KYBERNETIKA — VOLUME *48* (2012), NUMBER 4, PAGES 750–767

LIMITS OF BAYESIAN DECISION RELATED QUANTITIES OF BINOMIAL ASSET PRICE MODELS

Wolfgang Stummer and Wei Lao

In commemoration of our dear friend and wonderful colleague Igor Vajda.

We study Bayesian decision making based on observations $\left(X_{n,t} : t \in \{0, \frac{T}{n}, 2\frac{T}{n}, \ldots, n\frac{T}{n}\}\right)$ $(T > 0,\ n \in \mathbb{N})$ of the discrete-time price dynamics of a financial asset, when the hypothesis a special n-period binomial model and the alternative is a different n-period binomial model. As the observation gaps tend to zero (i. e. $n \to \infty$), we obtain the limits of the corresponding Bayes risk as well as of the related Hellinger integrals and power divergences. Furthermore, we also give an example for the "non-commutativity" between Bayesian statistical and optimal investment decisions.

Keywords: Bayesian decisions, power divergences, Cox–Ross–Rubinstein binomial asset price models

Classification: 62C10, 94A17, 91B25

1. INTRODUCTION

Let $X_{n,t} > 0$ be the value of a financial asset (resp. an economic quantity of interest) at time $t \in \tau_n := \{0, \frac{T}{n}, 2\frac{T}{n}, \ldots, n\frac{T}{n}, (n+1)\frac{T}{n}, (n+2)\frac{T}{n}, \ldots, N_n\frac{T}{n}\}\backslash\{\infty\}$ for some $T > 0$, $n \in \mathbb{N}$, as well as the integer part $N_n := \lfloor \widehat{T}\frac{n}{T} \rfloor$ of $\widehat{T}\frac{n}{T}$ with a (goal-dependent) *final* time horizon $\widehat{T} \in]T, \infty]$. Clearly, $N_n\frac{T}{n} \le \widehat{T}$ and $\lim_{n\to\infty} N_n\frac{T}{n} = \widehat{T}$. In the following, the parameter T will play the role of the fixed *observation* time horizon and the parameter n denotes the number of equally spaced observation times between 0 (e. g. now) and T; n will either be fixed or converge to infinity. We use the appropriate canonical sample path space Ω_n and the description $X_{n,0} := x > 0$, $X_{n,t} := x \cdot \prod_{i=1}^{tn/T} Y_{n,i}$ $(t \in \tau_n\backslash\{0\})$, where the capital-growth-factor describing random variables $Y_{n,i}$ can only take the two values $d_{n,i} > 0$ and $u_{n,i} > d_{n,i}$ $(i \in \mathbb{N})$. For the sake of consistency with arbitrage theory, we always choose $u_{n,i} > 1 + \rho_{n,i} > d_{n,i}$ for some realistic interest rate $\rho_{n,i} > 0$ for the ith time-period $](i-1)\frac{T}{n}, i\frac{T}{n}]$. Notice that this general setup covers also situations where the $u_{n,i}$ respectively $d_{n,i}$ depend on all previous capital-growth-factors $(u_{n,j})_{j=1,\ldots,i-1}$ respectively $(d_{n,j})_{j=1,\ldots,i-1}$.

Furthermore, suppose that for the corresponding quantification of the asset value dynamics at fixed $n \in \mathbb{N}$ we have the following two choices $(\mathcal{H})$ and $(\mathcal{A})$:

$(\mathcal{H})$ the probability law Q_n (on the sample path space) under which the random variables

$(Y_{n,i})_{i\in\mathbb{N}}$ are independent and distributed according to $Q_n[Y_{n,i} = u_{n,i}] = q_{n,i} = 1 - Q_n[Y_{n,i} = d_{n,i}]$ with $q_{n,i} \in]0,1[$ for all $i \in \{1,\ldots,N_n\}$.

$(\mathcal{A})$ the probability law P_n (on the sample path space) under which the random variables $(Y_{n,i})_{i\in\mathbb{N}}$ are independent and distributed according to $P_n[Y_{n,i} = u_{n,i}] = p_{n,i} = 1 - P_n[Y_{n,i} = d_{n,i}]$ with $p_{n,i} \in]0,1[$ for all $i \in \{1,\ldots,N_n\}$.

In order to avoid trivialities, we assume that $q_{n,i} \neq p_{n,i}$ holds for at least one $i \in \{1,\ldots,n\}$.

In the hypothesis model $\mathcal{H}$, the random process $(X_{n,t})_{t\in\tau_n}$ is a non-homogeneous Cox–Ross–Rubinstein-type (CRR [4]) model "evolving on a certain tree" with probabilities $q_{n,i}$ resp. $1-q_{n,i}$ in the ith time-period (independently of the other periods). In contrast, in the alternative model $\mathcal{A}$ the random process $(X_{n,t})_{t\in\tau_n}$ is a different non-homogeneous CRR model "evolving on the same tree" but with different probabilities $p_{n,i}$ resp. $1-p_{n,i}$ in the ith time-period.

Within such a framework, we study dichotomous Bayes decisions with possible actions $\theta_{\mathcal{H}}$ and $\theta_{\mathcal{A}}$ as well as loss functions

$$\begin{pmatrix} L\Big(\theta_{\mathcal{H}},\mathcal{H}\Big) & L\Big(\theta_{\mathcal{H}},\mathcal{A}\Big) \\ L\Big(\theta_{\mathcal{A}},\mathcal{H}\Big) & L\Big(\theta_{\mathcal{A}},\mathcal{A}\Big) \end{pmatrix} = \begin{pmatrix} \widetilde{L}_{\mathcal{H}} & L_{\mathcal{A}} \\ L_{\mathcal{H}} & \widetilde{L}_{\mathcal{A}} \end{pmatrix}, \tag{1}$$

with losses $L_{\mathcal{H}} > 0$, $L_{\mathcal{A}} > 0$, $\widetilde{L}_{\mathcal{H}} \in [0, L_{\mathcal{H}}[$, $\widetilde{L}_{\mathcal{A}} \in [0, L_{\mathcal{A}}[$ (the latter two upper bounds are not essential but rule out obvious cases).

More detailed, our Bayes decisions about the hypothesis $\mathcal{H}$ against the alternative $\mathcal{A}$ are based on the random asset value observations $\mathcal{X}_{n,T} := (X_{n,t} : t \in \{0, \frac{T}{n}, 2\frac{T}{n}, \ldots, n\frac{T}{n}\})$ between the $n-$*independent fixed* times 0 and $n\frac{T}{n} = T$, and thus can be formally considered as functions $\delta(\mathcal{X}_{n,T})$ of random paths $\mathcal{X}_{n,T}$ into the decision space $\mathcal{D} = \{\theta_{\mathcal{H}}, \theta_{\mathcal{A}}\}$. The Bayes decision function minimizes the risk (average loss)

$$\begin{aligned} &pr^{\mathcal{H}}\ L_{\mathcal{H}}\ Pr[\ \delta(\mathcal{X}_{n,T}) = \theta_{\mathcal{A}} \mid \mathcal{H}\ ] \ +\ pr^{\mathcal{A}}\ L_{\mathcal{A}}\ Pr[\ \delta(\mathcal{X}_{n,T}) = \theta_{\mathcal{H}} \mid \mathcal{A}\ ] \\ &+ pr^{\mathcal{H}}\ \widetilde{L}_{\mathcal{H}}\ Pr[\ \delta(\mathcal{X}_{n,T}) = \theta_{\mathcal{H}} \mid \mathcal{H}\ ] \ +\ pr^{\mathcal{A}}\ \widetilde{L}_{\mathcal{A}}\ Pr[\ \delta(\mathcal{X}_{n,T}) = \theta_{\mathcal{A}} \mid \mathcal{A}\ ] \end{aligned} \tag{2}$$

for given prior probabilities $pr^{\mathcal{H}} = Pr[\mathcal{H}] > 0$ for $\mathcal{H}$ and $pr^{\mathcal{A}} = Pr[\mathcal{A}] = 1 - pr^{\mathcal{H}} > 0$ for $\mathcal{A}$ (which describe the model risk knowledge at time $t = 0$, prior to the random asset value observations $\mathcal{X}_{n,T}$). According to the very nature of the underlying decision *goals*, we have assumed $T < \widehat{T}$ so that the process $X_{n,\cdot}$ lives beyond the observation time horizon T.

Within this setup, we compute the limits as $n \to \infty$ of the following quantities: decision-theoretic characteristics in form of Bayes factor moments and some related general functionals (Section 2), the Bayes risk (Section 4), as well as the related power divergences between the two laws P_n and Q_n at choice (Section 3); the total variation distance will be dealt with, too. These results differ from other statistical applications of power divergences (Cressie–Read measures, generalized cross-entropy family) and related quantities given e. g. the recent articles of Liese, Morales and Vajda [15], Vajda and Zvárová [29], Csiszár and Matúš [6], Harremoes and Vajda [11, 12], Hobza, Pardo and Morales [13], Broniatowski and Vajda [3], Morales and Vajda [19], Stummer and Vajda [27], Vajda and van der Meulen [28], Berlinet and Vajda [2], Gretton and Györfi [10], Pardo [22], and the numerous references therein; see also the the surveys of e. g. Maasoumi [18], Golan [9], Csiszár and Shields [7] and Liese and Vajda [17] as well as the books of e. g. Liese and Vajda [16], Read and Cressie [23], Stummer [25], Pardo [21] and Liese and Miescke [14].

Some of the abovementioned $n-$limits turn out to be consistent with the purely continuous-time investigations of Stummer and Vajda [26] about Bayesian decisions where the hypothesis $\mathcal{H}$ is a geometric Brownian motion with growth constant $c_{\mathcal{H}}$ and volatility σ, and the alternative $\mathcal{A}$ is another geometric Brownian motion with growth constant $c_{\mathcal{A}}$ and volatility σ. Moreover, we give several examples including one which shows the "non-commutativity" between Bayesian statistical and optimal investment decisions.

2. BAYES FACTOR MOMENTS

In a straightforward way, one can obtain from the prior binomial probabilities $pr^{\mathcal{H}}$ for $\mathcal{H}$ and $pr^{\mathcal{A}} = 1 - pr^{\mathcal{H}}$ for $\mathcal{A}$ the posterior probabilities

$$pr_{n,T}^{\mathcal{H},\text{post}} = \frac{pr^{\mathcal{H}}}{(1-pr^{\mathcal{H}})\ Z_{n,T}\ +\ pr^{\mathcal{H}}} \qquad \text{for } \mathcal{H}\ , \tag{3}$$

$$pr_{n,T}^{\mathcal{A},\text{post}} = \frac{(1-pr^{\mathcal{H}})\ Z_{n,T}}{(1-pr^{\mathcal{H}})\ Z_{n,T}\ +\ pr^{\mathcal{H}}} \qquad \text{for } \mathcal{A}\ , \tag{4}$$

with $Z_{n,T} = Z_{n,T}(\mathcal{X}_{n,T})$

$$= \begin{cases} \frac{p_{n,1}\cdot\ldots\cdot p_{n,n}}{q_{n,1}\cdot\ldots\cdot q_{n,n}}, & \text{if } X_{n,T} = x\,u_{n,1}\ldots u_{n,n}, \\ \frac{p_{n,1}\cdot\ldots\cdot p_{n,n-1}\,(1-p_{n,n})}{q_{n,1}\cdot\ldots\cdot q_{n,n-1}\,(1-q_{n,n})}, & \text{if } X_{n,T} = x\,u_{n,1}\ldots u_{n,n-1}\,d_{n,n}, \\ \frac{p_{n,1}\cdots p_{n,n-2}\,(1-p_{n,n-1})\,p_{n,n}}{q_{n,1}\cdots q_{n,n-2}\,(1-q_{n,n-1})\,q_{n,n}}, & \text{if } X_{n,T} = x\,u_{n,1}\ldots u_{n,n-2}\,d_{n,n-1}\,u_{n,n}, \\ \frac{p_{n,1}\cdots p_{n,n-2}\,(1-p_{n,n-1})\,(1-p_{n,n})}{q_{n,1}\cdots q_{n,n-2}\,(1-q_{n,n-1})\,(1-q_{n,n})}, & \text{if } X_{n,T} = x\,u_{n,1}\ldots u_{n,n-2}\,d_{n,n-1}\,d_{n,n}, \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ \frac{(1-p_{n,1})\ldots(1-p_{n,n})}{(1-q_{n,1})\ldots(1-q_{n,n})}, & \text{if } X_{n,T} = x\,d_{n,1}\ldots d_{n,n}. \end{cases} \tag{5}$$

Notice that the dependence of $pr_{n,T}^{\mathcal{H},\text{post}}$ and $pr_{n,T}^{\mathcal{A},\text{post}}$ on the observed asset value sample path $\mathcal{X}_{n,T}$ is not indicated explicitly here. As usual, the posterior odds ratio of $\mathcal{A}$ to $\mathcal{H}$ is obtained by

$$\frac{pr_{n,T}^{\mathcal{A},\text{post}}}{pr_{n,T}^{\mathcal{H},\text{post}}} = \frac{1-pr^{\mathcal{H}}}{pr^{\mathcal{H}}}\ Z_{n,T}\ .$$

Furthermore, the odds for $\mathcal{A}$ against $\mathcal{H}$ that are given by the asset-value sample paths $\mathcal{X}_{n,T}$ observed on the set of times $\widetilde{\tau}_{n,T} := \{0, \frac{T}{n}, 2\frac{T}{n}, \dots n\frac{T}{n}\}$ are reflected by the corresponding Bayes factor $\mathcal{B}_{n,T} := \frac{\text{posterior odds ratio of } \mathcal{A} \text{ to } \mathcal{H}}{\text{prior odds ratio of } \mathcal{A} \text{ to } \mathcal{H}} = Z_{n,T}$. It is obvious that for $\alpha \in \mathbb{R}$ the αth moment (with respect to the hypothesis measure Q_n) of the Bayes factor is nothing else but the appropriate Hellinger integral H_α; more precisely,

$$H_\alpha(P_{n,T} \| Q_{n,T}) := \int \left\{ \frac{\mathrm{d}P_n}{\mathrm{d}\mu}\Big|_{\widetilde{\tau}_{n,T}} \right\}^\alpha \left\{ \frac{\mathrm{d}Q_n}{\mathrm{d}\mu}\Big|_{\widetilde{\tau}_{n,T}} \right\}^{1-\alpha} \mathrm{d}\mu \;=\; E_{Q_{n,T}}\left[(\mathcal{B}_{n,T})^\alpha \right], \tag{6}$$

where $P_{n,T} := P_n\Big|_{\widetilde{\tau}_{n,T}}$ respectively $Q_{n,T} := Q_n\Big|_{\widetilde{\tau}_{n,T}}$ is the restriction of P_n respectively Q_n to the time-point set $\widetilde{\tau}_{n,T}$, and $\frac{\mathrm{d}P_n}{\mathrm{d}\mu}\Big|_{\widetilde{\tau}_{n,T}}$ respectively $\frac{\mathrm{d}Q_n}{\mathrm{d}\mu}\Big|_{\widetilde{\tau}_{n,T}}$ are the corresponding densities with respect to the specially chosen reference law $\mu = Q_{n,T}$. A definition which covers more general laws can be found in Liese and Vajda [16]. Similarly to (6), the Bayes factor moments with respect to the alternative law P_n are related to Hellinger integrals by $H_{\alpha+1}(P_{n,T} \| Q_{n,T}) = E_{P_{n,T}}\left[(\mathcal{B}_{n,T})^\alpha \right]$, $\alpha \in \mathbb{R}$. By combining (5) and (6) one gets

$$\begin{aligned}
&H_\alpha(P_{n,T} \| Q_{n,T}) \;=\; (p_{n,1} \cdot \ldots \cdot p_{n,n})^\alpha \cdot (q_{n,1} \cdot \ldots \cdot q_{n,n})^{1-\alpha} \\
&+ (p_{n,1} \cdot \ldots \cdot p_{n,n-1}\,(1-p_{n,n}))^\alpha \cdot (q_{n,1} \cdot \ldots \cdot q_{n,n-1}\,(1-q_{n,n}))^{1-\alpha} \\
&+ (p_{n,1} \ldots p_{n,n-2}\,(1-p_{n,n-1})\,p_{n,n})^\alpha \cdot (q_{n,1} \ldots q_{n,n-2}\,(1-q_{n,n-1})\,q_{n,n})^{1-\alpha} \\
&+ (p_{n,1} \ldots p_{n,n-2}\,(1-p_{n,n-1})\,(1-p_{n,n}))^\alpha \cdot (q_{n,1} \ldots q_{n,n-2}\,(1-q_{n,n-1})\,(1-q_{n,n}))^{1-\alpha} \\
&\qquad \vdots \qquad\qquad \vdots \\
&+ ((1-p_{n,1}) \ldots (1-p_{n,n}))^\alpha \cdot ((1-q_{n,1}) \ldots (1-q_{n,n}))^{1-\alpha} \\
&= \prod_{i=1}^n \left\{ p_{n,i}^\alpha \cdot q_{n,i}^{1-\alpha} + (1-p_{n,i})^\alpha \cdot (1-q_{n,i})^{1-\alpha} \right\} \;.
\end{aligned} \tag{7}$$

Clearly, this can be alternatively derived from the facts that the observations sequence $\mathcal{X}_{n,T} = (X_{n,t} : t \in \{0, \frac{T}{n}, 2\frac{T}{n}, \dots, n\frac{T}{n}\})$ can be represented as $\mathcal{X}_{n,T} = T \circ \Xi_{n,T}$ for some one-to-one measurable functional T (with measurable inverse) acting on the sequence $\Xi_{n,T} := (\xi_{n,i} : i \in \{0, 1, 2, \dots, n\})$ of independent (under either hypothesis respectively alternative law) Bernoulli random variables $\xi_{n,i}$, and thus (with a slight abuse of notation for the joint distributions)

$$\begin{aligned}
H_\alpha(P_{n,T} \| Q_{n,T}) &= H_\alpha\Big(\mathcal{L}(\xi_{n,1}, \dots, \xi_{n,n} | P_{n,T}) \;\|\; \mathcal{L}(\xi_{n,1}, \dots, \xi_{n,n} | Q_{n,T}) \Big) \\
&= \prod_{i=1}^n H_\alpha\Big(\mathcal{L}(\xi_{n,i} | P_{n,T}) \;\|\; \mathcal{L}(\xi_{n,i} | Q_{n,T}) \Big)
\end{aligned}$$

which leads immediately to (7). In order to achieve convergence as $n \to \infty$ for the Bayes factor moments (6) and thus for (7), we employ the following assumption which we impose for the rest of this paper, unless stated otherwise.

Assumption 2.1. There exists a constant $\widehat{p} \in]0,1[$ as well as real-number double arrays $(a_{n,i} : n \in \mathbb{N}, i \in \{1,2,\ldots,n\})$, $(b_{n,i} : n \in \mathbb{N}, i \in \{1,2,\ldots,n\})$ such that

$$p_{n,i} = \widehat{p} + a_{n,i} \in]0,1[, \quad q_{n,i} = \widehat{p} + b_{n,i} \in]0,1[, \quad \text{for all } n \in \mathbb{N},\ i \in \{1,2,\ldots,n\}, \tag{8}$$

$$\lim_{n\to\infty} \max_{1\le i\le n} |a_{n,i}| = 0, \qquad \lim_{n\to\infty} \max_{1\le i\le n} |b_{n,i}| = 0, \tag{9}$$

$$\sup_{n\in\mathbb{N}} \sum_{i=1}^{n} a_{n,i}^2 < \infty, \qquad \sup_{n\in\mathbb{N}} \sum_{i=1}^{n} b_{n,i}^2 < \infty, \tag{10}$$

$$A_T^2 := \lim_{n\to\infty} \sum_{i=1}^{n} (a_{n,i} - b_{n,i})^2 \quad \text{exists and is finite.} \tag{11}$$

As a side remark, notice that the validity of the three assumptions (8) to (10) does not imply (11). This can be exemplarily seen by taking $\widehat{p} := \frac{1}{2}$, $a_{n,i} := \frac{1}{4\sqrt{n}}\sqrt{1+(-1)^n}$ and $b_{n,i} := 0$. Furthermore, as indicated, the quantity A_T^2 depends in general on the observation time horizon T, due to the nature of the chosen setup (e. g., the choice $i = n$ corresponds “directly” to T).

Theorem 2.2. For each $\alpha \in \mathbb{R}$ there hold the Hellinger integral convergences

$$\lim_{n\to\infty} H_\alpha(P_{n,T}\|Q_{n,T}) \;=\; \exp\left\{\frac{\alpha(\alpha-1)}{2}\cdot\frac{A_T^2}{\widehat{p}\,(1-\widehat{p})}\right\} \;=\; H_\alpha(\widehat{P}_T\|\widehat{Q}) \;<\; \infty \tag{12}$$

and

$$\lim_{n\to\infty} H_\alpha(Q_{n,T}\|P_{n,T}) \;=\; \exp\left\{\frac{\alpha(\alpha-1)}{2}\cdot\frac{A_T^2}{\widehat{p}\,(1-\widehat{p})}\right\} \;=\; H_\alpha(\widehat{Q}\|\widehat{P}_T) \;<\; \infty \tag{13}$$

where $\widehat{P}_T := N(A_T, \widehat{p}\cdot(1-\widehat{p}))$, $\widehat{Q} := N(0, \widehat{p}\cdot(1-\widehat{p}))$ are two auxiliary Normal probability laws. Equivalently, the latter two can be replaced e. g. by the auxiliary LogNormal probability laws $\widehat{P}_T := logN(A_T, \widehat{p}\cdot(1-\widehat{p}))$, $\widehat{Q} := logN(0, \widehat{p}\cdot(1-\widehat{p}))$, or by any other two laws which arise from the abovementioned Normal probability laws via a synchronous transformation by means of any sufficient statistics.

Notice that for fixed n the Hellinger integrals $H_\alpha(P_{n,T}\|Q_{n,T})$ and $H_\alpha(Q_{n,T}\|P_{n,T})$ do generally not coincide (cf. (7)), however their limit as $n \to \infty$ always coincides. Before starting with the proof of Theorem 2.2, we first present some examples.

Example 2.3. Consider the following homogeneous special case SPH with final-time horizon $\widehat{T} = \infty$ (leading to $N_n = \infty$) and some constants $\sigma > 0$, $c_{\mathcal{H}} \in \mathbb{R}$, $c_{\mathcal{A}} \in \mathbb{R}$, $c_{\mathcal{H}} \neq c_{\mathcal{A}}$, $K \in \mathbb{N}$:

$$u_{n,i} := \exp\Big(\sigma\sqrt{\frac{T}{K+n}}\Big) =: u_n, \qquad d_{n,i} := \exp\Big(-\sigma\sqrt{\frac{T}{K+n}}\Big) =: d_n,$$

$$q_{n,i} := \frac{1}{2} + \frac{1}{2\sigma}\Big(c_{\mathcal{H}} - \frac{\sigma^2}{2}\Big)\sqrt{\frac{T}{K+n}} =: q_n, \quad p_{n,i} := \frac{1}{2} + \frac{1}{2\sigma}\Big(c_{\mathcal{A}} - \frac{\sigma^2}{2}\Big)\sqrt{\frac{T}{K+n}} =: p_n$$

($i \in \mathbb{N}$), where K is chosen large enough such that $q_n \in]0,1[$ and $p_n \in]0,1[$ for all $n \in \mathbb{N}$.

Here, the trees are "recombining" after every second period and the Bayes factor $\mathcal{B}_{n,T}$ is the same for all sample paths $\mathcal{X}_{n,T}$ which end at the same final value $X_{n,T}$ (which can be straightforwardly deduced from (5)). It is easy to see that Assumption 2.1 is satisfied, and hence from Theorem 2.2 we obtain for any $\alpha \in \mathbb{R}$ by means of $A_T = \frac{1}{2}\left|\frac{c_\mathcal{A}-c_\mathcal{H}}{\sigma}\right| \cdot \sqrt{T}$ the Hellinger-integral convergence

$$\lim_{n\to\infty} H_\alpha(P_{n,T}\|Q_{n,T}) \;=\; \exp\left\{\frac{\alpha(\alpha-1)}{2}\cdot\left(\frac{c_\mathcal{A}-c_\mathcal{H}}{\sigma}\right)^2\cdot T\right\} \;=:\; C_\alpha. \tag{14}$$

Notice the well-known fact that the corresponding discrete-time asset price process $(X_{n,t})_{t\in\tau_n}$ in this homogeneous CRR model converges (in distribution on the sample path space) to a geometric Brownian motion $(X_t)_{t\in\mathbb{R}_+}$ with volatility $\sigma > 0$ and growth constant $c_\mathcal{H}$ (in the hypothesis case $\mathcal{H}$) respectively $c_\mathcal{A}$ (in the alternative case $\mathcal{A}$); in other words, X_t satisfies the stochastic differential equation (SDE) $\mathrm{d}X_t = c_\mathcal{H}X_t\mathrm{d}t + \sigma X_t\mathrm{d}W_t$ in the hypothesis case $\mathcal{H}$ (under the limit law $Q := \lim_{n\to\infty} Q_n$) respectively the SDE $\mathrm{d}X_t = c_\mathcal{A}X_t\mathrm{d}t + \sigma X_t\mathrm{d}W_t$ in the alternative case $\mathcal{A}$ (under the limit law $P := \lim_{n\to\infty} P_n$). As usual, W_t denotes a standard Brownian motion at time t. By computing the corresponding αth moment of the logarithmic-normally distributed Girsanov [8] density Z_T between the two involved restricted laws $P|_{[0,T]}$ and $Q|_{[0,T]}$, one can consistently derive C_α also as the Hellinger integral within such a continuous-time framework (cf. Stummer and Vajda [26]).

Example 2.4. Let us examine the following inhomogeneous special case SPI with final-time horizon $\widehat{T} = \infty$ (leading to $N_n = \infty$) and some constants $\sigma > 0$, $c_\mathcal{H} \in \mathbb{R}$, $c_\mathcal{A} \in \mathbb{R}$, $c_\mathcal{H} \neq c_\mathcal{A}$, $K \in \mathbb{N}$:

$$u_{n,i} := 1 + \sigma\cdot\frac{T}{K+n}\cdot\sqrt{i}\ , \qquad d_{n,i} := 1 - \sigma\cdot\frac{T}{K+n}\cdot\sqrt{i}\ ,$$
$$q_{n,i} := \frac{1}{2} + \frac{c_\mathcal{H}}{2\sigma}\cdot\frac{T}{K+n}\cdot\sqrt{i}\ , \qquad p_{n,i} := \frac{1}{2} + \frac{c_\mathcal{A}}{2\sigma}\cdot\frac{T}{K+n}\cdot\sqrt{i}\ ,$$

where K is chosen large enough such that $d_{n,i} > 0$, $q_{n,i} \in]0,1[$, $p_{n,i} \in]0,1[$ for all $n \in \mathbb{N}$ and $i \in \mathbb{N}$. Here, the trees are generally "not recombining" after every second period. It is easy to see that Assumption 2.1 is satisfied, and hence from Theorem 2.2 we obtain for any $\alpha \in \mathbb{R}$ by means of $A_T = \frac{1}{\sqrt{8}}\left|\frac{c_\mathcal{A}-c_\mathcal{H}}{\sigma}\right| \cdot T$ the convergence

$$\lim_{n\to\infty} H_\alpha(P_{n,T}\|Q_{n,T}) \;=\; \exp\left\{\frac{\alpha(\alpha-1)}{2}\cdot\left(\frac{c_\mathcal{A}-c_\mathcal{H}}{\sigma}\right)^2\cdot\frac{T^2}{2}\right\} \;=:\; \overline{C}_\alpha.$$

Notice that in contrast with (14) of the previous Example 2.3, the Hellinger-integral limit now depends on the observation-time horizon T in an exponentially-quadratic (rather than exponentially-linear) way. Furthermore, one can straightforwardly show (e. g. by verifying the Assumptions 1 to 4 in Nelson and Ramaswamy [20]) that the corresponding discrete-time asset price process $(X_{n,t})_{t\in\tau_n}$ in the current inhomogeneous CRR model converges (in distribution on the sample path space) to a strong solution $(X_t)_{t\in\mathbb{R}_+}$ of the SDE $\mathrm{d}X_t = c_{\mathcal{H}} \cdot t \cdot X_t \,\mathrm{d}t + \sigma \cdot \sqrt{t} \cdot X_t \,\mathrm{d}W_t$ in the hypothesis case $\mathcal{H}$ (under the limit law $Q := \lim_{n\to\infty} Q_n$) respectively the SDE $\mathrm{d}X_t = c_{\mathcal{A}} \cdot t \cdot X_t \,\mathrm{d}t + \sigma \cdot \sqrt{t} \cdot X_t \,\mathrm{d}W_t$ in the alternative case $\mathcal{A}$ (under the limit law $P := \lim_{n\to\infty} P_n$). Analogously to Example 2.3, by computing the corresponding αth moment of the logarithmic normally distributed Girsanov density Z_T between the two involved restricted laws $P|_{[0,T]}$ and $Q|_{[0,T]}$, one can consistently derive $\overline{C}_\alpha$ also as the Hellinger integral within such a continuous-time framework.

Let us next present the

P r o o f of Theorem 2.2. The two cases $\alpha = 0$ and $\alpha = 1$ are obvious. Let $\alpha \in \mathbb{R}\backslash\{0,1\}$ and some auxiliary constants $0 < \underline{\varkappa} < \overline{\varkappa} < 1$ be arbitrary but fixed. By performing a Taylor expansion for the function $g(p) := p^\alpha \cdot q^{1-\alpha} + (1-p)^\alpha \cdot (1-q)^{1-\alpha}$ on the domain $0 < \underline{\varkappa} \le q < p \le \overline{\varkappa} < 1$ (for fixed q), we obtain

$$\begin{aligned} &\Big| p^\alpha \cdot q^{1-\alpha} + (1-p)^\alpha \cdot (1-q)^{1-\alpha} - \Big(1 + \frac{\alpha\cdot(\alpha-1)}{2} \cdot \frac{(p-q)^2}{q\cdot(1-q)}\Big)\Big| \\ &\le\; D(\alpha, \underline{\varkappa}, \overline{\varkappa}) \cdot |p-q|^3 \end{aligned} \tag{15}$$

for some "constant" $D(\alpha, \underline{\varkappa}, \overline{\varkappa}) < \infty$ depending only on $\alpha, \underline{\varkappa}, \overline{\varkappa}$. By interchanging the roles of p and q, one achieves the same bound (15) (with possibly different finite constant $D(\alpha, \underline{\varkappa}, \overline{\varkappa})$) for $0 < \underline{\varkappa} \le p < q \le \overline{\varkappa} < 1$. Hence, for all $0 < \underline{\varkappa} \le p, q \le \overline{\varkappa} < 1$ there holds (15) as well as

$$\begin{aligned} &\Big| \exp\Big\{ \frac{\alpha(\alpha-1)}{2} \cdot \frac{(p-q)^2}{q\cdot(1-q)} \Big\} - \Big(1 + \frac{\alpha\cdot(\alpha-1)}{2} \cdot \frac{(p-q)^2}{q\cdot(1-q)}\Big)\Big| \\ &\le\; \widetilde{D}(\alpha, \underline{\varkappa}, \overline{\varkappa}) \cdot \frac{(p-q)^4}{q^2\cdot(1-q)^2} \end{aligned} \tag{16}$$

for some constant $\widetilde{D}(\alpha, \underline{\varkappa}, \overline{\varkappa}) < \infty$, which follows easily by Taylor expansion for the exponential function. Because of the assumptions (8) and (9), in the inequalities (15) and (16) one can plug in $p_{n,i} = \widehat{p} + a_{n,i}$ for p and $q_{n,i} = \widehat{p} + b_{n,i}$ for q, for all large enough n, i such that (say) $0 < \underline{\varkappa} := \frac{\widehat{p}}{2} < p_{n,i}, q_{n,i} < \overline{\varkappa} := \frac{1+\widehat{p}}{2} < 1$. By using this together with (10), (11) we obtain

$$\begin{aligned} &\lim_{n\to\infty} \Big| \prod_{i=1}^{n} \exp\Big\{ \frac{\alpha(\alpha-1)}{2} \cdot \frac{(a_{n,i}-b_{n,i})^2}{(\widehat{p}+b_{n,i})\cdot(1-\widehat{p}-b_{n,i})} \Big\} \\ &\qquad - \prod_{i=1}^{n} (\widehat{p}+a_{n,i})^\alpha \cdot (\widehat{p}+b_{n,i})^{1-\alpha} + (1-\widehat{p}-a_{n,i})^\alpha \cdot (1-\widehat{p}-b_{n,i})^{1-\alpha} \Big| \\ &\le\; \lim_{n\to\infty} \prod_{i=1}^{n} \widetilde{D}(\alpha, \underline{\varkappa}, \overline{\varkappa}) \cdot \frac{(a_{n,i}-b_{n,i})^4}{(\widehat{p}+b_{n,i})^2\cdot(1-\widehat{p}-b_{n,i})^2} \\ &\quad +\; \lim_{n\to\infty} \prod_{i=1}^{n} D(\alpha, \underline{\varkappa}, \overline{\varkappa}) \cdot |a_{n,i}-b_{n,i}|^3 \\ &=\; 0 \end{aligned}$$

and thus

$$\lim_{n\to\infty}\prod_{i=1}^{n} p_{n,i}^{\alpha}\cdot q_{n,i}^{1-\alpha}+(1-p_{n,i})^{\alpha}\cdot(1-q_{n,i})^{1-\alpha} \;=\; \exp\left\{\frac{\alpha(\alpha-1)}{2}\cdot\frac{A_T^2}{\widehat{p}\,(1-\widehat{p})}\right\}$$

which by (7) gives the desired result (12). The second convergence (13) follows immediately from (12) by the well-known skew symmetry of Hellinger integrals. □

In the following, let us discuss how the above Hellinger-integral convergence results can be applied to establish limit assertions for other related quantities. To begin with, by the continuity theorem of Hellinger-/Mellin-transforms (see e. g. Strasser [24]), Liese and Miescke [14]) one can deduce from Theorem 2.2 the distributional convergences

$$\mathcal{L}\left(\frac{\mathrm{d}P_{n,T}}{\mathrm{d}Q_{n,T}}\,\Big|\,Q_{n,T}\right) \quad\overset{n\to\infty}{\Longrightarrow}\quad \mathcal{L}\left(\frac{\mathrm{d}\widehat{P}_T}{\mathrm{d}\widehat{Q}}\,\Big|\,\widehat{Q}\right) \tag{17}$$

$$\mathcal{L}\left(\frac{\mathrm{d}Q_{n,T}}{\mathrm{d}P_{n,T}}\,\Big|\,P_{n,T}\right) \quad\overset{n\to\infty}{\Longrightarrow}\quad \mathcal{L}\left(\frac{\mathrm{d}\widehat{Q}}{\mathrm{d}\widehat{P}_T}\,\Big|\,\widehat{P}_T\right) \tag{18}$$

and thus for every bounded continuous real-valued function f

$$E_{Q_{n,T}}\Big[f\Big(\frac{\mathrm{d}P_{n,T}}{\mathrm{d}Q_{n,T}}\Big)\Big] = \int f\Big(\frac{\mathrm{d}P_{n,T}}{\mathrm{d}Q_{n,T}}\Big)\,\mathrm{d}Q_{n,T} \quad\overset{n\to\infty}{\longrightarrow}\quad \int f\Big(\frac{\mathrm{d}\widehat{P}_T}{\mathrm{d}\widehat{Q}}\Big)\,\mathrm{d}\widehat{Q} = E_{\widehat{Q}}\Big[f\Big(\frac{\mathrm{d}\widehat{P}_T}{\mathrm{d}\widehat{Q}}\Big)\Big], \tag{19}$$

$$E_{P_{n,T}}\Big[f\Big(\frac{\mathrm{d}Q_{n,T}}{\mathrm{d}P_{n,T}}\Big)\Big] = \int f\Big(\frac{\mathrm{d}Q_{n,T}}{\mathrm{d}P_{n,T}}\Big)\,\mathrm{d}P_{n,T} \quad\overset{n\to\infty}{\longrightarrow}\quad \int f\Big(\frac{\mathrm{d}\widehat{Q}}{\mathrm{d}\widehat{P}_T}\Big)\,\mathrm{d}\widehat{P}_T = E_{\widehat{P}_T}\Big[f\Big(\frac{\mathrm{d}\widehat{Q}}{\mathrm{d}\widehat{P}_T}\Big)\Big]. \tag{20}$$

If f is unbounded and continuous, then (19) remains valid if in addition to (17) the two conditions

$$\int\Big|f\Big(\frac{\mathrm{d}\widehat{P}_T}{\mathrm{d}\widehat{Q}}\Big)\Big|\,\mathrm{d}\widehat{Q} \;<\; \infty \qquad \text{and} \tag{21}$$

$$\limsup_{M\to\infty}\,\limsup_{n\to\infty}\int\Big|f\Big(\frac{\mathrm{d}P_{n,T}}{\mathrm{d}Q_{n,T}}\Big)\Big|\cdot\mathbf{1}_{[M,\infty[}\Big(\frac{\mathrm{d}P_{n,T}}{\mathrm{d}Q_{n,T}}\Big)\,\mathrm{d}Q_{n,T} \;=\; 0 \tag{22}$$

hold, see e. g. Liese and Miescke [14]. Here, $\mathbf{1}_A(\cdot)$ denotes the indicator function on the set A. In face of (17), a sufficient condition for the uniform integrability (22) is

$$\sup_{n\in\mathbb{N}}\int\Big|f\Big(\frac{\mathrm{d}P_{n,T}}{\mathrm{d}Q_{n,T}}\Big)\Big|^{1+\varepsilon}\,\mathrm{d}Q_{n,T} \;<\; \infty \qquad \text{for some } \varepsilon>0. \tag{23}$$

By (12), the condition (23) is satisfied for all power functions. The derivation of sufficient conditions for the validity of (20) for unbounded continuous functions f works analogously. Altogether, we thus obtain

Proposition 2.5. Let f be a continuous real-valued function with the two properties (i) $f(x) = 0$ for all $x < 0$, and (ii) for all $x \geq 0$ there exist some nonnegative constants κ_1, κ_2, κ_3, κ_4, κ_5 such that $|f(x)| \leq \kappa_1 + \kappa_2 \cdot x^{\kappa_3} + \kappa_4 \cdot x^{-\kappa_5}$. Then the two convergences (19) as well as (20) hold.

3. POWER DIVERGENCES

Apart from the important Bayes factor, it is also useful to study the "distance" between the two non-homogeneous CRR models at choice. Along this line, let us investigate the power divergences – also known as Cressie-Read measures resp. generalized cross-entropy family – between the two corresponding probability laws $P_{n,T}$ and $Q_{n,T}$ (on the sample path space), defined by

$$I_\alpha(P_{n,T} \| Q_{n,T}) \ := \ \int \ f_\alpha \left(\frac{\mathrm{d}P_{n,T}}{\mathrm{d}Q_{n,T}} \right) \mathrm{d}Q_{n,T} \ ,$$

with the following nonnegative functions $f_\alpha : [0, \infty[\to [0, \infty[$:

$$f_\alpha(\rho) \ := \ \begin{cases} -\log \rho + \rho - 1, & \text{if} \quad \alpha = 0, \\ \frac{\alpha\rho + 1 - \alpha - \rho^\alpha}{\alpha(1-\alpha)}, & \text{if} \quad \alpha \in \mathbb{R} \backslash \{0, 1\}, \\ \rho \ \log \rho + 1 - \rho, & \text{if} \quad \alpha = 1. \end{cases} \tag{24}$$

For basic facts on power divergences of general laws, the reader is e. g. referred to Liese and Vajda [16]. As usual, prominent special cases are the relative entropy (Kullback–Leibler information measure) $I_1(P_{n,T} \| Q_{n,T})$, the squared Hellinger distance $\frac{1}{2}\, I_{\frac{1}{2}}(P_{n,T} \| Q_{n,T})$ and the χ^2-divergence $2\, I_2(P_{n,T} \| Q_{n,T})$.

In order to derive concrete expressions for $I_\alpha(P_{n,T} \| Q_{n,T})$, one can adapt from the general theory the formula

$$I_\alpha(P_{n,T} \| Q_{n,T}) = \frac{1 - H_\alpha(P_{n,T} \| Q_{n,T})}{\alpha(1 - \alpha)} \ , \qquad \alpha \in \mathbb{R} \setminus \{0, 1\} \ , \tag{25}$$

and accordingly make use of (7). Under our general Assumption 2.1, one gets

Proposition 3.1.

(a) $$I_0(P_{n,T}\|Q_{n,T}) = \sum_{i=1}^{n} q_{n,i}\log\left(\frac{q_{n,i}}{p_{n,i}}\right) + (1-q_{n,i})\log\left(\frac{1-q_{n,i}}{1-p_{n,i}}\right)$$
$$\overset{n\to\infty}{\longrightarrow} \quad \frac{A_T^2}{2\,\widehat{p}\,(1-\widehat{p})} = I_0(\widehat{P}_T\|\widehat{Q}) \ .$$

(b) for all $\alpha \in \mathbb{R}\setminus\{0,1\}$:
$$I_\alpha(P_{n,T}\|Q_{n,T}) = \frac{1}{\alpha(1-\alpha)}\Big[1-\prod_{i=1}^{n}\Big\{p_{n,i}^{\alpha}\cdot q_{n,i}^{1-\alpha} + (1-p_{n,i})^{\alpha}\cdot(1-q_{n,i})^{1-\alpha}\Big\}\Big]$$
$$\overset{n\to\infty}{\longrightarrow} \quad \frac{1}{\alpha(1-\alpha)}\Big[1-\exp\Big\{\frac{\alpha(\alpha-1)}{2}\cdot\frac{A_T^2}{\widehat{p}\,(1-\widehat{p})}\Big\}\Big] = I_\alpha(\widehat{P}_T\|\widehat{Q}) \ .$$

(c) $$I_1(P_{n,T}\|Q_{n,T}) = \sum_{i=1}^{n} p_{n,i}\log\left(\frac{p_{n,i}}{q_{n,i}}\right) + (1-p_{n,i})\log\left(\frac{1-p_{n,i}}{1-q_{n,i}}\right)$$
$$\overset{n\to\infty}{\longrightarrow} \quad \frac{A_T^2}{2\,\widehat{p}\,(1-\widehat{p})} = I_1(\widehat{P}_T\|\widehat{Q}) \ .$$

Notice that from (a), (b), (c) one can immediately deduce the non-obvious limit-interchangeabilities

$$\begin{aligned}\lim_{\alpha\to1}\lim_{n\to\infty} I_\alpha(P_{n,T}\|Q_{n,T}) &= \lim_{n\to\infty}\lim_{\alpha\to1} I_\alpha(P_{n,T}\|Q_{n,T}) \\ = \lim_{\alpha\to0}\lim_{n\to\infty} I_\alpha(P_{n,T}\|Q_{n,T}) &= \lim_{n\to\infty}\lim_{\alpha\to0} I_\alpha(P_{n,T}\|Q_{n,T}) \ .\end{aligned}$$

Furthermore, one can also straightforwardly derive versions of (a), (b), (c) for the $n-$limits of $I_\alpha(Q_{n,T}\|P_{n,T})$, $\alpha\in\mathbb{R}$, which leads to a "symmetry statement" of analogous form to the remark after Theorem 2.2.

P r o o f of Proposition 3.1. The assertion (b) follows immediately from (12) and (25), whereas (a) and (c) are straightforward applications of Proposition 2.5 to the functions $f_0(\cdot)$ and $f_1(\cdot)$ given in (24). Alternatively, (say) the assertion (a) can be derived from (c) by a skew symmetry argument. □

In particular, the assertions of Proposition 3.1 can be applied to the contexts of the Examples 2.3 and 2.4 to obtain power-divergence convergence results for the SPH and SPI models.

4. BAYESIAN DECISION PROCEDURES AND BAYES RISK

In formula (2) of Section 1, we introduced the risk (average loss) of a decision function $\delta(\mathcal{X}_{n,T})$ taking values in $\mathcal{D} = \{\theta_{\mathcal{H}}, \theta_{\mathcal{A}}\}$. If we reject the hypothesis-adjoint action $\theta_{\mathcal{H}}$ whenever the observed asset value sample path $\mathcal{X}_{n,T} = (X_{n,t} : t \in \widetilde{\tau}_{n,T})$ lies within a critical region $G = \delta^{-1}(\theta_{\mathcal{A}})$, we can rewrite this risk in the form

$$\begin{aligned}\mathcal{R}_{n,T}(G) \quad &= \quad pr^{\mathcal{H}}\ L_{\mathcal{H}}\ Q_{n,T}[G]\ +\ pr^{\mathcal{A}}\ L_{\mathcal{A}}\ P_{n,T}[\overline{\Omega} - G] \\ &\quad + pr^{\mathcal{H}}\ \widetilde{L}_{\mathcal{H}}\ Q_{n,T}[\overline{\Omega} - G]\ +\ pr^{\mathcal{A}}\ \widetilde{L}_{\mathcal{A}}\ P_{n,T}[G]\ ,\end{aligned}$$

where $\overline{\Omega}$ denotes the canonical space of all possible asset-value sample paths between the times 0 and T. By means of the parameters

$$\begin{aligned}\lambda_{\mathcal{H}} &:= pr^{\mathcal{H}}\ L_{\mathcal{H}}\ > 0\ , \quad \lambda_{\mathcal{A}}\ :=\ pr^{\mathcal{A}}\ L_{\mathcal{A}}\ > 0\ , \\ \widetilde{\lambda}_{\mathcal{H}} &:= pr^{\mathcal{H}}\ \widetilde{L}_{\mathcal{H}}\ \in [0, \lambda_{\mathcal{H}}[\ \ , \quad \widetilde{\lambda}_{\mathcal{A}}\ :=\ pr^{\mathcal{A}}\ \widetilde{L}_{\mathcal{A}}\ \in [0, \lambda_{\mathcal{A}}[\ ,\end{aligned}$$

which carry combined prior and loss information, we obtain the formula

$$\begin{aligned}\mathcal{R}_{n,T}(G) \quad &= \quad \lambda_{\mathcal{H}}\ Q_{n,T}[\,G\,]\ +\ \lambda_{\mathcal{A}}\ \Big(1\ -\ P_{n,T}[\,G\,]\ \Big) \\ &\quad + \widetilde{\lambda}_{\mathcal{H}}\ \Big(1 - Q_{n,T}[\,G\,]\Big)\ +\ \widetilde{\lambda}_{\mathcal{A}}\ P_{n,T}[\,G\,]\ .\end{aligned} \tag{26}$$

By definition, the Bayes risk $\mathcal{R}_{n,T}^{\min}$ minimizes the risk, i. e.

$$\mathcal{R}_{n,T}^{\min}\ :=\ \min \mathcal{R}_{n,T}(G)\ ,$$

where the minimum is taken over all measurable sets $G \subset \overline{\Omega}$ of asset-value sample paths between the times 0 and T. By (26),

$$\begin{aligned}\mathcal{R}_{n,T}(G)\ &=\ \int \Big[\widetilde{\lambda}_{\mathcal{H}} + \lambda_{\mathcal{A}}\,\mathcal{B}_{n,T} + \big[(\lambda_{\mathcal{H}} + \widetilde{\lambda}_{\mathcal{A}}\,\mathcal{B}_{n,T}) - (\widetilde{\lambda}_{\mathcal{H}} + \lambda_{\mathcal{A}}\,\mathcal{B}_{n,T})\big]\mathbf{1}_G\Big]\,\mathrm{d}Q_{n,T} \\ &\geq\ \int \Big[\widetilde{\lambda}_{\mathcal{H}} + \lambda_{\mathcal{A}}\,\mathcal{B}_{n,T} + \big[(\lambda_{\mathcal{H}} + \widetilde{\lambda}_{\mathcal{A}}\,\mathcal{B}_{n,T}) - (\widetilde{\lambda}_{\mathcal{H}} + \lambda_{\mathcal{A}}\,\mathcal{B}_{n,T})\big]\mathbf{1}_{G_n^{min}}\Big]\,\mathrm{d}Q_{n,T} \\ &=\ \int \min\Big[(\widetilde{\lambda}_{\mathcal{H}} + \lambda_{\mathcal{A}}\,\mathcal{B}_{n,T}),\ (\lambda_{\mathcal{H}} + \widetilde{\lambda}_{\mathcal{A}}\,\mathcal{B}_{n,T})\Big]\,\mathrm{d}Q_{n,T}\ =\ \mathcal{R}_{n,T}(G_n^{min})\ =\ \mathcal{R}_{n,T}^{\min}\end{aligned} \tag{27}$$

where

$$G_n^{min}\ =\ \Big\{\lambda_{\mathcal{H}} + \widetilde{\lambda}_{\mathcal{A}}\,\mathcal{B}_{n,T} \leq \widetilde{\lambda}_{\mathcal{H}} + \lambda_{\mathcal{A}}\,\mathcal{B}_{n,T}\Big\}\ =\ \Big\{\mathcal{B}_{n,T} \geq \frac{\lambda_{\mathcal{H}} - \widetilde{\lambda}_{\mathcal{H}}}{\lambda_{\mathcal{A}} - \widetilde{\lambda}_{\mathcal{A}}}\Big\}\ .$$

Thus, the Bayes risk is achieved by the decision rule $\delta(\mathcal{X}_{n,T})$ which rejects the hypothesis-adjoint action $\theta_{\mathcal{H}}$ (decides for the alternative-adjoint action $\theta_{\mathcal{A}}$) if the observed path $\mathcal{X}_{n,T}$ is contained in the sample path set G_n^{min} and rejects $\theta_{\mathcal{A}}$ (decides for $\theta_{\mathcal{H}}$) if $\mathcal{X}_{n,T}$ is contained in the complement of this set. This will be called the *first optimal decision procedure* henceforth. As usual, there is also a *second optimal decision procedure* between the two models $\mathcal{H}$ and $\mathcal{A}$, which works as follows: one rejects $\theta_{\mathcal{H}}$ (decides for $\theta_{\mathcal{A}}$) if

$$\begin{aligned}&pr_{n,T}^{\mathcal{H},\text{post}}\,L(\theta_{\mathcal{A}},\mathcal{H})\ +\ pr_{n,T}^{\mathcal{A},\text{post}}\,L(\theta_{\mathcal{A}},\mathcal{A}) \\ &\leq\ pr_{n,T}^{\mathcal{H},\text{post}}\,L(\theta_{\mathcal{H}},\mathcal{H})\ +\ pr_{n,T}^{\mathcal{A},\text{post}}\,L(\theta_{\mathcal{H}},\mathcal{A})\end{aligned} \tag{28}$$

(i. e. if the posterior expected loss of the decision $\theta_{\mathcal{A}}$ is less or equal than the posterior expected loss of the decision $\theta_{\mathcal{H}}$), and one rejects $\theta_{\mathcal{A}}$ (decides for $\theta_{\mathcal{H}}$) if

$$\begin{aligned} & pr_{n,T}^{\mathcal{H},\text{post}}\, L(\theta_{\mathcal{A}},\mathcal{H}) + pr_{n,T}^{\mathcal{A},\text{post}}\, L(\theta_{\mathcal{A}},\mathcal{A}) \\ & > pr_{n,T}^{\mathcal{H},\text{post}}\, L(\theta_{\mathcal{H}},\mathcal{H}) + pr_{n,T}^{\mathcal{A},\text{post}}\, L(\theta_{\mathcal{H}},\mathcal{A}) \,. \end{aligned} \tag{29}$$

Notice again that the dependence of $pr_{n,T}^{\mathcal{H},\text{post}}$ and $pr_{n,T}^{\mathcal{A},\text{post}}$ on the observed asset value sample path $\mathcal{X}_{n,T}$ is not indicated explicitly here. As usual, it can be seen in a straightforward manner that the second optimal decision procedure is equivalent to the first optimal decision procedure, and consequently the corresponding Bayes risk is also given by (27).

Except for some particular special cases (e. g. SPH), it will generally depend on the *whole* observed asset value sample path $\mathcal{X}_{n,T}$ which decision will actually be taken. Furthermore, in order to obtain an explicit expression for the corresponding Bayes risk in the general setup, one can plug (5) into (27).

Before we start with the according limit investigations, let us first illuminate an interesting "non-commutativity" between statistical and investment decisions, where for transparency we deal with a one-period context:

Example 4.1. Consider the special case SPH with the choices $T = \frac{1}{4}$ year, $n = 1$, $\sigma = 0.1$, $c_{\mathcal{H}} = 0.013$, $c_{\mathcal{A}} = 0.001$. This leads to (approximately) $u_{1,i} = u_1 = 1.05$, $d_{1,i} = d_1 = 0.95$, $q_{1,i} = q_1 = 0.52$, $p_{1,i} = p_1 = 0.49$ ($i \in \mathbb{N}$). Furthermore, let us choose equal prior probabilities $pr^{\mathcal{H}} = pr^{\mathcal{A}} = 0.5$. According to (5), (3) and (4), in case of observing an asset value sample path $\overline{\mathcal{X}_{1,T}} = (X_{1,t} : t \in \{0, \frac{1}{4}\}) = (x, x \cdot u_1)$ (i. e. an "up" in the first period) one gets (approximately) $\overline{Z_{1,T}} = Z_{1,T}(\overline{\mathcal{X}_{1,T}}) = 0.9423$, $\overline{pr_{1,T}^{\mathcal{H},\text{post}}} = pr_{1,T}^{\mathcal{H},\text{post}}(\overline{\mathcal{X}_{1,T}}) = 0.5149$ and $\overline{pr_{1,T}^{\mathcal{A},\text{post}}} = pr_{1,T}^{\mathcal{A},\text{post}}(\overline{\mathcal{X}_{1,T}}) = 0.4851$. Analogously, in case of observing a sample path $\overline{\overline{\mathcal{X}_{1,T}}} = (X_{1,t} : t \in \{0, \frac{1}{4}\}) = (x, x \cdot d_1)$ (i. e. a "down" in the first period) one gets (approximately) $\overline{\overline{Z_{1,T}}} = Z_{1,T}(\overline{\overline{\mathcal{X}_{1,T}}}) = 1.0625$, $\overline{\overline{pr_{1,T}^{\mathcal{H},\text{post}}}} = pr_{1,T}^{\mathcal{H},\text{post}}(\overline{\overline{\mathcal{X}_{1,T}}}) = 0.4848$ and $\overline{\overline{pr_{1,T}^{\mathcal{A},\text{post}}}} = pr_{1,T}^{\mathcal{A},\text{post}}(\overline{\overline{\mathcal{X}_{1,T}}}) = 0.5152$. If in the "classical" Bayesian *testing* setup $L_{\mathcal{H}} = L_{\mathcal{A}} = 1 = 1 - \widetilde{L}_{\mathcal{H}} = 1 - \widetilde{L}_{\mathcal{A}}$ one observes an "up" in the first period, then according to (29) one decides for the action $\theta_{\mathcal{H}}$ to accept the hypothesis model $\mathcal{H}$ (since $0.5149 \cdot 1 + 0.4851 \cdot 0 > 0.5149 \cdot 0 + 0.4851 \cdot 1$). If one observes a "down" in the first period, then (28) leads to the action $\theta_{\mathcal{A}}$ to accept the alternative model $\mathcal{A}$ (since $0.4848 \cdot 1 + 0.5152 \cdot 0 \leq 0.4848 \cdot 0 + 0.5152 \cdot 1$). The corresponding Bayes risk is $\mathcal{R}_{1,T}^{\min} = \min\{0.5 \cdot 1\,, 0.5 \cdot 1 \cdot 0.9423\} \cdot 0.52 + \min\{0.5 \cdot 1\,, 0.5 \cdot 1 \cdot 1.0625\} \cdot (1 - 0.52) = 0.4850$.

Example 4.2. In the setup of Example 4.1, let us consider a different loss structure which is derived from the following investment setup (adapted from Stummer and Vajda [26]): Suppose that we have observed the asset value sample path for one period and that now, at time $t = \frac{1}{4}$ (in years), we invest 10000 USD for the next period of $\frac{1}{4}$ year. If we invest all the money into this asset, then we buy $10000/X_{1,\frac{1}{4}}$ assets. Our corresponding expected wealth at time $t = \frac{1}{2}$ will be

$$E_{Q_{1,T}}\left[\frac{10000}{X_{1,\frac{1}{4}}} \cdot X_{1,\frac{1}{2}}\right] = 10000\,\{q_1 \cdot u_1 + (1-q_1)\cdot d_1\}$$
$$= 10000 \cdot \{0.52 \cdot 1.05 + (1-0.52)\cdot 0.95\} = 10020\,\mathrm{USD}$$

under the hypothesis model $\mathcal{H}$, and

$$E_{P_{1,T}}\left[\frac{10000}{X_{1,\frac{1}{4}}} \cdot X_{1,\frac{1}{2}}\right] = 10000\,\{p_1 \cdot u_1 + (1-p_1)\cdot d_1\}$$
$$= 10000 \cdot \{0.49 \cdot 1.05 + (1-0.49)\cdot 0.95\} = 9990\,\mathrm{USD}$$

under the alternative model $\mathcal{A}$. In contrast, if at time $t = \frac{1}{4}$ we invest all the money into a savings deposit for $\frac{1}{4}$ year with guaranteed (continuously compounded) annual growth rate of 0.19%, then our wealth at time $t = \frac{1}{2}$ will be $10000 \cdot e^{0.0019\cdot(\frac{1}{2}-\frac{1}{4})} = 10004.75$ USD. If the decision space $\mathcal{D} = \{\theta_{\mathcal{H}}, \theta_{\mathcal{A}}\}$ consists of the decisions $\theta_{\mathcal{H}}$ to invest (at time $t = \frac{1}{4}$) all the money into the asset and $\theta_{\mathcal{A}}$ to invest all the money into the savings deposit, then $\theta_{\mathcal{H}}$ leads to (assumingly) zero loss $\widetilde{L}_{\mathcal{H}} = L(\theta_{\mathcal{H}}, \mathcal{H}) = 0$ under $\mathcal{H}$ and to the expected loss $L_{\mathcal{A}} = L(\theta_{\mathcal{H}}, \mathcal{A}) = 10004.75 - 9990 = 14.75$ USD under $\mathcal{A}$. Similarly, $\theta_{\mathcal{A}}$ leads to (assumingly) zero loss $\widetilde{L}_{\mathcal{A}} = L(\theta_{\mathcal{A}}, \mathcal{A}) = 0$ under $\mathcal{A}$ but its expected loss under $\mathcal{H}$ is $L_{\mathcal{H}} = L(\theta_{\mathcal{A}}, \mathcal{H}) = 10020 - 10004.75 = 15.25$ USD. Thus, if one observes an "up" in the first period, then according to (29) one decides for the action $\theta_{\mathcal{H}}$ to invest (at time $t = \frac{1}{4}$) all the money into the asset (since $0.5149\cdot 15.25 + 0.4851\cdot 0 > 0.5149\cdot 0 + 0.4851\cdot 14.75$). If one observes a "down" in the first period, then (28) leads to the action $\theta_{\mathcal{A}}$ to invest all the money into the savings deposit (since $0.4848\cdot 15.25 + 0.5152\cdot 0 \leq 0.4848\cdot 0 + 0.5152\cdot 14.75$). The corresponding Bayes risk is $\mathcal{R}^{\min}_{1,T} = \min\{0.5\cdot 15.25\,, 0.5\cdot 14.75\cdot 0.9423\}\cdot 0.52 + \min\{0.5\cdot 15.25\,, 0.5\cdot 14.75\cdot 1.0625\}\cdot(1-0.52) = 7.27$ USD.

Notice that the decisions in Example 4.2 are "consistent" with the decisions taken in Example 4.1: if one would first perform a "classical" Bayes test (i. e. accept $\mathcal{H}$ in case of observing an "up" in the first period resp. accept $\mathcal{A}$ in case of observing a "down") and afterwards would decide about the investment, then one would end up with the same decisions as above (i. e. invest all the money into the asset in case of an "up" resp. invest all the money into the savings deposit in case of a "down"). However, if the guaranteed annual growth rate in Example 4.2 would be 0.22% (or larger) instead of 0.19%, then the decisions in Example 4.2 would be "inconsistent" with the decisions taken in Example 4.1: if one would now first perform a classical Bayes test (i. e. accept $\mathcal{H}$ in case of observing an "up" in the first period resp. accept $\mathcal{A}$ in case of observing a "down") and afterwards would decide about the investment, then one would still end up by investing all the money into the asset in case of an "up" resp. investing all the money into the savings deposit in case of a "down". In contrast, if one uses the (adapted) decision method of Example 4.2, one would end up with the decision to always invest all the money into the savings deposit. Indeed, the wealth at time $t = \frac{1}{2}$ from the savings deposit will now be $10000 \cdot e^{0.0022\cdot(\frac{1}{2}-\frac{1}{4})} = 10005.5$ USD. Hence, $L_{\mathcal{H}} = 10020 - 10005.5 = 14.5$ and $L_{\mathcal{A}} = 10005.5 - 9990 = 15.5$, and the proposed decision assertions follow from $0.5149 \cdot 14.5 + 0.4851 \cdot 0 \leq 0.5149 \cdot 0 + 0.4851 \cdot 15.5$ respectively from $0.4848 \cdot 14.5 + 0.5152 \cdot 0 \leq 0.4848 \cdot 0 + 0.5152 \cdot 15.5$.

The "opposite inconsistency" appears in the case where the growth rate in Example 4.2 would be 0.18% (or smaller) instead of 0.19%. With the (correspondingly adapted) decision method of Example 4.2, one would now always decide to invest all the money in the asset. Indeed, the wealth from the savings deposit will now be $10000 \cdot e^{0.0018 \cdot (\frac{1}{2} - \frac{1}{4})} = 10004.5$ USD, and consequently one gets $L_{\mathcal{H}} = 10020 - 10004.5 = 15.5$, $L_{\mathcal{A}} = 10004.5 - 9990 = 14.5$ and therefore $0.5149 \cdot 15.5 + 0.4851 \cdot 0 > 0.5149 \cdot 0 + 0.4851 \cdot 14.5$ respectively $0.4848 \cdot 15.5 + 0.5152 \cdot 0 > 0.4848 \cdot 0 + 0.5152 \cdot 14.5$.

The abovementioned investigations indicate that, in general, the – in practice often used – method to first perform a "classical" Bayesian statistical test and afterwards carry out an optimal investment decision, may lead to a different result than performing a "model-uncertainty-integrated" optimal investment decision (one might loosely call this effect a "non-commutativity" between Bayesian statistical and optimal investment decisions); also notice the "decision sensitivity" (even) within a very small range of interest rates.

In Section 2 resp. 3 we have given limits as n tends to infinity of Hellinger integrals (Bayes factor moments) $H_\alpha(P_{n,T} \| Q_{n,T})$ resp. power divergences $I_\alpha(P_{n,T} \| Q_{n,T})$. Let us now present the corresponding limits of the Bayes risk, under our general Assumption 2.1:

Theorem 4.3. There holds

$$
\begin{aligned}
\mathcal{R}^{\min}_{n,T} \quad &= \quad \widetilde{\lambda}_{\mathcal{H}} + \lambda_{\mathcal{A}} + (\lambda_{\mathcal{H}} - \widetilde{\lambda}_{\mathcal{H}}) \cdot Q_{n,T}\Big[\, \mathcal{B}_{n,T} \geq \frac{\lambda_{\mathcal{H}} - \widetilde{\lambda}_{\mathcal{H}}}{\lambda_{\mathcal{A}} - \widetilde{\lambda}_{\mathcal{A}}} \,\Big] \\
&\qquad + (\widetilde{\lambda}_{\mathcal{A}} - \lambda_{\mathcal{A}}) \cdot P_{n,T}\Big[\, \mathcal{B}_{n,T} \geq \frac{\lambda_{\mathcal{H}} - \widetilde{\lambda}_{\mathcal{H}}}{\lambda_{\mathcal{A}} - \widetilde{\lambda}_{\mathcal{A}}} \,\Big] \qquad (30) \\
&\overset{n \to \infty}{\longrightarrow} \quad \widetilde{\lambda}_{\mathcal{H}} + \lambda_{\mathcal{A}} + (\lambda_{\mathcal{H}} - \widetilde{\lambda}_{\mathcal{H}}) \cdot (1 - \Phi(\widehat{a_1})) + (\widetilde{\lambda}_{\mathcal{A}} - \lambda_{\mathcal{A}}) \cdot (1 - \Phi(\widehat{a_2})) \\
&= \quad \lambda_{\mathcal{H}} + \widetilde{\lambda}_{\mathcal{A}} + (\widetilde{\lambda}_{\mathcal{H}} - \lambda_{\mathcal{H}}) \cdot \Phi(\widehat{a_1}) \; + \; (\lambda_{\mathcal{A}} - \widetilde{\lambda}_{\mathcal{A}}) \cdot \Phi(\widehat{a_2})\,, \qquad (31)
\end{aligned}
$$

where we have used the quantities

$$
\begin{aligned}
\widehat{a_1} \; &:= \; \frac{\sqrt{\widehat{p}\,(1-\widehat{p})} \cdot \log\Big(\frac{\lambda_{\mathcal{H}} - \widetilde{\lambda}_{\mathcal{H}}}{\lambda_{\mathcal{A}} - \widetilde{\lambda}_{\mathcal{A}}}\Big)}{A_T} + \frac{A_T}{2\sqrt{\widehat{p}\,(1-\widehat{p})}}\,, \\
\widehat{a_2} \; &:= \; \frac{\sqrt{\widehat{p}\,(1-\widehat{p})} \cdot \log\Big(\frac{\lambda_{\mathcal{H}} - \widetilde{\lambda}_{\mathcal{H}}}{\lambda_{\mathcal{A}} - \widetilde{\lambda}_{\mathcal{A}}}\Big)}{A_T} - \frac{A_T}{2\sqrt{\widehat{p}\,(1-\widehat{p})}}\,,
\end{aligned}
$$

as well as the standard normal distribution function Φ.

For instance, one can apply Theorem 4.3 to the special setup SPI of Example 2.4 where $\widehat{p} = 1/2$ and $A_T = |\Delta| \cdot T/\sqrt{8}$ with $\Delta := \frac{c_{\mathcal{H}} - c_{\mathcal{A}}}{\sigma}$.

P r o o f of Theorem 4.3. The representation (30) follows immediately from (27). Furthermore, from the distributional convergence (17) we obtain with $\Upsilon := \frac{\lambda_{\mathcal{H}}-\widetilde{\lambda}_{\mathcal{H}}}{\lambda_{\mathcal{A}}-\widetilde{\lambda}_{\mathcal{A}}} > 0$ and the auxiliary (identity) random variable U with law $\widehat{Q} = N(0, \widehat{p}\cdot(1-\widehat{p}))$

$$\lim_{n\to\infty} Q_{n,T}\Big[\mathcal{B}_{n,T} \geq \Upsilon\Big] \;=\; \widehat{Q}\Big[\frac{\mathrm{d}\widehat{P}_T}{\mathrm{d}\widehat{Q}} \geq \Upsilon\Big] \;=\; \widehat{Q}\Big[\exp\Big\{\frac{2U\cdot A_T - A_T^2}{2\widehat{p}\,(1-\widehat{p})}\Big\} \geq \Upsilon\Big]$$
$$= \widehat{Q}\Big[\frac{U}{\sqrt{\widehat{p}\,(1-\widehat{p})}} \geq \widehat{a}_1\Big] \;=\; 1-\Phi(\widehat{a}_1)\,.$$

Similarly, from the distributional convergence (18) one can deduce by means of the auxiliary (identity) random variable U with law $\widehat{P}_T = N(A_T, \widehat{p}\cdot(1-\widehat{p}))$

$$\lim_{n\to\infty} P_{n,T}\Big[\mathcal{B}_{n,T} \geq \Upsilon\Big] \;=\; \widehat{P}_T\Big[\frac{\mathrm{d}\widehat{Q}}{\mathrm{d}\widehat{P}_T} \leq \frac{1}{\Upsilon}\Big] \;=\; \widehat{Q}\Big[\exp\Big\{\frac{-2U\cdot A_T + A_T^2}{2\widehat{p}\,(1-\widehat{p})}\Big\} \leq \frac{1}{\Upsilon}\Big]$$
$$= \widehat{P}_T\Big[\frac{U-A_T}{\sqrt{\widehat{p}\,(1-\widehat{p})}} \geq \widehat{a}_2\Big] \;=\; 1-\Phi(\widehat{a}_2)$$

which leads to the desired convergence assertion. □

An immediate consequence of Theorem 4.3 is the following result about the limit of the total variation distance V between the two corresponding probability laws $P_{n,T}$ and $Q_{n,T}$, defined by

$$V(P_{n,T}\|Q_{n,T}) \;:=\; 2\cdot\sup_G \Big\{P_{n,T}[G] - Q_{n,T}[G]\Big\} \;\in [0,2]\;.$$

Here, the supremum is taken over all measurable sets $G \subset \overline{\Omega}$ of asset-value sample paths between the times 0 and T.

Corollary 4.4. There holds

$$\frac{1}{2}\,V(P_{n,T}\|Q_{n,T}) \quad \stackrel{n\to\infty}{\longrightarrow} \quad 2\cdot\Phi\bigg(\frac{A_T}{2\sqrt{\widehat{p}\,(1-\widehat{p})}}\bigg) \;-\; 1\,.$$

To see this, one can first use the general representation formula (see e. g. Stummer and Vajda [26], adapted to the current context) $\mathcal{R}_{n,T}^{\min} = 1-\frac{1}{2}\,V(P_{n,T}\|Q_{n,T})$ for the special situation $\lambda_{\mathcal{H}} = \lambda_{\mathcal{A}} = 1$, $\widetilde{\lambda}_{\mathcal{H}} = \widetilde{\lambda}_{\mathcal{A}} = 0$. Then the assertion of Corollary 4.4 follows immediately from (31), by additionally employing an appropriate symmetry property of the standard normal distribution function Φ.

For the special time-homogeneous recombining-tree structure of Example 2.3 with the one-period-probabilities $q_n := \frac{1}{2} + \frac{1}{2\sigma}(c_{\mathcal{H}} - \frac{\sigma^2}{2})\sqrt{\frac{T}{K+n}} =: q_{n,i}$, $p_n := \frac{1}{2} + \frac{1}{2\sigma}(c_{\mathcal{A}} - \frac{\sigma^2}{2})\sqrt{\frac{T}{K+n}} =: p_{n,i}$ (with sufficiently large $K \in \mathbb{N}$), the $n-$th step Bayes risk $\mathcal{R}_{n,T}^{\min}$ in (30) (and hence, the total variation distance $V(P_{n,T}\|Q_{n,T})$) can be represented in a compact explicit way. Since in this setup there holds $\widehat{p} = 1/2$ as well as $A_T = |\Delta| \cdot \sqrt{T}/2$ with $\Delta := \frac{c_{\mathcal{H}}-c_{\mathcal{A}}}{\sigma}$, one can deduce

Corollary 4.5. Suppose that SPH and $\widetilde{L}_{\mathcal{H}} = \widetilde{L}_{\mathcal{A}} = 0$ holds. Then one gets for all $n \in \mathbb{N}$

$$\mathcal{R}^{\min}_{n,T} = \begin{cases} \lambda_{\mathcal{H}} \sum_{j=m_n}^{n} \binom{n}{j} (q_n)^j (1-q_n)^{n-j} \\ \qquad + \lambda_{\mathcal{A}} \left(1 - \sum_{j=m_n}^{n} \binom{n}{j} (p_n)^j (1-p_n)^{n-j}\right), & \text{if } c_{\mathcal{A}} > c_{\mathcal{H}} \\ \lambda_{\mathcal{H}} \sum_{j=0}^{\widetilde{m_n}} \binom{n}{j} (q_n)^j (1-q_n)^{n-j} \\ \qquad + \lambda_{\mathcal{A}} \left(1 - \sum_{j=0}^{\widetilde{m_n}} \binom{n}{j} (p_n)^j (1-p_n)^{n-j}\right), & \text{if } c_{\mathcal{A}} < c_{\mathcal{H}} \end{cases}$$

$$\overset{n\to\infty}{\longrightarrow} \quad \lambda_{\mathcal{H}} \cdot (1-\Phi(\widehat{a_1})) \; + \; \lambda_{\mathcal{A}} \cdot \Phi(\widehat{a_2})$$

with (the existing quantities)

$$m_n := \min\left\{ j \in \{0,\ldots,n\} : \frac{(p_n)^j (1-p_n)^{n-j}}{(q_n)^j (1-q_n)^{n-j}} \geq \frac{\lambda_{\mathcal{H}}}{\lambda_{\mathcal{A}}} \right\} ,$$

$$\widetilde{m_n} := \max\left\{ j \in \{0,\ldots,n\} : \frac{(p_n)^j (1-p_n)^{n-j}}{(q_n)^j (1-q_n)^{n-j}} \geq \frac{\lambda_{\mathcal{H}}}{\lambda_{\mathcal{A}}} \right\} ,$$

$$\widehat{a_1} \; := \; \frac{\log(\lambda_{\mathcal{H}}/\lambda_{\mathcal{A}})}{|\Delta|\sqrt{T}} \; + \; \frac{|\Delta|\sqrt{T}}{2} , \qquad \widehat{a_2} \; := \; \frac{\log(\lambda_{\mathcal{H}}/\lambda_{\mathcal{A}})}{|\Delta|\sqrt{T}} \; - \; \frac{|\Delta|\sqrt{T}}{2} .$$

Remark 4.6. Let us continue with the discussion launched in the second half of Example 2.3. If one starts within a continuous-time (rather than the here-used discrete-time) framework with the hypothesis model $\mathcal{H}$ being a geometric Brownian motion with growth constant $c_{\mathcal{H}}$ and volatility σ, and the alternative model $\mathcal{A}$ to be another geometric Brownian motion with growth constant $c_{\mathcal{A}}$ and volatility σ, then the corresponding Bayes risk turns out to be exactly $\lambda_{\mathcal{H}} \cdot (1-\Phi(\widehat{a_1})) + \lambda_{\mathcal{A}} \cdot \Phi(\widehat{a_2})$ (cf. Stummer and Vajda [26]). Hence, we have indirectly shown the nontrivial fact that the asset-value-process limit procedure and the above Bayes-risk limit procedure are consistent (in the associate sense).

ACKNOWLEDGEMENT

We are very much indebted to one referee whose suggestions lead to a considerably more general treatment with shorter proofs. Furthermore, we are very grateful to Norbert Henze for valuable comments and remarks.

(Received April 3, 2011)

REFERENCES

[1] A. K. Bera and Y. Bilias: The MM, ME, ML, EL, EF and GMM approaches to estimation: a synthesis. J. Econometrics *107* (2002), 51–86.

[2] A. Berlinet and I. Vajda: Selection rules based on divergences. Statistics *45* (2011), 479–495.

[3] M. Broniatowski and I. Vajda: Several applications of divergence criteria in continuous families. To appear in Kybernetika (2012). See also Research Report No. 2257, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague 2009; moreover, see arXiv:0911.0937v1 [math.ST].

[4] J. C. Cox, S. A. Ross, and M. Rubinstein: Option pricing: a simplified approach. J. Finan. Econ. *7* (1979), 229–263.

[5] N. Cressie and T. R. C. Read: Multinomial goodness-of-fit tests. J. Roy. Stat. Soc. Ser. B Stat. Methodol. *46* (1984), 440–464.

[6] I. Csiszár and F. Matúš: Generalized maximum likelihood estimates for exponential families. Probab. Theory Related Fields *141* (2008), 213–246.

[7] I. Csiszár and P. C. Shields: Information theory and statistics: a tutorial. Found. Trends Commun. Inform. Theory *1* (2004), 4, 417–528.

[8] I. V. Girsanov: On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. Theory Probab. Appl. *5* (1960), 285–301.

[9] A. Golan: Information and entropy econometrics – editor's view. J. Econometrics *107* (2002), 1–15.

[10] A. Gretton and L. Györfi: Consistent nonparametric tests of independence. J. Mach. Learn. Res. *11* (2010), 1391–1423.

[11] P. Harremoes and I. Vajda: On the Bahadur-effcient testing of uniformity by means of the entropy. IEEE Trans. Inform. Theory *54* (2008), 321–331.

[12] P. Harremoes and I. Vajda: On Bahadur efficiency of power divergence statistics. Preprint arXiv:1002.1493v1 [math.ST] (2010).

[13] T. Hobza, L. Pardo, and D. Morales: Rényi statistics for testing equality of autocorrelation coefficients. Statist. Methodol. *6* (2009), 424–436.

[14] F. Liese and K.-J. Miescke: Statistical Decision Theory. Springer–Verlag, New York 2008.

[15] F. Liese, D. Morales, and I. Vajda: Asymptotically sufficient partitions and quantizations. IEEE Trans. Inform. Theory *52* (2006), 5599–5606.

[16] F. Liese and I. Vajda: Convex Statistical Distances. Teubner, Leipzig 1987.

[17] F. Liese and I. Vajda: On divergences and informations in statistics and information theory. IEEE Trans. Inform. Theory *52* (2006), 4394–4412.

[18] E. Maasoumi: A compendium to information theory in economics and econometrics. Econometrics Rev. *12* (1993), 2, 137–181.

[19] D. Morales and I. Vajda: Generalized information criteria for optimal Bayes decisions. Research Report No. *2274*, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Prague 2010.

[20] D. B. Nelson and K. Ramaswamy: Simple binomial processes as diffusion approximations in financial models. Rev. Financ. Stud. *3* (1990), 393–430.

[21] L. Pardo: Statistical Inference Based on Divergence Measures. Chapman & Hall, Boca Raton 2005.

[22] M. C. Pardo: Testing equality restrictions in generalized linear models for multinomial data. Metrika *73* (2011), 231–253.

[23] T. R. C. Read and N. A. C. Cressie: Goodness-of-Fit Statistics for Discrete Multivariate Data. Springer–Verlag, New York 1988.

[24] H. Strasser: Mathematical Theory of Statistics. De Gruyter, Berlin 1985.

[25] W. Stummer: Exponentials, Diffusions, Finance, Entropy and Information. Shaker, Aachen 2004.

[26] W. Stummer and I. Vajda: Optimal statistical decisions about some alternative financial models. J. Econometrics *137* (2007), 441–471.

[27] W. Stummer and I. Vajda: On divergences of finite measures and their applicability in statistics and information theory. Statistics *44* (2010), 169–187.

[28] I. Vajda and E.C. van der Meulen: Goodness-of-Fit criteria based on observations quantized by hypothetical and empirical percentiles. In: Handbook of Fitting Statistical Distributions with R (Z.A. Katrian, E.J. Dudewicz eds.), Chapman & Hall / CRC, 2010, pp. 917–994.

[29] I. Vajda and J. Zvárová: On generalized entropies, Bayesian decisions and statistical diversity. Kybernetika *43* (2007), 675–696.

Wolfgang Stummer, Department of Mathematics, University of Erlangen-Nürnberg, Cauerstrasse 11*, 91058 Erlangen. Germany. (Corresponding author)*
e-mail: stummer@mi.uni-erlangen.de

Wei Lao, Institute for Stochastics, Karlsruhe Institute of Technology (KIT), Kaiserstrasse 89, 76133 Karlsruhe. Germany.
e-mail: wei.lao@kit.edu