AUSTRIAN JOURNAL OF STATISTICS
Volume 40 (2011), Number 1 & 2, 147–156

Robust Estimation of the Correlation Coefficient: An Attempt of Survey

Georgy Shevlyakov and Pavel Smirnov
St. Petersburg State Polytechnic University

Abstract: Various groups of robust estimators of the correlation coefficient are introduced. The performance of most prospective estimators is examined at contaminated normal distributions both on small and large samples, and the best of the proposed robust estimators are revealed.

Keywords: Correlation, Robustness, Contaminated Normal Distributions.

1 Introduction

The aim of robust methods is to ensure high stability of statistical inference under the deviations from the assumed distribution model. Less attention is devoted in the literature to robust estimators of association and correlation as compared to robust estimators of location and scale (Huber, 1981; Hampel, Ronchetti, Rousseeuw, and Stahel, 1986; Maronna and Yohai, 2006). However, it is necessary to study these problems due to their widespread occurrence (estimation of the correlation and covariance matrices in regression and multivariate analysis, estimation of the correlation functions of stochastic processes, etc.), and also because of the instability of classical methods of estimation in the presence of outliers in the data.

Consider the problem of estimation of the correlation coefficient ρ between the random variables X and Y. Given the observed sample $(x_1, y_1), \dots, (x_n, y_n)$ of a bivariate random variable (X, Y), the classical estimator of the correlation coefficient ρ is given by the *sample correlation coefficient*

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2\right]^{1/2}}, \tag{1}$$

where $\bar{x} = n^{-1}\sum_{i=1}^{n} x_i$ and $\bar{y} = n^{-1}\sum_{i=1}^{n} y_i$ are the sample means.

On the one hand, the sample correlation coefficient r is a statistical counterpart of the correlation coefficient ρ. On the other hand, it is the maximum likelihood estimator of ρ for the bivariate normal distribution density

$$\mathcal{N}(x, y; \mu_1, \mu_2, \sigma_1, \sigma_2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)} \times \left[\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right]\right\}, \tag{2}$$

where the parameters μ_1 and μ_2 are the means, σ_1 and σ_2 are the standard deviations of the random variables X and Y, respectively (Kendall and Stuart, 1963).

To illustrate the necessity in robust counterparts of the sample correlation coefficient, consider Tukey's gross error model (Tukey, 1960) described by the mixture of normal densities ($0 \le \varepsilon < 0.5$)

$$f(x,y) = (1-\varepsilon)\mathcal{N}(x,y;0,0,1,1,\rho) + \varepsilon\mathcal{N}(x,y;0,0,k,k,\rho')\,, \tag{3}$$

where the first and the second summands generate "good" and "bad" data, respectively; $0 \le \varepsilon < 0.5$, $k > 1$, $\mathrm{sgn}(\rho') = -\mathrm{sgn}(\rho)$. In general, the characteristics of "bad" data, namely their component means, standard deviations and especially the correlation ρ' may significantly differ from their counterparts in the first summand.

Further, we are mostly interested in estimation of the correlation coefficient ρ of "good" data regarding "bad" data as outliers. In model (3), the sample correlation coefficient is strongly biased with regard to ρ so that the presence of outliers in the data can completely destroy the sample correlation coefficient of "good" data up to the change of its sign (Gnanadesikan and Kettenring, 1972; Devlin, Gnanadesikan, and Kettenring, 1975).

The paper pursues two main goals: first, to give a brief overview of various approaches to robust estimation of correlation; second, to exhibit theoretical and experimental results on the performance of a selected subset of robust estimators generated by those approaches.

The paper is organized as follows. In Section 2, we successively describe various groups of robust estimators of the correlation coefficient focusing on the precise results concerned with Huber's minimax approach to robust estimation. In Section 3, the Monte Carlo performance of several prospective robust estimators on small and large samples is presented. In Section 4, some conclusions are drawn.

2 Main Approaches to Robust Correlation

2.1 Robust Correlation via Direct Robust Counterparts of the Sample Correlation Coefficient

A natural approach to robustify the sample correlation coefficient is to replace the linear procedures of averaging by the corresponding nonlinear robust counterparts (Gnanadesikan and Kettenring, 1972)

$$r_\alpha(\psi) = \frac{\Sigma_\alpha \psi(x_i - \widehat{x})\psi(y_i - \widehat{y})}{\left(\Sigma_\alpha \psi^2(x_i - \widehat{x})\Sigma_\alpha \psi^2(y_i - \widehat{y})\right)^{1/2}}\,, \tag{4}$$

where $\widehat{x}$ and $\widehat{y}$ are some robust estimators of location, for instance, the sample medians $\mathrm{med}(x)$ and $\mathrm{med}(y)$; $\psi = \psi(z)$ is a monotonic function, say, Huber's bounded linear score function: $\psi(z,k) = \max[-k, \min(z,k)]$; Σ_α is a robust analogue of a sum.

The latter transformation is based on trimming the outer order statistics with subsequent summation of the remaining ones:

$$\Sigma_\alpha z_i = nT_\alpha(z) = n(n-2r)^{-1}\sum_{i=r+1}^{n-r} z_{(i)}\,, \qquad 0 \le \alpha \le 0.5\,,\ r = [\alpha(n-1)]\,,$$

where $[\cdot]$ stands for the integer part. For $\alpha = 0$, the operations of ordinary and of robust summation coincide: $\Sigma_0 = \Sigma$. The following versions of estimator (4)

$$r_\alpha = \frac{\Sigma_\alpha(x_i - \mathrm{med}(x))(y_i - \mathrm{med}(y))}{[\Sigma_\alpha(x_i - \mathrm{med}(x))^2 \Sigma_\alpha(y_i - \mathrm{med}(y))^2]^{1/2}}$$

with $\alpha = 0.1, 0.2$ were used in Gnanadesikan and Kettenring (1972); Devlin et al. (1975); Shevlyakov and Vilchevsky (2002). For $\alpha = 0.5$, $\widehat{x} = \mathrm{med}(x)$, $\widehat{y} = \mathrm{med}(y)$, $\psi(z) = z$, formula (4) yields the *correlation median estimator* (Pasman and Shevlyakov, 1987; Falk, 1998)

$$r_{0.5} = r_{\mathrm{COMED}} = \frac{\mathrm{med}((x - \mathrm{med}(x))(y - \mathrm{med}(y)))}{\mathrm{MAD}(x)\mathrm{MAD}(y)},$$

where $\mathrm{MAD}(z) = \mathrm{med}(|z - \mathrm{med}(z)|)$ stands for the median absolute deviation.

2.2 Robust Correlation via Nonparametric Measures

An estimation procedure can be endowed with robustness properties by using a rank statistics. The best known of them are the *quadrant (sign) correlation coefficient* (Blomqvist, 1950)

$$r_{\mathrm{Q}} = n^{-1} \sum_{i=1}^{n} \mathrm{sgn}(x_i - \mathrm{med}(x))\mathrm{sgn}(y_i - \mathrm{med}(y)), \quad (5)$$

that is the sample correlation coefficient between the signs of deviations from medians, and the *Spearman rank correlation coefficient* r_{S} that is the sample correlation coefficient between the observation ranks (Spearman, 1904). It is noteworthy that the aforementioned nonparametric estimators can be also regarded as the representatives of the class (4) of estimators with the specific choices of the class parameters: say, choosing $\alpha = 0$ and $\psi(z) = \mathrm{sgn}(z)$ in (4), we get the r_{Q}.

2.3 Robust Correlation via Robust Regression

The problem of estimation of the correlation coefficient is directly related to the linear regression problem of fitting the straight line of the conditional expectation (Kendall and Stuart, 1963).

$$\mathrm{E}(X \mid Y = y) = \mu_1 + \beta_1(y - \mu_2), \qquad \mathrm{E}(Y \mid X = x) = \mu_2 + \beta_2(x - \mu_1).$$

For the bivariate normal distribution (2), $\rho^2 = \beta_1\beta_2$. Hence, using robust estimators of slope, we arrive at the robust estimator of the form (Pasman and Shevlyakov, 1987)

$$r_{\mathrm{REG}} = \sqrt{\widehat{\beta}_1 \widehat{\beta}_2}. \quad (6)$$

For instance, we may use the least absolute values (LAV) estimators and the least median squares (LMS) estimators (Rousseeuw, 1984).

2.4 Robust Correlation via Robust Principal Variables

Consider the following identity for the correlation coefficient ρ (Gnanadesikan and Kettenring, 1972)

$$\rho = \frac{\mathrm{var}(U) - \mathrm{var}(V)}{\mathrm{var}(U) + \mathrm{var}(V)}\,, \tag{7}$$

where $U = (X/\sigma_1 + Y/\sigma_2)/\sqrt{2}$, $V = (X/\sigma_1 - Y/\sigma_2)/\sqrt{2}$ are the principal variables such that $\mathrm{cov}(U, V) = 0$, $\mathrm{var}(U) = 1 + \rho$, $\mathrm{var}(V) = 1 - \rho$.

Introducing a robust scale functional $S(X) : S(aX + b) = |a|S(X)$, we may write $S^2(\cdot)$ for a robust counterpart of variance. Then a robust counterpart for (7) is given by

$$\rho^*(X, Y) = \frac{S^2(U) - S^2(V)}{S^2(U) + S^2(V)}\,. \tag{8}$$

By substituting the sample robust estimates for S into (8), we obtain robust estimates for ρ (Gnanadesikan and Kettenring, 1972)

$$\widehat{\rho} = \frac{\widehat{S}^2(U) - \widehat{S}^2(V)}{\widehat{S}^2(U) + \widehat{S}^2(V)}\,. \tag{9}$$

One of the possibilities is to use Huber's M-estimators of scale for $S(X)$ implicitly defined by the equation $\int \chi(x/S(X))\,dF(x) = 0$, where χ is a score function (Huber, 1981).

The choice of the median absolute deviation $\widehat{S} = \mathrm{MAD}(x)$ in (9) (the corresponding score function is $\chi_{\mathrm{MAD}}(x) = \mathrm{sgn}(|x| - 1)$) yields a remarkable robust estimator called the *MAD correlation coefficient* (Pasman and Shevlyakov, 1987)

$$r_{\mathrm{MAD}} = \frac{\mathrm{MAD}^2(u) - \mathrm{MAD}^2(v)}{\mathrm{MAD}^2(u) + \mathrm{MAD}^2(v)}\,, \tag{10}$$

where u and v are the robust principal variables

$$u = \frac{x - \mathrm{med}(x)}{\sqrt{2}\,\mathrm{MAD}(x)} + \frac{y - \mathrm{med}(y)}{\sqrt{2}\,\mathrm{MAD}(y)}\,, \qquad v = \frac{x - \mathrm{med}(x)}{\sqrt{2}\,\mathrm{MAD}(x)} - \frac{y - \mathrm{med}(y)}{\sqrt{2}\,\mathrm{MAD}(y)}\,. \tag{11}$$

Choosing Huber's trimmed standard deviation estimators as $\widehat{S}$ (see Huber, 1981, p. 121), we obtain the *trimmed correlation coefficient*:

$$r_{\mathrm{TRIM}} = \frac{\sum_{i=n_1+1}^{n-n_2} u_{(i)}^2 - \sum_{i=n_1+1}^{n-n_2} v_{(i)}^2}{\sum_{i=n_1+1}^{n-n_2} u_{(i)}^2 + \sum_{i=n_1+1}^{n-n_2} v_{(i)}^2}\,, \tag{12}$$

where $u_{(i)}^2$ and $v_{(i)}^2$ are the ith order statistics of the squared robust principal variables; n_1 and n_2 are the numbers of trimmed observations.

The general formula (12) yields the following limit cases: (i) the sample correlation coefficient r with $n_1 = 0$, $n_2 = 0$ and with the classical estimators (the sample means for location and the standard deviations for scale) in its inner structure; (ii) the *median correlation coefficient* with $n_1 = n_2 = [(n-1)/2]$

$$r_{\mathrm{MED}} = \frac{\mathrm{med}^2(|u|) - \mathrm{med}^2(|v|)}{\mathrm{med}^2(|u|) + \mathrm{med}^2(|v|)}\,. \tag{13}$$

Note that the estimators r_{MAD} (10) and r_{MED} (13) are asymptotically equivalent.

The other possibilities are connected with the use in (9) of the highly efficient and robust estimators of scale S_n and Q_n proposed by Rousseeuw and Croux (1993):

$$S_n = c_S \mathrm{med}_i \mathrm{med}_j(|x_i - x_j|)\,, \qquad Q_n = c_Q\{|x_i - x_j|;\ i < j\}_{(k)}\,, \tag{14}$$

where $k = C_h^2$, $h = [n/2] + 1$, c_S, and c_Q are the constants chosen to provide consistency of estimation of the standard deviation of a normal distribution (Rousseeuw and Croux, 1993). The corresponding robust estimators of correlation are denoted by r_{Sn} and r_{Qn}.

2.5 Minimax Variance Robust Estimation of Correlation

The class of robust estimators of correlation (9) based on robust principal variables (11) turned out to be one of most advantageous: Huber's minimax variance approach to robust estimation (Huber, 1981) is realized just in this class of estimators.

In Shevlyakov and Vilchevsky (2002) it is shown that the trimmed correlation coefficient r_{TRIM} (12) is asymptotically minimax with respect to variance for ε-contaminated bivariate normal distributions

$$f(x,y) \geq (1-\varepsilon)\,\mathcal{N}(x,y;0,0,1,1,\rho)\,, \qquad 0 \leq \varepsilon < 1\,. \tag{15}$$

This result holds under the underlying independent component distribution densities with unknown but equal variances (the parameters of location of the random variables X and Y are assumed known)

$$f(x,y) = \frac{1}{\sigma\sqrt{1+\rho}}\, g\left(\frac{u}{\sigma\sqrt{1+\rho}}\right) \frac{1}{\sigma\sqrt{1-\rho}}\, g\left(\frac{v}{\sigma\sqrt{1-\rho}}\right)\,, \tag{16}$$

where σ is the standard deviation; ρ is the correlation coefficient, u and v are the principal variables $u = (x+y)/\sqrt{2}$, $v = (x-y)/\sqrt{2}$; $g(x)$ is a symmetric density belonging to a certain class $\mathcal{G}$.

The idea of introducing class (16) is quite plain: for any pair (X,Y), the transformation $U = X + Y$, $V = X - Y$ gives the uncorrelated random principal variables (U,V), actually independent for densities (16). Thus, estimation of their scales $S(U) = \sigma\sqrt{1+\rho}$ and $S(V) = \sigma\sqrt{1-\rho}$ solves the problem of estimation of correlation between X and Y with the use of the estimators of Subsection 2.4, since $\rho = [S(U)^2 - S(V)^2]/[S(U)^2 + S(V)^2]$. Thus, class (9) of estimators entirely corresponds to class (16) of distributions, and this allows to extend Huber's results on minimax M-estimators of location and scale to estimation of the correlation coefficient.

Given the sample $(x_1,y_1),\ldots,(x_n,y_n)$ from the distribution with density (16), the following estimation procedure is considered:

1. the initial data $\{x_i, y_i\}_1^n$ are transformed to their principal components $u_i = (x_i + y_i)/\sqrt{2}$, $v_i = (x_i - y_i)/\sqrt{2}$, $i = 1, \ldots, n$;
2. M-estimates of scale $\widehat{S}(U)$ and $\widehat{S}(V)$ are computed as the solutions to the equations $\sum \chi(u_i/\widehat{S}(U)) = 0$ and $\sum \chi(v_i/\widehat{S}(V)) = 0$, where $\chi(\cdot)$ is a score function;
3. the estimator $\widehat{\rho}$ is taken in the form of (9).

In Shevlyakov and Vilchevsky (2002) it is shown that under regularity conditions the estimator $\widehat{\rho}$ is consistent and asymptotically normal with the following variance

$$var(\widehat{\rho}) = \frac{2(1-\rho^2)^2}{n} V(\chi, g)\,, \qquad V(\chi, g) = \frac{\int \chi^2(x) g(x)\, dx}{\left(\int x\chi'(x) g(x)\, dx\right)^2}\,. \tag{17}$$

where $n^{-1}V(\chi, g)$ is the asymptotic variance of M-estimators of scale.

Formula (17) for asymptotic variance has two factors: the first depends only on ρ, the second $n^{-1}V(\chi, g)$ is the asymptotic variance of M-estimators of scale (Huber, 1981). Thus Huber's minimax variance estimators of scale in the gross error model (15) can be directly applied for the minimax variance estimation of the correlation coefficient giving the trimmed correlation coefficient r_{TRIM}.

The levels of trimming n_1 and n_2 of the minimax trimmed correlation coefficient r_{TRIM} depend on the contamination parameter ε: $n_1 = n_1(\varepsilon)$ and $n_2 = n_2(\varepsilon)$ (Shevlyakov and Vilchevsky, 2002). In particular, the minimax variance estimator r_{TRIM} takes the following limit forms:

- as $\varepsilon \to 1$, it tends to the median correlation coefficient r_{MED};
- if $\varepsilon = 0$, it is equivalent to the sample correlation coefficient r.

Thus, the trimmed correlation coefficient r_{TRIM} may be regarded as a correlation analog of the classical Huber's robust estimators of location and scale, namely, the trimmed mean and standard deviation.

2.6 Minimax Bias Robust Estimation of Correlation

Monte Carlo experiments (Gnanadesikan and Kettenring, 1972; Devlin et al., 1975; Pasman and Shevlyakov, 1987) show that estimator's bias under contamination seems to be even a more informative characteristic of robustness than estimator's variance.

To the best of our knowledge, the first result in minimax bias robust estimation of the correlation coefficient belongs to Huber (1981) in the class of estimators r_0 (4), the quadrant correlation coefficient r_{Q} (5) is asymptotically minimax with respect to bias at the mixture $F = (1-\varepsilon)G + \varepsilon H$ with G and H centrosymmetric in $\mathbb{R}^2$.

Consider again the class of bivariate independent component distributions (16) with the corresponding estimator (9). Interestingly, its asymptotic bias has the same structure as its asymptotic variance:

$$\mathrm{E}(\widehat{\rho}_n) - \rho = b_n + o\,(1/n) \qquad \text{with} \qquad b_n(\chi, g) = -\frac{2\rho(1-\rho^2)}{n} V(\chi, g)\,, \tag{18}$$

where $V(\chi, g)$ is given by formula (17). Hence, due to the structure of formula (18), the problem of minimax bias estimation of ρ is equivalent to the problem of minimax variance estimation of a scale parameter. Since the latter problem is solved in Huber (1981), we directly arrive at the aforementioned robust estimators, namely, the trimmed correlation coefficient(12) and its limit form, the median correlation coefficient(13). Thus, the median correlation coefficient is simultaneously the asymptotically minimax bias and minimax variance estimator of the correlation coefficient at ε-contaminated bivariate normal distributions(15).

Furthermore, by the same reasons, the median correlation coefficient is the most B- and V-robust estimator in the sense of Theorems 9 and 10 (Hampel et al., 1986, pp. 142-143).

2.7 Robust Correlation via the Rejection of Outliers

The preliminary rejection of outliers from the data with the subsequent application of a classical estimator (for example, the sample correlation coefficient) to the rest of the observations defines the two-stage group of robust estimators of correlation. Their variety wholly depends on the variety of the rules for detection and/or rejection of multivariate outliers based on using discriminant, component, factor analysis, canonical correlation analysis, projection pursuit, etc. (for instance, see Atkinson and Riani, 2000; Hawkins, 1980; Rousseeuw and Leroy, 1987).

Each robust procedure of estimation inherently possesses the rule for rejection of outliers (for example, see Huber, 1981; Hampel et al., 1986), and it may seem that then there is no need for any independent procedure for rejection, at least if to aim at estimation, and therefore no need for two-stage procedures of robust estimation. However, a rejection rule may be quite informal, for example, based on a prior knowledge about the nature of outliers, and, in this case, its use can improve the efficiency of estimation.

3 Performance Evaluation: Monte Carlo Study

In this section, we compare the Monte Carlo performance (50000 trials) of the most prospective robust estimators of the correlation coefficient from Subsection 2.5 and Subsection 2.6 at the bivariate normal distribution and at the ε-contaminated bivariate normal distribution (3) both on small ($n = 20$) and large samples ($n = 1000$).

Since the quadrant correlation coefficient r_{Q} (5) is the minimax bias estimator in the class (4) of estimators, it is also included in Monte Carlo study. To provide its unbiasedness at the bivariate normal distribution, we use the following transformation: $r^*_{\mathrm{Q}} = \sin(\frac{1}{2}\pi r_{\mathrm{Q}})$ (Kendall and Stuart, 1963).

The r_{MAD} and r_{MED} correlation coefficients are defined by formulas (10) and (13), respectively; the trimmed correlation coefficient r_{TRIM} (12) is taken with the following values of trimming parameters $n_1 = n_2 = [0.2n]$.

The best performance values of estimators' expectations and variances are boldfaced starred in Tables 1 and 2.

Table 1: Normal distribution: $\rho = 0.9$.

	$n = 20$		$n = 1000$	
	$\mathrm{E}(r)$	$n\mathrm{var}(r)$	$\mathrm{E}(r)$	$n\mathrm{var}(r)$
r	$\mathbf{0.895}^*$	$\mathbf{0.049}^*$	0.899	$\mathbf{0.036}^*$
r_Q^*	0.858	0.352	0.899	0.233
r_{TRIM}	0.873	0.123	0.899	0.069
r_{MAD}	0.852	0.292	0.899	0.101
r_{MED}	0.832	0.311	0.899	0.101
r_{Sn}	0.871	0.164	0.900	0.062
r_{Qn}	0.881	0.103	0.900	0.045

Table 2: Contaminated normal distribution: $\rho = 0.9$, $\rho' = -0.9$, $k = 10$.

	$n = 20$		$n = 1000$	
	$\mathrm{E}(r)$	$n\mathrm{var}(r)$	$\mathrm{E}(r)$	$n\mathrm{var}(r)$
r	−0.330	8.771	−0.747	1.435
r_Q^*	0.710	$\mathbf{0.084}^*$	0.779	0.649
r_{TRIM}	0.810	0.210	0.812	0.104
r_{MAD}	0.838	0.322	$\mathbf{0.887}^*$	0.124
r_{MED}	0.795	0.434	$\mathbf{0.887}^*$	0.125
r_{Sn}	$\mathbf{0.844}^*$	0.189	0.880	0.100
r_{Qn}	$\mathbf{0.844}^*$	0.191	0.874	$\mathbf{0.084}^*$

4 Concluding Remarks

Normal distribution. From Table1 it follows that

1) on small and large samples, the best is the sample correlation coefficient r both in bias and variance;
2) the r_{MAD} and r_{MED} estimators are close to each other in performance, however, the r_{MAD} is better in bias on small samples;
3) on large samples, estimator's biases can be neglected, but not their variances;
4) the best estimator among robust estimators is the r_{Qn}.

Contaminated normal distribution. From Table 2 it follows that

1) the sample correlation coefficient r is catastrophically bad under contamination;
2) on small samples, the best estimators are the r_{Sn} and r_{Qn} both with respect to bias and to variance;
3) on large samples, the r_{Sn} and r_{Qn} are superior in variance, but the MAD and median correlation coefficients are better in bias confirming their asymptotic minimax bias properties;
4) under heavy contamination, estimator's bias is a more informative characteristic than its variance.

The former Monte Carlo studies of robust estimators of correlation (Gnanadesikan and Kettenring, 1972; Devlin et al., 1975; Pasman and Shevlyakov, 1987; Shevlyakov and Vilchevsky, 2002) show that the estimators based on robust principal variables, namely, r_{TRIM}, r_{MAD}, and r_{MED}, generally dominate over all the other robust estimators introduced in Section 2 including direct robust counterparts of the sample correlation coefficient (r_{COMED} and $r_{0.2}$), nonparametric r_{Q} and r_{S}, regression and two-stage estimators. In our Monte Carlo study, we have obtained that the r_{Qn} estimator is better than the others, so we may conclude that it is generally the best over the initially chosen set of estimators. However, its computation is much more time consuming than of its competitors.

References

Atkinson, A., and Riani, M. (2000). *Robust Diagnostics Regression Aalysis*. New York: Springer.

Blomqvist, N. (1950). On a measure of dependence between two random variables. *The Annals of Mathematical Statistics*, *21*, 593-600.

Devlin, S. J., Gnanadesikan, R., and Kettenring, J. R. (1975). Robust estimation and outlier detection with correlation coefficient. *Biometrika*, *62*, 531-545.

Falk, M. (1998). A note on the correlation median for elliptical distributions. *Journal of Multivariate Analysis*, *67*, 306-317.

Gnanadesikan, R., and Kettenring, J. R. (1972). Robust estimates, residuals and outlier detection with multiresponse data. *Biometrics*, *28*, 81-124.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., and Stahel, W. A. (1986). *Robust Statistics. The Approach Based on Influence Functions*. New York: John Wiley.

Hawkins, D. M. (1980). *The Identification of Outliers*. London: Chapman & Hall.

Huber, P. J. (1981). *Robust Statistics*. New York: John Wiley.

Kendall, M. G., and Stuart, A. (1963). *The Advanced Theory of Statistics. Inference and Relationship*. London: Griffin.

Maronna, R., and Yohai, D. M. V. (2006). *Robust Statistics. Theory and Methods*. New York: John Wiley & Sons.

Pasman, V. R., and Shevlyakov, G. L. (1987). Robust methods of estimation of a correlation coefficient. *Automation and Remote Control*, *??*, 332-340.

Rousseeuw, P. J. (1984). Least median of squares regression. *Journal of the American Statistical Association*, *79*, 871-880.

Rousseeuw, P. J., and Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, *88*, 1273-1283.

Rousseeuw, P. J., and Leroy, A. M. (1987). *Robust regression and outlier detection*. New York: John Wiley & Sons.

Shevlyakov, G. L., and Vilchevsky, N. O. (2002). Minimax variance estimation of a correlation coefficient for epsilon-contaminated bivariate normal distributions. *Statistics and Probability Letters*, *57*, 91-100.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, *15*, 88-93.

Tukey, J. W. (1960). A survey of sampling from contaminated distributions. In I. Olkin

(Ed.), *Contributions to Probability and Statistics* (p. 448-485). Stanford: Stanford University Press.

Authors' address:

Georgy Shevlyakov and Pavel Smirnov
Department of Applied Mathematics
St. Petersburg State Polytechnic University
Polytechnicheskaya, 29
St. Petersburg 195251, Russia

E-Mail: georgy.shevlyakov@gmail.com