

# STATISTICAL CLASSIFICATION BASED ON OBSERVATIONS OF RANDOM GAUSSIAN FIELDS

J. ŠALTYTĖ, K. DUČINSKAS

*Klaipėda University, Department of System Research*

H. Manto 84, Klaipėda Lt-5808

E-mail: jsaltyte@gmf.ku.lt, duce@gmf.ku.lt

Received October 10, 1999

## ABSTRACT

The problem of classification of objects located in domain  $D \subset R^2$  based on observations of random Gaussian fields with a factorized covariance function is considered. The first-order asymptotic expansion for the expected error regret is presented. Obtained numerical results allow us to compare suggested expansion for some widely applicable models of spatial covariance function.

## 1. INTRODUCTION

The notion that data close together in space are likely to be correlated is natural. And one of the most important (sometimes even unique) statistical characteristic of random field which describes the statistical spatial relationship between observations is a *spatial covariance function*  $\sigma(r, s) = E\{(X(r) - E(X(r)))(X(s) - E(X(s)))\}$ , where  $\{X(t), t \in D\}$  is an observed random field. We restrict our attention on covariance functions, which depend only on the distance  $h = r - s$  between points, i.e. we consider only *second-order stationary random fields*. When  $\sigma(r, s) = \sigma(h)$  is a function on both the magnitude and direction of  $h$ , the covariance function is said to be *anisotropic*; otherwise, it is said to be *isotropic* one.

In geostatistics literature, for the analysis of spatially correlated data the concept of a *variogram* is used, see, e.g., Matheron (1962), Cressie (1994) and others. This function is similar to the covariance function. By the definition, the variogram is  $var(X(r) - X(s)) = 2\gamma(r - s)$ ,  $r, s \in D$ . (The quantity

$2\gamma(\cdot)$  has been called a variogram (as  $\gamma(\cdot)$  - semivariogram) by Matheron (1962.) There is a simple relationship between the semivariogram and covariance function:

$$\gamma(h) = \sigma(0) - \sigma(h).$$

So, we interchangeably can use one of these concepts. In general, using variograms is better than using covariances because the estimator of variogram obtained by the method-of-moments (Matheron, 1962) is unbiased.

Obvious,  $\gamma(h) = \gamma(-h)$  and  $\gamma(0) = 0$ . If  $\gamma(h) \rightarrow \theta_0 > 0$ , as  $h \rightarrow 0$ , then  $\theta_0$  is measurement error, which has been called the *nugget effect*. If the semivariogram has the property  $\lim_{|h| \rightarrow \infty} \gamma(h) = \gamma_\infty < \infty$ , then  $\gamma_\infty$  has been called the *sill* of the semivariogram. The *range* of semivariogram is the distance after which semivariogram becomes constant, see, e.g., Cressie (1994).

Christensen (1989), Cressie (1994) present several covariance models, which are most often used in geostatistics. We consider three of them.

The *isotropic spherical covariance function* is given by expression

$$\sigma_s(|h|, \theta) = \begin{cases} \theta_1 \left(1 - \frac{3}{2} \frac{|h|}{\theta_2} + \frac{1}{2} \frac{|h|^3}{\theta_2^3}\right), & 0 \leq |h| \leq \theta_2, \\ \theta_0 + \theta_1, & |h| = 0, \\ 0, & |h| > \theta_2, \end{cases} \quad (1.1)$$

for nonnegative  $\theta_0, \theta_1, \theta_2$ . The nugget effect is  $\theta_0$  and the sill is  $\theta_0 + \theta_1$ . For this model, observations more than  $\theta_2$  units apart are uncorrelated, so the range is  $\theta_2$ .

The *exponential covariance function* is

$$\sigma_e(h, \theta) = \begin{cases} \theta_1 \exp\left(-\theta_2 \sqrt{t^2 h_1^2 + h_2^2}\right), & |h| > 0, \\ \theta_0 + \theta_1, & |h| = 0, \end{cases} \quad (1.2)$$

for nonnegative  $\theta_0, \theta_1, \theta_2$ . Here  $t$  is the parameter of anisotropy. When  $t = 1$ , the exponential covariance function becomes isotropic one; otherwise it is anisotropic. The nugget effect is  $\theta_0$ , the sill is  $\theta_0 + \theta_1$ , and the range is infinite. While the range is infinite, correlations decrease very rapidly as  $h$  increases. Of course, this phenomenon depends on the value of  $\theta_2$ .

The *Ornstein - Uhlenbeck covariance function* is defined as follows

$$\sigma_{ou}(h, \theta) = \begin{cases} \theta_1 \exp\left(-\theta_2 (t^2 h_1^2 + h_2^2)\right), & |h| > 0, \\ \theta_0 + \theta_1, & |h| = 0, \end{cases} \quad (1.3)$$

for  $\theta_0, \theta_1, \theta_2$  nonnegative. It is anisotropic covariance function, when  $t \neq 1$ . In the case of  $t = 1$  it becomes a well known isotropic covariance function often called the *Gaussian covariance function*. The behavior of the Ornstein - Uhlenbeck model is similar to that of the exponential model. However, the

covariances at distances greater than one approach zero much more rapidly than in the exponential model. Also, for small distances, the covariance approaches the value  $\theta_1$  much more rapidly than does the exponential.

In our paper we use the correlation functions, which can be easily defined from covariance function by the relation  $\rho(h) = \sigma(h) / \sigma(0)$ .

## 2. CLASSIFICATION PROBLEM

Suppose  $\Omega_1, \Omega_2$  are two mutually exclusive and exhaustive classes of objects. Let  $X$  be a  $p$ -dimensional feature vector, which is measured on each object. For objects randomly chosen from  $\Omega_l$ ,  $X$  follows the multivariate distribution with density function  $p_l(x; \theta_l) = p_l(x)$ , which belongs to the parametric family of regular densities  $F_l = \{p_l(x; \theta_l), \theta_l \in \Theta_l \subset R^m\}$ ,  $l = 1, 2$ .

Discriminant analysis deals with the problem of identifying the class of object for which  $X$  is measured. For a zero-one loss function, the Bayes classification rule (BCR)  $d_B(x)$  minimizing the probability of misclassification is equivalent to assigning  $X = x$  to  $\Omega_l$  if

$$\pi_l p_l(x) = \max_{k=1,2} \pi_k p_k(x),$$

where  $\pi_l$  is the prior probability of  $\Omega_l$ . Then BCR  $d_B(x)$  could be defined as

$$d_B(x) = \arg \max_{k=1,2} \pi_k p_k(x).$$

Let  $P_B$  denote the probability of misclassification for BCR  $d_B(x)$  or *Bayes error rate* (see, e.g., [1]).

In practical applications, the density functions  $\{p_l(x)\}$  are seldom completely known. Often they are only known up to the parameters  $\{\theta_l\}$ , i.e. we may only assert that  $p_l(x)$  is one element of a parametric family of density functions  $F_l$ . Under such conditions, it is customary to estimate  $\theta_l$  from the training sample  $T_l = \{X_{l1}, \dots, X_{lN_l}\}$  from  $\Omega_l$ , for  $l = 1, 2$ . Put  $T = T_1 \cup T_2$ ,  $N = N_1 + N_2$ .

Let  $\hat{\theta}_l$  be the maximum likelihood estimator (MLE) of  $\theta_l$  from  $T_l$  ( $l = 1, 2$ ). The estimator of rule  $d_B(x)$  is called a *plug-in rule*  $d_B(x, \hat{\theta}_1, \hat{\theta}_2)$  and is defined by

$$d_B(x, \hat{\theta}_1, \hat{\theta}_2) = \arg \max_{k=1,2} \pi_k p_k(x, \hat{\theta}_k).$$

The *actual error rate*  $P_A$  of  $d_B(x, \hat{\theta}_1, \hat{\theta}_2)$  is the probability of misclassifying a randomly selected object with feature  $X$  independent on  $T$  and is designated by

$$P_A = \sum_{l=1}^2 \pi_l \int \left(1 - \delta\left(l, d_B(x, \hat{\theta}_1, \hat{\theta}_2)\right)\right) p_l(x) dx,$$

where  $\delta(\cdot, \cdot)$  is Kronecker's delta.

DEFINITION 2.1. Expected error regret (EER) for  $d_B(\cdot, \widehat{\theta}_1, \widehat{\theta}_2)$  is the expectation of the difference between  $P_A$  and  $P_B$  with respect to the distribution of  $\widehat{\theta}_1, \widehat{\theta}_2$ , i.e.,

$$EER = E(P_A) - P_B.$$

The *purpose* of this article is to find an asymptotic expansion for EER. The case of independent normally distributed observations in training sample from one of two classes with  $\Sigma_l = \Sigma$ ,  $l = 1, 2$ , was considered in [2]. [3] has been made the generalization for the case of arbitrary number of classes ( $l \geq 2$ ) and regular class-conditional densities.

### 3. MAIN RESULTS

Suppose that any point  $r = (r_1, r_2) \in D \subset R^2$  can be assigned to one of two prescribed above classes  $\Omega_1, \Omega_2$  with positive prior probabilities  $\pi_1, \pi_2$ , respectively. Here we identify objects by points on  $D$ . The class of the point  $r$  is given by the random 2-dimensional vector  $Y_r^T = (Y_{1r}, Y_{2r})$  of zero-one variables. The  $l$ th component of  $Y$  is defined to be one or zero according as a class of point  $r$  is or not  $\Omega_l$  ( $l = 1, 2$ ). Then  $Y_r \sim Mult_2(1; (\pi_1, \pi_2))$ .

Suppose that  $X_r$  means the observation of  $X$  at point  $r \in D$ . A decision is to be made as to which class the randomly chosen point  $r \in D$  is assigned on the basis of observed value of  $X_r$ . Let

$$X_r = \sum_{l=1}^2 Y_{lr} \mu_l + \epsilon_r, \quad (3.1)$$

where  $\mu_1, \mu_2 \in R^p$ ,  $\mu_1 \neq \mu_2$  and the noise  $\epsilon_r = (\epsilon_r^1, \dots, \epsilon_r^p)$  is the observation of the second-order stationary multivariate random field at location  $r \in D$  with zero-mean vector.

The essential assumption is that  $\{\epsilon_r\}$  is Gaussian field with spatially factorized covariance. Hence, the common *class-conditional covariance* between any two observations  $X_r$  and  $X_s$  at points  $r, s \in D$  belonging to  $\Omega_l$  can be factorized as  $\text{cov}(X_r, X_s/r, s \in \Omega_l) = \rho^l(h)\Sigma$ , ( $r \neq s$ ), where  $\rho^l(\cdot)$  is the spatial correlation function ( $l = 1, 2$ ), and  $h = r - s$ ,  $\Sigma = \text{cov}(\epsilon_r, \epsilon_r)$ .

Also here we assume that the effect of *cross-correlation* between samples from different classes is *negligible*. In this paper we suppose, that it is equal to zero, i. e.,  $\text{cov}(X_r, X_s/r \in \Omega_1, s \in \Omega_2) = 0$ .

Let  $D_l = \{s_1^l, \dots, s_{N_l}^l\} \subset D$  be the set of points belonging to class  $\Omega_l$ ,  $l = 1, 2$ . Then  $X_{lj}$  means the observation of  $X$  at point  $s_j^l$ , i.e.  $X_{lj} = X(s_j^l)$ ,  $j = 1, \dots, N_l$ ,  $l = 1, 2$ .

Then the expectation for  $N_l p \times 1$  stacked vector  $T_l^V = (X_{l1}^l, \dots, X_{lN_l}^l)'$  is

$$\mu_l^+ = \mathbf{1}_{N_l} \otimes \mu_l, \quad (l = 1, 2), \tag{3.2}$$

where  $\mathbf{1}_{N_l}$  is the  $N_l$ -dimensional vector of ones, and  $\otimes$  is the Kronecker product. The covariance matrix of  $T_l^V$  is

$$\Sigma_l^+ = C_l \otimes \Sigma, \tag{3.3}$$

where  $C_l$  is the spatial correlation matrix of order  $N_l \times N_l$ , whose  $(i, j)$ th element is  $\rho(s_i^l - s_j^l)$  ( $i, j = 1, \dots, N_l$ ).

Suppose that  $\Sigma$  and  $C_l$  are known and  $\mu_l$  are unknown ( $l = 1, 2$ ). In this paper maximum likelihood estimators (MLE)  $\hat{\mu}_l$  of  $\mu_l$  based on  $T_l$  are used. Let  $C_l^{-1} = (c_l^{ij})$ .

**Lemma 3.1.** For  $l = 1, 2$  MLE of  $\mu_l$  based on  $T_l$  is

$$\hat{\mu}_l = \frac{1}{c_l^{\cdot\cdot}} \sum_{j=1}^{N_l} c_l^{ij} x_{lj},$$

where  $c_l^{ij} = \sum_{i=1}^{N_l} c_l^{ij}$  and  $c_l^{\cdot\cdot} = \sum_{i,j=1}^{N_l} c_l^{ij}$ .

*Proof.* The log-likelihood of  $T_l$  is

$$\begin{aligned} \ln L_l &= -const - \frac{1}{2} (N_l \ln |\Sigma| + p \ln |C|) - \\ &\quad - \frac{1}{2} \left( c_l^{\cdot\cdot} tr (\Sigma^{-1} S_l) + c_l^{\cdot\cdot} tr \left( \Sigma^{-1} (\mu_l - \bar{x}_l) (\mu_l - \bar{x}_l)' \right) \right), \end{aligned}$$

where  $\bar{x}_l = \frac{1}{c_l^{\cdot\cdot}} \sum_{j=1}^{N_l} c_l^{ij} x_{lj}$  and  $S_l = \frac{1}{c_l^{\cdot\cdot}} \sum_{i,j=1}^{N_l} c_l^{ij} (x_{lj} - \bar{x}_l) (x_{li} - \bar{x}_l)'$ .

Solving equation  $\frac{\partial \ln L_l}{\partial \mu_l} = 0$ , we complete the proof of Lemma. ■

MLE under spatial sampling of Gaussian random fields was studied by [4]. They gave the regularity conditions which ensure consistency and asymptotic normality of the parameter estimators. We assume that these conditions hold.

Put  $\gamma = \ln \frac{\pi_1}{\pi_2}$ ,  $\Delta \hat{\mu}_l = \hat{\mu}_l - \mu_l$  ( $l = 1, 2$ ) and let  $\Delta^2 = (\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 - \mu_2)$  be the Mahalanobis distance. Let  $\Phi(\cdot)$  and  $\varphi(\cdot)$  denote standard normal distribution and density functions, respectively.

The plug-in discriminant function can be written in the form

$$d_B(x; \hat{\mu}_1, \hat{\mu}_2) = \left( x - \frac{1}{2} (\hat{\mu}_1 + \hat{\mu}_2) \right)' (\hat{\mu}_1 - \hat{\mu}_2). \tag{3.4}$$

Then the actual error rate for  $d_B(x, \hat{\mu}_1, \hat{\mu}_2)$  (see [5]) is

$$P_A = \pi_1 \Phi \left( -\frac{(\mu_1 - \frac{1}{2}(\hat{\mu}_1 + \hat{\mu}_2))' (\hat{\mu}_1 - \hat{\mu}_2) + \gamma}{\sqrt{(\hat{\mu}_1 - \hat{\mu}_2)' \Sigma (\hat{\mu}_1 - \hat{\mu}_2)}} \right) \quad (3.5)$$

$$+ \pi_2 \Phi \left( \frac{(\mu_2 - \frac{1}{2}(\hat{\mu}_1 + \hat{\mu}_2))' (\hat{\mu}_1 - \hat{\mu}_2) + \gamma}{\sqrt{(\hat{\mu}_1 - \hat{\mu}_2)' \Sigma (\hat{\mu}_1 - \hat{\mu}_2)}} \right). \quad (3.6)$$

**Theorem 3.1.** *First-order asymptotic expansion of EER in terms of  $(c_i)^{-1}$  for  $d_B(x, \hat{\mu}_1, \hat{\mu}_2)$ , using MLE  $\hat{\mu}_1, \hat{\mu}_2$ , is*

$$\begin{aligned} EER &= \sum_{l=1}^2 \frac{1}{4c_l^i} \pi_l \varphi \left( -\frac{\gamma}{\Delta} + (-1)^l \frac{\Delta}{2} \right) \quad (3.7) \\ &\quad \left( \left( \left( -\frac{\gamma}{\Delta} + (-1)^l \frac{\Delta}{2} \right)^2 + (p-1) \right) / \Delta \right) + o \left( \frac{1}{\min(c_1^i, c_2^i)} \right) \end{aligned}$$

*Proof.* Since  $P_A$  is invariant under linear transformations of data we use the convenient canonical form of  $\Sigma = I$  and  $\mu_1 = -\mu_2 = \frac{1}{2}\Delta$ , where  $\Delta = (\Delta, 0, \dots, 0)'$  (see [6]). Expand  $P_A$  in Taylor series about the point  $\hat{\mu}_l = \mu_l$  and then averaging with respect to the distribution of  $\hat{\mu}_l$  ( $l = 1, 2$ ). Expansion for  $E(P_A)$  dropping the third order terms is as follows

$$E(P_A) \cong P_B + \sum_{l=1}^2 P_l^{(1)} E(\Delta \hat{\mu}_l) + \frac{1}{2} \sum_{l,k=1}^2 \text{tr} \left( P_{l,k}^{(2)} E(\Delta \hat{\mu}_l \Delta \hat{\mu}_k) \right), \quad (3.8)$$

where  $P_l^{(1)}$  is the vector of the first-order derivatives of  $P_A$  by  $\hat{\mu}_l$  evaluated at  $\mu_l$  ( $l = 1, 2$ ). Similarly,  $P_{l,k}^{(2)}$  denotes the matrix of the second-order derivatives of  $P_A$  by  $\hat{\mu}_l$  and  $\hat{\mu}_k$  evaluated at  $\mu_l$  and  $\mu_k$ , respectively, ( $l, k = 1, 2$ ). In considered situation there was obtained (see [8]) that

$$P_B = \pi_1 \Phi \left( -\frac{\Delta}{2} - \frac{\gamma}{\Delta} \right) + \pi_2 \Phi \left( -\frac{\Delta}{2} + \frac{\gamma}{\Delta} \right).$$

From Lemma and assumptions stated before we have

$$E(\Delta \hat{\mu}_l) = E(\Delta \hat{\mu}_l \Delta \hat{\mu}_k) = 0, \quad (3.9)$$

$$E\left((\Delta \hat{\mu}_l)^2\right) = \frac{1}{c_l^i}. \quad (3.10)$$

Then using (3.9) and (3.10) in (3.8) we complete the proof of the stated theorem. ■

*Corollary 3.1.* Whether  $T_l$  consists of statistically independent  $X_{lj}$ ,  $j = 1, \dots, N_l$ , then  $c_l = N_l$  in formula (3.7).

The corollary holds since  $C_l^{-1} = I$  for statistically independent  $X_{lj}$ ,  $j = 1, \dots, N_l$ .

The result of the proved theorem could be used in obtaining the optimal sampling design that ensures the minimum of asymptotic EER for the fixed training sample size  $N$ .

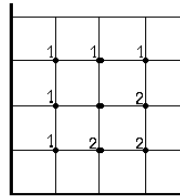
**4. EXAMPLE**

As an example we consider the integer regular 2-dimensional lattice and use the second-order neighborhood scheme for training sample.

Also we assume that there are two differently taken training samples: 1) 4 spatially symmetric observations in training sample for each class; 2) 5 observations in training sample for the first class and 3 for the second one.



**Figure 1.** Scheme 1.



**Figure 2.** Scheme 2.

Three spatial correlation functions obtained from the covariance functions defined by (1.1), (1.2) and (1.3) are considered.

The asymptotic expected error regret

$$AEER \triangleq \sum_{l=1}^2 \frac{1}{4c_l} \pi_l \varphi \left( -\frac{\gamma}{\Delta} + (-1)^l \frac{\Delta}{2} \right) \left( \left( \left( -\frac{\gamma}{\Delta} + (-1)^l \frac{\Delta}{2} \right)^2 + (p-1) \right) / \Delta \right)$$

for each correlation function is computed.

In Table 1 values of  $AEER$  with  $\pi_1 = \pi_2 = 0.5$  are presented. Here  $AEER_{ind}$  is  $AEER$  in the case of independent observations (considered for comparison),  $AEER_s$ ,  $AEER_e$ ,  $AEER_{ou}$  are  $AEER$  for spherical ( $\rho_s$ ), exponential ( $\rho_e$ ) and Ornstein-Uhlenbeck ( $\rho_{ou}$ ) correlation functions, respectively. As it was already mentioned, the spherical correlation function is isotropic one. We chose the range value  $\theta_2 = 3$  for this function. In general,  $\rho_e$  and  $\rho_{ou}$  are anisotropic functions, but by choosing the value of  $t = 1$  we obtain isotropic

**Table 1.**  
Values of *AEEER* for different correlation functions.

$\Delta$	$P_B$	$AEEER_{ind}$	$AEEER_s$	$AEEER_e^{is}$	$AEEER_e^{anis}$	$AEEER_{ou}^{is}$	$AEEER_{ou}^{anis}$
0,50	0,40129		0,28325	0,37089	0,37531	0,36293	0,37101
		0,09969	0,13625	0,17527	0,17804	0,17191	0,17635
		0,10633	0,26765	0,37034	0,37512	0,36274	0,37144
			0,14194	0,17952	0,18123	0,17663	0,17945
1,00	0,30854		0,14067	0,18421	0,18639	0,18025	0,18426
		0,04951	0,06767	0,08705	0,08842	0,08538	0,08759
		0,05281	0,13293	0,18393	0,18629	0,18015	0,18447
			0,07049	0,08915	0,09000	0,08772	0,08912
1,50	0,22663		0,09136	0,11963	0,12105	0,11706	0,11966
		0,03915	0,04395	0,05653	0,05743	0,05545	0,05688
		0,03429	0,08633	0,11945	0,12099	0,11699	0,11981
			0,04578	0,05789	0,05845	0,05697	0,05787
2,00	0,15866		0,06446	0,08441	0,08541	0,08259	0,08443
		0,02268	0,03101	0,03989	0,04052	0,03912	0,04013
		0,02419	0,06091	0,08427	0,08536	0,08254	0,08452
			0,03230	0,04085	0,04124	0,04019	0,04083
2,50	0,10565		0,04622	0,06052	0,06124	0,05922	0,06054
		0,01627	0,02223	0,02861	0,02905	0,02805	0,02878
		0,01735	0,04368	0,06043	0,06121	0,05919	0,06061
			0,02316	0,02929	0,02957	0,02882	0,02928
3,00	0,06681		0,03258	0,04267	0,04317	0,04175	0,04268
		0,01147	0,01567	0,02016	0,02048	0,01978	0,02029
		0,01223	0,03078	0,04260	0,04315	0,04173	0,04273
			0,01633	0,02065	0,02085	0,02032	0,02064

functions. In Table 1 there are presented both isotropic and anisotropic cases (denoted by index *is* and *anis*, respectively). Each cell contains 4 rows. The first row presents values of *AEEER* when there are no nugget effect ( $\theta_0 = 0$ ), i.e.  $\frac{\theta_1}{\theta_0 + \theta_1} = 1$  and the second one gives values of *AEEER*, when nugget effect  $\theta_0 = \frac{3}{4}$  is assumed (then  $\frac{\theta_1}{\theta_0 + \theta_1} = \frac{1}{4}$ ). To calculate quantities presented in the first and second rows the scheme of spatially symmetric observations (1.) was used. The third and fourth rows contain *AEEER* for  $\theta_0 = 0$  and  $\theta_0 = \frac{3}{4}$ , respectively, but in training sample 5 observations for the first class and 3 for the second one was taken.

For all described cases, the *AEEER* approaches zero when distances  $\Delta$  increases. As it was expected, *AEEER* for the case of independent observations is the smallest one.

The comparison of *AEEER* in the case of independent observations and in the case of dependent observations (three considered schemes of correlation functions) is presented in Table 2. In the first row of this table ratios for  $\theta_0 = 0$  (no nugget effect) and different training sample schemes (upper quantity in the cell is for the Scheme 1 and lower one for the Scheme 2) are presented, as in the second row the same situation is presented only the nugget effect



$\theta_0 = \frac{3}{4}$  is used.

It can be seen from Table 2, that the bigger the nugget effect, the closer ratio to one (see the second row of Table 2), because with increasing the nugget effect the situation approaches independent case.

Comparing columns (ratios for the same nuggets), we can determine which of the correlation functions gives smaller  $AEEER$ . For instance,  $\frac{AEEER_{IND}}{AEEER_S} = 0.3519$  and  $\frac{AEEER_{IND}}{AEEER_E^S} = 0.2688$  (for  $\theta_0 = 0$ ); the ratio of these two ratios is equal 1.31, so, Spherical correlation function is better (gives smaller  $AEEER$ ) than Exponential isotropic correlation function for the Scheme 1. In a similar way other functions can be compared. It is easy to see, that Spherical function gives smallest  $AEEER$  in all considered cases. Also it can be shown, that isotropic correlation functions give smaller  $AEEER$  than anisotropic do.

**Table 2.**  
Ratios of  $AEEER_{ind}$  and  $AEEER$  in the case of dependent observations.

Ratio	$\frac{AEEER_{IND}}{AEEER_S}$	$\frac{AEEER_{IND}}{AEEER_E^S}$	$\frac{AEEER_{IND}}{AEEER_E^{NIS}}$	$\frac{AEEER_{IND}}{AEEER_{OU}^S}$	$\frac{AEEER_{IND}}{AEEER_{OU}^{NIS}}$
$\theta_0 = 0$	0,3519 0,3973	0,2688 0,2871	0,2656 0,2835	0,2747 0,2931	0,2687 0,2863
$\theta_0 = \frac{3}{4}$	0,7317 0,7491	0,5688 0,5923	0,5599 0,5867	0,5799 0,6020	0,5653 0,5926

## REFERENCES

- [1] D.J. Hand. *Construction and Assessment of Classification Rules*. John Wiley & Sons, New York, 1997.
- [2] M. Okamoto. An asymptotic expansion for the distribution of the linear discriminant function. *Ann. Math. Statist.*, **34**, 1963, 1286 – 1301.
- [3] K. Dučinskas. An Asymptotic Analysis of the Regret Risk in Discriminant Analysis Under Various Training Schemes. *Lith. Math. J.*, **37** (4), 1997, 337 – 351.
- [4] K.V. Mardia and R.J. Marshall. Maximum Likelihood Estimation of Models for Residual Covariance and Spatial Regression. *Biometrika*, **71**, 1984, 135 – 146.
- [5] G.J. McLellan. The Asymptotic Distributions of the Conditional Error Rate and Risk in Discriminant Analysis. *Biometrika*, **61** ( 1), 1974, 131 – 135.
- [6] L.G. Malinovskyi. *Classification of Objects by Methods of Discriminant Analysis*. 1979, 162 – 163. (in Russian)
- [7] O.J. Dunn. Some Expected Values for Probabilities of Correct Classification in Discriminant Analysis. *Technometrics*, **13**, 1971, 345 – 353.
- [8] S. Ganesalingam and G.J. McLellan. The Efficiency of a Linear Discriminant Function Based on Unclassified Initial Samples. *Biometrika*, **65** (3), 1978, 658 – 662.
- [9] Z. Ying. Maximum Likelihood Estimation of Parameters under a Spatial Sampling Scheme. *The Annals of Statistics*, **21** (3), 1993, 1567 – 1590.

**STATISTINIS KLASIFIKAVIMAS REMIANTIS  
ATSITIKTINIŲ GAUSO LAUKŲ STEBĖJIM AIS**

J. ŠALTYTĖ, K. DUČINSKAS

Nagrinėjamas uždavinys apie objektų iš srities  $D \subset R^2$  klasifikavimą, remiantis atsitiktinių Gauso laukų stebėjimais. Pateikti asimptotiniai laukiamos paklaidos įverčiai. Atlikus skaitinio modeliavimo eksperimentą naujasis skleidinys lyginamas su kitais žinomais skleidiniais.