Kybernetika

Igor Vajda

Generalization of discrimination-rate theorems of Chernoff and Stein

Kybernetika, Vol. 26 (1990), No. 4, 273--288

Persistent URL: http://dml.cz/dmlcz/125435

KYBERENTIKA — VOLUME 26 (1990), NUMBER 4

GENERALIZATION OF DISCRIMINATION-RATE THEOREMS OF CHERNOFF AND STEIN

IGOR VAJDA

We consider a simple hypothesis and alternative about an abstract random observation parametrized by λ from a directed set Λ. The asymptotics over Λ is evaluated for mixed errors of the Bayes tests and second kind errors of the Neyman-Pearson tests. Similar asymptotics has been first evaluated by H. Chernoff and Ch. Stein when Λ is the set of naturals and the observation consists of the first λ terms of a sequence of i.i.d. r.v.'s (cf. [3], [4]). Extensions to the observation consisting of segments of random sequences, processes and fields have been studied by many authors. We show that the extension essentially depends on the existence of an asymptotic Rényi distance of the hypothesis and alternative and that this distance explicitly describes the discrimination rates attained by the Bayes and Neyman-Pearson tests. We do not use the theory of large deviations — all results are deduced from several simple properties of Rényi distances established by Liese and Vajda in [20] and from elementary inequalities established for Rényi distances by Kraft and Plachky and Vajda (cf. [15], [31]).

1. INTRODUCTION

Let Λ be a directed set containing a dominating sequence $(\lambda_n \mid n = 1, 2, \ldots)$. This means that Λ is ordered and contains $\lambda_1 \leqq \lambda_2 \leqq \ldots$ such that every finite subset of Λ is dominated by at least one λ_n. We consider for every $\lambda \in \Lambda$ a measurable space $(\mathcal{X}_\lambda, \mathcal{A}_\lambda)$ and probability measures P_λ, Q_λ, R_λ on it. We assume that on basis of a random observation X_λ from $(\mathcal{X}_\lambda, \mathcal{A}_\lambda, R_\lambda)$, one is testing the hypothesis H_λ: $R_\lambda = P_\lambda$ against the alternative K_λ: $R_\lambda = Q_\lambda$. By a test corresponding to $\lambda \in \Lambda$ we mean $\mathcal{A}_\lambda$-measurable real valued function φ: $\mathcal{X}_\lambda \to [0, 1]$. By Φ_λ we denote the set of all such functions. For each test $\varphi \in \Phi_\lambda$ we define probability of error of first and second kind

$$\alpha(\varphi) = \int \varphi \, \mathrm{d}P_\lambda \,, \quad \beta(\varphi) = \int (1 - \varphi) \, \mathrm{d}Q_\lambda \,. \tag{1.1}$$

Put finally

$$\Phi_\lambda(\alpha) = \{\varphi \in \Phi_\lambda \mid \alpha(\varphi) \leqq \alpha\} \,, \quad 0 < \alpha < 1 \,.$$

Hereafter we consider fixed $b > 0$ and $0 < \alpha < 1$. We define *optimum α-tests* φ_λ, $\lambda \in \Lambda$, by the condition

$$\beta(\varphi_\lambda) = \inf_{\varphi \in \Phi_\lambda(\alpha)} \beta(\varphi) \tag{1.2}$$

and *Bayes b-tests* ψ_λ, $\lambda \in \Lambda$, by the condition

$$\alpha(\psi_\lambda) + b\,\beta(\psi_\lambda) = \inf_{\varphi \in \Phi_\lambda} [\alpha(\varphi) + b\,\beta(\varphi)]\,. \tag{1.3}$$

In the terminology of Lehman [18], φ_λ is a most powerful test of level at most α. It follows from his version of the Neyman-Pearson lemma that such a test exists and that it is a likelihood ratio test of level α for $0 < \alpha \leqq 1 - P_{\lambda,s}(\mathscr{X}_\lambda)$, where $P_{\lambda,s}$ is the singular part of P_λ with respect to Q_λ. (If $1 - P_{\lambda,s}(\mathscr{X}_\lambda) < \alpha < 1$ then φ_λ is of level less than α and $\beta(\varphi_\lambda) = 0$.) Further, it is well known that ψ_λ is a Bayes test in the sense of Wald [33] for the zero-one loss function and for the prior probabilities $1/(1 + b)$, $b/(1 + b)$. It is also easy to verify that every likelihood ratio test for P_λ, Q_λ with the critical value b is ψ_λ.

In this paper we evaluate the asymptotics of $\beta(\varphi_\lambda)$ and of

$$\varepsilon(\psi_\lambda) = \alpha(\psi_\lambda) + b\,\beta(\psi_\lambda) \tag{1.4}$$

for λ growing above all bounds in Λ. This asymptotics will be given in term of distances of probability measures P_λ, Q_λ considered in the next section.

The following fact is frequently used in the sequel.

Lemma 1.1. For every real valued function γ defined on Λ and for every dominating sequence $(\lambda_n \mid n = 1, 2, \ldots)$, $\lim_{\Lambda} \gamma(\lambda)$ exists iff $\lim_{n\to\infty} \gamma(\lambda_n)$ exists, in which case the two limits are equal.

Proof. Clear in one direction. The opposite direction is easily proved by contradiction. □

2. DISTANCES OF HELLINGER, KULLBACK-LEIBLER AND RÉNYI

Let P, Q be probability measures on a measurable space $(\mathscr{X}, \mathscr{A})$ with Radon-Nikodym densities p, q with respect to a dominating σ-finite measure μ, and let us consider $a \geqq 0$. In accordance with [20] we define *Hellinger integrals*

$$H(a) = \begin{cases} \infty & \text{if } a > 1 \text{ and } P \not\ll Q \\ \int_{\{pq>0\}} (p/q)^a \,\mathrm{d}Q & \text{otherwise}\,. \end{cases} \tag{2.1}$$

We see that if $P \ll Q$ then $H(a)$ is nothing but the bilateral Laplace transform

$$H(a) = \int_{-\infty}^{\infty} \exp\{ay\}\,\mathrm{d}F(y) \tag{2.2}$$

of the distribution function

$$F(y) = Q(\{pq > 0, \ln(p/q) < y\})\,, \quad y \in \mathbb{R}\,, \tag{2.3}$$

which is complete (i.e. $F(\infty) = 1$) if also $Q \ll P$. This fact has interesting analytical consequences. These are not however considered in this paper which is based on inequalities and convexity rather than on derivatives and analyticity.

Lemma 2.1. It holds

$$H(0) = 1 - Q_s(\mathscr{X}), \quad H(1) = 1 - P_s(\mathscr{X})$$

where P_s, Q_s are singular parts of P, Q with respect to Q, P. Further, $H(a)$ is logarithmically convex, i.e. the function

$$\varrho(a) = \ln H(a) \tag{2.4}$$

is convex. Finally, $\varrho(a)$ is continuous on $[0, 1]$ and

$$\lim_{a \uparrow 1} \frac{\varrho(a)}{a - 1} = I(P \| Q) \tag{2.5}$$

where $I(P \| Q)$ is the *I-divergence* defined by

$$I(P \| Q) = \begin{cases} \infty & \text{if} \quad P \not\ll Q \\ \int_{\{pq>0\}} \ln \frac{p}{q} \, \mathrm{d}P & \text{otherwise} . \end{cases} \tag{2.6}$$

Proof. The first assertion is clear from (2.1). The second assertion is clear from (2.1) if $P \perp Q$ (singularity). For nonsingular P, Q it follows from Hölder's inequality

$$H(\alpha a_1 + (1 - \alpha) a_2) \leqq H^\alpha(a_1) H^{1-\alpha}(a_2)$$

for all $0 < \alpha < 1$ and all $a_1, a_2 \geqq 0$ so that the second assertion is clear from (2.4). The last assertion follows from (2.14) in [20]. □

The relation (2.5) motivates our definition of the *Rényi distance* (cf. [2], [30]):

$$R(a) = \begin{cases} \dfrac{\varrho(a)}{a - 1} & \text{if} \quad a \neq 1 \\ I(P \| Q) & \text{if} \quad a = 1 . \end{cases} \tag{2.7}$$

Lemma 2.2. The Rényi distance takes on values from $[0, \infty]$ and is nondecreasing. It is null at some point from $(0, 1)$ iff $P = Q$ in which case it is null everywhere. It is infinite at some point from $(0, 1)$ iff $P \perp Q$ in which case it is infinite everywhere. It is continuous on $[0, a_0]$ where $a_0 = \sup \{a > 0 \mid R(a) < \infty\}$ and where it holds $a_0 \geqq 1$ unless $P \perp Q$.

Proof. All stated properties follow directly from (2.11), (2.14) and (2.15) in [20]. □

Let us consider nonsingular P, Q and $0 < a < 1$. It holds $H(a) > 0$ and we define probability measures $P_a \ll \mu$ by the Radon-Nikodym derivatives

$$p_a = \frac{p^a q^{1-a}}{H(a)} \tag{2.8}$$

It is clear that $P_a \ll Q$. It follows from here and from (2.1) that for every $0 < t < 1/a$

the Hellinger integral

$$H(t \mid P_a, Q) = \int_{\{p_a q > 0\}} \left(\frac{p_a}{q}\right)^t \mathrm{d}Q\,, \tag{2.9}$$

which coincides with (2.1) for a, P replaced by t, P_a, satisfies the relation

$$H(t \mid P_a, Q) = \frac{H(at)}{H^t(a)}\,. \tag{2.10}$$

This relation together with the next two auxiliary results are needed in Section 6 only.

Lemma 2.3. Let φ be a likelihood ratio test of H: P against K: Q with the critical value 1 and φ_a a likelihood ratio test of H_a: P_a against K: Q with the critical value $1/H(a)$ for some $0 < a < 1$. Then the two tests coincide Q – a.s.

Proof. Clear from (2.8). □

The following assertion holds in particular for $\xi(a) = \ln H(a)$, or for $\xi(a) = H(a)$ which are convex on $(0, 1)$.

Lemma 2.4. Let ξ: $(0, 1) \to \mathbb{R}$ be convex. Then there exist right- and left-hand derivatives $\xi'_+(a)$, $\xi'_-(a)$, $0 < a < 1$, and they are continuous from the right and from the left, respectively.

Proof. Let us consider $0 < a < s < t$. The existence of

$$\xi'_+(a) = \lim_{s \downarrow a} \frac{\xi(s) - \xi(a)}{s - a}$$

is the well known consequence of Jensen inequality. Denote by $\xi'_{++}(a)$ the limit of $\xi'_+(s)$ for $s \downarrow a$. The well known monotonicity of the right-hand derivative implies

$$\xi'_+(a) \leqq \xi'_+(s) \leqq \frac{\xi(t) - \xi(s)}{t - s}\,.$$

For $s \downarrow a$ we obtain

$$\xi'_+(a) \leqq \xi'_{++}(a) \leqq \frac{\xi(t) - \xi(a)}{t - a}\,.$$

Sending $t \downarrow a$ we obtain from here the desired identity $\xi'_{++}(a) = \xi'_+(a)$. □

3. EXAMPLES

In this section we present several specifications of testing models of Section 1, and evaluate the distances of Section 2 for probability distributions P_λ, Q_λ considered in these specifications. In the rest of this paper we write

$$H_\lambda(a) = H(a \mid P_\lambda, Q_\lambda),\ \varrho_\lambda(a) = \varrho(a \mid P_\lambda, Q_\lambda),\ R_\lambda(a) = R(a \mid P_\lambda, Q_\lambda)\,, \tag{3.1}$$

i.e. the distances $H(a)$, $\varrho(a)$, $R(a)$ of measures $P = P_\lambda$, $Q = Q_\lambda$ are denoted by $H_\lambda(a)$, $\varrho_\lambda(a)$, $R_\lambda(a)$ for $\lambda \in \Lambda$.

The most important class of asymptotic testing models of Section 1 is represented by *martingale models*. In these models there is given a measurable space $(\mathscr{X}, \mathscr{A})$ and a generalized increasing sequence of sub-σ-algebras $(\mathscr{A}_\lambda^* \mid \lambda \in \Lambda)$ which is exhaustive, i.e. $\mathscr{A}_{\lambda_1}^* \subset \mathscr{A}_{\lambda_2}^*$ for $\lambda_1 \leqq \lambda_2$ and the union of all $\mathscr{A}_\lambda^*$ generates $\mathscr{A}$ (in symbols this is denoted by $\mathscr{A}_\lambda^* \uparrow \mathscr{A}$). Further, there are given two probability measures P, Q on $(\mathscr{X}, \mathscr{A})$ with densities p, q with respect to a dominating probability measure μ. If we denote by P_λ^*, Q_λ^*, μ_λ restriction of P, Q, μ on $\mathscr{A}_\lambda^*$ then the quadruples $(\mathscr{X}_\lambda, \mathscr{A}_\lambda, P_\lambda, Q_\lambda)$ of Section 1 are assumed to be defined by

$$(\mathscr{X}_\lambda, \mathscr{A}_\lambda, P_\lambda, Q_\lambda) = (\mathscr{X}, \mathscr{A}_\lambda^*, P_\lambda^*, Q_\lambda^*)\,. \tag{3.2}$$

Since P_λ, Q_λ are dominated by μ_λ, we may consider the corresponding Radon-Nikodym derivatives p_λ, q_λ. These derivatives are satisfying the conditional-expectation relations

$$p_\lambda = \mathsf{E}_\mu(p \mid \mathscr{A}_\lambda^*)\,, \quad q_\lambda = \mathsf{E}_\mu(q \mid \mathscr{A}_\lambda^*)\,, \tag{3.3}$$

i.e. they are martingales for $(\mathscr{A}^* \mid \lambda \in \Lambda)$. This fact considerably simplifies evaluation of distances $H_\lambda(a)$, $\varrho_\lambda(a)$, $R_\lambda(a)$ as well as of their limits over Λ.

The Hellinger integrals have been evaluated and applied in the martingale models of sequences and processes by many authors, e.g. by Kakutani [12], Chernoff [3], [4], Koopmans [14], Newman [23], [24], Perez [28], Newman and Stuck [25], Oosterhoff and van Zwett [26], Mémin and Shiryayev [21], Liese [19] and Kolomiets [13]. The distance $R_\lambda(1)$ has been studied in the martingale models of sequences and processes e.g. by Perez [27], Hájek [8, 9], Pinsker [29], Kullback et al. [16] and in the martingale models of random fields e.g. by Künsch [17]. The distances $R_\lambda(a)$ for martingale models of sequences and processes have been studied by Liese and Vajda [20] and Vajda [32], and for martingale models of random fields by Janžura [11].

Example 3.1. Let us consider the martingale model of stochastic processes on the time domain $[0, \infty)$, i.e. let $\Lambda = [0, \infty)$ with the natural ordering and let us consider probability measures P, Q induced by two random processes $X = (X(t) \mid 0 \leqq t < \infty)$, $Y = (Y(t) \mid 0 \leqq t < \infty)$ on the Kolmogorov Borel-lines product $(\mathscr{X}, \mathscr{A})$. The model of Section 1 is completely specified for every $0 \leqq \lambda < \infty$ by probability measures P_λ, Q_λ induced by processes $X_\lambda = (X(t) \mid 0 \leqq t \leqq \lambda)$ $Y_\lambda = (Y(t) \mid 0 \leqq t \leqq \lambda)$ on the sub-σ-algebra $\mathscr{A}_\lambda \subset \mathscr{A}$ corresponding to the time-subdomain $[0, \lambda] \subset [0, \infty)$.

Let us assume that it holds $X = m_1 + Z$, $Y = m_2 + Z$, where m_j: $[0, \infty) \to \mathbb{R}$ are continuous such that $m = m_1 - m_2$ satisfies the relations $m(0) = 0$ and

$$m(t) - m(s) = \int_s^t m_*(y)\,\mathrm{d}y\,, \quad 0 \leqq s < t < \infty\,,$$

for m_* measurable and bounded on each bounded subdomain, and where $Z = (Z(t) \mid 0 \leqq t < \infty)$ is a diffusion process satisfying the stochastic differential equation

$$dZ(t) = -A(t)\, Z(t)\, dt + B(t)\, dW(t)$$

with the initial condition $Z(0) = N(\mu, \sigma_0^2)$ for $\mu \in \mathbb{R}$, $\sigma_0^2 > 0$. In the equation it is assumed that A: $[0, \infty) \to \mathbb{R}$, B: $[0, \infty) \to (0, \infty)$ are continuous and that $W = W(t) \mid 0 \leqq t < \infty)$ is the standard Wiener process. We shall show that it holds for every $0 \leqq \lambda < \infty$

$$R_\lambda(a) = \frac{a}{2} \int_0^\lambda \eta^2(t)\, dt. \tag{3.4}$$

where

$$\eta(t) = -\frac{A(t)\, m(t) + m_*(t)}{B(t)}, \quad 0 \leqq t < \infty. \tag{3.5}$$

By Theorem 3 on p. 90 of [7], it holds $P_\lambda \ll Q_\lambda$ and the Radon-Nikodym derivative dP_λ/dQ_λ is given by

$$\frac{dP_\lambda}{dQ_\lambda} = \exp\{\int_0^\lambda \eta(t)\, dW(t) - \tfrac{1}{2}\int_0^\lambda \eta^2(t)\, dt\}.$$

Thus it follows from (2.1)

$$H_\lambda(a) = \int \left(\frac{dP_\lambda}{dQ_\lambda}\right)^a dQ_\lambda = \exp\left\{-\frac{a}{2}\int_0^\lambda \eta^2(t)\, dt\right\} \mathsf{E} \exp\{a \int_0^\lambda \eta(t)\, dW(t)\}.$$

The exponent behind the expectation has the normal density $\varphi(y/\sigma)/\sigma$ for

$$\sigma^2 = a^2 \int_0^\lambda \eta^2(t)\, dt.$$

Thus the expectation is equal to

$$\int_{-\infty}^{\infty} \exp\{y\} \frac{1}{\sigma} \psi\left(\frac{y}{\sigma}\right) dy = \exp\left\{\frac{a^2}{2}\int_0^\lambda \eta^2(t)\, dt\right\}.$$

Therefore

$$H_\lambda(a) = \exp\left\{\frac{a(a-1)}{2}\int_0^\lambda \eta^2(t)\, dt\right\}$$

and the rest follows from (2.4), (2.7) and from the continuity of $R_\lambda(a)$ from the left at 1 (see Lemma 2.2).

Example 3.2. In order to illustrate the relevance of results that follow in testing of hypotheses about random fields let us consider $\Lambda = \{1, 2, \ldots\} \times \{1, 2, \ldots\}$ ordered by the condition

$$\lambda_1 = (\lambda_1', \lambda_2'') \leqq \lambda_2 = (\lambda_2', \lambda_2'') \quad \text{iff} \quad \lambda_1' \leqq \lambda_2'' \quad \text{and} \quad \lambda_1' \leqq \lambda_2''.$$

Obviously, $(\lambda_n = (\lambda_n', \lambda_n'') \mid n = 1, 2, \ldots)$ is dominating in Λ iff

$$\lim_{n\to\infty} \min\{\lambda_n', \lambda_n''\} = \infty.$$

For example, $(\lambda_n = (n, n) \mid n = 1, 2, \ldots)$ is dominating in Λ.

Let us consider a martingale model of binary random field. In this case $\mathscr{X}$ is the set of all binary arrays $x = (\xi_{ij} \mid (i, j) \in \Lambda)$. For every $\lambda = (\lambda', \lambda'') \in \Lambda$ we consider the algebra $\mathscr{A}_\lambda$ of subsets of $\mathscr{A}$ generated by the sub-arrays $x_\lambda = (\xi_{ij} \mid i = 1, \ldots, \lambda', j = 1, \ldots, \lambda'')$. It suffices to define probability distributions P_λ, Q_λ on the sets $\mathscr{X}_\lambda$ of all arrays x_λ consistent in the usual Kolmogorov sense for different $\lambda_1, \lambda_2 \in \Lambda$, $\lambda_1 \leqq \lambda_2$. To this end decompose $\mathscr{X}_\lambda$ into the singleton $\mathscr{X}_\lambda^0$ containing the zero array x_λ consisting of pure 0's, the set $\mathscr{X}_\lambda^1$ containing all arrays x_Λ with $\xi_{11} = 1$, and the set $\mathscr{X}_\lambda^2 = \mathscr{X}_\lambda - (\mathscr{X}_\lambda^0 \cup \mathscr{X}_\lambda^1)$. Put

$$P_\lambda(\mathscr{X}_\lambda^0) = 1 - \alpha \quad (\text{cf. Sec. 1}), \quad P_\lambda(\mathscr{X}_\lambda^2) = 0,$$

and assume that P_λ is uniform on $\mathscr{X}_\lambda^1$ and Q_λ is uniform on the whole $\mathscr{X}_\lambda$. The second assumption implies that the Q_λ-probability of all arrays x_λ is

$$\beta_\lambda = \frac{1}{2^{\lambda'\lambda''}} \quad \text{for} \quad \lambda = (\lambda', \lambda'').$$

It is easy to see that $Q_\lambda(\mathscr{X}_\lambda^1) = \frac{1}{2}$. By (2.1)

$$H_\lambda(a) = \sum_{i=0}^{2} \left(\frac{P_\lambda(\mathscr{X}_\lambda^i)}{Q_\lambda(\mathscr{X}_\lambda^i)}\right)^a Q_\lambda(\mathscr{X}_\lambda^i) = \left(\frac{1-\alpha}{\beta_\lambda}\right)^a \beta_\lambda + (2\alpha)^a \tfrac{1}{2} + 0$$

and by (2.6) analogically

$$R_\lambda(1) = (1 - \alpha) \ln \frac{1-\alpha}{\beta_\lambda} + \alpha \ln 2\alpha .$$

Hence, by (2.4) and (2.7)

$$\lim_\Lambda R_\lambda(a) = \begin{cases} \ln 2 + \dfrac{a}{1-a} \ln \dfrac{1}{\alpha} & \text{if} \quad 0 \leqq a < 1 \\ \infty & \text{if} \quad a \geqq 1 \end{cases} \tag{3.6}$$

where the asymptotics in case $a \geqq 1$ may be expressed as follows

$$R_\lambda(a) \approx \begin{cases} (1-\alpha) \ln \dfrac{1}{\beta_\lambda} & \text{if} \quad a = 1 \\ \ln \dfrac{1}{\beta_\lambda} & \text{if} \quad a > 1 . \end{cases} \tag{3.7}$$

4. $\bar{R}$-DISTANCE

In this section we define an asymptotic distance for the hypothesis and alternative

$$\mathrm{H} = (H_\lambda \mid \lambda \in \Lambda) \quad \text{and} \quad \mathrm{K} = (K_\lambda \mid \lambda \in \Lambda)$$

considered in Section 1. This distance is based on the Rényi distance studied in Sections 2 and 3.

Let us consider $a_0 > 0$ and a neighborhood $A \subset [0, \infty)$ of a_0. We shall say that the asymptotic distance of H and K in the neighborhood A of a_0 exists if $R_\lambda(a_0)$ is positive and unboundedly increasing on Λ and if for every $a \in A$ there exists in the topological space $[0, \infty]$ the limit

$$\bar{R}(a) = \lim_{\Lambda} \frac{R_\lambda(a)}{R_\lambda(a_0)} . \tag{4.1}$$

The function $\bar{R}(a)$, $a \in A$, is then considered to be the asymptotic distance – it is called briefly $\bar{R}$-distance.

If there exists the $\bar{R}$-distance in the neighborhood $(0, 1)$ of $a_0 = \frac{1}{2}$ then we define

$$\bar{\varrho}(a) = (a - 1)\, \bar{R}(a), \quad 0 < a < 1 . \tag{4.2}$$

It follows from the relations $(a - 1)\, R_\lambda(a) = \varrho_\lambda(a)$, $R_\lambda(\frac{1}{2}) = -2\varrho_\lambda(\frac{1}{2})$ that the $\bar{R}$-distance in the neighborhood $(0, 1)$ of $\frac{1}{2}$ exists iff $-\varrho_\lambda(\frac{1}{2})$ is positive and unboudedly increasing on Λ and there exist in the topolgical space $[0, \infty]$ the limits

$$\bar{\varrho}(a) = \lim_{\Lambda} - \frac{\varrho_\lambda(a)}{2\varrho_\lambda(\frac{1}{2})}, \quad 0 < a < 1 . \tag{4.3}$$

In the positive case the limits (4.1) and (4.3) are related by (4.2).

Lemma 4.1. The function defined by (4.3) is finite, negative, continuous and convex on its domain of definition $(0, 1)$.

Proof. By Lemma 2.1, the function behind the limit in (4.3) is nonpositive, continuous and convex on $(0, 1)$. Consequently $\bar{\varrho}(a)$ is nonpositive and convex on $(0, 1)$. Since it is finite at $\frac{1}{2}$, it is finite and continuous on $(0, 1)$. The negativity of $\bar{\varrho}(a)$ follows from the fact that the nonnegative convex function behind the limit in (4.3) cannot exceed the piecewise linear function $f(a)$ the graph if which is passing through the points $(0, 0)$, $(\frac{1}{2}, -2)$, $(1, 0)$ of the (a, f)-plane. □

Example 4.1. Let us consider the diffusion processes of Example 3.1 and let the integral figuring in (3.4) satisfy the condition

$$\lim_{\lambda \to \infty} \int_0^\lambda \eta^2(t)\, \mathrm{d}t = \infty .$$

Then the $\bar{R}$-distance exists in the neighborhood $A = [0, \infty)$ of an arbitrary point $0 < a_0 < \infty$ and is linear on A,

$$\bar{R}(a) = a/a_0 . \tag{4.4}$$

The function (4.3) takes on the form

$$\bar{\varrho}(a) = -2a(1 - a), \quad 0 < a < 1 . \tag{4.5}$$

Example 4.2. Let us consider the random fields of Example 3.2. It follows from (3.6), (3.7) that the $\bar{R}$-distance exists in no nontrivial neighborhood of $0 < a_0 < 1$. The $\bar{R}$-distance in the neighborhood $A = [0, \infty)$ of $a_0 = 1$ is given by

$$\bar{R}(a) = \begin{cases} 0 & \text{if } 0 \leqq a < 1 \\ 1 & \text{if } a = 1 \\ \dfrac{1}{1 - \alpha} & \text{if } a > 1 . \end{cases} \tag{4.6}$$

5. GENERALIZED THEOREM OF STEIN

Let us consider the testing model of Section 1, in particular optimum α-tests φ_λ and corresponding second kind error probabilities $\beta(\varphi_\lambda)$, $\lambda \in \Lambda$. Results of this and of the following section are to the great extent based on the following fact proved as Theorem 1 in [15]: For every $0 < a < 1 < \tau$ it holds

$$(1 - \alpha)^{\tau/(\tau-1)} \exp\{-R_\lambda(\tau)\} \leqq \beta(\varphi_\lambda) \leqq (1 - a)(a/\alpha)^{a/(1-a)} \exp\{-R_\lambda(a)\}. \tag{5.1}$$

Theorem 5.1. Let the $\bar{R}$-distance exist in an open neighborhood $A \subset [0, \infty)$ of $a_0 = 1$ and let $\bar{R}(a)$ be continuous at $a = 1$. Then

$$\lim_{\Lambda} (\beta(\varphi_\lambda))^{1/R_\lambda(1)} = \exp\{-1\}. \tag{5.2}$$

Proof. Let $(\lambda_n \mid n = 1, \ldots)$ be a dominating sequence in Λ. It follows from the right-hand inequality in (5.1) and from (4.1) that it holds for every $a, \tau \in A$, $a < 1 < \tau$, (with φ_n, R_n written instead of φ_{λ_n}, R_{λ_n})

$$\limsup_{n\to\infty} (\beta(\varphi_n))^{1/R_n(1)} \leqq \exp\{-\bar{R}(a)\},$$

$$\limsup_{n\to\infty} (\beta(\varphi_n))^{1/R_n(1)} \geqq \exp\{-\bar{R}(\tau)\}.$$

Hence, by the continuity of $\bar{R}(a)$ at $a = 1$ and by the identity $\bar{R}(1) = 1$,

$$\limsup_{n\to\infty} (\beta(\varphi_n))^{1/R_n(1)} \geqq \exp\{-1\}.$$

The rest is clear from Lemma 1.1. □

Example 5.1. For diffusion processes of Example 4.1 the assumptions of Theorem 5.1 hold for $A = [0, \infty)$ so that in this case (5.2) holds for $R_\lambda(1)$ equal to half of the integral figuring in (3.4).

Example 5.2. We shall show that (5.2) need not be true without the continuity of $\bar{R}(a)$ at $a = 1$. This is an argument justifying also the assumption of existence of the $\bar{R}$-distance in Theorem 5.1.

Let us consider the random fields of Example 3.2. It follows the definition of P_λ, Q_λ that if $0 < \alpha < \frac{1}{2}$ then the optimum α-test φ_λ is nonrandomized, equal for every $\lambda = (\lambda', \lambda'') \in \Lambda$ to the characteristic function of $\mathscr{X}_\lambda - \mathscr{X}_\lambda^0$. It is clear that $\alpha(\varphi_\lambda) = P_\lambda(\mathscr{X}_\lambda - \mathscr{X}_\lambda^0) = 1 - (1 - \alpha) = \alpha$. It is also clear that $\beta(\varphi_\lambda) = Q_\lambda(\mathscr{X}_\lambda^0) = \beta_\lambda$. Therefore it follows from (3.7) that in this case

$$\lim_{\Lambda} (\beta(\varphi_\lambda))^{1/R_\lambda(1)} = \exp\{-1/(1 - \alpha)\}$$

so that (5.2) is not true. By (4.6), $\bar{R}(a)$ in this case exists in the neighborhood $[0, \infty)$ of 1, but it is discontinuous at 1.

A less general version of Theorem 5.1 appeared in [13] (cf. also the inner references to further more special versions). The proof presented here is essentially simpler. Another less general version appeared also as Theorem 11.19 in [32].

Now we present a variant of Theorem 5.1 with a weaker continuity assumption.

Theorem 5.2. Let us consider $0 < \alpha_\lambda < 1$ decreasing on Λ to zero and let φ_λ be the optimum α_λ-test. If

$$\lim_{\Lambda} (\alpha_\lambda)^{1/R_\lambda(1)} = 1 \tag{5.3}$$

and if there is $0 < s < 1$ such that the $\bar{R}$-distance exists in the neighborhood $(s, 1]$ of 1 and is continuous from left at 1, then

$$\lim_{\Lambda} (\beta(\varphi_\lambda))^{1/R_\lambda(1)} = \exp\{-1\}. \tag{5.4}$$

Proof. Analogically as in the proof of Theorem 5.1 we prove that for every $s < a < 1$

$$\limsup_{n\to\infty} (\beta(\varphi_n))^{1/R_n(1)} \leqq \exp\{-\bar{R}(a)\}$$

so that, by continuity,

$$\limsup_{n\to\infty} (\beta(\varphi_n))^{1/R_n(1)} \leqq \exp\{-1\}.$$

To prove the converse for lim inf take first into account that the family of probabilities $((\varphi_\lambda(x), 1 - \varphi_\lambda(x)) \mid x \in \mathscr{X}_\lambda)$ defines a Markov kernel from $(\mathscr{X}_\lambda, \mathscr{A}_\lambda)$ to the binary set $\{0, 1\}$. The images of measures P_λ, Q_λ defined by this kernel on $\{0, 1\}$ are $(\alpha(\varphi_\lambda), 1 - \alpha(\varphi_\lambda))$, $(1 - \beta(\varphi_\lambda), \beta(\varphi_\lambda))$. Therefore Theorem (1.24) of [20] yields the inequality

$$R_\lambda(1) \geqq \alpha(\varphi_\lambda) \ln \frac{\alpha(\varphi_\lambda)}{1 - \beta(\varphi_\lambda)} + (1 - \alpha(\varphi_\lambda)) \ln \frac{1 - \alpha(\varphi_\lambda)}{\beta(\varphi_\lambda)}$$

where $0 \ln (0/\beta) = 0$ for $\beta \geqq 0$ and $\alpha \ln (\alpha/0) = \infty$ for $\alpha > 0$. Since $R_\lambda(1)$ in (4.1) is assumed to be finite, it holds $P_\lambda \ll Q_\lambda$. Thus the singular part $P_{\lambda,s}$ considered in Section 1 below (1.3) is zero and, consequently, $\alpha(\varphi_\lambda) = \alpha_\lambda$. Since $0 < \alpha_\lambda < 1$, it follows from the above inequality $0 < \beta(\varphi_\lambda) < 1$. Therefore the above inequality yields

$$R_\lambda(1) \geqq \alpha_\lambda \ln \alpha_\lambda + (1 - \alpha_\lambda) \ln (1 - \alpha_\lambda) - (1 - \alpha_\lambda) \ln \beta(\varphi_\lambda).$$

The desired inequality

$$\liminf_{n\to\infty} (\beta(\varphi_n))^{1/R_n(1)} \geqq \exp\{-1\}$$

follows obviously from here. □

Example 5.3. It follows from the formula for $R_\lambda(1)$ in Example 3.2 that if $\alpha = \alpha_\lambda$ is satisfying (5.3) then (5.4) holds in spite of that the $\bar{R}$-distance (4.6) is discontinuous at $a = 1$.

To illustrate the applicability of Theorem 5.2 let us take the Borel line for $(\mathscr{X}_\lambda, \mathscr{A}_\lambda)$,

$\lambda \in \Lambda$ arbitrary, and consider on $\mathbb{R}$ the densities

$$P_\lambda(x) = \tfrac{1}{2} \exp\{-|x|\}, \quad q_\lambda(x) = \frac{1}{\sigma(\lambda)} \varphi\left(\frac{x}{\sigma(\lambda)}\right),$$

where φ is the standard normal density and $\sigma\colon \Lambda \to (0, \infty)$ is increasing to ∞. After some calculations one obtains that the $\bar{R}$-distance in the neighborhood $[0, \infty)$ of arbitrary $0 < a_0 \leqq 1$ is given by

$$\bar{R}(a) = \begin{cases} 0 & \text{if } a = 0 \\ 1 & \text{if } 0 < a \leqq 1 \\ \infty & \text{if } a > 1 . \end{cases} \tag{5.5}$$

(Note that a martingale model with independent observations satisfying (5.5) is easily implanted here.) We see that Theorem 5.1 is not able to say anything about the asymptotics of $\beta(\varphi_\lambda)$. Theorem 5.2. says that if (5.3) holds then also (5.4) holds with

$$R_\lambda(1) = \tfrac{1}{2} \ln \frac{\pi\sigma^2(\lambda)}{2} + \frac{1}{\sigma^2(\lambda)} - 1, \tag{5.6}$$

i.e. in fact with $R_\lambda(1) = \ln \sigma(\lambda)$. The formula (5.6) is obtained by a routine integration after substitution to (2.6). □

Let us note that the case where the $\bar{R}$-distance considered in Theorem 5.2 exists and is null for $0 \leqq a < 1$ (and consequently discontinuous at 1 from left) has been studied by Janssen [10]. He restricted the attention to special models with sequences of i.i.d. observations.

Let us also note that if Λ is the set of natural numbers, i.e. if we restrict ourselves to the statistical observation of segments of random sequences, then one easily obtains from the Shannon-McMillan-Breiman theorem (cf. [1] that, in the model of Theorem 5.1,

$$\lim_{\Lambda} (\beta(\varphi_\lambda))^{1/R_\lambda(1)} \leqq \exp\{-1\}$$

provided the limit of $R_\lambda(1)/\lambda$ is non-zero and finite (cf. [28]). The assumption here is weaker than in Theorem 5.1, but the assertion is weaker too.

6. GENERALIZED THEOREM OF CHERNOFF

In this section we consider the testing model of Section 1, in particular a Bayes b-test ψ_λ and its mixed error $\varepsilon(\psi_\lambda)$, $\lambda \in \Lambda$. It is easy to see that, up to a multiplicative factor not depending on λ, $\varepsilon(\psi_\lambda)$ is the Bayes risk (in the sense of Wald [33]) in a Bayes testing model with the hypothesis H_λ and alternative K_λ.

In the following theorem we consider the negative function $\bar{\varrho}(a)$ defined on (0, 1) by (4.2) under the assumption that the $\bar{R}$-distance exists in the neighborhood (0, 1) of $\frac{1}{2}$. We also assume that the subset argmin $\bar{\varrho}(a) \subset (0, 1)$ is nonempty (Examples 4.1

or 5.3 presents the situations where this assumption is or is not satisfied respectively). By Lemma 4.1, $\bar{\varrho}(a)$ is continuous and convex. It follows from here that argmin $\bar{\varrho}(a)$ is an interval. The closure of this interval we denote by $[a_-, a_+]$ where $0 \leqq a \leqq \leqq a_+ \leqq 1$. It also follows from here that the right- and left-hand derivatives $\bar{\varrho}'_+(a)$ and $\bar{\varrho}'_-(a)$ exist on $(0, 1)$.

Theorem 6.1. Let the $\bar{R}$-distance exists in the neighborhood $(0, 1)$ of $\frac{1}{2}$ and let either $a_+ < 1$ and $\bar{\varrho}'_+(a_+) = 0$ or $a_- > 0$ and $\bar{\varrho}'_-(a_-) = 0$. Then for any $a_* \in \operatorname{argmin} \bar{\varrho}(a)$

$$\lim_{\Lambda} (\varepsilon(\psi_\lambda))^{1/R_\lambda(1/2)} = \exp\{\bar{\varrho}(a_*)\} . \tag{6.1}$$

Proof. Let us denote by $\varepsilon_1(\psi_\lambda)$ the value of $\varepsilon(\psi_\lambda)$ for $b = 1$ (cf. (1.4)). By Theorem 2 in [31]

$$\frac{2 \min\{1, b\}}{1 + b} \varepsilon_1(\psi_\lambda) \leqq \varepsilon(\psi_\lambda) \leqq \frac{2 \max\{1, b\}}{1 + b} \varepsilon_1(\psi_\lambda) .$$

Thus it suffices to prove (6.1) for $b = 1$. Further, by Theorem 1 ibid., the mixed error $\alpha(\varphi) + b\,\beta(\varphi)$ of every Bayes b-test of H: P against K: Q is bounded above for every $a \in (0, 1)$ by $b^{1-a} H(a)$ where $H(a)$ is given by (2.1). Therefore

$$\varepsilon(\psi_\lambda) \leqq b^{1-a} \exp\{(a - 1) R_\lambda(a)\} , \quad a \in (0, 1) . \tag{6.2}$$

Let $(\lambda_n \mid n = 1, 2, \ldots)$ be a dominating sequence in Λ and let us write in the sequel ψ_n, $R_n(a)$ instead of ψ_{λ_n}, $R_{\lambda_n}(a)$. The last inequality together with (4.1) (where a_0 is replaced by $\frac{1}{2}$) and (4.2) imply

$$\limsup_{n \to \infty} (\varepsilon(\psi_n))^{1/R_n(1/2)} \leqq \exp\{\bar{\varrho}(a)\} , \quad a \in (0, 1) .$$

In particular,

$$\limsup_{n \to \infty} (\varepsilon(\psi_n))^{1/R_n(1/2)} \leqq \exp\{\bar{\varrho}(a_*)\}$$

for $a_* \in \operatorname{argmin} \bar{\varrho}(a)$. Thus in view of Lemma 1.1 it suffices to prove for $b = 1$ that there exists $a_* \in \operatorname{argmin} \bar{\varrho}(a)$ such that

$$\liminf_{n \to \infty} (\varepsilon(\psi_n))^{1/R_n(1/2)} \geqq \exp\{\bar{\varrho}(a_*)\} . \tag{6.3}$$

Let us consider $b = 1$ and let $a_+ < 1$ and $\bar{\varrho}'_+(a_+) = 0$. In this case $a_+ \in \operatorname{argmin} \bar{\varrho}(a)$. We shall prove that (6.3) holds for $a_* = a_+$ (if $a_- > 0$ and $\bar{\varrho}'_-(a_-) = 0$ then one can prove analogically that (6.3) holds for $a_* = a_-$). Let us consider $a \in (a_+, 1)$, $t \in (0, 1)$, $\tau \in (1, 1/a)$, and $\lambda \in \Lambda$. The existence of $\bar{R}$-distance implies that $R_\lambda(\frac{1}{2})$ is finite, i.e. that P_λ, Q_λ are nonsingular. Therefore $H_\lambda(a) > 0$ and there exists the measure $P_{\lambda,a}$ defined by (2.8) for $(P, Q) = (P_\lambda, Q_\lambda)$. Let φ_a be a likelihood ratio test of $\mathrm{H}_{\lambda,a}$: $P_{\lambda,a}$ against K_λ: Q_λ with the critical value $1/H_\lambda(a)$ and let

$$\alpha(\varphi_a) = \int \varphi_a \, \mathrm{d}P_{\lambda,a} , \quad \beta(\varphi_a) = \int (1 - \varphi_a) \, \mathrm{d}Q_\lambda$$

be its errors. In view of what has been said in Section 1, φ_a is a Bayes $1/H_\lambda(a)$-test.

Hence, using the inequality which (6.2) is based on, we obtain

$$\alpha(\varphi_a) + (1/H_\lambda(a))\,\beta(\varphi_a) \leqq (1/H_\lambda(a))^{1-t}\, H(t \mid P_{\lambda,a}, Q_\lambda) \quad (\text{cf. } (2.9))$$
$$= H_\lambda(at)/H_\lambda(a) \quad (\text{cf. } (2.10))\,.$$

It follows from here

$$\alpha(\varphi_a) \leqq \frac{H_\lambda(at)}{H_\lambda(a)}\,. \tag{6.4}$$

Further, we know from Section 1 that ψ_λ is a likelihood ratio test of H_λ: P_λ against K_λ: Q_λ with the critical value $b = 1$. Therefore it follows from Lemma 2.3 that ψ_λ and φ_a coincide Q_λ – a.s. so that

$$\beta(\psi_\lambda) = \beta(\varphi_a)$$

Let us consider at last an optimum $\alpha(\varphi_a)$-level test of $\mathrm{H}_{\lambda,a}$: $P_{\lambda,a}$ against K_λ: Q_λ. We denote this test by φ_a^*. By definition in Section 1,

$$\beta(\varphi_a^*) = \int (1 - \varphi_a^*)\, \mathrm{d}Q_\lambda \leqq \beta(\varphi_a)\,.$$

Hence it follows from the left-hand part of (5.1) and from (6.4)

$$\beta(\psi_\lambda) \geqq \beta(\varphi_a^*) \geqq \left(\frac{(1 - \alpha(\varphi_a))^\tau}{H(\tau \mid P_{\lambda,a}, Q_\lambda)}\right)^{1/(\tau-1)} \quad (\text{cf. } (2.9))$$
$$\geqq \left(1 - \frac{H_\lambda(at)}{H_\lambda(a)}\right)^{\tau/(\tau-1)} \left(\frac{H_\lambda^\tau(a)}{H_\lambda(a\tau)}\right)^{1/(\tau-1)} \quad (\text{cf. } (2.10))\,. \tag{6.5}$$

Putting $t = a_+/a$ and taking into account that $\varepsilon(\psi_\lambda) \geqq \beta(\psi_\lambda)$ we get from (6.5)

$$\varepsilon(\psi_\lambda) \geqq H_\lambda(a)\left(1 - \frac{H_\lambda(a_+)}{H_\lambda(a)}\right)^{\tau/(\tau-1)} \left(\frac{H_\lambda(a)}{H_\lambda(a\tau)}\right)^{1/(\tau-1)} =$$
$$= \exp\{(a-1)\, R_\lambda(a) - A_\lambda(a,\tau) - B_\lambda(a,\tau)\}$$

where

$$A_\lambda(a,\tau) = -\frac{\tau}{\tau - 1} \ln\left(1 - \frac{H_\lambda(a_+)}{H_\lambda(a)}\right),$$

$$B_\lambda(a,\tau) = \frac{1}{\tau - 1}\left[(a\tau - 1)\, R_\lambda(a\tau) - (a-1)\, R_\lambda(a)\right].$$

The easily verified relation

$$\lim_{\Lambda} \left(\frac{H_\lambda(a_+)}{H_\lambda(a)}\right)^{1/R_\lambda(1/2)} = \exp\{\bar{\varrho}(a_+) - \bar{\varrho}(a)\} \quad (\text{cf. } (4.3), (2.4))$$

together with the obvious inequality $\bar{\varrho}(a_+) < \bar{\varrho}(a)$ implies that

$$\lim_{\Lambda} A_\lambda(a,\tau) = 0\,.$$

This and the following easily verified fact

$$\lim_{\Lambda} \frac{B_\lambda(a, \tau)}{R_\lambda(\frac{1}{2})} = a \frac{\bar{\varrho}(a\tau) - \bar{\varrho}(a)}{a\tau - a}$$

imply that in the above introduced notation

$$\liminf_{n \to \infty} (\varepsilon(\psi_n))^{1/R_n(1/2)} \geqq \exp\{\bar{\varrho}(a) - a\,\bar{\varrho}'_+(a)\}\,.$$

Here $\bar{\varrho}(a)$ is continuous by Lemma 4.1 and $\bar{\varrho}'_+(a)$ is continuous from the right by Lemma 4.1 and 2.4. Thus the right-hand side of last inequality tends for $a \downarrow a_+$ to $\bar{\varrho}(a_+) - a_+\,\bar{\varrho}'_+(a_+) = \bar{\varrho}(a_+)$. Therefore (6.3) holds for $a_* = a_+$. □

It is clear that if the function $\bar{\varrho}$ exists and has a unique point of minima a_* at which it is differentiable then all assumptions of Theorem 6.1 hold. Using the results (2.2), (2.3), one can establish differentiability of $\bar{\varrho}$ on (0, 1) under wide assumptions about (P_λ, Q_λ), $\lambda \in \Lambda$.

Example 6.1. Let us consider the diffusion processes of Example 4.1. By (4.5), argmin $\bar{\varrho}(a) = \{\frac{1}{2}\}$ and all assumptions of Theorem 6.1 hold. Therefore Bayes tests ψ_λ, $\lambda \in [0, \infty)$, satisfy the asymptotic relation

$$\lim_{\lambda \to \infty} (\varepsilon(\psi_\lambda))^{1/R_\lambda(1/2)} = \exp\{-\tfrac{1}{2}\}$$

where $R_\lambda(\frac{1}{2})$ is given by (3.4) for $a = \frac{1}{2}$.

Example 6.2. The random fields model of Example 3.2 is not satisfying the assumptions of Theorem 6.1. By what is said in Example 4.2, the $\bar{R}$-distance required in the present section does not exist. The model of Example 5.3 has the desired $\bar{R}$-distance but it is constant on (0, 1), equal to 1. Therefore the function

$$\bar{\varrho}(a) = a - 1\,, \quad 0 < a < 1\,,$$

has argmin $\bar{\varrho}(a)$ empty and nothing can be said here about Bayes tests.

Let us note that results similar to Theorem 6.1 can also be deduced from known general large deviation theorems (e.g. Ellis [5] or Theorem VII.6.1. in Ellis [6]). Similar approach was in fact used already in the fundamental paper of Chernoff [3]. The result of Chernoff valid for sequences of i.i.d. observations has been extended to Markov chains and processes by Koopmans [14], Nemetz [22], and Newman and Stuck [25].

REFERENCES

[1] P. H. Algoet and T. M. Cover: A sandwich proof of the Shannon-McMillan-Breiman theorem. Ann. Probab. *16* (1988), 899—909.

[2] A. Bhattacharyya: On some analogous to the amount of information and their uses in statistical estimation. Sankhya *8* (1946), 1—14.

[3] H. Chernoff: A measure of asymptotic efficiency for tests of a hypothesis based on a sum of observations. Ann. Math. Statist. *23* (1952), 493—507.
[4] H. Chernoff: Large sample theory: Parametric case. Ann. Math. Statist. *27* (1956), 1—22.
[5] R. S. Ellis: Large deviations for a general class of random vectors. Ann. Probab. *12* (1984), 1—12.
[6] R. S. Ellis: Entropy, Large Deviations, and Statistical Mechanics. Springer-Verlag, Berlin—Heidelberg—New York 1985.
[7] I. I. Gikhman and A. V. Skorokhod: Stochastic Differential Equations (in Russian). Naukova Dumka, Kiev 1986.
[8] J. Hájek: A property of J-divergences or marginal probability distributions. Czechoslovak Math. J. *8* (1958), 460—463.
[9] J. Hájek: On a property of normal distributions of an arbitrary stochastic process. Czechoslovak Math. J. *8* (1958), 610—618.
[10] A. Janssen: Asymptotic properties of Neyman-Pearson test for infinite Kullback-Leibler information. Ann. Statist. *14* (1986), 1068—1079.
[11] M. Janžura: Divergences of Gauss-Markov random fields with application to statistical inference. Kybernetika *24* (1988), 6, 401—412.
[12] S. Kakutani: On equivalence of infinite product measures. Ann. of Math. *49* (1948), 214—226.
[13] E. I. Kolomietz: Asymptotic behaviour of type II errors of Neyman-Pearson test (in Russian). Teor. Veroyatnost. i Primenen. *33* (1986), 503—522.
[14] S. H. Koopmans: Asymptotic rate of discrimination for Markov processes. Ann. Math. Statist. *31* (1960), 982—994.
[15] O. Kraft and D. Plachky: Bounds for the power of likelihood ratio test and their asymptotic properties. Ann. Math. Statist. *41* (1970), 1646—1654.
[16] S. Kullback, J. C. Keegel, and J. H. Kullback: Topics in Statistical Information Theory. Springer-Verlag, Berlin—Heidelberg—New York 1987.
[17] H. Künsch: Thermodynamics and statistical analysis of Gaussian random fields. Z. Wahrsch. verw. Geb. *58* (1981) 407—421.
[18] E. L. Lehman: Testing Statistical Hypotheses. J. Wiley, New York 1959.
[19] F. Liese: Hellinger integrals of diffusion processes. Statistics *17* (1986), 63—78.
[20] F. Liese and I. Vajda: Convex Statistical Distances. Teubner, Leipzig 1987.
[21] J. Mémin and A. N. Shiryayev: Distances de Hellinger-Kakutani des lois correspondant à acroissements indépendants. Z. Wahrsch. verw. Geb. *70* (1985), 67—89.
[22] T. Nemetz: On the α-divergence rate for Markov-dependent hypotheses. Problems Control Inform. Theory *3* (1974), 147—155.
[23] C. M. Newman: The inner product of path space measures corresponding to random processes with independent increments. Bull. Amer. Math. Soc. *78* (1982), 268—272.
[24] C. M. Newman: The orthogonality of independent increment processes. In: Topics in Probability Theory (D. W. Strook, S. R. S. Varadhan, eds.), Convent Inst. of Math. Sciences, New York 1973, pp. 93—111.
[25] C. M. Newman and B. W. Stuck: Chernoff bounds for discriminating between two Markov processes. Stochastics *2* (1979), 139—153.
[26] J. Oosterhooff and W. R. van Zwett: A note on contiguity and Hellinger distance. In: Contributions to Statistics (J. Hájek Memorial Volume, J. Jurečková, ed.), Reidel, Dordrecht 1979.
[27] A. Perez: Notions generalisées d'incertitude, d'entropie et d'information du point de vue de la theorie de martingales. In: Trans. 1st Prague Conf. on Inform. Theory, Statist. Dec. Functions, Random Processes. Publ. House Czechosl. Acad. Sci., Prague 1957, pp. 183—206.
[28] A. Perez: Generalization of Chernoff's result on the asymptotic discernibility of two random processes. Colloq. Math. Soc. János Bolyai *9* (1974), 619—632.

[29] M. S. Pinsker: Information and Information Stability of Random Variables and Processes. Holden-Day, San Francisco 1964.

[30] A. Rényi: On measures of entropy and information. In: Proc. 4th Berkeley Symp. on Probab. Theory Math. Statist., Vol. 1, Berkeley Univ. Press, Berkeley 1961, pp. 447—561.

[31] I. Vajda: Limit theorems for total variation of cartesian product measures. Studia Sci. Math. Hungar. *6* (1971), 317—333.

[32] I. Vajda: Theory of Statistical Inference and Information. Kluwer, Dordrecht—Boston 1989.

[33] A. Wald: Statistical Decision Functions. J. Wiley, New York 1961.

Ing. Igor Vajda, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8. Czechoslovakia.