VERSITA

TATRA MOUNTAINS
Mathematical Publications
DOI: 10.2478/v10127-012-0015-y
Tatra Mt. Math. Publ. **51** (2012), 151–160

ON PARALLELISM OF REGRESSION LINES

František Rublík

ABSTRACT. Testing of parallelism of regression lines is discussed and the classical F-test is compared by means of simulations with a rank test computed by means of results of Sen's 1969 AMS paper. A rank based multiple comparisons rule for detecting regression lines with different slopes is constructed and asymptotic distribution of the underlying statistic is derived. This rule is compared by means of simulations with the rule derived from the assumption that the random fluctuations are Gaussian.

1. Introduction and the main results

Let $k > 1$ be a fixed integer and

$$Y_{ij} = \alpha_i + \beta_i X_{ij} + \varepsilon_{ij}, \qquad i = 1, \dots, k, \quad j = 1, \dots, n_i, \tag{1.1}$$

where i denotes the index of the regression line, X_{ij} are known constants, α_i, β_i are unknown regression constants, $\{\varepsilon_{ij};\, i = 1, \dots, k,\, j = 1, \dots, n_i\}$ are i.i.d random variables and the distribution function $F(z) = P(\varepsilon_{ij} < z)$ is continuous. The topic of this paper is inference on the null hypothesis

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_k, \tag{1.2}$$

proofs can be found in Section 2. Rank test statistics of this hypothesis and their asymptotic distribution were constructed in [8] under the following assumptions. Put

$$\overline{X}_{i.} = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}, \quad D_i^2 = \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_{i.})^2, \quad D^2 = \sum_{i=1}^{k} D_i^2, \quad \hat{\gamma}_i = \frac{D_i^2}{D^2}. \tag{1.3}$$

2010 Mathematics Subject Classification: Primary 62G10; Secondary 62J05.
Keywords: rank test statistic, equality of slopes of regression lines, hypotheses testing, multiple comparisons procedures, asymptotic distribution.
This research was supported by the VEGA grant No. 2/0038/12.

(A 1) For $i = 1, \dots, k$ the controlled variables $\mathbf{X_i} = (X_{i1}, X_{i2}, \dots, X_{in_i})$, $X_{i1} \le X_{i2} \le \dots \le X_{in_i}$, where at least one of the inequalities is strict,

$$\max_{j=1,\dots,n_i} \frac{(X_{ij} - \overline{X}_{i.})^2}{D_i^2} \to 0, \qquad i = 1, \dots, k$$

and

$$n_i \to \infty, \quad D_i^2 \to \infty, \quad \hat{\gamma}_i \to g_i > 0, \qquad i = 1, \dots, k. \tag{1.4}$$

(A 2) The distribution function F posseses a density $f(x) = \frac{dF(x)}{dx}$, the derivative $f'(x) = \frac{\mathrm{d}f(x)}{\mathrm{d}x}$ exists (with the possible exception of finitely many numbers x),

$$I(F) = \int_{-\infty}^{+\infty} \left(\frac{f'(x)}{f(x)} \right)^2 \mathrm{d}F(x) < +\infty,$$

$$\int_{-\infty}^{+\infty} \frac{f'(x)}{f(x)} \mathrm{d}F(x) = 0$$

and the quantity

$$\Psi(u) = -\frac{f'(F^{-1}(u))}{f(F^{-1}(u))}$$

fulfils the inequalities $0 < \int_0^1 \Psi(u) u \mathrm{d}u < +\infty$.

For $\mathbf{X_i}$, $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{in_i})$ and $b \in \mathbb{R}^1$ let

$$T_i(\mathbf{Y}_i - b\mathbf{X_i}) = \frac{\sqrt{12}}{D_i} \sum_{j=1}^{n_i} \left(X_{ij} - \overline{X}_{i.} \right) \frac{R(Y_{ij} - bX_{ij})}{n_i + 1}, \tag{1.5}$$

where $R(Y_{ij} - bX_{ij})$, $j = 1, \dots, n_i$ are ranks of the vector $\mathbf{Y}_i - b\mathbf{X_i}$. Let $T^*(\mathbf{Y} - b\mathbf{X}) = \sum_{i=1}^k \frac{D_i}{D} T_i(\mathbf{Y_i} - b\mathbf{X_i})$, $\beta^*_{(1)} = \sup\{b;\, T^*(\mathbf{Y} - b\mathbf{X}) > 0\}$, $\beta^*_{(2)} = \inf\{b;\, T^*(\mathbf{Y} - b\mathbf{X}) < 0\}$. It follows from Theorem 2.1 that this star quantities are well defined and are real numbers. To construct a test of the hypothesis (1.2), Sen used in [8] the pooled estimate of the slope under the validity of (1.2), defined by the equality $\beta^* = \frac{\beta^*_{(1)} + \beta^*_{(2)}}{2}$. Let

$$\hat{T}_i = T_i(\mathbf{Y}_i - \beta^* \mathbf{X_i}). \tag{1.6}$$

If (1.2) holds, then according of [8, Theorem 3.2] the statistic $Q = \sum_{i=1}^k \hat{T}_i^2$ converges in distribution to chi-square distribution with $k-1$ degrees of freedom provided that the assumptions (A 1), (A 2) are fulfilled. Hence the test rejecting (1.2) if $Q > \chi^2_{k-1}(1-\alpha)$, where $\chi^2_{k-1}(1-\alpha)$ denotes the $(1-\alpha)$th quantile of chi--square distribution with $k-1$ degrees of freedom, is test of H_0 at the asymptotic significance level α. The statistics (1.6), and consequently the statistic Q, can be easily computed by means of the following Lemma.

Lemma 1.1. *Let*

$$S_i = \left\{ \frac{Y_{ij_1} - Y_{ij_2}}{X_{ij_1} - X_{ij_2}};\, 1 \le j_1 < j_2 \le n_i,\, X_{ij_1} < X_{ij_2} \right\}, \qquad S = \bigcup_{i=1}^{k} S_i \tag{1.7}$$

and $S = \{s_1 < s_2 < \cdots < s_{N^*}\}$ denote the ordering of the values of this set according to the magnitude. Put

$$j_h = \min\bigl\{j;\, T^*(\mathbf{Y} - s_j\mathbf{X}) < 0\bigr\}, \qquad j_d = \max\bigl\{j;\, T^*(\mathbf{Y} - s_j\mathbf{X}) > 0\bigr\}. \tag{1.8}$$

Then $s_{j_h} = \beta^*_{(2)}$, $s_{j_d+1} = \beta^*_{(1)}$ and therefore the statistic (1.6)

$$\hat{T}_i = T_i\left(\mathbf{Y}_i - \frac{(s_{j_d+1} + s_{j_h})}{2}\mathbf{X_i}\right). \tag{1.9}$$

If the test rejects the null hypothesis, then usually the next step is to find out which of the regression lines have different slopes. For this purpose we construct the following theorem and the multiple comparisons rule (1.12).

Theorem 1.1. *Put (cf.* (1.9), (1.3)*)*

$$M_i = \frac{1}{\sqrt{\hat{\gamma}_i}}\hat{T}_i = \sqrt{12\frac{D^2}{D_i^4}}\sum_{j=1}^{n_i}\bigl(X_{ij} - \overline{X}_i\bigr)\frac{R(Y_{ij} - \beta^* X_{ij})}{n_i + 1}, \tag{1.10}$$

$$D_{i_1 i_2} = \frac{M_{i_1} - M_{i_2}}{\sqrt{\frac{D^2}{D_{i_1}^2} + \frac{D^2}{D_{i_2}^2}}}\sqrt{2}\,. \tag{1.11}$$

Let $t(k, 1-\alpha)$ denote the $(1-\alpha)$th quantile of the maximum modulus of $N_k(\mathbf{0}, \mathbf{I_k})$ distribution, i.e., $P\bigl(\max_{i,j}|y_i - y_j| \le t(k, 1-\alpha)\,|\,N_k(\mathbf{0}, \mathbf{I_k})\bigr) = 1 - \alpha$. If the assumptions (A 1), (A 2) are fulfilled and (1.2) *holds then $P\bigl(\max_{i_1,i_2}|D_{i_1 i_2}| > t(k, 1-\alpha)\bigr) \longrightarrow \beta \le \alpha$. This inequality holds with the equality sign if for the limits* (1.4) *the equality $g_1 = \cdots = g_k$ holds.*

In accordance with this theorem we declare the regression coefficients β_{i_1}, β_{i_2} to be different if the quantity (1.11) fulfils the inequality

$$|D_{i_1 i_2}| > t(k, 1-\alpha). \tag{1.12}$$

The inference on H_0 can be carried out also by procedures derived from the assumption that F is the distribution function of $N(0, \sigma^2)$ distribution with unknown positive variance. In this setting

$$\hat{\beta}_i = \frac{\sum_{j=1}^{n_i}(X_{ij} - \overline{X}_{i.})Y_{ij}}{D_i^2}, \qquad \hat{\alpha}_i = \overline{Y}_{i.} - \hat{\beta}_i\overline{X}_{i.}$$

denote LS estimates of the coefficients of the ith regression line,

$$\hat{\sigma}^2 = \frac{1}{n-2k}\sum_{i=1}^{k}\left[\sum_{j=1}^{n_i}\big(Y_{ij}^2 - Y_{ij}(\hat{\alpha}_i + \hat{\beta}_i X_{ij})\big)\right], \qquad n = \sum_{i=1}^{k} n_i$$

denotes estimate of σ^2 and $\overline{\beta} = \sum_{i=1}^{k}\hat{\gamma}_i\hat{\beta}_i$ is the estimate of the slope under the validity of H_0. If (1.2) holds then under these normality assumptions the statistic

$$Q_F = \frac{1}{(k-1)\hat{\sigma}^2}\sum_{i=1}^{k}\left[\hat{\beta}_i - \overline{\beta}\right]^2 D_i^2$$

has F distribution with $k-1$, $n-2k$ degrees of freedom (the proof can be found, e.g., in [3, pp. 285–290]). Hence under these normality assumptions the test rejecting (1.2) if $Q_F > F(k-1, n-2k, 1-\alpha)$, where $F(k-1, n-2k, 1-\alpha)$ denotes the $(1-\alpha)$th quantile of F distribution with $k-1$, $n-2k$ degrees of freedom, is test of H_0 at the significance level α. The regression coefficients β_{i_1}, β_{i_2} are declared to be different if

$$\frac{|\hat{\beta}_{i_1} - \hat{\beta}_{i_2}|}{\sqrt{\frac{1}{D_{i_1}^2} + \frac{1}{D_{i_2}^2}}} > t\sqrt{\hat{\sigma}^2}. \tag{1.13}$$

If

$$t = \sqrt{(k-1)F(k-1, n-2k, 1-\alpha)},$$

then the rule (1.13) can be derived by means of an application of the Scheffé theorem [1, p. 147]. If $t = t\big(n-2k, 1-\frac{\alpha}{k(k-1)}\big)$ is the $\big(1-\frac{\alpha}{k(k-1)}\big)$th quantile of the Student distribution with $n-2k$ degrees of freedom, then the rule (1.13) can be derived by means of an application of the Bonferroni inequality [1, p. 24].

Since in practice the distribution function F is usually not known, we shall illustrate the behaviour of the mentioned procedures by means of simulations for various types of the distribution. All simulation estimates in this paper are based on 5000 trials and because of the space limitations deal with the situation when $k=3$, $\beta_1 = \beta_2 = 2$ and β_3 has the values mentioned in the particular table. The caption *Normal distribution* means that the simulation estimates are computed for distribution function F of the $N(0,1)$ distribution, *Cauchy distribution* means that F is distribution function of the Cauchy $C(0,1)$ distribution and *Lognormal distribution* means that F is distribution function of the random variable $\exp(\xi)$, where ξ is $N(0,1)$ distributed.

In what follows P_{rejQ} denotes the simulation estimate of the probability $P\big(Q > \chi_{k-1}^2(1-\alpha)\big)$ of rejection of H_0 and P_{rejQ_F} is the simulation estimate of $P\big(Q_F > F(k-1, n-2k, 1-\alpha)\big)$. Since simulations suggest that for cases where $n_i < 5$ for some i the Q test is insensitive due to the fact that the set S_i from (1.7) contains few points, we shall deal with the setting

$$\begin{aligned} \mathbf{X_1} &= (1, 2, 4, 6, 9, 10), \\ \mathbf{X_2} &= (-3, -1, 0, 5, 6, 9), \\ \mathbf{X_3} &= (7, 11, 22, 23, 24, 25), \qquad n_1 = n_2 = n_3 = 6. \end{aligned} \tag{1.14}$$

All the simulations are carried out for $\alpha = 0.1$, the results for $\alpha = 0.05$ are similar. For values (1.14) of controlled variables we obtain the results:

	Normal distribution					
β_3	2	2.1	2.2	2.3	2.4	2.5
P_{rejQ_F}	0.100	0.211	0.513	0.824	0.963	0.997
P_{rejQ}	0.062	0.145	0.394	0.698	0.908	0.983

	Cauchy distribution							
β_3	2	2.1	2.2	2.3	2.4	2.5	2.6	2.8
P_{rejQ_F}	0.118	0.126	0.152	0.192	0.246	0.300	0.359	0.452
P_{rejQ}	0.059	0.088	0.145	0.224	0.319	0.408	0.495	0.624

	Lognormal distribution							
β_3	2	2.1	2.2	2.3	2.4	2.5	2.6	2.8
P_{rejQ_F}	0.104	0.150	0.314	0.504	0.677	0.779	0.854	0.929
P_{rejQ}	0.073	0.172	0.385	0.608	0.742	0.834	0.888	0.951

These simulations suggest that for symmetric distribution with light tails (like the normal distribution) the classical F-test is better than the test based on Q. In the case of symmetric heavy tailed distribution (like the Cauchy distribution) the situation is different. The power of the Q test is (with the exception of regions close to H_0 where the power is weak) in favor of the Sen statistic. The results of these simulations suggest that for asymmetric distributions (like the lognormal) the Q test is better than the F test.

In multiple comparisons methods constructed under normality assumptions for $k = 3$, $15 \le n \le 30$ the ratio

$$r(\alpha, n, k) = \frac{t(n-2k, 1-0.1/(k(k-1))}{\sqrt{(k-1)F(k-1, n-2k, 1-a)}}$$

attains the values $1.0234 \ge r(0.1, n, k) \ge 1.002$. Because of this the difference between the comparison based on the Scheffé theorem and that based on the Bonferroni inequality is for the considered sampling settings small and not influencing the efficiency with respect to the proposed nonparametric method. Therefore in the following simulation results we include only the method based on the Scheffé inequality. In these results $P_{Sch}(+)$ denotes the simulation estimate of the probability of the correct detection of regression lines with different slopes by multiple comparisons rule based on the Scheffé theorem, i.e., of the probability that the inequality (1.13) occurs for $t = \sqrt{(k-1)F(k-1, n-2k, 1-a)}$ and some i_1, i_2 such that $\beta_{i_1} \neq \beta_{i_2}$. Similarly $P_D(+)$ denotes simulation estimate of the probability that (1.12) occurs for some i_1, i_2 such that $\beta_{i_1} \neq \beta_{i_2}$. Since the probabilities $P_{Sch}(-)$, $P_D(-)$ of the false detection of different slopes attain negligible values, their simulation estimates are not included in the following tables.

For the controlled values (1.14) we obtain the results:

	Normal distribution					
β_3	2	2.1	2.2	2.3	2.4	2.5
$P_{Sch}(+)$	0	0.161	0.432	0.755	0.931	0.991
$P_D(+)$	0	0.088	0.271	0.554	0.808	0.936

	Cauchy distribution							
β_3	2	2.1	2.2	2.3	2.4	2.5	2.6	2.8
$P_{Sch}(+)$	0	0.095	0.117	0.149	0.198	0.252	0.303	0.402
$P_D(+)$	0	0.048	0.091	0.157	0.241	0.322	0.412	0.557

	Lognormal distribution							
β_3	2	2.1	2.2	2.3	2.4	2.5	2.6	2.8
$P_{Sch}(+)$	0	0.112	0.253	0.432	0.601	0.724	0.814	0.906
$P_D(+)$	0	0.106	0.290	0.497	0.661	0.772	0.842	0.928

For the values of the controlled variables $\mathbf{X_1} = (1, 2, 4, 6, 9, 10, 15, 21, 23)$, $\mathbf{X_2} = (-3, -1, 0, 5, 6, 9, 13, 19, 23)$, and $\mathbf{X_3} = (7, 11, 22, 23, 24, 25, 29, 33, 39)$, i.e., $n_1 = n_2 = n_3 = 9$ we obtain the results:

	Normal distribution			
β_3	2	2.1	2.2	2.3
$P_{Sch}(+)$	0	0.516	0.970	1.000
$P_D(+)$	0	0.440	0.931	0.999

	Lognormal distribution				
β_3	2	2.1	2.2	2.3	2.4
$P_{Sch}(+)$	0	0.260	0.628	0.845	0.932
$P_D(+)$	0	0.462	0.799	0.926	0.965

	Cauchy distribution							
β_3	2	2.1	2.2	2.3	2.4	2.5	2.6	2.8
$P_{Sch}(+)$	0	0.114	0.185	0.287	0.378	0.464	0.528	0.622
$P_D(+)$	0	0.170	0.408	0.615	0.748	0.817	0.864	0.910

Thus in the case of normal distribution the rule based on the Scheffé theorem is better than the method based on (1.12), but for larger sample sizes the difference will diminish. In the case of symmetric heavy tailed distributions (Cauchy distribution) or in the case of asymmetric distributions (lognormal distribution) is mostly the difference between these methods either mild or in favor of the method based on (1.12).

2. Proofs

Theorem 6.1 of [8, p. 1681] can be reformulated in the following way.

Theorem 2.1. *Let $c_1 \le c_2 \le \cdots \le c_n$ be real numbers and $c_j < c_{j+1}$ for some j. Suppose that $Y_1, \dots, Y_n$ are independent random variables with continuous distribution functions. Let $\mathbf{R}(b) = \big(R_1(b), R_2(b), \dots, R_n(b)\big)$ denote midranks of the numbers $Z_i(b) = Y_i - bc_i$, $i = 1, \dots, n$. Put*

$$T_n(b) = \sum_{i=1}^{n} (c_i - \overline{c}) R_i(b), \qquad \overline{c} = \frac{1}{n} \sum_{i=1}^{n} c_i.$$

Let $b_1 \le b_2 \le \cdots \le b_{n^*}$ denote ordering of the arguments of the intersections $\big\{Z_i(b) \cap Z_j(b);\ c_i < c_j\big\}$ of these lines. Then $1 \le n^* \le \binom{n}{2}$ and with the notation $b_0 = -\infty$, $b_{n^*+1} = +\infty$ the following assertions are true with probability 1.

(I) *$b_1 < \cdots < b_{n^*}$ and for $s = 0, \dots, n^*$ on the interval (b_s, b_{s+1}) the vector $\mathbf{R}(b)$ is constant and a permutation of the set $\{1, \dots, n\}$.*

(II) *Let $T_n(b_s - 0)$, $T_n(b_s + 0)$ denote the limit from the left and from the right, respectively. Then $T_n(b_s - 0) \ge T_n(b_s) \ge T_n(b_s + 0)$, $s = 1, \dots, n^*$, $T_n(b) > 0$ for for $b < b_1$ and $T_n(b) < 0$ for $b > b_{n^*}$.*

(III) *$T_n(b)$ is a nondecreasing function of $b \in (-\infty, +\infty)$.*

P r o o f o f L e m m a 1.1. The Lemma easily follows from Theorem 2.1. □

Similarly as in [2], or [5] or in [6] (cf. also [7]), the following proof uses the Hayter theorem from [4].

P r o o f o f T h e o r e m 1.1. According to [8, formula (3.26)] for statistics (1.6) the equality

$$\hat{T}_i = \rho \sqrt{I(F) D_i^2} \big(\beta_i^* - \beta_i - (\beta^* - \beta_i)\big) + o_P(1) \tag{2.1}$$

holds. Here

$$\beta_i^* = \frac{\beta_{(1)}^{*(i)} + \beta_{(2)}^{*(i)}}{2}, \qquad \beta_{(1)}^{*(i)} = \sup\big\{b; T_i(\mathbf{Y} - b\mathbf{X_i}) > 0\big\},$$

$$\beta^{*(i)}_{(2)} = \inf\big\{b; T_i(\mathbf{Y} - b\mathbf{X_i}) < 0\big\} \quad \text{and} \quad \rho = \sqrt{12/I(F)}\int_0^1 \Psi(u)u du.$$

But

$$\left|\sqrt{D^2}\left(\beta^* - \sum_{i=1}^k \hat{\gamma}_i \beta_i^*\right)\right| = o_P(1)$$

by [8, Lemma 3.3], and since $0 < D_i < D$, we see that

$$\begin{aligned} D_i(\beta^* - \beta_i) &= D_i\left(\beta^* - \sum_{j=1}^k \hat{\gamma}_j \beta_j^*\right) + D_i \sum_{j=1}^k \hat{\gamma}_j(\beta_j^* - \beta_j) \\ &= o_P(1) + D_i \sum_{j=1}^k \hat{\gamma}_j(\beta_j^* - \beta_j), \end{aligned} \tag{2.2}$$

because $\sum_{j=1}^k \hat{\gamma}_j = 1$ and $\beta_1 = \cdots = \beta_k$. Put

$$\eta_i = \rho\sqrt{I(F)}D_i(\beta_i^* - \beta_i), \qquad \boldsymbol{\eta} = (\eta_1, \ldots, \eta_k)' \tag{2.3}$$

and use the equality $D_i\hat{\gamma}_j = \sqrt{\hat{\gamma}_j\hat{\gamma}_i}D_j$. This together with (2.2) yields

$$\begin{aligned} &\rho\sqrt{I(F)}D_i\big(\beta_i^* - \beta_i - (\beta^* - \beta_i)\big) \\ &= \rho\sqrt{I(F)}D_i(\beta_i^* - \beta_i) + o_P(1) - \rho\sqrt{I(F)}\sum_{j=1}^k D_i\hat{\gamma}_j(\beta_j^* - \beta_j) \\ &= \eta_i - \sum_{j=1}^k \sqrt{\hat{\gamma}_i\hat{\gamma}_j}\eta_j + o_P(1) \\ &= \eta_i - \sum_{j=1}^k \sqrt{g_i g_j}\eta_j + o_P(1), \end{aligned} \tag{2.4}$$

because according to [8, Lemma 3.4] the convergence in distribution $\boldsymbol{\eta} \to N_k(\mathbf{0}, \mathbf{I_k})$ holds and therefore $\boldsymbol{\eta} = \mathcal{O}_P(1)$ holds. But (2.1) and (2.4) imply that for $\hat{\mathbf{T}} = \big(\hat{T}_1, \ldots, \hat{T}_k\big)'$ the equality $\hat{\mathbf{T}} = (\mathbf{I_k} - \mathbf{\Gamma})\boldsymbol{\eta} + o_P(1)$ holds. Here

$$\mathbf{\Gamma} = \sqrt{\gamma}(\sqrt{\gamma})', \qquad \sqrt{\gamma} = (\sqrt{g_1}, \ldots, \sqrt{g_k})' \tag{2.5}$$

and $\mathbf{I_k} = \operatorname{diag}(1, \ldots, 1)$ is the $k \times k$ unit matrix. Since $\sum_{j=1}^k g_j = 1$, the matrix $\mathbf{I_k} - \mathbf{\Gamma}$ is idempotent and we see that

$$\hat{\mathbf{T}} \to N_k(\mathbf{0}, \mathbf{I_k} - \mathbf{\Gamma}) \tag{2.6}$$

in distribution.

Now use (1.10) and put $\mathbf{M} = (M_1, \ldots, M_k)'$. Employing (2.6) and (2.5) one can easily prove the convergence in distribution $\mathbf{M} \to N_k(\mathbf{0}, \mathbf{W})$, $\mathbf{W} = \operatorname{diag}(1/g_1, \ldots, 1/g_k) - \mathbf{1}\mathbf{1}'$. Hence the random vector $\mathbf{S} = (D_{ij}; 1 \le i < j \le k)$ converges in distribution to the normal $N_{(k-1)k/2}(\mathbf{0}, \mathbf{\Sigma})$ distribution, where for $1 \le i_1 < j_1 \le k$, $1 \le i_2 < j_2 \le k$ the asymptotic covariance is

$$2\operatorname{cov}(M_{i_1} - M_{j_1}, M_{i_2} - M_{j_2}) \Big/ \sqrt{\left(\frac{1}{g_{i_1}} + \frac{1}{g_{j_1}}\right)\left(\frac{1}{g_{i_2}} + \frac{1}{g_{j_2}}\right)}$$

and $\operatorname{cov}(M_{i_1} - M_{j_1}, M_{i_2} - M_{j_2})$ is the asymptotic covariance of $(M_{i_1} - M_{j_1}, M_{i_2} - M_{j_2})$. Hence after some computation one finds out that the asymptotic covariance of $\mathbf{S}$ coincides with the covariance matrix of $\mathbf{U} = (U_{ij}; 1 \le i < j \le k)$, where

$$U_{ij} = \sqrt{2}(\xi_i - \xi_j) \Big/ \sqrt{\frac{1}{g_i} + \frac{1}{g_j}} \quad \text{and} \quad \boldsymbol{\xi} = (\xi_1, \ldots, \xi_k)'$$

has normal distribution with mean $\mathbf{0}$ and covariance matrix $\operatorname{diag}\left(\frac{1}{g_1}, \ldots, \frac{1}{g_k}\right)$. Therefore

$$P\left(\max_{i<j} |D_{ij}| > t(k, 1-\alpha)\right) \longrightarrow \gamma = P\left(\max_{i<j} |U_{ij}| > t(k, 1-\alpha)\right). \tag{2.7}$$

But according to [4, Theorem] if z_i has normal $N(0, \sigma_i^2)$ distribution and $z_1, \ldots, z_m$ are independent, then for every real number t

$$P\left[\max_{i<j} \frac{|z_i - z_j|}{\sqrt{\sigma_i^2 + \sigma_j^2}} > t\right] \le P\left[\max_{i<j} \frac{|x_i - x_j|}{\sqrt{2}} > t \,\middle|\, \mathbf{x} \sim N_k(\mathbf{0}, \mathbf{I_k})\right]$$

and this inequality becomes an equality, if $\sigma_1^2 = \cdots = \sigma_m^2$. An application of this to (2.7) yields the assertion of the Theorem. $\square$

REFERENCES

[1] ANDĚL, J.: *Matematická statistika.* SNTL, Praha, 1978. (In Czech)

[2] CRITCHLOW, D. E.—FLIGNER, M. A.: *On distribution free multiple comparisons in the one-way analysis of variance*, Comm. Statist. Theory Methods **20** (1991), 127–139.

[3] GRAYBILL, F. A.: *Theory and Application of the Linear Model.* Duxbury Press, Belmont, 1976.

[4] HAYTER, A. J.: *A proof of the conjecture that the Tukey-Kramer multiple comparison procedure is conservative,* Ann. Math. Statist. **12** (1984), 61–75.

[5] RUBLÍK, F.: *The multisample version of the Lepage test,* Kybernetika **43** (2007), 279–306.

[6] RUBLÍK, F.: *A note on simultaneous confidence intervals for ratio of variances,* Comm. Statist. Theory Methods **39** (2010), 1038–1045.

[7] RUBLÍK, F.: *Neparametrické metódy*. Veda, Bratislava, 2011. (In Slovak)

[8] SEN, P. K.: *On a class of rank order tests for the parallelism of several regression lines,* Ann. Math. Stat. **40** (1969), 1668–1683.

Received October 20, 2011

Institute of Measurement Science
Slovak Academy of Sciences
Dúbravská cesta 9
SK–841-04 Bratislava
SLOVAKIA
E-mail: umerrubl@savba.sk