Eva Lešanská

Optimization of the size of nonsensitiveness regions

Applications of Mathematics, Vol. 47 (2002), No. 1, 9–23

Persistent URL: http://dml.cz/dmlcz/134482

OPTIMIZATION OF THE SIZE OF NONSENSITIVENESS REGIONS*

Eva Lešanská, Olomouc

(Received October 11, 1999)

Abstract. The problem is to determine the optimum size of nonsensitiveness regions for the level of statistical tests. This is closely connected with the problem of the distribution of quadratic forms.

Keywords: linear model with inaccurate variance components, nonsensitiveness regions

MSC 2000: 62J05, 62E17

1. Introduction

In solution of standard statistical problems in a linear regression model with inaccurate variance components a question arises whether and how much these inaccuracies influence estimators of unknown parameters, the position and shape of the confidence ellipsoids, the level of statistical tests and their power function. Let us consider the problem connected with the risk of a test. It is obvious that differences $\delta\boldsymbol{\vartheta}$ between approximate and true values of variance components can cause decrease or increase in the value of the risk of the test. Let us admit the risk of the test α to be worse by a value ε, i.e. the level of the test is equal to $\alpha+\varepsilon$. Then we want to find the region $\mathcal{R}_\varepsilon$ of points $\delta\boldsymbol{\vartheta}$ such that for all $\delta\boldsymbol{\vartheta} \in \mathcal{R}_\varepsilon$ the risk of the test is not greater than the value $\alpha+\varepsilon$. This region $\mathcal{R}_\varepsilon$ is called the nonsensitiveness region. Evidently, the greater the region $\mathcal{R}_\varepsilon$, the greater are the differences $\delta\boldsymbol{\vartheta}$ and thus so much the more these differences $\delta\boldsymbol{\vartheta}$ can be neglected. How to determine each nonsensitiveness region, i.e. for the estimators of unknown parameters, for confidence ellipsoids, etc. is given in [3].

* This work was supported by the Council of the Czech Government J14/98: 153 1000 11.

The aim of the paper is to offer a procedure how to optimize the size of the nonsensitiveness region $\mathcal{R}_\varepsilon$.

2. The distribution of quadratic forms

Let $\boldsymbol{Y}$ be normally distributed n-dimensional random vector with the mean equal to $\mathbf{0}$ and with the positive definite covariance matrix $\boldsymbol{\Sigma}$, i.e. $\boldsymbol{Y} \sim N_n(\mathbf{0}, \boldsymbol{\Sigma})$. Consider a quadratic form $\mathcal{Q} = (\boldsymbol{Y}+\boldsymbol{\mu})'\boldsymbol{A}(\boldsymbol{Y}+\boldsymbol{\mu})$, where $\boldsymbol{\mu} \in \mathbb{R}^n$ and $\boldsymbol{A}$ is a symmetric matrix. Then there exists a non-singular linear transformation of the expression $\mathcal{Q}$ in the form

$$\mathcal{Q} = \sum_{r=1}^{m} \lambda_r \chi^2_{f_r}(\delta_r), \tag{2.1}$$

where λ_r are distinct non-zero eigenvalues of $\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{A}\boldsymbol{\Sigma}^{\frac{1}{2}}$, f_r are their respective multiplicities. The noncentrality parameter δ_r is given by $\delta_r = \sum\limits_{i=1}^{n} \gamma_i$, where

$$\gamma_i = \begin{cases} \boldsymbol{g}_i'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{\mu}\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\boldsymbol{g}_i, & \lambda_i = \lambda_r, \\ 0, & \lambda_i \neq \lambda_r, \end{cases}$$

$\boldsymbol{g}_i$ are eigenvectors from the spectral decomposition

$$\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{A}\boldsymbol{\Sigma}^{\frac{1}{2}} = \sum_{i=1}^{n} \lambda_i \boldsymbol{g}_i \boldsymbol{g}_i', \quad \boldsymbol{g}_i'\boldsymbol{g}_j = \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases}$$

and $\chi^2_{f_r}(\delta_r)$, $r = 1, \ldots, m$, are independent χ^2-variables with f_r degrees of freedom.

Lemma 2.1. *The characteristic function of* (2.1) *is*

$$\Psi(t) = \prod_{r=1}^{m} (1 - 2\mathrm{i}\lambda_r t)^{-f_r/2} \exp\left\{ \mathrm{i} \sum_{r=1}^{m} \frac{\delta_r \lambda_r t}{1 - 2\mathrm{i}\lambda_r t} \right\}. \tag{2.2}$$

P r o o f. The characteristic function of the random variable $\chi^2_k(\delta)$ is

$$\Psi(t) = (1 - 2\mathrm{i}t)^{-k/2} \exp\left\{ \frac{\mathrm{i}\delta t}{1 - 2\mathrm{i}t} \right\}.$$

The result follows from properties of characteristic functions:

$$Y = \lambda X, \ \ \lambda \in \mathbb{R}^1 \Rightarrow \Psi_Y(t) = \Psi_X(\lambda t)$$

and

$$X_1, \ldots, X_n \text{ independent random variables } \Rightarrow \Psi_{X_1+\ldots X_n}(t) = \prod_{i=1}^{n} \Psi_{X_i}(t).$$

□

To obtain the probability density by integration of the inversion formula appears hopeless, except for the particular case $m = 1$. Thus, the only way how to compute the distribution of the variable $\mathcal{Q}$ is to use numerical integration of the inversion formula. For the cumulative distribution function $F(x)$ we have

$$F(x) = \frac{1}{2} - \frac{1}{\pi} \int_0^{+\infty} t^{-1} \operatorname{Im}\{e^{-itx}\Psi(t)\}\, dt.$$

Using some geometric relations this equation can be rewritten, after the substitution $2t = u$, to the expression

$$P\{\mathcal{Q} \leqslant x\} = \frac{1}{2} - \frac{1}{\pi} \int_0^{+\infty} \frac{\sin \tau(u)}{u\varrho(u)}\, du, \tag{2.3}$$

where

$$\tau(u) = \frac{1}{2}\sum_{r=1}^{m}\left[f_r \arctan(\lambda_r u) + \frac{\delta_r \lambda_r u}{1 + \lambda_r^2 u^2}\right] - \frac{1}{2}xu,$$

$$\varrho(u) = \prod_{r=1}^{m}(1 + \lambda_r^2 u^2)^{f_r/4} \exp\left\{\frac{1}{2}\sum_{r=1}^{m}\frac{\delta_r \lambda_r^2 u^2}{1 + \lambda_r^2 u^2}\right\}.$$

Since the function $u\varrho(u)$ increases monotonically towards $+\infty$, in numerical work the integration in (2.3) will be carried out over a finite range $0 \leqslant u \leqslant U$ only. The upper bound U results from the assumption that the error of integration must be less than ω. Let t_U denote the error of truncation, i.e. we get the assumption

$$t_U = \frac{1}{\pi}\int_U^{+\infty} \frac{\sin \tau(u)}{u\varrho(u)}\, du \leqslant \frac{\omega}{2}.$$

The value $|t_U|$ can be bounded above by T_U, where

$$T_U^{-1} = \pi k U^k \prod_{r=1}^{m} |\lambda_r|^{\frac{1}{2}f_r} \exp\left\{\frac{1}{2}\sum_{r=1}^{m}\frac{\delta_r \lambda_r^2 U^2}{1 + \lambda_r^2 U^2}\right\}, \tag{2.4}$$

with $k = \frac{1}{2}\sum\limits_{r=1}^{m} f_r$.

The probability density $g(x)$ of the quadratic form $\mathcal{Q}$ could be computed by using a formula analogous to (2.3). In fact,

$$g(x) = \frac{1}{\pi}\int_0^{+\infty} \frac{\cos \tau(u)}{\varrho(u)}\, du. \tag{2.5}$$

Due to the absence of the factor u^{-1} in the integrand in the expression (2.5), numerical integration, for the same accuracy, can be expected to require a slightly larger number of steps than is needed to compute the distribution function. For more detail cf. [1].

If $\boldsymbol{A}$ is a p.s.d. matrix, then $\lambda_i \geqslant 0$, $i = 1\ldots,n$. In such a case, another way how to compute the distribution of the quadratic form $\mathcal{Q}$ is on the basis of the approximation of the linear combination of the independent noncentral χ^2-variables.

Lemma 2.2. *Let* $\lambda_1 > 0$ *and* $\lambda_2 > 0$*. Let* $\chi^2_{f_1}(\delta_1)$ *and* $\chi^2_{f_2}(\delta_2)$ *be stochastically independent. Then*

$$\lambda_1\chi^2_{f_1}(\delta_1) + \lambda_2\chi^2_{f_2}(\delta_2) \approx \lambda\chi^2_f(\delta), \tag{2.6}$$

where

$$\begin{aligned} f &= \frac{(\lambda_1^2 f_1 + \lambda_2^2 f_2)(\lambda_1 f_1 + \lambda_1\delta_1 + \lambda_2 f_2 + \lambda_2\delta_2)^2}{(\lambda_1^2 f_1 + \lambda_1^2\delta_1 + \lambda_2^2 f_2 + \lambda_2^2\delta_2)^2}, \\ \lambda &= \frac{\lambda_1^2 f_1 + \lambda_1^2\delta_1 + \lambda_2^2 f_2 + \lambda_2^2\delta_2}{\lambda_1 f_1 + \lambda_1\delta_1 + \lambda_2 f_2 + \lambda_2\delta_2}, \\ \delta &= \frac{(\lambda_1^2 \delta_1 + \lambda_2^2 \delta_2)(\lambda_1 f_1 + \lambda_1\delta_1 + \lambda_2 f_2 + \lambda_2\delta_2)^2}{(\lambda_1^2 f_1 + \lambda_1^2\delta_1 + \lambda_2^2 f_2 + \lambda_2^2\delta_2)^2}. \end{aligned}$$

P r o o f. The basic characteristics for the random variable $\lambda\chi^2_f(\delta)$ are

$$\begin{aligned} E(\lambda\chi^2_f(\delta)) &= \lambda(f+\delta), \\ \operatorname{var}(\lambda\chi^2_f(\delta)) &= \lambda^2(2f+4\delta). \end{aligned}$$

Results follow from the solution of the system

$$\begin{aligned} \lambda_1(f_1+\delta_1) + \lambda_2(f_2+\delta_2) &= \lambda(f+\delta), \\ \lambda_1^2(2f_1+4\delta_1) + \lambda_2^2(2f_2+4\delta_2) &= \lambda^2(2f+4\delta), \\ \lambda_1^2\delta_1 + \lambda_2^2\delta_2 &= \lambda^2\delta. \end{aligned}$$

□

R e m a r k 2.3. The approximation (2.6) is not unique. Another type of the approximation can be derived for example under assumptions that the first three moments of χ^2-variables must be equal. The quality of the approximation (2.6) is shown in Example 4.1.

R e m a r k 2.4. The assumption of the matrix $\boldsymbol{A}$ to be p.s.d. is necessary in the above mentioned approximation. If $\boldsymbol{A}$ is not p.s.d. we could use the numerical way for computing the probability of the quadratic form $\mathcal{Q}$ only.

Finally, we can approximate the noncentral χ^2-distribution by the central χ^2-distribution.

Lemma 2.5.

$$\chi_g^2(\delta) \approx \frac{g+2\delta}{g+\delta}\chi^2_{\frac{(g+\delta)^2}{g+2\delta}}(0). \tag{2.7}$$

P r o o f. It can be proved analogously to the proof of Lemma 2.2 (cf. also in [2]). □

3. Nonsensitiveness regions

Let

$$\boldsymbol{Y} \sim N_n(\boldsymbol{X\beta}, \boldsymbol{\Sigma}(\boldsymbol{\vartheta})),\ \ \boldsymbol{\beta} \in \mathbb{R}^k,\ \ \boldsymbol{\vartheta} \in \underline{\boldsymbol{\vartheta}} = \{\boldsymbol{\vartheta}\colon\ \boldsymbol{\vartheta} \in \mathbb{R}^p,\ \ \vartheta_1 > 0, \ldots, \vartheta_p > 0\}, \tag{3.1}$$

where $\boldsymbol{Y}$ is an n-dimensional random vector (observation vector), $\boldsymbol{X}_{n\times k}$ a known matrix (design matrix), $\boldsymbol{\beta}$ an unknown vector (parameter of the first order), $\boldsymbol{\Sigma}(\boldsymbol{\vartheta}) = \sum_{i=1}^{p} \vartheta_i \boldsymbol{V}_i$ a covariance matrix, $\boldsymbol{\vartheta}$ an unknown vector (parameter of the second order) and $\boldsymbol{V}_1, \ldots, \boldsymbol{V}_p$ known positive semidefinite matrices of the type $n \times n$.

In the sequel the matrix $\boldsymbol{X}$ will be supposed to be of the full rank in columns, i.e. $r(\boldsymbol{X}) = k < n$, and $\boldsymbol{\Sigma}(\boldsymbol{\vartheta})$ is positive definite for all $\boldsymbol{\vartheta} \in \underline{\boldsymbol{\vartheta}}$. In such a case, the mixed linear model (3.1) is called regular.

Let $\boldsymbol{\vartheta}^*$ be the true value of the parameter $\boldsymbol{\vartheta}$. Let the null hypothesis concerning the parameter $\boldsymbol{\beta}$ be

$$H_0\colon\ \boldsymbol{H\beta} + \boldsymbol{h} = \boldsymbol{0}, \tag{3.2}$$

where $\boldsymbol{H}_{q\times k}$ is a given matrix with the rank $r(\boldsymbol{H}) = q \leqslant k$ and $\boldsymbol{h}$ is a known q-dimensional vector. Let the alternative hypothesis be

$$H_a\colon\ \boldsymbol{H\beta} + \boldsymbol{h} \neq \boldsymbol{0}. \tag{3.3}$$

Lemma 3.1. *Let the regular mixed linear model* (3.1) *and hypotheses* (3.2) *and* (3.3) *be under consideration. Let*

$$(3.4)\; T_H(\boldsymbol{Y},\boldsymbol{\vartheta}^*) = (\boldsymbol{H}\hat{\boldsymbol{\beta}}(\boldsymbol{Y},\boldsymbol{\vartheta}^*)+\boldsymbol{h})'[\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{X})^{-1}\boldsymbol{H}']^{-1}(\boldsymbol{H}\hat{\boldsymbol{\beta}}(\boldsymbol{Y},\boldsymbol{\vartheta}^*)+\boldsymbol{h}),$$

where

$$\hat{\boldsymbol{\beta}}(\boldsymbol{Y},\boldsymbol{\vartheta}^*) = (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{Y}.$$

(i) *If H_0 is true, then the statistic $T_H(\mathbf{Y},\boldsymbol{\vartheta}^\star)$ has the central chi-square distribution with q degrees of freedom.*
(ii) *If H_0 is not true, then $T_H(\boldsymbol{Y},\boldsymbol{\vartheta}^*)$ has the noncentral chi-square distribution with q degrees of freedom and the parameter of its noncentrality is*

$$\delta = (\boldsymbol{H}\boldsymbol{\beta}+\boldsymbol{h})'[\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{X})^{-1}\boldsymbol{H}']^{-1}(\boldsymbol{H}\boldsymbol{\beta}+\boldsymbol{h}).$$

P r o o f. Both statements follow from the second fundamental theorem of the least squares theory given in [4], p. 155. □

The statistic $T_H(\boldsymbol{Y},\boldsymbol{\vartheta}^*)$ has been used for testing the hypothesis H_0 against H_a. If $T_H(\boldsymbol{y},\boldsymbol{\vartheta}^*) \geqslant \chi^2_q(0,1-\alpha)$, where $\boldsymbol{y}$ means a realization of $\boldsymbol{Y}$, then H_0 is rejected with the risk α. Here $\chi^2_q(0,1-\alpha)$ denotes $(1-\alpha)$-quantile of $\chi^2_q(0)$.

Let $\boldsymbol{\vartheta}^*$ be changed into $\boldsymbol{\vartheta}^*+\delta\boldsymbol{\vartheta}$. We will study how the change $\delta\boldsymbol{\vartheta}$ influences the risk of the test. That is why in the following we will suppose H_0 to be true.

Theorem 3.2. *Let the regular mixed linear model* (3.1) *and the hypothesis* (3.2) *be under consideration. Let H_0 be true and let*

$$\delta T_H = \delta\boldsymbol{\vartheta}'\frac{\partial T_H(\boldsymbol{Y},\boldsymbol{\vartheta})}{\partial\boldsymbol{\vartheta}}\Big|_{\boldsymbol{\vartheta}=\boldsymbol{\vartheta}^*}.$$

Then

$$(3.5)\quad \begin{aligned}\delta T_H = &-2[\boldsymbol{H}\hat{\boldsymbol{\beta}}(\boldsymbol{Y},\boldsymbol{\vartheta}^*)+\boldsymbol{h}]'\boldsymbol{C}_H\boldsymbol{F}_H\boldsymbol{\Sigma}(\delta\boldsymbol{\vartheta})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)(\boldsymbol{Y}-\boldsymbol{X}\hat{\boldsymbol{\beta}}(\boldsymbol{Y},\boldsymbol{\vartheta}^*))\\ &-[\boldsymbol{H}\hat{\boldsymbol{\beta}}(\boldsymbol{Y},\boldsymbol{\vartheta}^*)+\boldsymbol{h}]'\boldsymbol{C}_H\boldsymbol{F}_H\boldsymbol{\Sigma}(\delta\boldsymbol{\vartheta})\boldsymbol{F}_H'\boldsymbol{C}_H[\boldsymbol{H}\hat{\boldsymbol{\beta}}(\boldsymbol{Y},\boldsymbol{\vartheta}^*)+\boldsymbol{h}],\end{aligned}$$

where

$$\boldsymbol{F}_H = \boldsymbol{H}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)$$

and

$$\boldsymbol{C}_H = (\boldsymbol{H}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{X})^{-1}\boldsymbol{H}')^{-1}.$$

The mean value of δT_H is

$$E(\delta T_H|\boldsymbol{\beta},\boldsymbol{\vartheta}^*) = -\delta\boldsymbol{\vartheta}'[\mathrm{Tr}(\boldsymbol{U}_H\boldsymbol{V}_1),\ldots,\mathrm{Tr}(\boldsymbol{U}_H\boldsymbol{V}_p)]', \tag{3.6}$$

where $\boldsymbol{U}_H = \boldsymbol{F}_H'\boldsymbol{C}_H\boldsymbol{F}_H$ and $\mathrm{Tr}(\boldsymbol{U}_H)$ means the trace of the matrix $\boldsymbol{U}_H$. The variance of δT_H is

$$\begin{aligned}\mathrm{var}(\delta T_H|\boldsymbol{\beta},\boldsymbol{\vartheta}^*) = {} & 4\,\mathrm{Tr}\{\boldsymbol{U}_H\boldsymbol{\Sigma}(\delta\boldsymbol{\vartheta})[\boldsymbol{M}_X\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\boldsymbol{M}_X]^+\boldsymbol{\Sigma}(\delta\boldsymbol{\vartheta})\} \\ & + 2\,\mathrm{Tr}\{\boldsymbol{U}_H\boldsymbol{\Sigma}(\delta\boldsymbol{\vartheta})\boldsymbol{U}_H\boldsymbol{\Sigma}(\delta\boldsymbol{\vartheta})\},\end{aligned} \tag{3.7}$$

where

$$[\boldsymbol{M}_X\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\boldsymbol{M}_X]^+ = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*) - \boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{X}[\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{X}]^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*).$$

Proof. Proof can be found in [3]. □

The mean value $E(\delta T_H)$ depends on $\delta\boldsymbol{\vartheta}$ linearly and the term $t\sqrt{\mathrm{var}(\delta T_H)}$ depends linearly on the norm $\|\delta\boldsymbol{\vartheta}\| = \sqrt{(\delta\boldsymbol{\vartheta})'(\delta\boldsymbol{\vartheta})}$. Let the function $\Phi(\delta\boldsymbol{\vartheta})$, $\delta\boldsymbol{\vartheta} \in \mathbb{R}^p$, be defined as follows

$$\Phi(\delta\boldsymbol{\vartheta}) = -\delta\boldsymbol{\vartheta}'\boldsymbol{a}_0 + t\sqrt{\delta\boldsymbol{\vartheta}'\boldsymbol{A}_0\delta\boldsymbol{\vartheta}}, \tag{3.8}$$

where for $i,j = 1,\ldots,p$

$$\{\boldsymbol{A}_0\}_{i,j} = 2\,\mathrm{Tr}(\boldsymbol{U}_H\boldsymbol{V}_i\boldsymbol{U}_H\boldsymbol{V}_j) + 4\,\mathrm{Tr}(\boldsymbol{U}_H\boldsymbol{V}_i[\boldsymbol{M}_X\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)\boldsymbol{M}_X]^+\boldsymbol{V}_j), \tag{3.9}$$

$$\boldsymbol{a}_0 = [\mathrm{Tr}(\boldsymbol{U}_H\boldsymbol{V}_1),\ldots,\mathrm{Tr}(\boldsymbol{U}_H\boldsymbol{V}_p)]'. \tag{3.10}$$

Definition 3.3. Let

$$\mathcal{R}_\varepsilon = \{\delta\boldsymbol{\vartheta}\colon\ \delta\boldsymbol{\vartheta} \in \mathbb{R}^p,\ \ \Phi(\delta\boldsymbol{\vartheta}) \leqslant \delta_\varepsilon\}, \tag{3.11}$$

where δ_ε is given by

$$P\{\chi_q^2(0) \geqslant \chi_q^2(0, 1-\alpha) - \delta_\varepsilon\} = \alpha + \varepsilon.$$

The set $\mathcal{R}_\varepsilon$ is called the nonsensitiveness region for the risk of the test.

Lemma 3.4. *Let the regular mixed linear model* (3.1) *and the hypothesis* (3.2) *be under consideration. Let* H_0 *be true and let* $\boldsymbol{a}_0$ *and* $\boldsymbol{A}_0$ *be given by* (3.10) *and* (3.9), *respectively. The boundary of the set* $\mathcal{R}_\varepsilon$ *is*

$$\overline{\mathcal{R}}_\varepsilon = \left\{\delta\boldsymbol{\vartheta}\colon\ (\delta\boldsymbol{\vartheta} - \boldsymbol{x}_0)'(t^2\boldsymbol{A}_0 - \boldsymbol{a}_0\boldsymbol{a}_0')(\delta\boldsymbol{\vartheta} - \boldsymbol{x}_0) = \frac{\delta_\varepsilon^2 t^2}{t^2 - \boldsymbol{a}_0'\boldsymbol{A}_0^- \boldsymbol{a}_0}\right\}, \tag{3.12}$$

where $\boldsymbol{x}_0 = \frac{\delta_\varepsilon}{t^2 - \boldsymbol{a}_0'\boldsymbol{A}_0^-\boldsymbol{a}_0}\boldsymbol{A}_0^-\boldsymbol{a}_0$, $\delta_\varepsilon = \chi_q^2(0, 1-\alpha) - \chi_q^2(0, 1-\alpha-\varepsilon)$ *and* ε, t *are chosen positive numbers. Here* $\boldsymbol{A}_0^-$ *means* g*-inverse of the matrix* $\boldsymbol{A}_0$.

Proof. It follows from the solution of the equation $\Phi(\delta\boldsymbol{\vartheta}) = \delta_\varepsilon$ from Definition (3.3) of the nonsensitiveness region. For details see [3]. □

Theorem 3.5. *Let the regular mixed linear model* (3.1) *and hypotheses* (3.2), (3.3) *be under consideration. If* H_0 *is true, then*

$$\delta\boldsymbol{\vartheta} \in \mathcal{R}_\varepsilon \Rightarrow P\{T_H(\boldsymbol{Y}, \boldsymbol{\vartheta}^* + \delta\boldsymbol{\vartheta}) \geqslant \chi_q^2(0, 1-\alpha)\} \leqslant \alpha + \varepsilon. \tag{3.13}$$

Proof. Proof can be found in [3]. □

It is easy to see that $\mathcal{R}_\varepsilon$ is greater for lower values of $t \in (0, \infty)$. The minimum value of the parameter t, i.e. the maximum size of the nonsensitiveness region, can be determined from the natural condition

$$E(\delta T_H) + t\sqrt{\mathrm{var}(\delta T_H)} = q(1-\alpha),$$

where $q(1-\alpha)$ is a $(1-\alpha)$-quantile of a distribution of the random variable δT_H with a sufficiently small α. Thus the problem is to determine the distribution of δT_H, which enables us to determine the optimum value of the parameter t.

Lemma 3.6. *Let the regular mixed linear model* (3.1) *and the hypothesis* (3.2) *be under consideration. Let the null hypothesis be true. Then*

$$\delta T_H = -\sum_{r=1}^{m} \lambda_r \chi_{f_r}^2(0),$$

where λ_r *are distinct non-zero eigenvalues of the matrix* $\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{A}\boldsymbol{W}^{\frac{1}{2}}$ *with multiplicity* f_r*, random variables* $\chi^2_{f_r}(0)$ *are independent and*

$$\boldsymbol{W} = \begin{pmatrix} \boldsymbol{H}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{X})^{-1}\boldsymbol{H}', & \boldsymbol{0} \\ \boldsymbol{0}, & \boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*) - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{X})^{-1}\boldsymbol{X}' \end{pmatrix},$$

$$\boldsymbol{U}_i = \begin{bmatrix} \boldsymbol{C}_H\boldsymbol{F}_H\boldsymbol{V}_i\boldsymbol{F}_H'\boldsymbol{C}_H, & \boldsymbol{C}_H\boldsymbol{F}_H\boldsymbol{V}_i\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*) \\ \boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{V}_i\boldsymbol{F}_H'\boldsymbol{C}_H, & \boldsymbol{0} \end{bmatrix}, \quad i = 1, \dots, p,$$

$$\boldsymbol{A} = \sum_{i=1}^{p} \delta\vartheta_i \boldsymbol{U}_i.$$

Proof. Let us denote

$$\boldsymbol{\eta} = \boldsymbol{H}\hat{\boldsymbol{\beta}}(\boldsymbol{Y}, \boldsymbol{\vartheta}^*) + \boldsymbol{h},$$
$$\boldsymbol{v} = \boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}(\boldsymbol{Y}, \boldsymbol{\vartheta}^*).$$

Then, taking H_0 into consideration, we have

$$\begin{pmatrix} \boldsymbol{\eta} \\ \boldsymbol{v} \end{pmatrix} \sim N_{q+n}\left[\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}; \boldsymbol{W}\right],$$

where

$$\boldsymbol{W} = \begin{pmatrix} \boldsymbol{H}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{X})^{-1}\boldsymbol{H}', & \boldsymbol{0} \\ \boldsymbol{0}, & \boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*) - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{X})^{-1}\boldsymbol{X}' \end{pmatrix}.$$

Now we can rewrite the expression (3.5) of the correction term δT_H in the form

$$\delta T_H = -\sum_{i=1}^{p} \delta\vartheta_i \left\{ (\boldsymbol{\eta}', \boldsymbol{v}')\, \boldsymbol{U}_i \begin{pmatrix} \boldsymbol{\eta} \\ \boldsymbol{v} \end{pmatrix} \right\} = -(\boldsymbol{\eta}', \boldsymbol{v}')\, \boldsymbol{A} \begin{pmatrix} \boldsymbol{\eta} \\ \boldsymbol{v} \end{pmatrix},$$

where

$$\boldsymbol{U}_i = \begin{bmatrix} \boldsymbol{C}_H\boldsymbol{F}_H\boldsymbol{V}_i\boldsymbol{F}_H'\boldsymbol{C}_H, & \boldsymbol{C}_H\boldsymbol{F}_H\boldsymbol{V}_i\boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*) \\ \boldsymbol{\Sigma}^{-1}(\boldsymbol{\vartheta}^*)\boldsymbol{V}_i\boldsymbol{F}_H'\boldsymbol{C}_H, & \boldsymbol{0} \end{bmatrix},$$

$$\boldsymbol{A} = \sum_{i=1}^{p} \delta\vartheta_i \boldsymbol{U}_i.$$

Since the vector $(\boldsymbol{\eta}', \boldsymbol{v}')'$ is normally distributed, using the transformation (2.1) we obtain

$$\delta T_H = -\sum_{r=1}^{m} \lambda_r \chi^2_{f_r}(0),$$

where λ_r are distinct non-zero eigenvalues of the matrix $\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{A}\boldsymbol{W}^{\frac{1}{2}}$ with multiplicity f_r, $\delta_r = 0$ and $\chi^2_{f_r}(0)$, $r = 1, \dots, m$, are independent. □

Now can determine the optimum value t^* of the parameter t maximizing the size of the nonsensitiveness region $\mathcal{R}_\varepsilon$.

Theorem 3.7. *Let the regular mixed linear model* (3.1) *and the hypothesis* (3.2) *be under consideration. Let matrices* $\boldsymbol{W}$ *and* $\boldsymbol{A}$ *be defined as in Lemma* 3.6. *Let* H_0 *be true. The optimum value* t^**, which maximize the size of* $\mathcal{R}_\varepsilon$*, is*

$$t^* = \max\{t_{\delta\vartheta} \colon \ \|\delta\boldsymbol{\vartheta}\| = 1\},$$

where $t_{\delta\vartheta}$ *is a solution of the equation*

$$E(\delta T_H|\delta\boldsymbol{\vartheta}) + t_{\delta\vartheta}\sqrt{\mathrm{var}(\delta T_H|\delta\boldsymbol{\vartheta})} = q(1-\alpha), \quad \|\delta\boldsymbol{\vartheta}\| = 1 \tag{3.14}$$

and

$$E(\delta T_H|\delta\boldsymbol{\vartheta}) = -\sum_{r=1}^{m} \lambda_r f_r,$$

$$\mathrm{var}(\delta T_H|\delta\boldsymbol{\vartheta}) = 2\sum_{r=1}^{m} \lambda_r^2 f_r,$$

where λ_j *are distinct non-zero eigenvalues of the matrix* $\boldsymbol{W}^{\frac{1}{2}}\boldsymbol{A}\boldsymbol{W}^{\frac{1}{2}}$ *with multiplicity* f_j, $j = 1, \ldots, m$.

P r o o f. With respect to Lemma 3.6, it is sufficient to prove that the value the parameter t is independent on the norm $\|\delta\boldsymbol{\vartheta}\|$ in the fixed direction $\delta\boldsymbol{\vartheta}$. Let $\widetilde{\delta\boldsymbol{\vartheta}} = k\delta\boldsymbol{\vartheta}$, $k > 0$, $\|\delta\boldsymbol{\vartheta}\| = 1$. Then $\|\widetilde{\delta\boldsymbol{\vartheta}}\| = k\|\delta\boldsymbol{\vartheta}\|$ and $\delta T_H(\widetilde{\delta\boldsymbol{\vartheta}}) = k\delta T_H(\delta\boldsymbol{\vartheta})$. Hence the mean value, the standard deviation and quantiles of δT_H are proportional to the norm $\|\delta\boldsymbol{\vartheta}\|$ and the equation (3.14) implies $t_{\widetilde{\delta\vartheta}} = t_{\delta\vartheta}$. $\square$

4. Numerical results

At the beginning of this section the results for computing the distribution of quadratic forms will be given. An example of the determination of the maximum size of the nonsensitiveness region will be given in the second part.

Simpson's rule is used for numerical integration. The accuracy of computation of numerical integration is $\omega = 0.001$. In Tab. 1 a comparison of the probability value of quadratic forms $\mathcal{Q}_i$, $i = 1, \ldots, 4$, and their corresponding approximate noncentral χ^2-variables from Example 4.1 are given. The α-quantiles, denoted by x_α, of the approximate χ^2-variable are also shown. In Tab. 2 the quadratic form $\mathcal{Q}_5$, when $\boldsymbol{A}$ is not p.s.d., is shown. In this case, the results are compared with [1].

E x a m p l e 4.1. The quadratic forms and their approximation.

$$\mathcal{Q}_1 = \chi_3^2(0.2) + 2\chi_5^2(3.3) \approx 1.8384\chi_{6.8054}^2(3.9649),$$
$$\mathcal{Q}_2 = 5\chi_{10}^2(4) + 3\chi_6^2(7) \approx 4.2844\chi_{16.5612}^2(8.8799),$$
$$\mathcal{Q}_3 = 2\chi_{15}^2(1.5) + 3\chi_4^2(4.7) \approx 2.4416\chi_{16.1031}^2(8.1019),$$
$$\mathcal{Q}_4 = 1.4\chi_9^2(0.8) + 8.5\chi_7^2(6.3) \approx 7.7316\chi_{8.7556}^2(7.6407),$$
$$\mathcal{Q}_5 = 0.1693\chi_{9.9426}^2(0) - 0.253\chi_{7.8947}^2(0).$$

	$\mathcal{Q}_1$		$\mathcal{Q}_2$		$\mathcal{Q}_3$		$\mathcal{Q}_4$	
α	x_α	P	x_α	P	x_α	P	x_α	P
0.01	3.7227	0.0086	42.7505	0.0100	22.6837	0.0086	33.3823	0.0076
0.05	6.4723	0.0466	57.3855	0.0496	30.6599	0.0471	51.7346	0.0459
0.25	12.4584	0.2498	83.4492	0.2499	44.9574	0.2501	87.6815	0.2512
0.50	18.2406	0.5033	105.4013	0.5008	57.0617	0.5036	119.9470	0.5046
0.75	25.4544	0.7522	130.6387	0.7507	71.0263	0.7525	158.4509	0.7523
0.95	38.4565	0.9488	172.8972	0.9496	94.4923	0.9490	225.1027	0.9466
0.99	49.4923	0.9887	206.8763	0.9881	113.4154	0.9890	280.0111	0.9886

Table 1. Probability $P = P(\mathcal{Q}_i \leqslant x_\alpha)$, $i = 1, 2, 3, 4$.

x	P	[1]
-2	0.0918	0.0898
0	0.5927	0.5939
2.5	0.9901	0.9902

Table 2. Probability $P = P(\mathcal{Q}_5 \leqslant x)$.

How we can see in Tabs. 1 and 2, the accuracy 0.005 is obtained in computing the probability value by the above mentioned numerical algorithm and by the approximation from Lemma 2.2. A greater accuracy has not been reached yet.

E x a m p l e 4.2. Let a straight line be given in a plane. We have four measurements at points $x = 1, 2, 3, 4$. The accuracy of the measurement is characterized by the standard deviation $\sigma_1 = 0.004$ (at points $x = 1$ and $x = 2$) and $\sigma_2 = 0.001$ (at points $x = 3$ and $x = 4$). Let the null hypothesis be "the coefficients in the equation of the straight line are equal". The problem is to maximizing the size of the nonsensitiveness region for the risk of the test.

The process of measurement if the error vector is assumed to be normally distributed can be modelled by

$$\boldsymbol{Y} \sim N_4[\boldsymbol{X\beta}, \boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*)],$$

where

$$\boldsymbol{X} = \begin{pmatrix} 1, & 1 \\ 1, & 2 \\ 1, & 3 \\ 1, & 4 \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix},$$

and

$$\boldsymbol{\Sigma}(\boldsymbol{\vartheta}^*) = \begin{pmatrix} 16\cdot 10^{-6}, & 0, & 0, & 0 \\ 0, & 16\cdot 10^{-6}, & 0, & 0 \\ 0, & 0, & 1\cdot 10^{-6}, & 0 \\ 0, & 0, & 0, & 1\cdot 10^{-6} \end{pmatrix}.$$

The null hypothesis is

$$H_0\colon\ (1,-1)\boldsymbol{\beta} = 0.$$

In this case, we obtain the correction term δT_H in the form

$$\delta T_H = \lambda_1 \chi_1^2(0) + \lambda_2 \chi_1^2(0).$$

Numerical results of coefficients λ_1, λ_2 and of values $t_{\delta\vartheta}$ for different directions $\delta\boldsymbol{\vartheta} = (\cos\gamma, \sin\gamma)'$, i.e. for different angles γ, are given in Tab. 3 and in Figs. 1, 2. Symbols $q(0.95)$ and $q(0.99)$ describe the corresponding quantiles of the distribution of the random variable δT_H, parameters t_1 and t_2 denote values of $t_{\delta\vartheta}$ for the risk of the test $\alpha_1 = 0.05$ and $\alpha_2 = 0.01$, respectively.

γ	λ_1	λ_2	$q(0.95)$	$q(0.99)$	t_1	t_2
0	−50675	18632	49380	100700	1.07	1.74
$\pi/16$	−154502	28004	61100	137000	0.84	1.19
$\pi/8$	−300908	84815	210300	443000	0.96	1.49
$\pi/4$	−557706	190460	496000	1020000	1.04	1.66
$3\pi/8$	−730488	267999	709800	1448000	1.07	1.74
$\pi/2$	−792371	305049	830000	1687500	1.10	1.81
$5\pi/8$	−733849	259884	801000	1668750	1.11	1.88
$3\pi/4$	−563886	241955	664000	1331000	1.14	1.90
$7\pi/8$	−308746	151859	429000	848000	1.20	2.07
$15\pi/16$	−161829	98184	288900	561000	1.32	2.33
π	50675	−18632	179400	321000	1.93	3.78
$17\pi/16$	154502	−28004	568400	1018300	1.99	4.02
$9\pi/8$	300908	−84815	1083500	1925000	1.96	3.87
$5\pi/4$	557706	−190460	1984000	3540000	1.94	3.81
$11\pi/8$	730488	−267999	2586200	4631050	1.93	3.79
$3\pi/2$	792371	−305049	2793000	4920000	1.92	3.69
$13\pi/8$	733849	−295884	2580000	4626000	1.91	3.74
$7\pi/4$	563886	−241955	1973000	3547000	1.90	3.71
$15\pi/8$	308746	−151859	1068000	1928000	1.87	3.64
$31\pi/16$	161829	−98184	548500	1016500	1.81	3.56

Table 3. Values of t_1 and t_2 in dependence on the direction $\delta\boldsymbol{\vartheta}$.

From Table 3, the minimum value of parameter t which optimizes the size of the nonsensitiveness region $\mathcal{R}_\varepsilon$ for the risk of the test α is

$$\alpha = 0.05: \quad t_1^* = \max\{t_1 \colon \|\delta\boldsymbol{\vartheta}\| = 1\} \doteq 1.99,$$
$$\alpha = 0.01: \quad t_2^* = \max\{t_2 \colon \|\delta\boldsymbol{\vartheta}\| = 1\} \doteq 4.02.$$

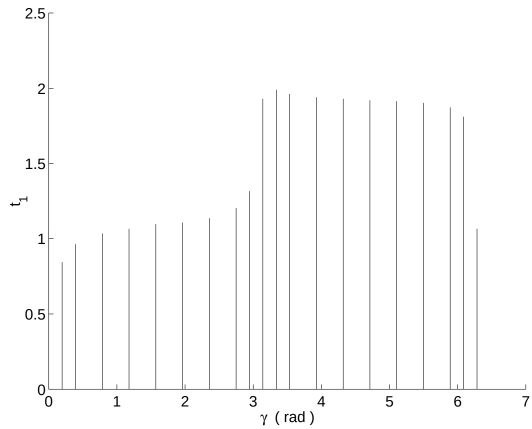

Figure 1. Dependence of the parameter t_1 on the angle γ.

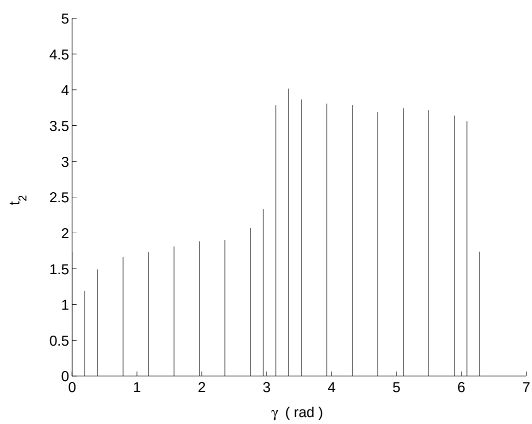

Figure 2. Dependence of the parameter t_2 on the angle γ.

At the end we determine the set $\mathcal{R}_\varepsilon$ in dependence on the risk of the test α and on the given number ε. Obviously,

$$T_H(\boldsymbol{Y}, \boldsymbol{\vartheta}^*) \sim \chi_1^2(0).$$

From definition (3.3) we obtain

$$\delta_\varepsilon = \chi_1^2(0, 1-\alpha) - \chi_1^2(0, 1-\alpha-\varepsilon).$$

Values of δ_ε for some α and ε are given in Tabs. 4 and 5. Further, we have to express the center of $\mathcal{R}_\varepsilon$. This point $\boldsymbol{x}_0$ depends on δ_ε and on the parameter t. In our case, the boundary of $\mathcal{R}_\varepsilon$ given by (3.12) can be characterized by the ellipse. Hence the nonsensitiveness region $\mathcal{R}_\varepsilon$ for the risk α of the test can be expressed as the set of points

$$\mathcal{R}_\varepsilon = \{\delta\boldsymbol{\vartheta}\colon\ \delta\boldsymbol{\vartheta} \in \mathbb{R}^2,\ \ (\delta\boldsymbol{\vartheta} - \boldsymbol{x}_0)'\boldsymbol{K}(t^*)(\delta\boldsymbol{\vartheta} - \boldsymbol{x}_0) \leqslant s\},$$

where

$$\boldsymbol{K}(t_1^*) = \begin{bmatrix} 0.0221 \cdot 10^{13}, & -0.1312 \cdot 10^{13} \\ -0.1312 \cdot 10^{13}, & 5.4723 \cdot 10^{13} \end{bmatrix},$$
$$\boldsymbol{K}(t_2^*) = \begin{bmatrix} 0.0093 \cdot 10^{13}, & -0.0487 \cdot 10^{13} \\ -0.0487 \cdot 10^{13}, & 2.3017 \cdot 10^{13} \end{bmatrix}$$

and the other terms are given in Tabs. 4 and 5.

ε	δ_ε	$\{x_0\}_{1,1}$	$\{x_0\}_{2,1}$	s
0.015	0.339	$7.8378 \cdot 10^{-7}$	$0.4899 \cdot 10^{-7}$	0.1315
0.005	0.113	$2.6126 \cdot 10^{-7}$	$0.1633 \cdot 10^{-7}$	0.0146
0.002	0.045	$1.0404 \cdot 10^{-7}$	$0.0650 \cdot 10^{-7}$	0.0023
0.001	0.023	$0.5318 \cdot 10^{-7}$	$0.0332 \cdot 10^{-7}$	0.0006

Table 4. Expressions in $\mathcal{R}_\varepsilon$ for $\alpha = 0.05$, $t_1^* = 1.99$.

ε	δ_ε	$\{x_0\}_{1,1}$	$\{x_0\}_{2,1}$	s
0.015	1.61	$8.2413 \cdot 10^{-7}$	$0.5151 \cdot 10^{-7}$	2.6750
0.005	0.537	$2.7488 \cdot 10^{-7}$	$0.1718 \cdot 10^{-7}$	0.2976
0.002	0.215	$1.1005 \cdot 10^{-7}$	$0.0688 \cdot 10^{-7}$	0.0477
0.001	0.107	$0.5477 \cdot 10^{-7}$	$0.0342 \cdot 10^{-7}$	0.0118

Table 5. Expressions in $\mathcal{R}_\varepsilon$ for $\alpha = 0.01$, $t_2^* = 4.02$.

All nonsensitiveness regions $\mathcal{R}_\varepsilon$ are demonstrated in Figs. 3 and 4, where Fig. 3, Fig. 4 correspond to the above mentioned results in Tab. 4 and Tab. 5, respectively. In Figs. 3 and 4 we can see natural properties of the nonsensitiveness regions $\mathcal{R}_\varepsilon$ for the risk of the test:

- the lower given number ε, the lower $\mathcal{R}_\varepsilon$,
- the greater value of the parameter t, the lower $\mathcal{R}_\varepsilon$,
- the lower the value of the risk of the test α, the greater $\mathcal{R}_\varepsilon$.

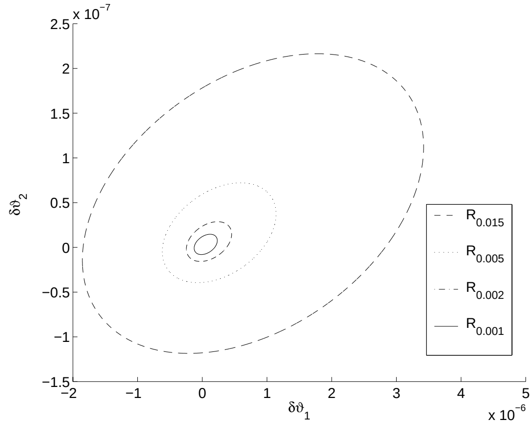

Figure 3. Nonsensitiveness regions $\mathcal{R}_\varepsilon$ for $\alpha = 0.05$ and $t^* = 1.99$.

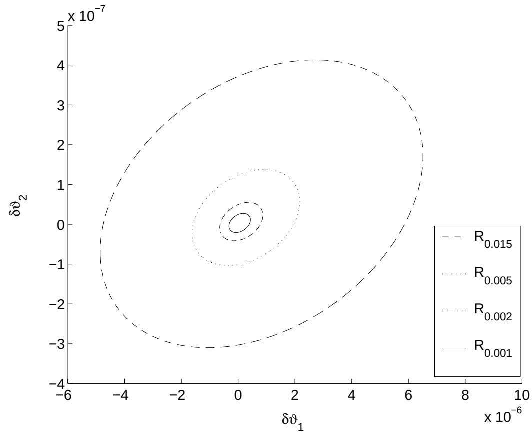

Figure 4. Nonsensitiveness regions $\mathcal{R}_\varepsilon$ for $\alpha = 0.01$ and $t^* = 4.02$.

References

[1] *J. P. Imhof*: Computing the distribution of quadratic forms in normal variables. Biometrika *48* (1961), 419–426.

[2] *J. Janko*: Statistical Tables. Academia, Praha, 1958. (In Czech.)

[3] *L. Kubáček*: Linear model with inaccurate variance components. Appl. Math. *41* (1996), 433–445.

[4] *C. R. Rao*: Statistical Inference and Its Applications. J. Wiley, New York-London-Sydney, 1965.

Author's address: *Eva Lešanská*, Department of Mathematical Analysis and Applied Mathematics, Faculty of Science, Palacký University, Tomkova 40, 779 00 Olomouc, Czech Republic, e-mail: `lesanske@aix.upol.cz`.