KYBERNETIKA — VOLUME 23 (1987), NUMBER 3

ESTIMATES OF HELLINGER INTEGRALS OF INFINITELY DIVISIBLE DISTRIBUTIONS

FRIEDRICH LIESE

It is shown that the f-divergence is a lower semi-continuous functional on the space of probability measures if this space equipped with a topology similar to the weak topology. This statement is used for estimating Hellinger integrals and the variational distance of infinitely divisible distributions.

0. INTRODUCTION

Hellinger integrals are useful tools of evaluation a distinction between two probability distributions. They are applied to many problems both in information theory and statistics. Hellinger integrals are closely related to the variational distance [3], [11] and to error probabilities in the problem of testing statistical hypotheses [1], [7].

Up to the sign Hellinger integrals are special f-divergences introduced by Csiszar [2]. f-divergences behave monotonously if the distribution laws are transformed by a stochastic kernel. This and the fact that the Hellinger integral of product measures is the product of the Hellinger integrals of product components give the possibility to estimate the Hellinger integral of the distributions of statistics connected with independent observations.

A further common method to investigate observations is to study the limit behavior of the corresponding distributions. Although f-divergences are continuous with respect to increasing sequences of sub-σ-algebras [12] they are, in general, not continuous with respect to the variational distance. On the other hand it is known [5] that the Kullback-Leibler I-divergence, being an f-divergence for $f = x \ln x$, is lower semi-continuous with respect to the setwise convergence of the distributions on all sets. In the first part of this paper we introduce a concept of weak convergence of probability measures which includes as special cases both the setwise convergence and the common weak convergence on metric spaces. In Theorem 1·8 it will be shown that f-divergences are lower semi-continuous with respect to the weak topology introduced in this paper. In the second part Hellinger integrals of infinitely divisible distributions are considered. Theorem 1.8 gives the possibility to estimate Hellinger integrals of distributions which can be approximated by compound Poisson distributions in the sense of weak topology. These estimates are applied to infinitely divisible distributions on Hilbert spaces enabling to estimate variational distance of these distributions in terms of their characteristic triplets. Practical importance of such estimates in the statistics of stochastic processes, in particular in the non-stationary signal detection, is evident.

1. f-DIVERGENCES AND WEAK CONVERGENCE

Denote by $\mathbb{R}$ the real line. Let $f: (0, \infty) \mapsto \mathbb{R}$ be a continuous convex function. Then the limits $\lim_{u \downarrow 0} f(u)$ and $\lim_{u \to \infty} f(u)/u$ exist and take on values in $(-\infty, +\infty]$. We introduce a function $\mathsf{f}: [0, \infty)^2 \mapsto (-\infty, +\infty]$ by

$$\mathsf{f}(u, v) = \begin{cases} 0 & u = 0, \quad v = 0 \\ \lim_{u \downarrow 0} f(u)\, v & u = 0, \quad v > 0 \\ f(u/v)\, v & u > 0, \quad v > 0 \\ u . \lim_{t \to \infty} [f(t)/t] & u > 0, \quad v = 0 \end{cases}$$

f is a convex function which is lower semi-continuous on $[0, \infty)^2$ and continuous on $(0, \infty)^2$.

Let $(X, \mathfrak{X})$ be a measurable space. Denote by $\mathscr{S}_{\mathfrak{X}}$, $\mathscr{F}_{\mathfrak{X}}$ and $\mathscr{P}_{\mathfrak{X}}$ the set of all σ-finite, finite and probability measures on $(X, \mathfrak{X})$, respectively.

Given any real valued function h we set $h^+ = \max(h, 0)$, $h^- = -\min(h, 0)$.

Suppose $\mu, \nu \in \mathscr{S}_{\mathfrak{X}}$ and denote by $\gamma \in \mathscr{S}_{\mathfrak{X}}$ any dominating measure. Put

$$p = \frac{d\mu}{d\gamma}, \quad q = \frac{d\gamma}{d\gamma}$$

and set

$$\mathsf{I}_f(\mu, \nu) = \int \mathsf{f}(p, q)\, d\gamma$$

if the integral on the right hand side is well-defined, i.e. if at least one of the integrals

$$\int [\mathsf{f}(p, q)]^+ \, d\gamma, \quad \int [\mathsf{f}(p, q)]^- \, d\gamma$$

is finite. $\mathsf{I}_f(\mu, \nu)$ is called the f-divergence of μ and ν. The definition of $\mathsf{I}_f(\mu, \nu)$ is independent of the choice of the dominating measure γ. The functional I_f has been introduced by Csiszar in [2]. As f is a convex function there are constants a, b so that $f(u) \geqq au + b$ and consequently $\mathsf{f}(u, v) \geqq au + bv$. This means that $\mathsf{I}_f(\mu, \nu)$ is defined for every $\mu, \nu \in \mathscr{F}_{\mathfrak{X}}$. Taking special convex functions f one obtains a large number of information measures.

1.2. Example. Put for $0 < x < \infty$

$$f_\alpha(x) = \begin{cases} \dfrac{1}{\alpha(\alpha-1)}(x^\alpha - \alpha x - (1-\alpha)), & -\infty < \alpha < \infty, \quad \alpha \neq 0, \quad \alpha \neq 1 \\ \ln x + x - 1 & \alpha = 0 \\ x \ln x - x + 1 & \alpha = 1 \end{cases}$$

The family f_α is constructed in such a way that f_α depends continuously on α. Furthermore the convex functions f_α are non-negative. Hence $\mathsf{J}_{f_\alpha}(\mu, \nu)$ is defined for every $\mu, \nu \in \mathscr{S}_{\mathfrak{X}}$. We set

$$\mathsf{I}_\alpha(\mu, \nu) = \mathsf{J}_{f_\alpha}(\mu, \nu).$$

For $\mu, \nu \in \mathscr{P}_{\mathfrak{X}}$ it holds

$$\begin{aligned} \mathsf{I}_1(\mu, \nu) &= \int (p \ln (p/q) - p + q)\, \mathrm{d}\gamma \\ &= \int p \ln (p/q)\, \mathrm{d}\gamma \end{aligned}$$

i.e. I_1 is the Kullback-Leibler I-divergence.

1.2. Example. Consider the family $h_\alpha(x)$, $-\infty < \alpha < \infty$, $\alpha \neq 0$, $\alpha \neq 1$ of convex functions defined by

$$h_\alpha(x) = \operatorname{sign}(\alpha(\alpha - 1))\, x^\alpha, \quad 0 < x < \infty.$$

$\mathsf{H}_\alpha(\mu, \nu) = \operatorname{sign}(\alpha(\alpha-1))\, \mathsf{J}_{h_\alpha}(\mu, \nu)$ is called the Hellinger integral of order α. Put

$$\mathsf{R}_\alpha(\mu, \nu) = \begin{cases} \dfrac{1}{\alpha(\alpha-1)} \ln (\mathsf{H}_\alpha(\mu, \nu)) & -\infty < \alpha < \infty, \quad \alpha \neq 0, \quad \alpha \neq 1 \\ \mathsf{I}_0(\mu, \nu) & \alpha = 0 \\ \mathsf{I}_1(\mu, \nu) & \alpha = 1 \end{cases}$$

where the conventions $\ln 0 = -\infty$, $\ln \infty = \infty$ are used. The functional R_α has been introduced by Renyi in [9].

We need a simple convexity property of f-divergences.

1.3. Lemma. Suppose $\mu_i, \nu_i \in \mathscr{F}_{\mathfrak{X}}$, $\alpha_i \geqq 0$, $i = 1, 2, \ldots, \sum_{i=1}^\infty \alpha_i = 1$ and assume $\sum_{i=1}^\infty \alpha_i \mu_i(X) < \infty$, $\sum_{i=1}^\infty \alpha_i \nu_i(X) < \infty$.

Then

$$\mathsf{J}_f\Big(\sum_{i=1}^\infty \alpha_i \mu_i, \sum_{i=1}^\infty \alpha_i \nu_i\Big) \leqq \sum_{i=1}^\infty \alpha_i\, \mathsf{J}_f(\mu_i, \nu_i).$$

Proof. Let $\gamma \in \mathscr{S}_{\mathfrak{X}}$ be a measure which dominates μ_i, ν_i, $i = 1, 2, \ldots$. Put $p_i = \mathrm{d}\mu_i/\mathrm{d}\gamma$, $q_i = \mathrm{d}\nu_i/\mathrm{d}\gamma$ and choose constants a, b so that $f(t) \geqq at + b$. Introduce $\bar{f}$ by $\bar{f}(t) = f(t) - at - b$. Then $\bar{f}$ is a non-negative convex function. Hence

$$\begin{aligned} \mathsf{J}_{\bar{f}}\Big(\sum_{i=1}^\infty \alpha_i \mu_i, \sum_{i=1}^\infty \alpha_i \nu_i\Big) &= \int \bar{f}\Big(\sum_{i=1}^\infty \alpha_i p_i, \sum_{i=1}^\infty \alpha_i q_i\Big)\, \mathrm{d}\gamma \leqq \\ &\leqq \int \sum_{i=1}^\infty \alpha_i \bar{f}(p_i, q_i)\, \mathrm{d}\gamma \leqq \sum_{i=1}^\infty \alpha_i\, \mathsf{J}_{\bar{f}}(\mu_i, \nu_i) \end{aligned}$$

where the last relation follows from the theorem of Fubini. To finish the proof we note that

$$\mathsf{J}_f\Big(\sum_{i=1}^{\infty} \alpha_i\mu_i, \sum_{i=1}^{\infty} \alpha_i\nu_i\Big) = \mathsf{J}_{\bar f}\Big(\sum_{i=1}^{\infty} \alpha_i\mu_i, \sum_{i=1}^{\infty} \alpha_i\nu_i\Big) + $$
$$+ a\sum_{i=1}^{\infty} \alpha_i\, \mu_i(X) + b\sum_{i=1}^{\infty} \alpha_i\, \nu_i(X)\,. \qquad \square$$

In the sequel we also make use of the fact that the Hellinger integral of product measures is the product of Hellinger integrals of product components.

1.4. Lemma. Suppose $(X_i, \mathfrak{X}_i)$, $i = 1, \ldots, n$, are measurable spaces and $\mu_i, \nu_i \in \mathscr{F}_{\mathfrak{X}_i}$. Denote by $\prod_{i=1}^{n} \mu_i, \prod_{i=1}^{n} \nu_i$ the product measures. Then

(1) $$\mathsf{H}_\alpha\Big(\prod_{i=1}^{n} \mu_i, \prod_{i=1}^{n} \nu_i\Big) = \prod_{i=1}^{n} \mathsf{H}_\alpha(\mu_i, \nu_i) \quad \alpha \neq 0\,, \quad \alpha \neq 1$$

(2) $$\mathsf{I}_\alpha\Big(\prod_{i=1}^{n} \mu_i, \prod_{i=1}^{n} \nu_i\Big) = \sum_{i=1}^{n} \mathsf{I}_\alpha(\mu_i, \nu_i) \quad \alpha = 0\,, \quad \alpha = 1\,.$$

f-divergences have a large number of properties useful for applications in probability theory and statistics. They do not increase under a transformation of μ and ν by a stochastic kernel [2]. Furthermore they are closely related to the variational distance [3]. This means that for strictly convex f the variational distance of the probability measure P and Q is small provided $\mathsf{J}_f(\mathsf{P}, \mathsf{Q}) - f(1)$ is small. The converse implication holds only if $\lim_{u\downarrow 0} f(u) + \lim_{u\to\infty} f(u)/u < \infty$, see [11].

Consequently, in general J_f is not continuous with respect to the variational distance. But we will show that J_f is lower semi-continuous even with respect to topologies weaker than that generated by the variational distance.

Denote by $\mathbb{R}^n$ the n-dimensional Euclidean space and by $\mathfrak{L}^n$ the σ-algebra of Borel sets of $\mathbb{R}^n$. $\mathbb{C}(\mathbb{R}^n)$ denotes the space of all real valued bounded continuous functions on $\mathbb{R}^n$. The weakest topology $\boldsymbol{\tau}[\mathbb{C}(\mathbb{R}_n)]$ for $\mathscr{P}_{\mathfrak{L}^n}$ for which all mappings $\mathsf{P} \mapsto \int \varphi \, \mathrm{d}\mathsf{P}$, $\varphi \in \mathbb{C}(\mathbb{R}^n)$, are continuous is referred as to the weak topology for $\mathscr{P}_{\mathfrak{L}^n}$. It is well-known that $\boldsymbol{\tau}[\mathbb{C}(\mathbb{R}^n)]$ is metrizable by the Prokhorov metric ϱ and $(\mathscr{P}_{\mathfrak{L}^n}, \varrho)$ is a Polish space (cf. [9] Theorem 1.11). Given any measurable real valued functions $\varphi_1, \varphi_2, \ldots \ldots, \varphi_n$ defined on $(X, \mathfrak{X})$ we introduce the mapping

$$\mathsf{T}_{\varphi_1,\varphi_2,\ldots,\varphi_n} : \mathscr{P}_{\mathfrak{X}} \mapsto \mathscr{P}_{\mathfrak{L}^n}$$

by

$$(\mathsf{T}_{\varphi_1,\varphi_2,\ldots,\varphi_n}\mathsf{P})(B) = \mathsf{P}(\{(\varphi_1, \ldots, \varphi_n) \in B\})\,, \quad B \in \mathfrak{L}^n\,.$$

Next we generalize the concept of weak converegence in such a way that different concepts of convergence as the weak convergence of measures on metric spaces, the setwise convergence and the weak convergence of finite dimensional distributions are included as special cases.

1.4. Definition. Let $\mathbf{\Phi}$ be a set of real valued measurable functions on $(X, \mathfrak{X})$. The weak topology $\tau[\mathbf{\Phi}]$ for $\mathscr{P}_{\mathfrak{X}}$ generated by $\mathbf{\Phi}$ is the weakest topology for which the mappings

$$\mathsf{T}_{\varphi_1,\ldots,\varphi_n}: \mathscr{P}_{\mathfrak{X}} \mapsto (\mathscr{P}_{\mathfrak{L}^n}, \tau[\mathbb{C}(\mathbb{R}^n)])$$

are continuous for every choice $\varphi_1, \ldots, \varphi_n \in \mathbf{\Phi}$. $\tau(\mathbf{\Phi})$ is said to be complete if the smallest σ-algebra $\sigma(\mathbf{\Phi})$ for which all $\varphi \in \mathbf{\Phi}$ are measurable coincides with $\mathfrak{X}$.

Going back to the definition of $\tau[\mathbb{C}(\mathbb{R}^n)]$ we can also say that $\tau[\mathbf{\Phi}]$ is the weakest topology for $\mathscr{P}_{\mathfrak{X}}$ for which the mappings $\mathsf{P} \mapsto \int \varphi(\varphi_1, \ldots, \varphi_n) \, \mathrm{d}\mathsf{P}$ are continuous for every $\varphi_1, \ldots, \varphi_n \in \mathbf{\Phi}$ and every $\varphi \in \mathbb{C}(\mathbb{R}^n)$.

1.5. Example. Let X be a metric space and $\mathfrak{X}$ the σ-algebra of Borel sets. Denote by $\mathbb{C}(X)$ the space of all real valued bounded continuous functions on X and put $\mathbf{\Phi} = \mathbb{C}(X)$. Since the system

$$\{\psi: \psi = \varphi(\varphi_1, \ldots, \varphi_n)\,, \quad \varphi_i \in \mathbf{\Phi}\,, \quad \varphi \in \mathbb{C}(\mathbb{R}^n),\, n = 1, 2, \ldots\}$$

coincides with $\mathbb{C}(X)$ we see that $\tau[\mathbb{C}(X)]$ is the common weak topology for $\mathscr{P}_{\mathfrak{X}}$. $\tau[\mathbb{C}(X)]$ is complete as in a metric space X the σ-algebra of Borel sets is generated by functions from $\mathbb{C}(X)$.

1.6. Example. Suppose T is any non-empty set. Denote by $\mathbb{R}^T$ the set of all real valued functions on T. Let $\mathbf{\Phi} = \{\varphi_t, t \in T\}$ be the family of all projections. Denote by $\mathfrak{L}^T$ the σ-algebra generated by $\mathbf{\Phi}$, i.e. $\mathfrak{L}^T = \sigma(\mathbf{\Phi})$. Obviously $\tau[\mathbf{\Phi}]$ is complete. It is clear that the convergence with respect to $\tau[\mathbf{\Phi}]$ means the common-sense weak convergence of finite dimensional distributions.

1.7. Example. Let $(X, \mathfrak{X})$ be a measurable space and $\mathfrak{X}_0 \subseteq \mathfrak{X}$ is a subalgebra of $\mathfrak{X}$. Denote by $\mathbb{S}(\mathfrak{X}_0)$ the set of all functions φ of the form

$$\varphi = \sum_{i=1}^{n} c_i \, 1_{A_i}\,, \quad A_i \in \mathfrak{X}_0\,, \quad c_i \in \mathbb{R}\,, \quad n = 1, 2, \ldots$$

where 1_A denotes the indicator function of A. It is easy to see that $\tau[\mathbb{S}(\mathfrak{X}_0)]$ is the weakest topology for $\mathscr{P}_{\mathfrak{X}}$ for which every mapping $\mathsf{P} \mapsto \mathsf{P}(A)$, $A \in \mathfrak{X}_0$, is continuous. $\tau[\mathbb{S}(\mathfrak{X}_0)]$ is complete iff $\mathfrak{X}_0$ generates $\mathfrak{X}$.

It is known (cf. [5]) that the Kullback-Leiber I-divergence is lower semi-continuous with respect to $\tau[\mathbb{S}(\mathfrak{X})]$. This property is employed in the theory of large deviations [5]. We will show that this property continuous to hold for every J_f and every weak topology which is complete.

1.8. Theorem. Let f be a continuous convex function on $(0, \infty)$ and $\tau[\mathbf{\Phi}]$ a weak topology for $\mathscr{P}_{\mathfrak{X}}$ which is complete. Then the f-divergence J_f is a lower semi-continuous function on $(\mathscr{P}_{\mathfrak{X}} \times \mathscr{P}_{\mathfrak{X}}, \tau[\mathbf{\Phi}] \times \tau[\mathbf{\Phi}])$.

Proof. Let Λ be a directed index set and $(\mathsf{P}_\lambda, \mathsf{Q}_\lambda)$, $\lambda \in \Lambda$ be a net which tends to (P, Q) in the sense of $\tau[\Phi] \times \tau[\Phi]$. This means

(3) $$\lim_{\lambda \in \Lambda} \mathsf{P}_\lambda = \mathsf{P}\,, \quad \lim_{\lambda \in \Lambda} \mathsf{Q}_\lambda = \mathsf{Q}$$

in the sense of $\tau[\Phi]$. Put $\mathsf{R} = \frac{1}{2}(\mathsf{P} + \mathsf{Q})$ and

$$\mathfrak{X}_0 = \{A\colon A \in \mathfrak{X} \text{ there exist } n \in \{1, 2, \ldots\}\,, \quad B \in \mathfrak{L}^n$$
$$\varphi_i \in \Phi,\ i = 1, 2, \ldots, n,\ A = \mathsf{T}^{-1}_{\varphi_1,\ldots,\varphi_n}(B),\ \mathsf{R}(\mathsf{T}^{-1}_{\varphi_1,\ldots,\varphi_n}(\partial B) = 0\}$$

where ∂B denotes the boundary of B. It is clear that $A \in \mathfrak{X}_0$ implies $A^c \in \mathfrak{X}_0$. Assume $A_i \in \mathfrak{X}_0$, $i = 1, 2$,

$$A_i = \mathsf{T}^{-1}_{\varphi_{i,1},\ldots,\varphi_{i,n_i}}(B_i)\,, \quad B_i \in \mathfrak{L}^{n_i}\,, \quad \varphi_{i,k} \in \Phi$$

and put $B = B_1 \times B_2$. Then

$$\mathsf{R}(\mathsf{T}^{-1}_{\varphi_{1,1},\ldots,\varphi_{1,n_1},\varphi_{2,1},\ldots,\varphi_{2,n_2}}(\partial B)) \leqq \mathsf{R}(\mathsf{T}^{-1}_{\varphi_{1,1},\ldots,\varphi_{2,n_2}}(((\partial B_1) \times \mathbb{R}^{n_2}) \cup (\mathbb{R}^{n_1} \times (\partial B_2))) \leqq$$
$$\leqq \mathsf{R}(\mathsf{T}^{-1}_{\varphi_{1,1},\ldots,\varphi_{1,n_1}}(\partial B_1)) + \mathsf{R}(\mathsf{T}^{-1}_{\varphi_{2,1},\ldots,\varphi_{2,n_2}}(\partial B_2) = 0$$

Hence $A_1 \cap A_2 \in \mathfrak{X}_0$. This means that $\mathfrak{X}_0$ is a subalgebra of $\mathfrak{X}$. For every $\tilde{\mathsf{Q}} \in \mathscr{P}_{\mathfrak{L}^n}$ and $a_i \in \mathbb{R}$ it holds

$$(\tilde{\mathsf{Q}}\partial((-\infty, a_1) \times \ldots \times (-\infty, a_n)) \leqq \sum_{i=1}^{n} \tilde{\mathsf{Q}}_i(\{a_i\})$$

where $\tilde{\mathsf{Q}}_i$ is the ith marginal distribution of $\tilde{\mathsf{Q}}$. Since there are at most countably many a_i such that the right hand term is non-zero we see $\sigma(\{B\colon B \in \mathfrak{L}^n, \tilde{\mathsf{Q}}(\partial B) = 0\}) = \mathfrak{L}^n$. Puting $\tilde{\mathsf{Q}} = \mathsf{R} \circ \mathsf{T}^{-1}_{\varphi_1,\ldots,\varphi_n}$ we obtain that the system of sets

$$\{B : B \in \mathfrak{L}^n, \mathsf{R}(\mathsf{T}^{-1}_{\varphi_1,\ldots,\varphi_n}(\partial B)) = 0\}$$

generates $\mathfrak{L}^n$. As it is assumed $\mathfrak{X} = \sigma(\Phi)$, $\mathfrak{X}_0$ generates $\mathfrak{X}$. The weak convergence of probability measures on $(\mathbb{R}^n, \mathfrak{L}^n)$ implies the convergence of the values of the measures on every Borel set with boundary of limit measure zero.

Consequently, by (3) and $\mathsf{P} \leqq 2\mathsf{R}$, $\mathsf{Q} \leqq 2\mathsf{R}$,

$$\lim_{\lambda \in \Lambda} \mathsf{P}_\lambda(A) = \mathsf{P}(A)\,, \quad \lim_{\lambda \in \Lambda} \mathsf{Q}_\lambda(A) = \mathsf{Q}(A)$$

for every $A \in \mathfrak{X}_0$. Denote by $\mathscr{Z}$ the set of all partitions of X into sets of $\mathfrak{X}_0$. A simple modification of the approximation theorem for f-divergence (see Theorem 5 in [12]) shows that

$$\mathsf{J}_f(\mathsf{P}, \mathsf{Q}) = \sup \sum_A \mathsf{f}(\mathsf{P}(A), \mathsf{Q}(A))\,.$$

where A is taken from $\mathbb{Z}$ and sup is considered for all $\mathbb{Z} \in \mathscr{Z}$. f is a lower semi-continuous function. Hence

$$\liminf_{\lambda \in \Lambda} \mathsf{J}_f(\mathsf{P}_\lambda, \mathsf{Q}_\lambda) = \liminf_{\lambda \in \Lambda} (\sup \sum_A \mathsf{f}(\mathsf{P}_\lambda(A), \mathsf{Q}_\lambda(A)) \geqq$$
$$\geqq \sup (\liminf_{\lambda \in \Lambda} \sum_A \mathsf{f}(\mathsf{P}_\lambda(A), \mathsf{Q}_\lambda(A)) \geqq \sup \sum_A \mathsf{f}(\mathsf{P}(A), \mathsf{Q}(A)) = \mathsf{J}_f(\mathsf{P}, \mathsf{Q})$$

which proves the lower semi-continuity of J_f. □

2. INFINITELY DIVISIBLE DISTRIBUTIONS

Suppose $(X, +)$ is a commutative semigroup with zero element 0.

2.1. Definition. Let $\mathfrak{X}$ be a σ-algebra of subsets of X. $(X, \mathfrak{X})$ is said to be a measurable semigroup if the mapping

$$\chi: (x, y) \mapsto x + y$$

is $\mathfrak{X} \otimes \mathfrak{X} - \mathfrak{X}$ measurable.

Denote by $\mathbb{B}$ the Banach space of all real valued finite measures on $(X, \mathfrak{X})$ equipped with the total variation norm $\| \ \|$. For every $\mu, \nu \in \mathbb{B}$ we introduce the convolution

$$\mu * \nu = (\mu \times \nu) \circ \chi^{-1}$$

where $\mu \times \nu$ is the product measure. Denote by δ_x the δ-distribution concentrated at x. Similarly as in 1.4, 1.5 in [6] one can show that $(\mathbb{B}, *, \| \ \|)$ is a real Banach algebra with the unit element δ_0.

Denote by μ^{n*} the convolution powers for $n = 1, 2, \ldots$ and put $\mu^0 = \delta_0$.

2.2. Definition. Let $(X, \mathfrak{X})$ be a measurable semigroup. $\mathsf{P} \in \mathscr{P}_{\mathfrak{X}}$ is called infinitely divisible if for every $n = 1, 2, \ldots$ there exists $\mathsf{P}_n \in \mathscr{P}_{\mathfrak{X}}$ such that $\mathsf{P} = \mathsf{P}_n^{n*}$.

It is easy to see that for every $\mu \in \mathscr{F}_{\mathfrak{X}} \subseteq \mathbb{B}$ the series $\sum_{k=0}^{\infty} \mu^{k*}/k!$ converges in the norm of $\mathbb{B}$, i.e. in variational distance. Put

$$U_\mu = e^{-\mu(X)} \sum_{k=0}^{\infty} \frac{\mu^{k*}}{k!}.$$

Then $U_\mu \in \mathscr{P}_{\mathfrak{X}}$. As in the case of complex valued power series one obtains the functional equation

$$U_\mu * U_\nu = U_{\mu+\nu}.$$

Hence $U_\mu = (U_{\mu/n})^{n*}$ which shows that U_μ is infinitely divisible.

We now deal with Hellinger integrals or with other R_α-related functionals of the distribution laws approximable by distributions of the type U_μ.

2.3. Theorem. Let $(X, \mathfrak{X})$ be a measurable semigroup and $\tau[\Phi]$ a complete weak topology for $\mathscr{P}_{\mathfrak{X}}$. For $\mathsf{P}, \mathsf{Q} \in \mathscr{P}_{\mathfrak{X}}$ it is assumed that there are $\mu, \nu \in \mathscr{S}_{\mathfrak{X}}$, a net $x_\lambda \in X$ and a net $A_\lambda \in \mathfrak{X}$ such that $\mu(A_\lambda) < \infty$, $\nu(A_\lambda) < \infty$ and

$$\mathsf{P} = \lim_{\lambda \in \Lambda} \delta_{x_\lambda} * U_{\mu(\cdot \cap A_\lambda)}\,; \quad \mathsf{Q} = \lim_{\lambda \in \Lambda} \delta_{x_\lambda} * U_{\nu(\cdot \cap A_\lambda)}$$

in the sense of $\tau[\Phi]$. Then it holds

$$\mathsf{R}_\alpha(\mathsf{P}, \mathsf{Q}) \leqq \mathsf{I}_\alpha(\mu, \nu) \quad \text{for every} \quad -\infty < \alpha < \infty\,.$$

2.4. Corollary. Under the assumptions of Theorem 2.3

$$\|\mathsf{P} - \mathsf{Q}\| \leqq \sqrt{[4(1 - \exp\{-\tfrac{1}{2} \,.\, \mathsf{I}_{1/2}(\mu, \nu)\})]}\,.$$

Proof. Put $\mu_\lambda = \mu(\cdot \cap A_\lambda)$, $\nu_\lambda = \nu(\cdot \cap A_\lambda)$. For every $x \in X$ the mapping ψ:

$y \mapsto y + x$ is $\mathfrak{X}$-measurable and for every $\varrho \in \mathscr{F}_{\mathfrak{X}}$ it holds $\varrho \circ \psi^{-1} = \delta_x * \varrho$. Since f-divergences do not increase under measurable mappings [2] we get

(4) $$\mathrm{J}_{h_\alpha}(\delta_{x_\lambda} * U_{\mu_\lambda}, \delta_{x_\lambda} * U_{\nu_\lambda}) \leqq \mathrm{J}_{h_\alpha}(U_{\mu_\lambda}, U_{\nu_\lambda}) .$$

We consider at first the case $0 < \alpha < 1$. Then by $\mathrm{H}_\alpha = -\mathrm{J}_{h\alpha}$ and (4)

(5) $$\mathrm{H}_\alpha(\delta_{x_\lambda} * U_{\mu_\lambda}, \delta_{x_\lambda} * U_{\nu_\lambda}) \geqq \mathrm{H}_\alpha(U_{\mu_\lambda}, U_{\nu_\lambda})$$

In view of Lemma 1.3

$$\mathrm{H}_\alpha(U_{\mu_\lambda}, U_{\nu_\lambda}) \geqq \sum_{k=0}^{\infty} e^{-1} \frac{1}{k!} \mathrm{H}_\alpha(e^{1-\mu_\lambda(X)} \mu_\lambda^{k*}, e^{1-\nu_\lambda(X)} \nu_\lambda^{k*}) .$$

It holds $\mathrm{H}_\alpha(c\mu, d\nu) = c^\alpha d^{1-\alpha} \mathrm{H}_\alpha(\mu, \nu)$, $c, d > 0$. This means

(6) $$\mathrm{H}_\alpha(U_{\mu_\lambda}, U_{\nu_\lambda}) \geqq \exp\{-\alpha\mu_\lambda(X) - (1 - \alpha)\nu_\lambda(X)\} \sum_{k=0}^{\infty} \frac{1}{k!} \mathrm{H}_\alpha(\mu_\lambda^{k*}, \nu_\lambda^{k*})$$

As $\mathrm{H}_\alpha = -\mathrm{J}_{h\alpha}$, $0 < \alpha < 1$, the functional H_α increases under the measurable mapping $(x_1, \dots, x_n) \mapsto x_1 + \dots + x_n$, $x_i \in X$. Hence

$$\mathrm{H}_\alpha(\mu_\lambda^{k*}, \nu_\lambda^{k*}) \geqq \mathrm{H}_\alpha(\mu_\lambda \times \dots \times \mu_\lambda, \nu_\lambda \times \dots \times \nu_\lambda)$$

and, by Lemma 1.4,

$$\mathrm{H}_\alpha(\mu_\lambda^{k*}, \nu_\lambda^{k*}) \geqq (\mathrm{H}_\alpha(\mu, \nu))^k .$$

Inserting the term on the right-hand side into (6) we arrive at

(7) $$\mathrm{H}_\alpha(U_{\mu_\lambda}, U_{\nu_\lambda}) \geqq \exp\{-\alpha\,\mu_\lambda(X) - (1 - \alpha)\,\nu_\lambda(X) + \mathrm{H}_\alpha(\mu_\lambda, \nu_\lambda)\} \geqq$$
$$\geqq \exp\{\alpha(\alpha - 1)\, \mathrm{I}_\alpha(\mu_\lambda, \nu_\lambda)\} .$$

$f_\alpha \geqq 0$ implies

$$\mathrm{J}_\alpha(\mu_\lambda, \nu_\lambda) = \mathrm{J}_{f_\alpha}(\mu(\cdot \cap A_\lambda), \nu(\cdot \cap A_\lambda)) \leqq \mathrm{J}_{f_\alpha}(\mu, \nu) = \mathrm{I}_\alpha(\mu, \nu) .$$

The inequalities (5) and (7) yield

$$\liminf_{\lambda \in A} \mathrm{R}_\alpha(\delta_{x_\lambda} * U_{\mu_\lambda}, \delta_{x_\lambda} * U_{\nu_\lambda}) = \liminf_{\lambda \in A} \frac{1}{\alpha(\alpha - 1)} \ln \mathrm{H}_\alpha(\delta_{x_\lambda} * U_{\mu_\lambda}, \delta_{x_\lambda} * U_{\nu_\lambda}) \leqq \mathrm{I}_\alpha(\mu, \nu).$$

R_α is lower semi-continuous in accordance with Theorem 1.8. As $\delta_{x_\lambda} * U_{\mu_\lambda}$ and $\delta_{x_\lambda} * U_{\nu_\lambda}$ tend to P and Q, respectively, in the sense of the complete weak topology $\tau[\mathbf{\Phi}]$ Theorem 2.3 is proved for $0 < \alpha < 1$. The case $\alpha \notin (0, 1)$ may be treated analogously. □

The corollary follows from $\mathrm{H}_{1/2}(\mu, \nu) = \exp\{-\frac{1}{4}\mathrm{R}_{1/2}(\mu, \nu)\}$, Theorem 2.3, and the inequality

(8) $$\|\mathrm{P} - \mathrm{Q}\| \leqq \sqrt{[4(1 - \mathrm{H}_{1/2}^2(\mathrm{P}, \mathrm{Q})]}$$

established in [11]

We apply Theorem 2.3 to infinitely divisible distributions on separable Hilbert spaces.

Let X be a real separable Hilbert space with the scalar product $(\cdot, \cdot)$. Denote by $\mathfrak{X}$

the σ-algebra of Borel sets. Then $(X, \mathfrak{X})$ is obviously a measurable semigroup in the sense of Definition 2.1. Put $\mathbf{\Phi} = X$, i.e. $\mathbf{\Phi}$ is the family of mappings $y \mapsto (y, x)$, $x \in X$. As X is separable one easily concludes that $\mathfrak{X} = \sigma(X)$. Hence the weak topology $\boldsymbol{\tau}[X]$ for $\mathscr{P}_{\mathfrak{X}}$ is complete. Convergence in the sense of $\boldsymbol{\tau}[X]$ means the weak convergence of all finite-dimensional distributions.

Put

$$S = \{x \in X : \|x\| < 1\}.$$

The characteristic functional of $\mathsf{P} \in \mathscr{P}_{\mathfrak{X}}$ is defined by

$$\varphi_{\mathsf{P}}(x) = \int \exp\{i(x, y)\}\, \mathsf{P}(\mathrm{d}y), \quad x \in X.$$

If P is infinitely divisible then φ_{P} admits a Levy-Khintchin representation [4]. This means that there exist an element $a \in X$, a nuclear operator A and a σ-finite measure μ with

$$(9) \qquad \int (\|x\|^2\, 1_S(x) + 1_{X \setminus S}(x))\, \mu(\mathrm{d}x) < \infty, \quad \mu(\{0\}) = 0$$

such that

$$\varphi_{\mathsf{P}}(x) = \exp\{i(a, x) - \tfrac{1}{2}(Ax, x) + \int (\exp\{i(x, y)\} - 1 - i(x, y)\, 1_S(y))\, \mu(\mathrm{d}y)\}.$$

μ is called the canonical measure and (a, A, μ) the characteristic triplet. If the canonical measure μ is the zero measure then the corresponding infinitely distribution is a Gaussian distribution with expectation a and covariance operator A. This distribution is also denoted by $\mathsf{N}(a, \mathsf{A})$.

In order to prove the next theorem we need a technical lemma.

2.5. Lemma. Suppose that $\mu, \nu \in \mathscr{S}_{\mathfrak{X}}$ fulfil condition (9). If $\mathsf{I}_{1/2}(\mu, \nu) < \infty$ then

$$(10) \qquad \int 1_S(x)\, \|x\|\, |p(x) - q(x)|\, \varrho \mathrm{d}x) < \infty$$

where $\varrho = \mu + \nu$, $p = \mathrm{d}\mu/\mathrm{d}\varrho$, $q = \mathrm{d}\nu/\mathrm{d}\varrho$.

Proof.

$$\int 1_S(x)\, \|x\|\, |p(x) - q(x)|\, \varrho(\mathrm{d}x) =$$
$$= \int 1_S(x)\, \|x\|\, |\sqrt{(p(x))} - \sqrt{(q(x))}|\, |\sqrt{(p(x))} + \sqrt{(q(x))}|\, \varrho(\mathrm{d}x)$$

and by the Schwarz inequality

$$\leqq \left(\int (\sqrt{(p(x))} - \sqrt{(q(x))})^2\, \varrho(\mathrm{d}x)\right)^{1/2} \left(\int 1_S(x)\, \|x\|^2\, (\sqrt{(p(x))} + \sqrt{(q(x))})^2\, \varrho(\mathrm{d}x)\right)^{1/2}$$

Using the relation $\int (\sqrt{(p(x))} - \sqrt{(q(x))})^2\, \varrho(\mathrm{d}x) = \frac{1}{2}\, \mathsf{I}_{1/2}(\mu, \nu)$ and the simple inequality $(\sqrt{(u)} + \sqrt{(v)})^2 \leqq 2(u + v)$, $u, v \geqq 0$ we get

$$\int 1_S(x)\, \|x\|\, p(x) - q(x)|\, \varrho(\mathrm{d}x) \leqq (\tfrac{1}{2}\, \mathsf{I}_{1/2}(\mu, \nu))^{1/2}\, (2 \int 1_S(x)\, \|x\|^2\, \varrho(\mathrm{d}x))^{/12}.$$

In view of $\mathsf{I}_{1/2}(\mu, \nu) < \infty$ and (9), for both μ and ν, the stated inequality holds. □

If the conditions of Lemma 2.5 are satisfied then there exists an element $a_{\mu,\nu} \in X$ such that

$$(11) \qquad (x, a_{\mu,\nu}) = \int 1_S(z)\, (x, z)\, (p(z) - q(z)\, \varrho(\mathrm{d}z)$$

for every $x \in X$. Put $a_{\mu,\nu} = 0$ if $\mathsf{I}_{1/2}(\mu, \nu) = \infty$.

2.6. Theorem. Let P, Q be infinitely divisible distribution laws on the Borel sets of the separable Hilbert space X. If (a, A, μ) and (b, B, ν) are the characteristic triplets of P and Q respectively then

$$\mathsf{R}_\alpha(\mathsf{P}, \mathsf{Q}) \leqq \mathsf{R}_\alpha(\mathsf{N}(a, \mathsf{A}), \mathsf{N}(b + a_{\mu,\nu}, \mathsf{B})) + \mathsf{I}_\alpha(\mu, \nu).$$

2.7. Corollary. Under the assumptions of Theorem 2.6 the inequality

$$\|\mathsf{P} - \mathsf{Q}\| \leqq [4(1 - \exp\{-\tfrac{1}{2}(\mathsf{R}_{1/2}(\mathsf{N}(a, \mathsf{A}), \mathsf{N}(b + a_{\mu,\nu}, \mathsf{B})) + \mathsf{I}_{1/2}(\mu, \nu))\})]^{1/2}$$

holds.

Proof. It is easy to see that for every $-\infty < \alpha < \infty$ there exists a constant c_α such that

$$f_{1/2}(u) \leqq c_\alpha f_\alpha(u) \quad \text{for every} \quad u \in \mathbb{R}.$$

Consequently, if $\mathsf{I}_{1/2}(\mu, \nu) = \infty$ then $\mathsf{I}_\alpha(\mu, \nu) = \infty$ for every $-\infty < \alpha < \infty$ and the stated inequality is trivially fulfilled. Assume now $\mathsf{I}_{1/2}(\mu, \nu) < \infty$ and put $C_n = \{n^{-1} \leqq \|x\|\}$. Condition (9) implies $\int 1_{S\cap C_n}(x)\, \|x\|\, \mu(\mathrm{d}x) < \infty$. Consequently there exists $a_n \in X$ with

$$(x, a_n) = -\int 1_{S\cap C_n}(z)\,(x, z)\,\mu(\mathrm{d}z).$$

Let $\varrho \in \mathscr{F}_{\mathfrak{X}}$ be given. The characteristic functional of U_ϱ is

$$\varphi_{U_\varrho}(x) = \exp\{\int(\exp\{i(x, y)\} - 1)\,\varrho(\mathrm{d}y)\}. \tag{12}$$

Denote by ψ_n and χ_n the characteristic functionals of $\delta_{a_n} * U_{\mu(\cdot\cap C_n)}$ and $\delta_{a_n} * U_{\nu(\cdot\cap C_n)}$. Then, by (12),

$$\psi_n(x) = \exp\{\int_{C_n}(\exp\{i(x, y)\} - 1 - i(x, y)\,1_S(y))\,\mu(\mathrm{d}y)\}$$

$$\chi_n(x) = \exp\{-\int_{C_n} i(x, y)\,1_S(y)\,(\mu - \nu)(\mathrm{d}y) + \int_{C_n}(\exp\{i(x, y)\} - 1 - i(x, y)\,1_S(y))\,\nu(\mathrm{d}y)\}.$$

Let P′ and Q′ be infinitely divisible distributions on $(X, \mathfrak{X})$ with the characteristic triplets $(0, 0, \mu)$ and $(-a_{\mu,\nu}, 0, \nu)$, respectively. Inequality (10) and property (9) which holds for both μ and ν yield

$$\lim_{n\to\infty} \psi_n(x) = \varphi_{\mathsf{P}'}(x), \quad \lim_{n\to\infty} \chi_n(x) = \varphi_{\mathsf{Q}'}(x)$$

for every $x \in X$. Consequently $\delta_{a_n} * U_{\mu(\cdot\cap C_n)}$ and $\delta_{a_n} * U_{\nu(\cdot\cap C_n)}$ converge to P′ and Q′ in the sense of the topology $\tau[X]$.

Using Theorem 2.3 we get

$$\mathsf{R}_\alpha(\mathsf{P}', \mathsf{Q}') \leqq \mathsf{I}_\alpha(\mu, \nu). \tag{13}$$

f-divergences do not increase under measurable mappings. The definition of R_α shows that this property remains true for R_α. Hence

$$\mathsf{R}_\alpha(\mathsf{P}_1 * \mathsf{P}_2, \mathsf{Q}_1 * \mathsf{Q}_2) \leqq \mathsf{R}_\alpha(\mathsf{P}_1 \times \mathsf{P}_2, \mathsf{Q}_1 \times \mathsf{Q}_2)$$

and, by Lemma 1.4,

$$\mathsf{R}_\alpha(\mathsf{P}_1 * \mathsf{P}_2, \mathsf{Q}_1 * \mathsf{Q}_2) \leqq \mathsf{R}_\alpha(\mathsf{P}_1, \mathsf{Q}_1) + \mathsf{R}_\alpha(\mathsf{P}_2, \mathsf{Q}_2). \tag{14}$$

Put $P_1 = N(a, A)$, $P_2 = P'$, $Q_1 = N(b + a_{\mu,\nu}, B)$, $Q_2 = Q'$. Then $P_1 * P_2 = P$, $Q_1 * * Q_2 = Q$ and our statement follows from the inequalities (13) and (14). □

Corollary 2.7 follows from Theorem 2.6 and inequality (8). The calculation of the first term on the right-hand side of the inequality in Theorem 2.6 requires to answer the question under which contitions $N(a, A)$ and $N(b + a_{\mu,\nu}, B)$ are mutually singular or equivalent. We refer to [4] for details. For normal distributions on the real line, however, the values of R_α can be easily calculated.

Put for $y, z \geqq 0$, $x \in \mathbb{R}$,

$$\psi_\alpha(x, y, z) = \begin{cases} 1 & x = 0\,, \quad \max(y, z) = 0 \\ 0 & x \neq 0\,, \quad \max(y, z) = 0 \\ \dfrac{y^{1-\alpha} z^\alpha}{\alpha z + (1-\alpha)\, y} \exp\left\{-\alpha(1-\alpha)\dfrac{x^2}{\alpha z + (1-\alpha)\, y}\right\}, & \max(y, z) > 0 \end{cases}$$

Denote by $N(a, \sigma^2)$ a normal distribution on the Borel sets of the real line with expectation a and variance $\sigma^2 \geqq 0$, where $N(a, 0) = \delta_a$. In order to avoid a complicated formulation we restrict ourselves to the case $0 < \alpha < 1$. An easy calculation of the corresponding integrals shows

$$\text{(15)} \qquad H_\alpha^2(N(a_1, \sigma_1^2), N(a_2, \sigma_2^2)) = \psi_\alpha(a_1 - a_2, \sigma_1^2, \sigma_2^2)\,.$$

Corollary 2.7 and related (15) yield

2.8. Proposition. Suppose P_1, P_2 are infinitely divisible distribution laws on $(\mathbb{R}, \mathfrak{L})$ with the characteristic triplets (a_i, σ_i^2, μ_i), $i = 1, 2$. Then for $0 < \alpha < 1$

$$\text{(16)} \qquad R_\alpha(P_1, P_2) \leqq \frac{1}{2\alpha(1-\alpha)} \ln \psi_\alpha(a_1 - a_2 - a_{\mu_1,\mu_2}, \sigma_1^2, \sigma_2^2) + I_\alpha(\mu_1, \mu_2)$$

and

$$\|P_1 - P_2\| \leqq [4(1 - \psi_{1/2}(a_1 - a_2 a_{\mu_1,\mu_2}, \sigma_1^2, \sigma_2^2) \exp\{-\tfrac{1}{2} I_{1/2}(\mu_1, \mu_2)\}]^{1/2}\,.$$

2.9. Example. Denote by π_λ the Poisson distribution on $\{0, 1, \ldots\}$ with parameter $\lambda > 0$. Suppose $P_i = \pi_{\lambda_i}$, $i = 1, 2$, $\lambda_i > 0$. Then an easy calculation leads to

$$H_\alpha(P_1, P_2) = \exp\{-(\alpha\lambda_1 - (1-\alpha)\lambda_2 - \lambda_1^\alpha \lambda_2^{1-\alpha})\}\,.$$

Hence

$$R_\alpha(P_1, P_2) = \frac{1}{\alpha(1-\alpha)}(\alpha\lambda_1 + (1-\alpha)\lambda_2 - \lambda_1^\alpha\lambda_2^{1-\alpha}) = f_\alpha\left(\frac{\lambda_1}{\lambda_2}\right)\lambda_2\,.$$

Otherwise P_i is infinitely divisible and the characteristic triplet is $(0, 0, \lambda_i\delta_1)$. Consequently,

$$I_\alpha(\mu_1, \mu_2) = f_\alpha\left(\frac{\lambda_1}{\lambda_2}\right)\lambda_2\,.$$

This means equality in inequality (16). Furthermore, it is clear that the equality in (16) holds for normal distributions as well.

2.10. Example. We will illustrate by an example that the estimate (16) may be very rough if the distributions P_i are too far from normal and Poisson distributions.

Let P_i be an infinitely divisible distribution on $(\mathbb{R}, \mathfrak{L})$ which has the characteristic triplet $(0, 0, \mu_i)$ where

$$\mu_i(B) = \int_{B \setminus \{0\}} |u|^{-(1+\alpha_i)} \, \mathrm{d}u \,, \quad 0 < \alpha_i < 2 \,.$$

By definition, P_i is symmetric and stable. It is known that P_i is absolutely continuous with respect to the Lebesgue measure and the density is symmetric, increasing for $x < 0$ and decreasing for $x > 0$. Hence the measures P_1 and P_2 are not mutually singular and

(18) $$\mathsf{R}_\alpha(\mathsf{P}_1, \mathsf{P}_2) < \infty \quad \text{for} \quad 0 < \alpha < 1 \,.$$

On the other hand

$$\mathsf{I}_\alpha(\mu_1, \mu_2) = \int_{\mathbb{R} \setminus \{0\}} \frac{1}{\alpha(1-\alpha)} \left(\alpha |u|^{-(1+\alpha_1)} + (1-\alpha)\, u^{-(1+\alpha_2)} - \right.$$
$$\left. - |u|^{\alpha(1+\alpha_1)+(1-\alpha)(1+\alpha_2)}\right) \mathrm{d}u = \infty$$

for $\alpha_1 \neq \alpha_2$. This means that the right hand side of (16) is infinite while the left hand side in view of (18) is finite.

(Received September 18, 1985.)

REFERENCES

[1] H. Chernoff: A measure of efficiency for tests of a hypothesis based on sum of observations. Ann. Math. Statist. *23* (1952), 493–507.

[2] J. Csiszar: Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. Magyar Tud. Akad. Mat. Kutato Int. Közl. *8* (1964), 85–108.

[3] J. Csiszar: Information type measures of difference of probability distributions. Stud. Sci. Math. Hungar. *2* (1967), 229–318.

[4] I. I. Gichman and A. B. Skorochod: Theory of Stochastic Processes (in Russian). Nauka, Moscow 1973.

[5] P. Groneboom, J. Oosterhoff and F. H. Ruymgaart: Large deviation theorems for empirical probability measures, Ann. Probab. *7* (1979), 553–586.

[6] J. Kerstan, K. Matthes and J. Mecke: Unbegrenzt teilbare Punktprozesse, Akademie Verlag, Berlin 1974.

[7] O. Krafft and D. Plachky: Bounds for the power of likelihood ratio tests and their asymptotic properties. Ann. Math. Statist. *41* (1970), 5, 1646–1654.

[8] C. M. Newman: On the orthogonality of independent increment processes. In: Topics in Probability Theory (D. W. Strock, S. R. S. Varadhan, eds.). Courant Institute of Mathematical Sciences, N. Y. U. 1973, pp. 93–111.

[9] J. V. Prokhorov: Convergence of stochastic processes and limit theorems of probability theory (in Russian). Prob. Theory and Applic. *1*, (1956) 177–237.

[10] A. Renyi: On measures of entropy and information. In: Proc. 4th Berkeley Symposium Vol. 1, Berkeley, California 1961, pp. 547–561.

[11] I. Vajda: Limit theorems for total variation of Cartesian product measures. Stud. Sci. Math. Hungar. *6* (1973), 317–333.

[12] I. Vajda: On the f-divergence and singularity of probability measures. Period. Math. Hungar. *2* (1972), 223–234.

Prof. Dr. Friedrich Liese, Wilhelm-Pieck-Universität Rostock, Sektion Mathematik, DDR-2500 Rostock, Universitätsplatz. German Democratic Republic.