Martin Janžura

Divergences of Gauss-Markov random fields with application to statistical inference

Kybernetika, Vol. 24 (1988), No. 6, 401--412

Persistent URL: http://dml.cz/dmlcz/124184

KYBERNETIKA — VOLUME 24 (1988), NUMBER 6

DIVERGENCES OF GAUSS-MARKOV RANDOM FIELDS WITH APPLICATION TO STATISTICAL INFERENCE

MARTIN JANŽURA

In addition to the previous asymptotic theory of parameter estimation (cf. [2]) further asymptotic properties of the Gauss-Markov random fields are studied in the present paper. The explicit formulas for the entropy rate, the I-divergence, and the α-divergence are obtained. Applications to parameter estimation and hypotheses testing are included.

1. INTRODUCTION

The Gauss-Markov random fields are used as the probability models for the statistical analysis of spatial data. In the preceding paper [2] a convenient way of their parameter description was given, and a method for the parameter estimation was proposed.

The method and its asymptotic properties are closely connected with those characteristics of distributions which are studied in frame of thermodynamics or, parallelly, in frame of information theory. Namely, we mean the entropy rate, the I-divergence (information gain), and the α-divergence (for the definitions see below).

Thus, the present paper is devoted to deriving the explicit form of these characteristics for the Gauss-Markov random fields. The main results are obtained in Section 3 with the proofs and some auxiliary results in the following Section 4.

As an application in Section 5 we investigate the connection between the estimator proposed in [2] are the so called minimum distance methods (cf. [6]). This connection is based on the considered notions, and it seems to be interesting and fruitful from both the computational and the methodological aspects.

Section 6 contains an application of the results to testing hypotheses, namely the appropriate versions of the Stein and the Chernoff theorems on the asymptotic behaviour of the error probabilities are introduced.

Some of the asymptotic results were attained by Künsch [3] with the aid of a bit different methods. But only the Gauss-Markov fields with zero mean value were considered, and the question of α-divergence was not studied at all.

The divergences of probability measures and related topics being concerned, we follow [8] as the main reference.

2. PRELIMINARIES

We shall only briefly recall some basic definitions and results concerning Gauss-Markov random fields (for more details cf. [2] – Section 2 and Section 3).

By a Gauss-Markov random field we mean a stochastic process $\{X_t\}_{t\in\mathscr{Z}^d}$ on a d-dimensional lattice $\mathscr{Z}^d$ with

i) translation invariant distribution $\mathbf{P}$;

ii) Gaussian finite-dimensional marginals $\mathbf{P}^{\mathscr{V}}$,

$$\mathscr{V} \in \mathfrak{K} = \{\mathscr{V} \subset \mathscr{Z}^d;\, 0 < |\mathscr{V}| = \operatorname{card} \mathscr{V} < \infty\}\,;$$

iii) spectral density given by

$$f_U(\lambda) = [2\,.\sum_{k\in\mathscr{M}} U(k)\cos k\lambda]^{-1} \quad \text{for every} \quad \lambda \in \mathscr{I}_d = [-\pi, \pi]^d\,,$$

where

$$U = \{U(k)\}_{k\in\mathscr{M}} \in \mathfrak{D}_{\mathscr{M}} = \{U \in \mathscr{R}^{|\mathscr{M}|};\, \sum_{k\in\mathscr{M}} U(k)\cos k\lambda > 0 \text{ for every } \lambda \in \mathscr{I}_d\}\,,$$

$\mathscr{M} \in \mathfrak{K}$, $\mathscr{M} \subset \mathscr{Z}^d_+ = \{t \in \mathscr{Z}^d;\, t \geqq 0\}$ ("$\geqq$" is the lexicographical ordering);

iv) constant mean value given by

$$\mu_{h,U} = -h\,.\,f_U(0)\,, \quad h \in \mathscr{R}\,.$$

Thus, we can see that the distribution $\mathbf{P}$ depends on a $(1 + |\mathscr{M}|)$-dimensional parameter $\boldsymbol{\theta} = (h, U) \in \mathscr{R} \times \mathfrak{D}_{\mathscr{M}} = \boldsymbol{\Theta}$.

In what follows we shall use the term "random field" for the distribution and we shall deal with the (locally asymptotically normal – cf. [2]) parameter family $\mathbb{P} = \{\mathbf{P}_\theta\}_{\theta\in\Theta}$ of Gauss-Markov random fields.

Let us note that the Gauss-Markov random field $\mathbf{P}_\theta$ may be understood as a Gibbs field with a finite range pair potential U given by

$$U_{\{t\}}(x_t) = U(0)\, x_t^2 + h x_t \quad \text{for every} \quad t \in \mathscr{Z}^d\,,$$

and

$$U_{\{s,t\}}(x_s, x_t) = U_{\{t,s\}}(x_t, x_s) = U(t - s)\, x_t x_s \quad \text{for} \quad t - s \in \mathscr{M} \setminus \{0\}$$

(for detailed treatment of this approach cf. [1] and [3]).

For every $U \in \mathfrak{D}_{\mathscr{M}}$ the corresponding covariance function

$$R_U(t) = (2\pi)^{-d} \int_{\mathscr{I}_d} e^{it\lambda} f_U(\lambda)\, d\lambda\,, \quad t \in \mathscr{Z}^d\,,$$

is absolutely convergent: $\sum_{t\in\mathscr{Z}^d} |R_U(t)| < \infty$.

Let a function $\boldsymbol{a}: \mathscr{Z}^d \to \mathscr{R}$ be given by

$$\boldsymbol{a}(0) = 2 \,.\, \boldsymbol{U}(0)$$
$$\boldsymbol{a}(t) = \boldsymbol{U}(t) \quad \text{for} \quad t \in \mathscr{M} \setminus \{0\} \quad \text{or} \quad -t \in \mathscr{M} \setminus \{0\}$$
$$\boldsymbol{a}(t) = 0 \quad \text{otherwise}\,.$$

Then we can easily verify that

$$\boldsymbol{a}(t) = (2\pi)^{-d} \int_{\mathscr{I}_d} e^{-it\lambda} [f_U(\lambda)]^{-1} \, d\lambda \quad \text{for every} \quad t \in \mathscr{Z}^d$$

and the infinite matrix $\mathbf{A} = (\boldsymbol{a}(t - s))_{t,s \in \mathscr{Z}^d}$ is inverse to the infinite covariance matrix $\mathbf{R}_U = (\boldsymbol{R}_U(t - s))_{t,s \in \mathscr{Z}^d}$.

3. ENTROPY RATE, I-DIVERGENCE, AND α-DIVERGENCE

Suppose a stationary random field $\mathbf{P}$ to be given by its densities $p^{\mathscr{V}}$, $\mathscr{V} \in \mathfrak{K}$.

Then we define the entropy rate as the limit

$$\mathsf{S}(\mathbf{P}) = \lim_{\mathscr{V} \nearrow \mathscr{Z}^d} |\mathscr{V}|^{-1} \, \mathsf{E}_{\mathbf{P}}\{-\log p^{\mathscr{V}}\}$$

which always exists (it can be equal to $-\infty$, cf. [4]).

For a pair $\mathbf{P}$, $\mathbf{Q}$ of stationary random fields with densities $p^{\mathscr{V}}$ and $q^{\mathscr{V}}$, $\mathscr{V} \in \mathfrak{K}$, respectively, we define the I-divergence of $\mathbf{P}$ with respect to $\mathbf{Q}$ by

$$\mathsf{H}_I(\mathbf{P} \mid \mathbf{Q}) = \lim_{\mathscr{V} \nearrow \mathscr{Z}^d} |\mathscr{V}|^{-1} \, \mathsf{E}_{\mathbf{P}}\{\log p_{\mathscr{V}} / q_{\mathscr{V}}\}$$

whenever the integrals and the limit exist. Otherwise we set $\mathsf{H}_I(\mathbf{P} \mid \mathbf{Q}) = \infty$.

The convergence $\mathscr{V} \nearrow \mathscr{Z}^d$ is defined in order to satisfy

$$|\mathscr{V}|^{-1} \, |\mathscr{V}_k| \to 1 \quad \text{for every} \quad k \in \mathscr{Z}^d\,,$$

where $\mathscr{V}_k = \mathscr{V} \cap (\mathscr{V} + k)$.

We denote by $\mathbb{M}$ the family of all stationary random fields with finite second moments, finite entropy rate, and with all the marginal densities, i.e.

$$\mathbf{P} \in \mathbb{M} \quad \text{iff} \quad \mathsf{E}_{\mathbf{P}}[X_0^2] < \infty\,, \quad \mathsf{S}(\mathbf{P}) > -\infty\,, \quad \text{and} \quad p^{\mathscr{V}} \quad \text{exists for every} \quad \mathscr{V} \in \mathfrak{K}.$$

Theorem 3.1. Let $\mathbf{Q} \in \mathbb{M}$, $\mathbf{P}_\theta \in \mathbb{P}$. Then

i) $\mathsf{S}(\mathbf{P}_\theta) = \frac{1}{2}(1 + \log(2\pi) + (2\pi)^{-d} \int_{\mathscr{I}_d} \log f_U(\lambda) \, d\lambda)$;

ii) $\mathsf{H}_I(\mathbf{Q} \mid \mathbf{P}_\theta) = \frac{1}{2}\{\log(2\pi) + (2\pi)^{-d} \int_{\mathscr{I}_d} \log f_U(\lambda) \, d\lambda + 2 \sum_{k \in \mathscr{M}} \boldsymbol{U}(k) \,.\, [\boldsymbol{R}_{\mathbf{Q}}(k) + (\mu_{h,U} - \nu_{\mathbf{Q}})^2]\} - \mathsf{S}(\mathbf{Q})$,

where $\nu_{\mathbf{Q}} = \mathsf{E}_{\mathbf{Q}}[X_0]$ and $\boldsymbol{R}_{\mathbf{Q}}(k) = \mathsf{E}_{\mathbf{Q}}[(X_0 - \nu_{\mathbf{Q}})(X_k - \nu_{\mathbf{Q}})]$ are the mean value and the covariance function, respectively, of the random field $\mathbf{Q}$.

The proof of the theorem is given in Section 4. Here we continue with some easy consequences.

Corollary 3.2. i) Let $\mathbf{P}_\theta$, $\mathbf{P}_{\theta^*} \in \mathbb{P}$. Then

$$\mathsf{H}_I(\mathbf{P}_{\theta^*} \mid \mathbf{P}_\theta) = \frac{1}{2}\left\{(2\pi)^{-d} \int_{\mathscr{I}^d} \log \frac{f_U(\lambda)}{f_{U^*}(\lambda)} \, \mathrm{d}\lambda - 1\right\} +$$
$$+ \sum_{k \in \mathscr{M}} U(k) \left[\mathbf{R}_{U^*}(k) + (\mu_{h,U} - \mu_{h^*,U^*})^2\right].$$

ii) For $\mathbf{Q} \in \mathbb{M}$ it holds

$$\mathsf{H}_I(\mathbf{Q} \mid \mathbf{P}_{\theta^0}) = \min_{\theta \in \Theta} \mathsf{H}_I(\mathbf{Q} \mid \mathbf{P}_\theta) \quad \text{iff}$$
$$\nu_{\mathbf{Q}} = \mu_{h^0,U^0} \text{ and } \boldsymbol{R}_{\mathbf{Q}}(k) = \boldsymbol{R}_{U^0}(k) \text{ for every } k \in \mathscr{M}.$$

Proof. The first statement is straightforward if we properly substitute for the terms from Theorem 3.1.

If the minimum in the second statement is reached at θ^0 we obtain the claimed identities by differentiation of the I-divergence with respect to the parameters.

From the other side if the condition is satisfied we have

$$\mathsf{H}_I(\mathbf{Q} \mid \mathbf{P}_\theta) - \mathsf{H}_I(\mathbf{Q} \mid \mathbf{P}_{\theta^0}) = \mathsf{H}_I(\mathbf{P}_{\theta^0} \mid \mathbf{P}_\theta) \geqq 0$$

for every $\theta \in \Theta$. □

It is not difficult to see that for a pair $\mathbf{P}$, $\mathbf{Q}$ of stationary random fields with densities $p^{\mathscr{V}}$ and $q^{\mathscr{V}}$, $\mathscr{V} \in \mathfrak{K}$, respectively, we have

$$\mathsf{E}_{\mathbf{P}}\left[\log \frac{p^{\mathscr{V}}}{q^{\mathscr{V}}}\right] = \lim_{\alpha \to 0} \frac{1}{\alpha(\alpha - 1)} \log \mathsf{E}_{\mathbf{P}}\left[\left(\frac{q^{\mathscr{V}}}{p^{\mathscr{V}}}\right)^\alpha\right].$$

whenever the expressions make sense, and similarly

$$\mathsf{E}_{\mathbf{Q}}\left[\log \frac{q^{\mathscr{V}}}{p^{\mathscr{V}}}\right] = \lim_{\alpha \to 1} \frac{1}{\alpha(\alpha - 1)} \log \mathsf{E}_{\mathbf{P}}\left[\left(\frac{q^{\mathscr{V}}}{p^{\mathscr{V}}}\right)^\alpha\right].$$

Therefore, we can understand the I-divergence as a special case of the α-divergence defined by

$$\mathsf{H}_\alpha(\mathbf{P} \mid \mathbf{Q}) = \frac{1}{\alpha(\alpha - 1)} \lim_{\mathscr{V} \nearrow \mathscr{Z}^d} |\mathscr{V}|^{-1} \log \mathsf{E}_{\mathbf{P}}\left[\left(\frac{q^{\mathscr{V}}}{p^{\mathscr{V}}}\right)^\alpha\right]$$

if the integrals and the limit exist.

For the sake of brevity let us denote

$$w(c) = (1 - c)^{-1} c \quad \text{for} \quad c < 1, \quad \text{and}$$
$$w(c) = +\infty \qquad \text{for} \quad c \geqq 1.$$

Theorem 3.3. Let $\mathbf{P}_\theta$, $\mathbf{P}_{\theta^*} \in \mathbb{P}$,

$$\alpha \in \left(-w\left(\min_{\lambda \in \mathscr{I}^d} \frac{f_U(\lambda)}{f_{U^*}(\lambda)}\right), \quad 1 + w\left(\min_{\lambda \in \mathscr{I}^d} \frac{f_{U^*}(\lambda)}{f_U(\lambda)}\right)\right).$$

Then

$$\mathsf{H}_\alpha(\mathbf{P}_{\theta^*} \mid \mathbf{P}_\theta) = \tfrac{1}{2}\{(2\pi)^{-d} \int_{\mathscr{I}^d} [(1 - \alpha)^{-1} \log f_U(\lambda) + \alpha^{-1} \log f_{U^*}(\lambda) -$$
$$- (\alpha(1 - \alpha))^{-1} \log f_{\alpha U + (1-\alpha) U^*}(\lambda)] \, \mathrm{d}\lambda +$$
$$+ [f_U(0) \cdot f_{U^*}(0)]^{-1} \cdot f_{\alpha U + (1-\alpha) U^*}(0) \cdot (\mu_{h,U} - \mu_{h^*,U^*})^2\}.$$

The proof of the theorem is given again in the next section.

Corollary 3.4. Let $\mathbf{P}_{\theta}, \mathbf{P}_{\theta^*} \in \mathbb{P}$. Then

$$\mathsf{H}_I(\mathbf{P}_{\theta^*} \mid \mathbf{P}_{\theta}) = \lim_{\alpha \to 0} \mathsf{H}_{\alpha}(\mathbf{P}_{\theta^*} \mid \mathbf{P}_{\theta})$$

and

$$\mathsf{H}_I(\mathbf{P}_{\theta} \mid \mathbf{P}_{\theta^*}) = \lim_{\alpha \to 1} \mathsf{H}_{\alpha}(\mathbf{P}_{\theta^*} \mid \mathbf{P}_{\theta}) .$$

Proof. We may write

$$\lim_{\alpha \to 0} \mathsf{H}_{\alpha}(\mathbf{P}_{\theta^*} \mid \mathbf{P}_{\theta}) = \frac{1}{2}\left\{(2\pi)^{-d} \int_{\mathscr{I}_d} \left[\log \frac{f_U(\lambda)}{f_{U^*}(\lambda)} + \frac{f_{U^*}(\lambda)}{f_U(\lambda)} - 1\right] \mathrm{d}\lambda + \right.$$

$$\left. + \left[f_U(0)\right]^{-1} (\mu_{h,U} - \mu_{h^*,U^*})^2 \right\} ,$$

and, since $\left[f_U(0)\right]^{-1} = 2 \sum_{k \in \mathscr{M}} U(k)$ and

$$(2\pi)^{-d} \int_{\mathscr{I}_d} \frac{f_{U^*}(\lambda)}{f_U(\lambda)} \mathrm{d}\lambda = 2 \sum_{k \in \mathscr{M}} U(k) R_{U^*}(k) ,$$

we obtain the claimed statement for $\alpha \to 0$.

For $\alpha \to 1$ we proceed in the same way. □

4. PROOFS AND AUXILIARY RESULTS

In this section we intend to prove the theorems introduced in the preceding section. In fact, we shall prove something more general.

Let $\mathfrak{G}$ be the class of bounded, positive, real valued and differentiable functions f defined on $\mathscr{I}_d$ by

$$f(\lambda) = \sum_{t \in \mathscr{Z}^d} r(t)\, \mathrm{e}^{it\lambda} \quad \text{for every} \quad \lambda \in \mathscr{I}_d .$$

Then, according to Corollary VII 1.9 in [5] the Fourier coefficients are absolutely summable, i.e.

$$\sum_{t \in \mathscr{Z}^d} |r(t)| < \infty ,$$

and the same is true for the Fourier coefficients of the reciprocal function, i.e.

$$\sum_{t \in \mathscr{Z}^d} |a(t)| < \infty$$

where

$$[f(\lambda)]^{-1} = \sum_{t \in \mathscr{Z}^d} a(t)\, \mathrm{e}^{it\lambda} \quad \text{for every} \quad \lambda \in \mathscr{I}_d .$$

It is easy to see that the infinite matrix $\mathbf{A} = \left(a(t-s)\right)_{t,s \in \mathscr{Z}^d}$ is inverse to the matrix $\mathbf{R} = \left(r(t-s)\right)_{t,s \in \mathscr{Z}^d}$.

Let us denote $f^{\max} = \max_{\lambda \in \mathscr{I}_d} f(\lambda)$, $f^{\min} = \min_{\lambda \in \mathscr{I}_d} f(\lambda)$. For an $m \times m$ matrix $\mathbf{D}$ with the eigenvalues $c_j, j = 1, \ldots, m$, let us denote $\varrho(\mathbf{D}) = \max_{j=j,\ldots,m} |c_j|$. Further, we introduce some useful basic results concerning positive definite matrices.

Lemma 4.1. i) For $\mathbf{C}, \mathbf{D} > 0$ it holds

$$\frac{\mathrm{d}}{\mathrm{d}\gamma} \log \operatorname{Det}(\gamma\mathbf{C} + (1-\gamma)\,\mathbf{D}) = \operatorname{Tr}\{(\gamma\mathbf{C} + (1-\gamma)\,\mathbf{D})^{-1}(\mathbf{C} - \mathbf{D})\}.$$

ii) For $\mathbf{C}, \mathbf{D} \geqq 0$ it holds

$$0 \leqq \operatorname{Tr}(\mathbf{CD}) \leqq \varrho(\mathbf{C}) \operatorname{Tr}(\mathbf{D}).$$

The results are well-known and need not be proved.

Lemma 4.2. Let $\mathbf{R}_*$ be an arbitrary infinite covariance matrix, i.e. $\mathbf{R}_*$ is positive semidefinite and $\mathbf{R}_* = (\mathbf{R}_*(t, s) = r_*(t-s))_{t,s\in\mathscr{Z}^d}$, and $\mathbf{A}$ corresponds to some $f \in \mathfrak{G}$. Then

i) $\lim\limits_{\mathscr{V}\nearrow\mathscr{Z}^d} |\mathscr{V}|^{-1} \operatorname{Tr}(\mathbf{R}_*^{\mathscr{V}\mathscr{V}}\mathbf{A}^{\mathscr{V}\mathscr{V}}) = \sum\limits_{t\in\mathscr{Z}^d} r_*(t)\, a(t)$;

ii) $\lim\limits_{\mathscr{V}\nearrow\mathscr{Z}^d} |\mathscr{V}|^{-1} \operatorname{Tr}(\mathbf{R}_*^{\mathscr{V}\mathscr{V}}\mathbf{A}^{\mathscr{V}\mathscr{V}^c}(\mathbf{A}^{\mathscr{V}^c\mathscr{V}^c})^{-1}\,\mathbf{A}^{\mathscr{V}^c\mathscr{V}}) = 0$.

Proof. We may write

$$|\mathscr{V}|^{-1} \sum_{t,s\in\mathscr{V}} r_*(t-s)\, a(s-t) = \sum_{k\in\mathscr{Z}^d} r_*(k)\, a(k) \,.\, |\mathscr{V}|^{-1} \,.\, |\mathscr{V}_k|\,.$$

Since $|r_*(k)| \leqq r_*(0)$ and $\sum_{k\in\mathscr{Z}^d} |a(k)| < \infty$, the convergence is dominated and we obtain the first statement.

In order to prove the second one we observe $\mathbf{R}^{\mathscr{V}^c\mathscr{V}^c} - (\mathbf{A}^{\mathscr{V}^c\mathscr{V}^c})^{-1} \geqq 0$ and therefore

$$0 \leqq |\mathscr{V}|^{-1} \operatorname{Tr}[\mathbf{R}_*^{\mathscr{V}\mathscr{V}}\mathbf{A}^{\mathscr{V}\mathscr{V}^c}(\mathbf{A}^{\mathscr{V}^c\mathscr{V}^c})^{-1}\,\mathbf{A}^{\mathscr{V}^c\mathscr{V}}] \leqq |\mathscr{V}|^{-1} \operatorname{Tr}[\mathbf{R}_*^{\mathscr{V}\mathscr{V}}\mathbf{A}^{\mathscr{V}\mathscr{V}^c}\mathbf{R}^{\mathscr{V}^c\mathscr{V}^c}\mathbf{A}^{\mathscr{V}^c\mathscr{V}}] \leqq$$
$$\leqq r_*(0) \,.\, \sum_{k,l\in\mathscr{Z}^d} |a(k)|\,|a(l)| \,.\, \sum_{m\in\mathscr{Z}^d} |r(m)| \,.\, |\mathscr{V}|^{-1} \,.\, |(\mathscr{V} + k) \cap (\mathscr{V} + l + m) \cap$$
$$\cap\, \mathscr{V}^c \cap (\mathscr{V}^c + m)| \to 0$$

again by the dominated convergence arguments. □

Lemma 4.3. For $f_1, f_2, f_3 \in \mathfrak{G}$ it holds

$$\lim_{\mathscr{V}\nearrow\mathscr{Z}^d} |\mathscr{V}|^{-1}\, \mathbf{1}_{\mathscr{V}}^{\mathrm{T}}[\mathbf{R}_1^{\mathscr{V}\mathscr{V}}]^{-1}\, \mathbf{R}_2^{\mathscr{V}\mathscr{V}}[\mathbf{R}_3^{\mathscr{V}\mathscr{V}}]^{-1}\, \mathbf{1}_{\mathscr{V}} = [f_1(0) f_3(0)]^{-1} f_2(0)\,.$$

Proof. We have

$$[\mathbf{R}_1^{\mathscr{V}\mathscr{V}}]^{-1}\, \mathbf{R}_2^{\mathscr{V}\mathscr{V}}[\mathbf{R}_3^{\mathscr{V}\mathscr{V}}]^{-1} - \mathbf{A}_1^{\mathscr{V}\mathscr{V}}\mathbf{R}_2^{\mathscr{V}\mathscr{V}}\mathbf{A}_3^{\mathscr{V}\mathscr{V}} =$$
$$= [\mathbf{R}_1^{\mathscr{V}\mathscr{V}}]^{-1}\, \mathbf{R}_2^{\mathscr{V}\mathscr{V}}([\mathbf{R}_3^{\mathscr{V}\mathscr{V}}]^{-1} - \mathbf{A}_3^{\mathscr{V}\mathscr{V}}) + ([\mathbf{R}_1^{\mathscr{V}\mathscr{V}}]^{-1} - \mathbf{A}_1^{\mathscr{V}\mathscr{V}})\, \mathbf{R}_2^{\mathscr{V}\mathscr{V}}\mathbf{A}_3^{\mathscr{V}\mathscr{V}}\,.$$

Now, due to Lemma 4.1 ii) and Lemma 4.2 ii) it follows

$$0 \geqq |\mathscr{V}|^{-1}\, \mathbf{1}_{\mathscr{V}}^{\mathrm{T}}[\mathbf{R}_1^{\mathscr{V}\mathscr{V}}]^{-1}\, \mathbf{R}_2^{\mathscr{V}\mathscr{V}}([\mathbf{R}_3^{\mathscr{V}\mathscr{V}}]^{-1} - \mathbf{A}_3^{\mathscr{V}\mathscr{V}})\, \mathbf{1}_{\mathscr{V}} \geqq$$
$$\geqq f_2^{\max}/f_1^{\min}|\mathscr{V}|^{-1} \operatorname{Tr}(\mathbf{1}_{\mathscr{V}\mathscr{V}}\mathbf{A}_3^{\mathscr{V}\mathscr{V}^c}(\mathbf{A}_3^{\mathscr{V}^c\mathscr{V}^c})^{-1}\, \mathbf{A}_3^{\mathscr{V}^c\mathscr{V}}) \to 0\,,$$

and similarly for the second term.

Thus, we finish the proof by

$$|\mathscr{V}|^{-1}\, \mathbf{1}_{\mathscr{V}}^{\mathrm{T}} \mathbf{A}_1^{\mathscr{V}\mathscr{V}} \mathbf{R}_2^{\mathscr{V}\mathscr{V}} \mathbf{A}_3^{\mathscr{V}\mathscr{V}}\, \mathbf{1}_{\mathscr{V}} \to \sum_{t\in\mathscr{Z}^d} a_1(t) \sum_{u\in\mathscr{Z}^d} r_2(u) \sum_{v\in\mathscr{Z}^d} a_3(v)\,. \qquad \square$$

Lemma 4.4. It holds

$$\lim_{\mathscr{V}\nearrow\mathscr{Z}^d} |\mathscr{V}|^{-1} \log \mathrm{Det}\,(\mathbf{R}^{\mathscr{V}\mathscr{V}}) = (2\pi)^{-d} \int_{\mathscr{I}_d} \log\,(f(\lambda))\,\mathrm{d}\lambda\,.$$

Proof. Due to Lemma 4.1 i) we may write

$$G_{\mathscr{V}} = |\mathscr{V}|^{-1} (\log \mathrm{Det}\,(\mathbf{A}^{\mathscr{V}\mathscr{V}}) - \log \mathrm{Det}\,([\mathbf{R}^{\mathscr{V}\mathscr{V}}]^{-1})) =$$
$$= |\mathscr{V}|^{-1}\, \mathrm{Tr}\,(\gamma^* \mathbf{A}^{\mathscr{V}\mathscr{V}} + (1-\gamma^*)[\mathbf{R}^{\mathscr{V}\mathscr{V}}]^{-1})^{-1} (\mathbf{A}^{\mathscr{V}\mathscr{V}} - [\mathbf{R}^{\mathscr{V}\mathscr{V}}]^{-1})$$

for some $\gamma^* \in [0, 1]$.

Since $\mathbf{A}^{\mathscr{V}\mathscr{V}} - [\mathbf{R}^{\mathscr{V}\mathscr{V}}]^{-1} = \mathbf{A}^{\mathscr{V}\mathscr{V}^c}[\mathbf{A}^{\mathscr{V}^c\mathscr{V}^c}]^{-1}\, \mathbf{A}^{\mathscr{V}^c\mathscr{V}}$ it holds

$$0 \leqq G_{\mathscr{V}} \leqq \bar{f}\,.\, |\mathscr{V}|^{-1}\, \mathrm{Tr}\,(\mathbf{A}^{\mathscr{V}\mathscr{V}^c}[\mathbf{A}^{\mathscr{V}^c\mathscr{V}^c}]^{-1}\, \mathbf{A}^{\mathscr{V}^c\mathscr{V}}) \to 0$$

according to Lemma 4.1 ii) and Lemma 4.2 ii). Therefore

$$\lim_{\mathscr{V}\nearrow\mathscr{Z}^d} |\mathscr{V}|^{-1} \log \mathrm{Det}\,(\mathbf{R}^{\mathscr{V}\mathscr{V}}) = -\lim_{\mathscr{V}\nearrow\mathscr{Z}^d} |\mathscr{V}|^{-1} \log \mathrm{Det}\,(\mathbf{A}^{\mathscr{V}\mathscr{V}}) =$$
$$= (2\pi)^{-d} \int_{\mathscr{I}_d} \log\,(f(\lambda))\,\mathrm{d}\lambda \qquad \square$$

by Theorem 2.5 in [3].

Now, we may prove the main results of the preceding section.

Proof of Theorem 3.1. We may write

$$\mathsf{E}_{\mathbf{Q}}[-\log p_{\boldsymbol{\theta}}^{\mathscr{V}}] = \tfrac{1}{2}[|\mathscr{V}| \log\,(2\pi) + \log \mathrm{Det}\,(\mathbf{R}_U^{\mathscr{V}\mathscr{V}}) +$$
$$+ \mathrm{Tr}\,(\mathbf{R}_{\mathbf{Q}}^{\mathscr{V}\mathscr{V}}[\mathbf{R}_U^{\mathscr{V}\mathscr{V}}]^{-1}) + (\mu_{h,U} - \nu_{\mathbf{Q}})^2\, \mathbf{1}_{\mathscr{V}}^{\mathrm{T}}[\mathbf{R}_U^{\mathscr{V}\mathscr{V}}]^{-1}\, \mathbf{1}_{\mathscr{V}}]\,,$$

where

$$|\mathscr{V}|^{-1} \log \mathrm{Det}\,(\mathbf{R}_U^{\mathscr{V}\mathscr{V}}) \to (2\pi)^{-d} \int_{\mathscr{I}_d} \log f_U(\lambda)\,\mathrm{d}\lambda$$

by Lemma 4.4;

$$|\mathscr{V}|^{-1}\, \mathrm{Tr}\,(\mathbf{R}_{\mathbf{Q}}^{\mathscr{V}\mathscr{V}}[\mathbf{R}_U^{\mathscr{V}\mathscr{V}}]^{-1}) \to 2 \sum_{k\in\mathscr{M}} U(k)\, R_{\mathbf{Q}}(k)$$

by Lemma 4.2. i) and ii);

and
$$|\mathscr{V}|^{-1}\, \mathbf{1}_{\mathscr{V}}^{\mathrm{T}}[\mathbf{R}_U^{\mathscr{V}\mathscr{V}}]^{-1}\, \mathbf{1}_{\mathscr{V}} \to [f_U(0)]^{-1} = 2\,.\sum_{k\in\mathscr{M}} U(k)$$

by Lemma 4.3.

Since for $\mathbf{Q} = \mathbf{P}_{\boldsymbol{\theta}}$ we have $2.\sum_{k\in\mathscr{M}} U(k)\, R_{\mathbf{Q}}(k) = 1$ and $\nu_{\mathbf{Q}} = \mu_{h,U}$ we obtain

$$\mathsf{S}(\mathbf{P}_{\boldsymbol{\theta}}) = \lim_{\mathscr{V}\nearrow\mathscr{Z}^d} |\mathscr{V}|^{-1}\, \mathsf{E}_{\mathbf{P}_{\boldsymbol{\theta}}}[-\log p_{\boldsymbol{\theta}}^{\mathscr{V}}] = \tfrac{1}{2}[1 + \log\,(2\pi) + (2\pi)^{-d} \int_{\mathscr{I}_d} \log f_U(\lambda)\,\mathrm{d}\lambda]\,.$$

For general $\mathbf{Q} \in \mathbb{M}$ we have

$$\mathsf{H}_I(\mathbf{Q} \mid \mathbf{P}_{\boldsymbol{\theta}}) = \lim_{\mathscr{V}\nearrow\mathscr{Z}^d} |\mathscr{V}|^{-1}\, \mathsf{E}_{\mathbf{Q}}[-\log p_{\boldsymbol{\theta}}^{\mathscr{V}}] - \mathsf{S}(\mathbf{Q}) =$$
$$= \tfrac{1}{2}\{\log\,(2\pi) + (2\pi)^{-d} \int_{\mathscr{I}_d} \log f_U(\lambda)\,\mathrm{d}\lambda + 2 \sum_{k\in\mathscr{M}} U(k)\,[R_{\mathbf{Q}}(k) + (\mu_{h,U} - \nu_{\mathbf{Q}})^2]\} -$$
$$- \mathsf{S}(\mathbf{Q})\,. \qquad \square$$

The remaining Theorem 3.3 we obtain as a corollary to the more general following proposition.

Proposition 4.5. For $j = 1, 2$ let $\mathbf{Q}_j$ be the stationary Gaussian random field with the spectral density $f_j \in \mathfrak{G}$ and a mean value μ_j.

We fix $\alpha, \beta \in \mathscr{R}$ satisfying

$$\beta > -\alpha(f_2/f_1)^{\min} \quad \text{if} \quad \alpha \geqq 0 ;$$
$$\beta > -\alpha(f_2/f_1)^{\max} \quad \text{if} \quad \alpha < 0 .$$

Then

$$\lim_{\mathscr{V} \nearrow \mathscr{Z}^d} |\mathscr{V}|^{-1} \log \mathsf{E}_{\mathbf{Q}_1}\{[q_1^{\mathscr{V}}]^{\alpha-1} [q_2^{\mathscr{V}}]^{\beta}\} =$$
$$= \tfrac{1}{2}\{(1 - \alpha - \beta) \log (2\pi) + (2\pi)^{-d} \int_{\mathscr{I}_d} [(1 - \alpha) \log f_1(\lambda) + (1 - \beta) \log f_2(\lambda) -$$
$$- \log (\alpha f_2(\lambda) + \beta f_1(\lambda))] \, d\lambda - \alpha\beta(\mu_1 - \mu_2)^2 [\alpha f_2(0) + \beta f_1(0)]^{-1}\} .$$

Proof. Let us denote $f_3 = \alpha f_2 + \beta f_1$. Then we have

$$|\mathscr{V}|^{-1} \log \mathsf{E}_{\mathbf{Q}_1}\{[q_1^{\mathscr{V}}]^{\alpha-1} [q_2^{\mathscr{V}}]^{\beta}\} = \tfrac{1}{2}\{(1 - \alpha - \beta) \log (2\pi)\} + |\mathscr{V}|^{-1} .$$
$$. \tfrac{1}{2}\{(1 - \alpha) \log \operatorname{Det} (\mathbf{R}_1^{\mathscr{V}\mathscr{V}}) + (1 - \beta) \log \operatorname{Det} (\mathbf{R}_2^{\mathscr{V}\mathscr{V}}) - \log \operatorname{Det} (\mathbf{R}_3^{\mathscr{V}\mathscr{V}})\} +$$
$$+ |\mathscr{V}|^{-1} \tfrac{1}{2}\{(\alpha\mu_1[\mathbf{R}_1^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{1}_{\mathscr{V}} + \beta\mu_2[\mathbf{R}_2^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{1}_{\mathscr{V}})^{\mathrm{T}} (\alpha[\mathbf{R}_1^{\mathscr{V}\mathscr{V}}]^{-1} + \beta[\mathbf{R}_2^{\mathscr{V}\mathscr{V}}]^{-1})^{-1} .$$
$$. (\alpha\mu_1[\mathbf{R}_1^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{1}_{\mathscr{V}} + \beta\mu_2[\mathbf{R}_2^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{1}_{\mathscr{V}}) - \alpha\mu_1^2 \mathbf{1}_{\mathscr{V}}^{\mathrm{T}}[\mathbf{R}_1^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{1}_{\mathscr{V}} - \beta\mu_2^2 \mathbf{1}_{\mathscr{V}}^{\mathrm{T}}[\mathbf{R}_2^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{1}_{\mathscr{V}}\} .$$

By Lemma 4.4 we obtain

$$\lim_{\mathscr{V} \nearrow \mathscr{Z}^d} |\mathscr{V}|^{-1} . \tfrac{1}{2}\{(1 - \alpha) \log \operatorname{Det} (\mathbf{R}_1^{\mathscr{V}\mathscr{V}}) + (1 - \beta) \log \operatorname{Det} (\mathbf{R}_2^{\mathscr{V}\mathscr{V}}) - \log \operatorname{Det} (\mathbf{R}_3^{\mathscr{V}\mathscr{V}})\} =$$
$$= \tfrac{1}{2}(2\pi)^{-d} \int_{\mathscr{I}_d} \{(1 - \alpha) \log f_1(\lambda) + (1 - \beta) \log f_2(\lambda) - \log f_3(\lambda)\} \, d\lambda .$$

The last term may be rewritten as

$$|\mathscr{V}|^{-1} \tfrac{1}{2}\{\alpha^2\mu_1^2 \mathbf{1}_{\mathscr{V}}^{\mathrm{T}}[\mathbf{R}_3^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{R}_2^{\mathscr{V}\mathscr{V}}[\mathbf{R}_1^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{1}_{\mathscr{V}} + 2\alpha\beta\mu_1\mu_2 \mathbf{1}_{\mathscr{V}}^{\mathrm{T}}[\mathbf{R}_3^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{1}_{\mathscr{V}} +$$
$$+ \beta^2\mu_2^2 \mathbf{1}_{\mathscr{V}}^{\mathrm{T}}[\mathbf{R}_3^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{R}_1^{\mathscr{V}\mathscr{V}}[\mathbf{R}_2^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{1}_{\mathscr{V}} - \alpha\mu_1^2 \mathbf{1}_{\mathscr{V}}^{\mathrm{T}}[\mathbf{R}_1^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{1}_{\mathscr{V}} -$$
$$- \beta\mu_2^2 \mathbf{1}_{\mathscr{V}}^{\mathrm{T}}[\mathbf{R}_2^{\mathscr{V}\mathscr{V}}]^{-1} \mathbf{1}_{\mathscr{V}}\} \to \tfrac{1}{2}\{\alpha^2\mu_1^2[f_1(0) f_3(0)]^{-1} f_2(0) +$$
$$+ 2\alpha\beta\mu_1\mu_2[f_3(0)]^{-1} + \beta^2\mu_2^2[f_2(0) f_3(0)]^{-1} f_1(0) - \alpha\mu_1^2[f_1(0)]^{-1} -$$
$$- \beta\mu_2^2[f_2(0)]^{-1}\} = -\tfrac{1}{2}\alpha\beta(\mu_1 - \mu_2)^2 [f_3(0)]^{-1}$$

by Lemma 4.3. □

Proof of Theorem 3.3. If we substitute $\beta = 1 - \alpha$, $\mathbf{Q}_1 = \mathbf{P}_{\boldsymbol{\theta}}$, $\mathbf{Q}_2 = \mathbf{P}_{\boldsymbol{\theta}^*}$, and realize

$$f_{\alpha \boldsymbol{U} + (1-\alpha)\boldsymbol{U}^*} = (\alpha[f_{\boldsymbol{U}}]^{-1} + \beta[f_{\boldsymbol{U}^*}]^{-1})^{-1} ,$$

we obtain directly the statement of Theorem 3.3. □

5. APPLICATION TO ESTIMATION

At first let us try to give a brief sketch of the basic idea of the so called "minimum distance method" used in statistical decision (for detailed explication see [6]).

Thus, suppose we are given (in some sense regular) parameter family of probability distributions

$$\mathbb{P} = \{\mathbf{P}_\theta\}_{\theta \in \Theta}$$

and a collection of observed data $\{\bar{x}_t\}_{t \in \mathscr{V}}$. On the basis of the given data we intend to estimate an unknown parameter $\theta^0 \in \Theta$.

We suppose the data to generate some "empirical distribution" $\hat{P}$ which need not be from $\mathbb{P}$. Therefore we seek for the distribution $\mathbf{P}_{\hat{\theta}} \in \mathbb{P}$ with minimal distance from $\hat{P}$:

$$\mathscr{D}(\hat{P}, \mathbf{P}_{\hat{\theta}}) = \min_{\theta \in \Theta} \mathscr{D}(\hat{P}, \mathbf{P}_\theta),$$

where $\mathscr{D}$ is some suitably chosen measure of distance. And $\hat{\theta}$ is considered to be the estimate of θ^0.

There are two obvious questions, namely what distance $\mathscr{D}$ to choose and what to understand under the "empirical distribution" in a considered situation.

If the parameter family is of an "exponential-like" type, and the Gauss-Markov random fields represent such a case, the I-divergence seems to be the convenient measure of distance.

Let $\hat{P} \in \mathbb{M}$. According to Corollary 3.2. ii) the minimization of $\mathsf{H}_I(\hat{P} \mid \mathbf{P}_\theta)$ is equivalent to solving the system of equations

$$\mu_{\hat{h},\hat{U}} = \nu_{\hat{P}}$$
$$\boldsymbol{R}_{\hat{U}}(k) = \boldsymbol{R}_{\hat{P}}(k) \quad \text{for} \quad k \in \mathscr{M}.$$

Therefore it is obvious that we may not construct any "empirical distribution" $\hat{P}$, but the above mentioned moments are all what we actually need to know.

If we set

$$\nu_{\hat{P}} = \hat{\mu}_{\mathscr{V}}(\bar{x}_{\mathscr{V}})$$

and

$$\boldsymbol{R}_{\hat{P}}(k) = \widehat{\boldsymbol{M}}_{\mathscr{V},k}(\bar{x}_{\mathscr{V}}) - (\hat{\mu}_{\mathscr{V}}(\bar{x}_{\mathscr{V}}))^2 \quad \text{for} \quad k \in \mathscr{M},$$

where

$$\hat{\mu}_{\mathscr{V}}(\bar{x}_{\mathscr{V}}) = |\mathscr{V}|^{-1} \sum_{t \in \mathscr{V}} \bar{x}_t,$$
$$\widehat{\boldsymbol{M}}_{\mathscr{V},k}(\bar{x}_{\mathscr{V}}) = |\mathscr{V}_k|^{-1} \sum_{t \in \mathscr{V}_k} \bar{x}_t \bar{x}_{t-k} \quad \text{for} \quad k \in \mathscr{M},$$

then the solution $\hat{\theta} = (\hat{h}, \hat{U})$ coincides with the estimate introduced in [2].

Thus, we know that there is at most one solution, and the solution exists with a probability tending to one for growing $\mathscr{V}$. The other asymptotic properties were also derived in [2].

Moreover, Corollary 3.2. ii) represents also the key for the implementation of the method since we obtain the estimate as the solution of the minimization problem which, under some reasonable assumptions on the dimension d and the "range" $\mathscr{M}$, is numerically solvable.

The only possible defect of the described method may consists in the absence of robustness. It is well known that the estimates of the "maximum likelihood type" are highly efficient but not robust enough. The proposed estimator is not exactly the maximum likelihood one, but asymptotically it coincides with such a one and therefore it may be considered as an approximate maximum likelihood estimator.

From the general theory (cf. e.g. [7]) it follows that, using the α-divergence instead of the I-divergence, we should obtain a more robust estimator. Nevertheless, there are several new problems connected with this approach, namely,

i) what α to choose;

ii) how to construct the "empirical spectral density" $\hat{f} \in \mathfrak{G}$ which is explicitly needed for expressing the α-divergence formula (cf. Proposition 4.5), and which should be a consistent estimate of the unknown spectral density;

iii) how many (if any) local minima there are, i.e. the question of existence and uniqueness of the solution;

iv) what are the (asymptotic) properties of the obtained estimate.

Some of the indicated problems might be (with some additional assumptions) solved in a satisfactory way but the necessary effort does not seem worthwhile to compare with the possible gain.

Let us realize that the main weak point of the maximum likelihood estimation, i.e. the nonuniqueness of the estimate, does not occur in the proposed method based on the minimum I-divergence. And if we feel some doubt about the contamination of the given data we may apply some known robust estimators of the first and the second moments instead of $\hat{\mu}_{\mathscr{V}}$ and $\{\hat{\boldsymbol{M}}_{\mathscr{V},k}\}_{k\in\mathscr{M}}$, respectively.

Thus, we may keep the I-divergence as the proper distance.

6. APPLICATION TO TESTING HYPOTHESES

In this section we intend to show the role of the divergences for a characterization of the asymptotic behaviour of the error probabilities in testing simple statistical hypotheses.

Suppose we are given a collection of observed data $\bar{x}_{\mathscr{V}} = \{\bar{x}_t\}_{t\in\mathscr{V}}$. Testing the hypothesis $\mathbf{H}_0: \mathbf{P} = \mathbf{P}_{\theta^0}$ against the alternative $\mathbf{H}_1: \mathbf{P} = \mathbf{P}_{\theta^1}$ $(\boldsymbol{\theta}^0, \boldsymbol{\theta}^1 \in \boldsymbol{\Theta})$, we reject the hypothesis $\mathbf{H}_0$ whenever

$$p^{\mathscr{V}}_{\theta^0}(\bar{x}_{\mathscr{V}}) \leqq c_{\mathscr{V}} \, . \, p^{\mathscr{V}}_{\theta^1}(\bar{x}_{\mathscr{V}})$$

with some constant $c_{\mathscr{V}} > 0$ called the critical value.

Thus, the test is given by the critical region

$$\mathscr{C}_{\mathscr{V}}(c_{\mathscr{V}}) = \{x_{\mathscr{V}} \in \mathscr{R}^{\mathscr{V}};\ p^{\mathscr{V}}_{\theta^0}(x_{\mathscr{V}}) \leqq c_{\mathscr{V}}\, p^{\mathscr{V}}_{\theta^1}(x_{\mathscr{V}})\} \, .$$

The probabilities of the errors of the first kind and of the second kind are given by

$$\mathbf{e}_1^{\mathscr{V}} = \mathbf{P}_{\mathbf{0}^0}^{\mathscr{V}}(\mathscr{C}_{\mathscr{V}}(c_{\mathscr{V}})) \quad \text{and} \quad \mathbf{e}_2^{\mathscr{V}} = \mathbf{P}_{\mathbf{0}^1}^{\mathscr{V}}((\mathscr{C}_{\mathscr{V}}(c_{\mathscr{V}}))^c),$$

respectively.

We are interested in the asymptotic behaviour of the error probabilities for growing $\mathscr{V}$. We shall treat two basic possibilities of choice of the critical values $\{c_{\mathscr{V}}\}_{\mathscr{V} \in \mathfrak{K}}$.

At first let the critical value be fixed, i.e. $c_{\mathscr{V}} = c > 0$ for every $\mathscr{V} \in \mathfrak{K}$. Then the test corresponds to the optimal Bayes test which minimizes the mixed errors

$$q_1 \mathbf{e}_1^{\mathscr{V}} + q_2 \mathbf{e}_2^{\mathscr{V}},$$

where the prior probabilities are given by $q_1 = (1 + c)^{-1}$ and $q_2 = c(1 + c)^{-1}$, respectively.

In the second case let the critical value $c_{\mathscr{V}}$ be given by

$$\mathbf{e}_1^{\mathscr{V}} = \mathbf{P}_{\mathbf{0}^0}^{\mathscr{V}}(\mathscr{C}_{\mathscr{V}}(c_{\mathscr{V}})) = l$$

for some fixed $l \in (0, 1)$, i.e. the test is optimal on the level l.

Theorem 6.1. i) Let $c_{\mathscr{V}} = c > 0$ for every $\mathscr{V} \in \mathfrak{K}$. Then for $j = 1, 2$

$$\lim_{\mathscr{V} \nearrow \mathscr{Z}^d} \{-|\mathscr{V}|^{-1} \log \mathbf{e}_j^{\mathscr{V}}\} = \max_{\alpha \in [0,1]} \{\alpha(1 - \alpha)\, \mathsf{H}_\alpha(\mathbf{P}_{\mathbf{0}^0} \mid \mathbf{P}_{\mathbf{0}^1})\}.$$

ii) Let $c_{\mathscr{V}}$ be for every $\mathscr{V} \in \mathfrak{K}$ given by

$$\mathbf{e}_1^{\mathscr{V}} = \mathbf{P}_{\mathbf{0}^0}^{\mathscr{V}}((\mathscr{C}_{\mathscr{V}}(c_{\mathscr{V}})) = l \in (0, 1).$$

Then

$$\lim_{\mathscr{V} \nearrow \mathscr{Z}^d} \{-|\mathscr{V}|^{-1} \log \mathbf{e}_2^{\mathscr{V}}\} = \mathsf{H}_l(\mathbf{P}_{\mathbf{0}^0} \mid \mathbf{P}_{\mathbf{0}^1}).$$

Proof. With the aid of Theorem 3.3, Theorem 3.1 and Corollary 3.4 the statements follow from Theorem 12.19 and 12.20 in [8]. □

7. CONCLUDING REMARK

Some other applications of the divergences are introduced in [8] where the above mentioned problems are treated in general. The reader can find many interesting results and consequences there as well as a lot of useful references relevant for the topic.

(Received April 5, 1988.)

REFERENCES

[1] R. L. Dobrushin: Gaussian random fields — Gibbsian point of view. In: Multicomponent random systems (R. L. Dobrushin, Ya. G. Sinai, eds.). M. Dekker, New York 1980, 119—151.

[2] M. Janžura: Asymptotic theory of parameter estimation for Gauss-Markov random fields. Kybernetika *24* (1988) 3, 161—176.

[3] H. Künsch: Thermodynamics and statistical analysis of Gaussian random fields. Z. Wahrsch. verw. Gebiete *58* (1981), 407—421.

[4] C. Preston: Random Fields. (Lecture Notes in Mathematics 534.) Springer-Verlag, Berlin—Heidelberg—New York 1976.

[5] E. M. Stein and G. Weiss: Introduction to Fourier Analysis on Euclidean Spaces. Princeton University Press, Princeton, N. J. 1971.

[6] I. Vajda: A new general approach to minimum distance estimation. In: Trans. Ninth Prague Conference on Inform. Theory, Statist. Dec. Functions, Random Processes. Academia, Prague 1983, 103—112.

[7] I. Vajda: Efficiency and robustness control via distorted maximum likelihood estimation. Kybernetika *22* (1986), 47—67.

[8] I. Vajda: Theory of Statistical Inference and Information. Reidel, Amsterdam 1988.

RNDr. Martin Janžura, 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.