13. DISCRIMINATION OF TWO POPULATIONS WITH COMMON UNKNOWN COVARIANCE MATRIX. G_{13}-ANDERSON-FISHER STATISTICS ESTIMATOR

Let $\vec{x}_i,\ \vec{y}_j;\ i=1,...,n_1;\ j=1,...,n_2$ be independent observations of m-dimensional random vectors $\vec{\xi}=\vec{a}_1+\sqrt{R}\vec{\mu},\ \vec{\mu}^T=\{\mu_i,\ i=1,\dots,m\}$ and $\vec{\eta}=\vec{a}_2+\sqrt{R}\vec{\nu},\ \vec{\nu}^T=\{\nu_i,\ i=1,\dots,m\}$ respectively and suppose that random variables $\mu_i,\ \nu_i,\ i=1,\dots,m$ are independent for every n.

For the observation classification in the case of two Normal populations, the so-called discriminant function is used:

$$U(\vec{x})=\left\{\vec{x}-\frac{1}{2}(\vec{\alpha}_1+\vec{\alpha}_2)\right\}^T R^{-1}(\vec{\alpha}_1-\vec{\alpha}_2),\vec{x}^T=(x_1,\cdots,x_m).$$

We use empirical mean value vectors and the covariance matrix R,

$$\hat{\vec{a}}_1=n_1^{-1}\sum_{i=1}^{n_1}\vec{x}_i,\quad \hat{\vec{a}}_2=n_2^{-1}\sum_{i=1}^{n_2}\vec{y}_i,$$

$$\hat{R}=\frac{1}{n_1+n_2-2}\left\{\sum_{i=1}^{n_1}\left(\vec{x}_i-\hat{\vec{a}}_1\right)\left(\vec{x}_i-\hat{\vec{a}}_1\right)^T+\sum_{i=1}^{n_2}\left(\vec{y}_i-\hat{\vec{a}}_2\right)\left(\vec{y}_i-\hat{\vec{a}}_2\right)^T\right\}.$$

We shall refer to the expression

$$G_{13}(\vec{x})=\left\{\vec{x}-\frac{1}{2}\left(\hat{\vec{a}}_1+\hat{\vec{a}}_2\right)^T\hat{R}^{-1}\left(\hat{\vec{a}}_1-\hat{\vec{a}}_2\right)\right\}\frac{n_1+n_2-2-m}{n_1+n_2-2}$$

as the $G_{13}(\vec{x})$ -estimator of the discriminant function.

Theorem 13.1. [Gir54, p.611] *If in addition to the conditions of Theorem 11.1*

$$\lim_{n_1,n_2\to\infty}(\vec{a}_1+\vec{a}_2)^T R^{-1}(\vec{a}_1+\vec{a}_2)\left[n_1^{-1}+n_2^{-1}\right]=0,$$

then

$$p\lim_{n\to\infty}\{G_{13}(\nu_i)-U(\nu_i)\}=0,$$

where $\vec{\nu}_i$ is an observation of vector $\vec{\xi}$ or $\vec{\eta}$ which does not depend on $\vec{x}_i,\ \vec{y}_j;\ i=1,...,n_1;$ $j=1,...,n_2$ and distributed as $N\{\vec{a}_1,\ R\}$ or $N\{\vec{a}_2,\ R\}$.

Random Matrices. Kiev University Publishing, Ukraine, 1975, 448pp. (in Russian).

Theory of Random Determinants. Kiev University Publishing, Ukraine, 1975, 368pp. (in Russian).

Limit Theorems for Functions of Random Variables. "Higher School" Publishing, Kiev, Ukraine, 1983, 207pp. (in Russian).

Multidimensional Statistical Analysis. "Higher School" Publishing, Kiev, Ukraine, 1983, 320pp. (in Russian).

Spectral Theory of Random Matrices. "Science" Publishing, Moscow, Russia, 1988, 376pp. (in Russian).

Theory of Random Determinants. Kluwer Academic Publishers, The Netherlands, 1990, 677pp.

Theory of Systems of Empirical Equations. "Lybid" Publishing, Kiev, Ukraine, 1990, 264pp. (in Russian).

Statistical Analysis of Observations of Increasing Dimensions. Kluwer Academic Publishers, The Netherlands, 1995, 286pp.

Theory of Linear Algebraic Equations with Random Coefficients. Allerton Press, Inc, New York, U.S.A. 1996, 320pp.

An Introduction to Statistical Analysis of Random Arrays. VSP, Utrecht, The Netherlands. 1998, 673pp.

Theory of Stochastic Canonical equations. Volume I and II. Kluwer Academic Publishers, (xxiv + 497, xviii + 463).(2001).