{Y}) = \phi(Y^1) + \cdots + \phi(Y^d).$$

It is related to the Hellinger integral of order α as in the assertion of corollary 5.7 where the g-mean measure G_α is specified as follows: under G_α the intensity density (with respect to the same A) of N^i is

$$\frac{g(Y^i)}{1-\Delta h(\alpha)} = \frac{g(Y^i)}{g(\Delta\mathbf{\Lambda}) + g(1-\Delta A)} \quad \text{with } g(\Delta\mathbf{\Lambda}) = g(\Delta\Lambda^1) + \cdots + g(\Delta\Lambda^d). \tag{6.5}$$

Since under the a-mean measure $\bar{P}_\alpha$ the intensities are $(\bar{\Lambda}^1, \ldots, \bar{\Lambda}^d)$, with the densities $(\bar{Y}^1, \ldots, \bar{Y}^d)$ (w.r.t. A; cf (4.4)), we have

$$\frac{d\alpha^{\cdot}}{d\alpha}(\theta) = e^{-\Lambda(\theta)^c + \bar{\Lambda}^c} \prod_{s\le\cdot} \left(\frac{1-\Delta\Lambda_s(\theta)}{1-\Delta\bar{\Lambda}_s}\right)^{1-\Delta N_s} \prod_{i=1}^{d} \left(\frac{Y^i_s(\theta)}{\bar{Y}^i_s}\right)^{\Delta N^i_s}$$

with $N = N^1 + \cdots + N^d$ and $\bar{\Lambda} = \bar{\Lambda}^1 + \cdots + \bar{\Lambda}^d$. Finally, by (6.5)

$$e^{-I(\alpha^{\cdot}|\alpha)} = e^{\phi(\Lambda)^c} \mathcal{E}(-h(\alpha)) \prod_{s\le\cdot} \frac{1}{1-\Delta h_s(\alpha)}\, g_s\left(\frac{1-\Delta\Lambda}{1-\Delta\bar{\Lambda}}\right)^{1-\Delta N_s} \prod_{i=1}^{d} g_s\left(\frac{Y^i}{\bar{Y}^i}\right)^{\Delta N^i_s}.$$

7. Cumulative information in the posterior given a prior

7.1 Definitions; discrete time

Parallel to (6.4), it seems instructive to trace how the information accumulates in case of the discrete time parameter, by first defining the amount of information provided by a single observation X_i, for any $i = 1, 2, \ldots$, that is the Kullback-Leibler information in the posterior α^i, given α^{i-1}:

$$I(\alpha^i|\alpha^{i-1}) = E_{\alpha^{i-1}} \log \frac{d\alpha^{i-1}}{d\alpha^i}(\vartheta), \tag{7.1}$$

$\alpha^0 = \alpha$ is understood as the *a priori* measure, of course. Then the cumulative information in the n^{th} experiment $\bar{I}(\alpha^n|\alpha)$ is defined by summing up the amount of information provided by each individual observation $X_i, i = 1, \ldots, n$, i.e.

$$\bar{I}(\alpha^n|\alpha) = \sum_{i=1}^{n} I(\alpha^i|\alpha^{i-1}). \tag{7.2}$$

In view of (6.3), the Kullback-Leibler information defined by (7.1) may be rewritten in terms of the densities $z_i(P_\theta, \bar{P}_\alpha)$ as follows $I(\alpha^i|\alpha^{i-1}) = -E_\alpha\{z_{i-1}(P_\vartheta, \bar{P}_\alpha)\, \Delta \log z_i(P_\vartheta, \bar{P}_\alpha)\}$. Thus the expression (7.2) for the cumulative information may be abbreviated to

$$\bar{I}(\alpha^n|\alpha) = -\sum_{i=1}^{n} E_\alpha\{z_{i-1}\, \Delta \log z_i\} = -E_\alpha(z_{\_} \cdot \log z)_n \tag{7.3}$$

with $z_i = z_i(P_\vartheta, \bar{P}_\alpha)$. The stochastic integral (truly the sum, as time is discrete) on the right hand side of (7.3) is written in the form usual in stochastic calculus - it proves useful in the general setting we are going to treat next.

7.2 Multiplicative decomposition
The considerations of the preceding section extend to the general case of filtered statistical experiments (3.1). For each $t > 0$ and $\theta \in \Theta$ define the density

$$\frac{d\alpha^{t-}}{d\alpha^t}(\theta) = \frac{z_{t-}(\theta, \bar{P}_\alpha)}{z_t(\theta, \bar{P}_\alpha)}$$

that agrees with (3.7). Then the Kullback-Leibler information in the posterior measure α^t with respect to α^{t-} is given by

$$I(\alpha^t|\alpha^{t-}) = E_{\alpha^{t-}} \log \frac{d\alpha^{t-}}{d\alpha^t}(\vartheta) \tag{7.4}$$

Analogously to (7.3), we define the cumulative information $\bar{I}(\alpha^T|\alpha)$ at any stopping time T by

$$\bar{I}(\alpha^T|\alpha) = -E_\alpha(z_{\_} \cdot \log z)_T \tag{7.5}$$

with $z_t = z_t(\vartheta, \bar{P}_\alpha)$. We intend to show that the cumulative information $\bar{I}(\alpha^T|\alpha)$ satisfies a relation similar to (5.23) - a certain multiplicative decomposition. Assume therefore the conditions of section 5.4. The predictable part of the new multiplicative decomposition will involve a new (dynamic version of) Hellinger process $\bar{h}(\alpha)$ that is defined similarly to $h(\alpha)$ of (5.10), however with the expectations E_{α^-} instead of E_α. In order to carry out these substitutions we will extend as follows our previous notations concerning an arbitrary parametric family of processes $\{X(\theta)\}_{\theta\in\Theta}$: put $\bar{X} \doteq \bar{a}(X) = E_{\alpha^-}X(\vartheta)$ and $\bar{v}(X) \doteq var_{\alpha^-}X(\vartheta) = \bar{a}(|X - \bar{X}|^2)$ for the first two predictable posterior moments (when they exist, of course). For positive valued processes put $\bar{g}(X) = e^{\bar{a}(\log X)}$ and $\bar{\phi}(X) = \bar{a}(X) - \bar{g}(X)$. Then the predictable finite-valued increasing process $\bar{h}(\alpha)$ is defined by

$$\bar{h}(\alpha) = \frac{1}{2}\bar{v}(\beta) \cdot C + \bar{\phi}(Y) * \nu + \sum_{s\leq \cdot} \bar{\phi}_s(1 - \hat{Y}), \tag{7.6}$$

cf (5.10). We will need also a new (dynamic version of) g-mean measure $\bar{G}_\alpha$ that is defined analogously to G_α, with the same substitutions as above. In the situation described in the beginning of section 5.4, suppose that the measure $\bar{G}_\alpha$ prescribes to semimartingale observations X the triplet of predictable characteristics $T^{\bar{G}} = (B^{\bar{G}}, C^{\bar{G}}, \nu^{\bar{G}})$

$$\begin{cases} B^{\bar{G}} &= \bar{B} + (Y^{\bar{G}} - \bar{Y})\hbar * \nu \\ C^{\bar{G}} &= C \\ \nu^{\bar{G}} &= Y^{\bar{G}} \cdot \nu \ \text{ with } \ Y^{\bar{G}} = \frac{\bar{g}(Y)}{\bar{g}(1-\hat{Y}) + \hat{\bar{g}}(Y)}. \end{cases} \tag{7.7}$$

Theorem 7.1. *Under the conditions of theorem 5.2 the cumulative information $\bar{I}(\alpha^T|\alpha)$ at a stopping time $T > 0$, defined by (7.5), can be presented as follows:*

$$e^{-\bar{I}(\alpha^T|\alpha)} = z_T(\bar{G}_\alpha, \bar{P}_\alpha)\, \mathcal{E}\left(-\bar{h}(\alpha)\right)_T \tag{7.8}$$

with the Hellinger process $\bar{h}(\alpha)$ given by (7.6) and the g-mean measure $\bar{G}_\alpha$ prescribing to observations the triplet given by (7.7). The density process $z(\bar{G}_\alpha, \bar{P}_\alpha)$ of the g-mean measure $\bar{G}_\alpha$ with respect to the a-mean measure $\bar{P}_\alpha$ is the Doléans-Dade exponential

$$z(\bar{G}_\alpha, \bar{P}_\alpha) = \mathcal{E}\left(\frac{1}{1 - \Delta\bar{h}(\alpha)} \cdot \bar{N}(\alpha, \bar{P}_\alpha)\right) \tag{7.9}$$

with

$$\bar{N}(\alpha, \bar{P}_\alpha) = \left(\frac{\bar{g}(Y)}{\bar{Y}} - \frac{\bar{g}(1-\hat{Y})}{1-\hat{\bar{Y}}} \right) * (\mu^X - \bar{\nu}) \tag{7.10}$$

where $\bar{\beta}$, $\bar{Y}$ and $\bar{\nu}$ are predictable characteristics of the observed process X with respect to the arithmetic mean measure $\bar{P}_\alpha$, as defined in section 4.2.

The proof of this theorem is preceded by the following two lemmas.

Lemma 7.2. *Under the conditions of theorem 7.1*

$$e^{-\bar{I}(\alpha^T|\alpha)} = \mathcal{E}(-S(\alpha, \bar{P}_\alpha))_T \tag{7.11}$$

where

$$S(\alpha, \bar{P}_\alpha) = \frac{1}{2}\bar{v}(\beta) \cdot C + \sum_{s \le \cdot} \bar{\phi}_s(1 + \Delta m(\cdot, \bar{P}_\alpha)) \tag{7.12}$$

with

$$m(\vartheta, \bar{P}_\alpha) = (\beta(\vartheta) - \bar{\beta}) \cdot X^c + \left(\frac{Y(\vartheta)}{\bar{Y}} - 1 + \frac{\hat{Y}(\vartheta) - \hat{\bar{Y}}}{1 - \hat{\bar{Y}}} \right) * (\mu^X - \bar{\nu}). \tag{7.13}$$

Proof. Since $\langle S^c \rangle = 0$ and $\Delta S = \bar{\phi}(1+\Delta m(\cdot, \bar{P}_\alpha)) = 1 - \bar{g}(1+\Delta m(\cdot, \bar{P}_\alpha))$ (here the shorthand notation $S = S(\alpha, \bar{P}_\alpha)$ is used), we have by the usual exponential formula that

$$\mathcal{E}(-S) = e^{-S - \frac{1}{2}\langle S^c \rangle} \prod_{s \le \cdot} (1 - \Delta S_s) e^{\Delta S_s} = e^{-\frac{1}{2}\bar{v}(\beta) \cdot C} \prod_{s \le \cdot} \bar{g}_s(1 + \Delta m(\cdot, \bar{P}_\alpha)).$$

So, we need to prove that

$$e^{-\bar{I}(\alpha^T|\alpha)} = e^{-\frac{1}{2}\bar{v}(\beta) \cdot C_T} \prod_{s \le T} \bar{g}_s(1 + \Delta m(\cdot, \bar{P}_\alpha)). \tag{7.14}$$

The definition of the cumulative information (7.5) involves the integral with respect to $\log z(\vartheta, \bar{P}_\alpha) = \log \mathcal{E}\left(m(\vartheta, \bar{P}_\alpha)\right)$, i.e. with respect to

$$m(\vartheta, \bar{P}_\alpha) - \frac{1}{2}\langle m(\vartheta, \bar{P}_\alpha)^c \rangle + \sum_{s \le \cdot} \{\log(1 + \Delta m_s(\vartheta, \bar{P}_\alpha)) - \Delta m_s(\vartheta, \bar{P}_\alpha)\}. \tag{7.15}$$

But the integral with respect to $m(\vartheta, \bar{P}_\alpha)$ vanishes, since by $z(\vartheta, \bar{P}_\alpha) = 1 + z_-(\vartheta, \bar{P}_\alpha) \cdot m(\vartheta, \bar{P}_\alpha)$ and by the identity $E_\alpha z(\vartheta, \bar{P}_\alpha) = a(\alpha, \bar{P}_\alpha) \equiv 1$ of section 3.2, we have $E_\alpha z_-(\vartheta, \bar{P}_\alpha) \cdot m(\vartheta, \bar{P}_\alpha) = E_\alpha z(\vartheta, \bar{P}_\alpha) - 1 = 0$ (note the following direct way to verify this: evaluate $E_\alpha z_-(\vartheta, \bar{P}_\alpha) \cdot m(\vartheta, \bar{P}_\alpha)$ to get 0 by using (4.4) and (7.13)). Besides, we have

$$E_\alpha z_-(\vartheta, \bar{P}_\alpha) \cdot \langle m(\vartheta, \bar{P}_\alpha)^c \rangle = \bar{a}(|\beta - \bar{\beta}|^2) \cdot \langle X^c \rangle = \bar{v}(\beta) \cdot C \tag{7.16}$$

and

$$\sum_{s \le \cdot} \bar{a}(\log(1 + \Delta m(\cdot, \bar{P}_\alpha)) = \sum_{s \le \cdot} \log \bar{g}_s(1 + \Delta m(\cdot, \bar{P}_\alpha)). \tag{7.17}$$

On determining thus the cumulative information (by evaluating the expectation E_α of the integral in (7.5) with respect to (7.15)) we get only the contributions corresponding to (7.16) and (7.17):

$$\bar{I}(\alpha^T|\alpha) = \frac{1}{2}\bar{v}(\beta)\cdot C_T - \sum_{s\le T}\log \bar{g}_s(1+\Delta m(\cdot,\bar{P}_\alpha)).$$

This yields (7.14). □

Lemma 7.3. *Under the conditions of theorem 7.1 the $(\bar{P},F)$-supermartingale $S(\alpha,\bar{P}_\alpha)$ of class (D), defined by (7.12), has the following Doob-Meyer decomposition*

$$S(\alpha,\bar{P}_\alpha) = -\bar{N}(\alpha,\bar{P}_\alpha) + \bar{h}(\alpha), \tag{7.18}$$

cf (7.6) and (7.10).

Proof. In view of (7.6), (7.10) and (7.12), the decomposition (7.18) is equivalent to

$$\begin{aligned}\sum_{s\le\cdot}\bar{\phi}_s(1+\Delta m(\cdot,\bar{P}_\alpha)) &= \left\{\frac{\bar{\phi}(Y)}{\bar{Y}} - \frac{\bar{\phi}(1-\hat{Y})}{1-\hat{\bar{Y}}}\right\} * (\mu^X - \bar{\nu}) \\ &+ \bar{\phi}(Y)*\nu + \sum_{s\le\cdot}\bar{\phi}_s(1-\hat{Y}).\end{aligned} \tag{7.19}$$

Note that in the first term (and in (7.10)) the substitution of $\bar{\phi}$ by $-\bar{g}$ is allowed. Compare the latter equation with (5.14) to conclude that by the same arguments as in the course of proving theorem 5.2, we only need to show the following relationship between the jumps of the $(\bar{P}_\alpha,F)$-local martingale $m(\vartheta,\bar{P}_\alpha)$ and the observed process X:

$$\begin{aligned}1+\Delta m(\vartheta,\bar{P}_\alpha) &= 1+\left\{\frac{Y(\vartheta;\cdot,\Delta X)}{\bar{Y}(\cdot,\cdot,\Delta X)} - 1\right\} I_{\{\Delta X\neq 0\}} - \frac{\hat{Y}(\vartheta)-\hat{\bar{Y}}}{1-\hat{\bar{Y}}} I_{\{\Delta X=0\}} \\ &= \frac{Y(\cdot,\cdot,\Delta X)}{\bar{Y}(\cdot,\cdot,\Delta X)} I_{\{\Delta X\neq 0\}} + \frac{1-\hat{Y}(\vartheta)}{1-\hat{\bar{Y}}} I_{\{\Delta X=0\}},\end{aligned}$$

which is derived similarly to (5.13). Thus we have

$$\bar{\phi}(1+\Delta m(\cdot,\bar{P}_\alpha)) = \frac{\bar{\phi}(Y(\cdot,\cdot,\Delta X))}{\bar{Y}(\cdot,\cdot,\Delta X)} I_{\{\Delta X\neq 0\}} + \frac{\bar{\phi}(1-\hat{Y})}{1-\hat{\bar{Y}}} I_{\{\Delta X=0\}},$$

which implies (7.19) exactly in the same manner as (5.12) implies (5.14). □

7.3 Proof of theorem 7.1

The lemmas 7.2 and 7.3 provide key arguments in the course of the following

Proof. Note that in view of (4.3) and (7.7) the drift coefficient of X^c is $\bar{\beta}$ under both $\bar{P}_\alpha$ or $\bar{G}_\alpha$. Therefore, there is no continuous part in the expression (7.10) for the $(\bar{P}_\alpha,F)$-uniformly integrable martingale $\bar{N}(\alpha,\bar{P}_\alpha)$. Note also that by the same considerations as above, cf corollary 5.3, we have $1-\Delta\bar{h}(\alpha) = \bar{g}(1-\hat{Y}) + \hat{\bar{g}}(Y)$. Hence $Y^{\bar{G}} = \frac{\bar{g}(Y)}{1-\Delta\bar{h}(\alpha)}$ and

$\hat{Y}^{\bar{G}} = \frac{\hat{\bar{g}}(Y)}{1-\Delta\bar{h}(\alpha)}$. Taking these equalities into consideration, one can easily reduce the usual exponential representation for the density process $z(\bar{G}_\alpha, \bar{P}_\alpha) = \mathcal{E}\left(m(\bar{G}_\alpha, \bar{P}_\alpha)\right)$ with

$$m(\bar{G}_\alpha, \bar{P}_\alpha) = \left(\frac{Y^{\bar{G}}}{\bar{Y}} - 1 + \frac{\hat{Y}^{\bar{G}} - \hat{\bar{Y}}}{1 - \hat{\bar{Y}}}\right) * (\mu^X - \bar{\nu})$$

to (7.9) with (7.10). In view of (5.5), (7.9) and (7.11) the desired equation (7.8) is equivalent to

$$\mathcal{E}\left(\frac{1}{1-\Delta\bar{h}(\alpha)} \cdot \bar{N}(\alpha, \bar{P}_\alpha)\right) = \mathcal{E}(-S(\alpha, \bar{P}_\alpha))\,\mathcal{E}\left(\frac{1}{1-\Delta\bar{h}(\alpha)} \cdot \bar{h}(\alpha)\right) \tag{7.20}$$

i.e. to

$$\frac{1}{1-\Delta\bar{h}(\alpha)} \cdot \bar{N}(\alpha, \bar{P}_\alpha) = -S(\alpha, \bar{P}_\alpha) + \frac{1}{1-\Delta\bar{h}(\alpha)} \cdot (\bar{h}(\alpha) - [S(\alpha, \bar{P}_\alpha), \bar{h}(\alpha)]), \tag{7.21}$$

since the product of exponentials in (7.20) is itself an exponential, namely the exponential of the process on the right. But (7.21) is an easy consequence of the decomposition (7.18). □

References

1. K. Dzhaparidze, P.J.C. Spreij and E. Valkeila (1997), On Hellinger processes for parametric families of experiments. *Statistics and Control of Stochastic Processes*. The Liptser Festschrift (eds. Kabanov et al.), World Scientific, 41-62.

2. B. Grigelionis (1990), Hellinger integrals and Hellinger processes for solutions of martingale problems, *Prob. Theory and Math. Stat.* vol 1 (eds. Grigelionis et al.), VSP/Mokslas., 446-454.

3. J. Jacod (1989), Filtered statistical models and Hellinger processes, *Stochastic Processes and their Appl.* **32**, 3-45.

4. J. Jacod (1989), Convergence of filtered statistical models and Hellinger processes, *Stochastic Processes and their Appl.* **32**, 47-68.

5. J. Jacod and A.N. Shiryaev (1987), *Limit Theorems for Stochastic Processes*, Springer.

6. E.I. Kolomiets (1984), Relations between triplets of local characteristics of semimartingales, *Russian Math. Surveys* **39**, 123-124.

7. R.Sh. Liptser and A.N. Shiryayev (1986), *Theory of Martingales*, Kluwer.

8. M. Loève (1977, 1978), *Probability Theory*, Springer.

9. H. Strasser (1985), *Mathematical Theory of Statistics*, de Gruyter.