A New Statistical Tool: Scalar Score Function

Fabián, Zdeněk
2010

Institute of Computer Science
Academy of Sciences of the Czech Republic

A New Statistical Tool: Scalar Score Function

Zdeněk Fabián

Technical report No. 1076

September 2010

Pod Vodárenskou věží 2, 182 07 Prague 8, phone: +420 266 053 210, fax: +420 286 585 789,
e-mail:zdenek@cs.cas.cz

Institute of Computer Science
Academy of Sciences of the Czech Republic

A New Statistical Tool: Scalar Score Function[1]

Zdeněk Fabián

Technical report No. 1076

September 2010

Abstract:

The basic inference function of mathematical statistics, the score function, is a vector function. We have introduced the scalar score, a scalar inference function, which reflects main features of a continuous probability distribution and which is simple. Its simplicity makes it possible to introduce new relevant numerical characteristics of continuous distributions. The t-mean and score variance are descriptions of distributions without the drawbacks of the mean and variance, which may not exist even in cases of regular distributions. Their sample counterparts appear to be alternative descriptions of the observed data. The scalar score itself appears to be a new mathematical tool, which could be used in solving traditional statistical problems for models far from the normal one, skewed and heavy-tailed.

Keywords:
point estimation, score function, basic characteristics of distributions, description of data samples

[1]I am grateful to J. Antoch, M. Stehlík and I. Vajda for support and valuable comments. The research presented was supported by projects AV0Z10300504 and GACR 205/09/1079.

1 Introduction

In statistical estimation problems, estimates are obtained as the solutions to equations of the form

$$\Psi(\text{data}\theta) = 0$$

where $\theta = (\theta_1, ..., \theta_m)$ is a vector of parameters of interest and Ψ is an inference function, the meaning of which is to adapt the data to the assumed model.

A parametric model is a parametric family of distributions F_θ with probability densities $f(x;\theta)$. The commonly used inference function is the score function, the vector $\Psi = (\Psi_{\theta_1}, ..., \Psi_{\theta_m})$ of derivatives of $\log f(x;\theta)$ with respect to the components of θ. The data $x_1, ..., x_n$ are considered as realizations of random variables $X_1, ..., X_n$, independent and identically distributed according to F_{θ_0}, a member of the assumed model F_θ. The solution of the system of equations

$$\sum_{i=1}^{n} \Psi_{\theta_k}(x_i;\theta) = 0, \qquad k = 1, ..., m$$

is the maximum likelihood estimate $\hat{\theta}_{ML}$ of θ_0. Since $\hat{\theta}_{ML}$ has the lowest possible variance, the density $f(x;\hat{\theta}_{ML})$ is considered as the best result of the estimation process.[1]

Apart from the fact that $\hat{\theta}_{ML}$ is influenced by observations far from the bulk of the data (outliers: this problem is solved by robust statistics), this result has some other, usually not accented drawbacks:

i) Parameters of different models are, as a rule, of different nature (location, scale, shape, frailty). A comparison of the precision (variance) of estimates in differently parametrized models is a difficult task.

ii) Instead of $f(x;\hat{\theta}_{ML})$, a few numbers characterizing the data would be often more useful in further analysis. A desirable description of the data seems to be something like 'center' and 'radius', perhaps skewness of the data. The commonly used numerical characteristic of distributions are the mean $m = EX = \int x f(x)\ dx$, variance $m_2 = E(X - m)^2$ and higher central moments $m_k = E(X - EX)^k$; it seems that $\hat{m} = m(\hat{\theta}_{ML})$ and $\hat{m}_k = m_k(\hat{\theta}_{ML})$ should be the values we search for. However, such an approach is not used in statistical practice. Moments are often queer expressions containing special functions, and moments of heavy-tailed distributions (distributions whose probability densities are approaching zero too slowly) do not exist (the corresponding integrals are infinite).

iii) More complex statistical problems, such as estimation of the degree of similarity of random variables X and Y, are usually solved by using 'pure' data without adapting them to the assumed model. A well-known example is the measure of association of two random variables, the Pearson correlation coefficient $\rho(X,Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - m_X)(y_i - m_Y)$, which is not able to make clear which part of the dependence stems from the real dependence and which part stems from the properties of marginal distributions of X and Y.

The cause is that the score function is a vector function, suitable for estimation of parameters, but too complicated to afford useful proposals of sensible numeric characteristics of distributions and too complicated to be used in more complex problems.

In the paper, we describe a scalar inference function proposed by the author.[2–5] It reflects basic features of the model distribution, being simultaneously a simple scalar function. We outline new possibilities of description of probability distributions and data taken from them, and suggest possible use of the function for solution of traditional statistical problems, which appears to be particularly useful if the underlying distribution is skewed and/or heavy-tailed.

2 Scalar score

As a scalar score of random variable X with distribution G and differentiable density g positive on the whole real line $\mathbb{R}$ was identified, using lesson drawn from Hampel et.al.,[6] function describing the relative change of the density,

$$S_G(x) = -\frac{g'(x)}{g(x)}. \tag{1}$$

The reason is that if G is a location distribution (that is, if it is in the form $G(x-\mu)$ where $\mu \in \mathbb{R}$ is the location parameter, indicating position of the maximum of density), function

$$S_G(x-\mu) = \Psi_\mu(x-\mu)$$

equals the score function for μ. The solution of equation

$$S_G(x;\theta) = 0, \tag{2}$$

where $S_G(x;\theta)$ is a parametric form of (1), is the coordinate of the maximum of density (mode) even if G is not a location distribution.

However, (1) is not a suitable description of distributions with densities positive only on a part $\mathcal{X} \neq \mathbb{R}$ of the real line (that is, with support $\mathcal{X} \neq \mathbb{R}$),

since the maximum of the density can lie at the edge of the support and equation $S_G(x) = 0$ has no solution. Based on the old idea of Johnson,[7] we suggested[2] to view any random variable X with distribution F supported by $\mathcal{X} \neq \mathbb{R}$ as a transformed random variable $X = \eta^{-1}(Y)$, where $Y = \eta(X)$ has a 'prototype' distribution G with support $\mathbb{R}$. By using a suitable mapping $\eta : \mathcal{X} \to \mathbb{R}$, an interesting function of distribution $F(x) = G(\eta(x))$ was identified the transformed scalar score of the prototype,

$$T(x) = S_G(\eta(x)). \tag{3}$$

Function (3) can be expressed by means of the density f of F as[2]

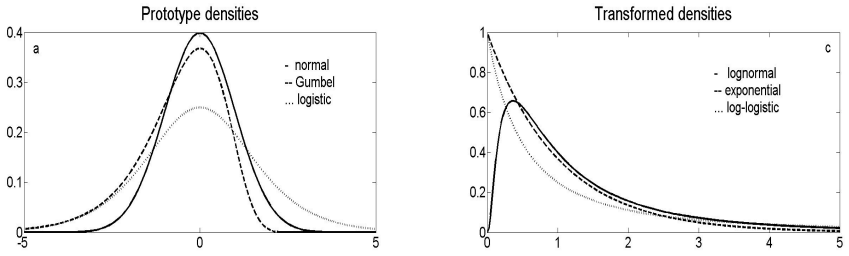
$$T(x) = -\frac{1}{f(x)} \frac{d}{dx} \left(\frac{1}{\eta'(x)} f(x) \right). \tag{4}$$

(4) expresses the relative change of a 'basic component of the density', which appeared to be the density divided by the Jacobian of the mapping used.

For comparison of properties of functions (4) of different distributions, it is necessary to use consistently one concrete mapping for a given support. We used that one providing simple mathematical forms of (3) for a large amount of commonly used distributions. According to Johnson,[7] we put

$$\eta(x) = \begin{cases} \log(x - a) & \text{if} \quad \mathcal{X} = (a, \infty) \\ \log \frac{x}{1-x} & \text{if} \quad \mathcal{X} = (0, 1) \end{cases} \tag{5}$$

(a generalization to common intervals is straightforward). The result, (4) with η given by (5), is called the *transformation-based score* or shortly the *t-score*. Scalar scores of some 'prototypes' and t-scores of the corresponding transformed distributions on $(0, \infty)$ are given in Fig. 1.



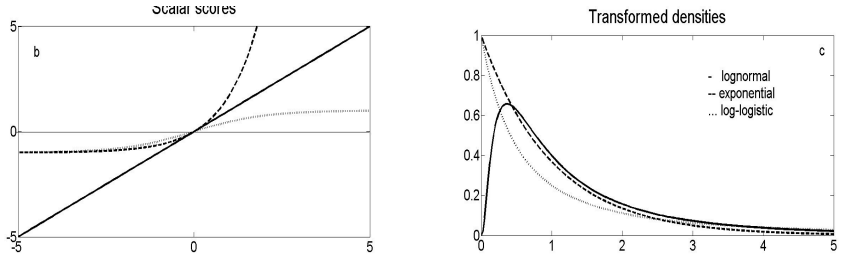


Figure 1|Transformed distributions. a. Densities and **b**. scalar scores of distributions with support $\mathbb{R}$. **c**. Densities and **d**. t-scores of corresponding transformed distributions with support $(0, \infty)$.

A location-type distribution $G(y-\mu)$ is mapped into $(0,\infty)$ as distribution

$$F(x;\tau) = G(\eta(x) - \eta(\tau)). \tag{6}$$

The parameter $\tau = \eta^{-1}(\mu)$ is the *transformed location parameter*,[8] the 'image' on $\mathcal{X}$ of the location of the prototype. It can be taken as a 'center' of $F(x;\tau)$, as it is, according to (6) and (2), the solution of equation $T(x;\tau) = 0$, where $T(x;\tau)$ is the corresponding t-score. It was proved[2] for these particular class of distributions that it holds true

$$\eta'(\tau)T(x;\tau) = \Psi_\tau(x;\tau). \tag{7}$$

We thus obtained a function which equals the score function for parameter expressing the position of distribution F on the x-axis. The score function Ψ_τ is decomposed into a product of two terms: an individual member characterizing the distribution (obtained by differentiating with respect to the variable) and the value of the Jacobian of the transformation at τ.

A generalization for arbitrary continuous distribution is the following[3]:

As a measure of central tendency of distribution $F(x;\theta)$ we suggest the zero of its t-score,

$$x^*(\theta): \qquad T(x;\theta) = 0,$$

called the *transformation-based mean* or shortly the *t-mean* (actually, the transformed mode of the prototype). Function (7) was generalized by using the t-mean instead of τ. The result,

$$S(x;\theta) = \eta'(x^*)T(x;\theta), \tag{8}$$

is called the *scalar score* (actually the 'score function for t-mean'). We suggest $S(x;\theta)$ as a scalar inference function of distribution F, whether

the t-mean is a parameter of the distribution or not. For particular class of distributions with location and transformed location parameter it is identical with the score function for this parameter. In other cases (such as the inverted gamma distribution, see Table 1 and Fig. 2) it is a new function.

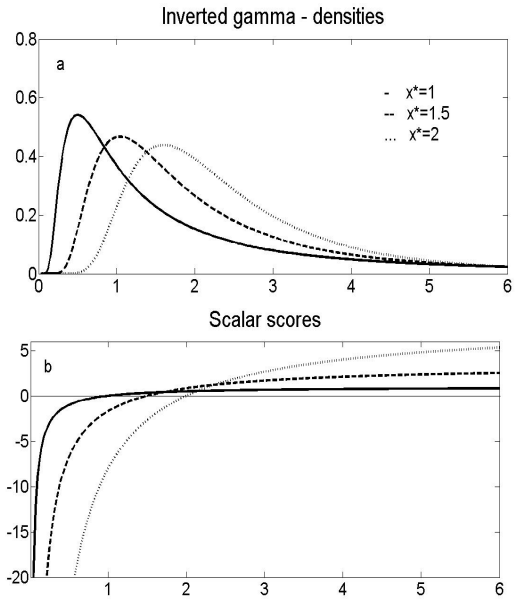


Figure 2| Numerical characteristics of distributions. a. Densities and **b**. scalar scores of inverted gamma distributions with parameters chosen in such a way that t-means are $x^* = 1, 1.5, 2$ and score variance $\omega^2 = 1$. All the plotted distributions have neither mean nor variance. It is apparent that t-mean characterizes a position of distributions on the $x-$axis and that all three distributions have a similar character of variability. Scalar scores of inverted gamma distributions are not sensitive to large values, but exhibit a large sensitivity to observations near zero.

3 Description of probability distributions

Besides the cumulative distribution function $F(x)$ and probability density $f(x)$, probability distribution F can be described by the scalar score $S(x) = \eta'(x^*)T(x)$. Since (7) is proportional to the influence function of the maximum likelihood estimator of τ, value $S(x)$ can be analogically understood as a relative influence of $x \in \mathcal{X}$ on an estimate of the t-mean.

If the probability density of prototype distribution goes to $\pm\infty$ slowly as e^{-x} (the distribution is heavy-tailed), S_G is bounded. Due to (3), scalar scores of heavy-tailed distributions with arbitrary support are bounded. This property implies the existence of score moments

$$M_k = ES^k(X) = \int_{\mathcal{X}} S^k(x) f(x)\, dx, \tag{9}$$

which can be used as numeric characteristics of distributions. It is easy to see that $M_1 = 0$ (scalar score is centered around the t-mean). The t-mean exists and is unique for distributions with unimodal prototypes (in other cases, some additional convention is necessary).

The value $M_2 = ES^2$ of location and transformed location distributions is, respectively, the Fisher information for the location and transformed location parameter. Accordingly, ES^2 is the Fisher information for the t-mean which could be taken as information of distribution F. The usual regularity conditions secure that it is finite. The reciprocal value

$$\omega^2 = \frac{1}{ES^2}, \tag{10}$$

appears to be a natural measure of the variability (dispersion) of the distribution even in cases in which the usual variance does not exist.[3] We call it a *score variance*. For distributions with support $(0, \infty)$, (10) turns into

$$\omega^2 = \frac{(x^*)^2}{ET^2}$$

. t-means and score variances of some distributions are given in Table 1.

Table 1|t-mean and score variance of some distributions

Distribution	$f(x)$	$T(x)$	x^*	ω^2
exponential	$\frac{1}{\tau}e^{-x/\tau}$	$\frac{x}{\tau}-1$	τ	τ^2
gamma	$\frac{\gamma^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\gamma x}$	$\gamma x-\alpha$	α/γ	α/γ^2
lognormal	$\frac{c}{\sqrt{2\pi x}}e^{-\frac{1}{2}\log^2(\frac{x}{\tau})^c}$	$c\log(\frac{x}{\tau})^c$	τ	τ^2/c^2
Pareto	c/x^{c+1}	$c-\frac{c+1}{x}$	$\frac{c+1}{c}$	$\frac{c+2}{c^3}$
Lomax	$\frac{\alpha}{(1+x)^{\alpha+1}}$	$\frac{\alpha x-1}{x+1}$	$\frac{1}{\alpha}$	$\frac{\alpha+2}{\alpha^3}$
inverted gamma	$\frac{\gamma^\alpha}{\Gamma(\alpha)}x^{-(\alpha+1)}e^{-\gamma/x}$	$\alpha-\gamma/x$	γ/α	γ^2/α^3

The exponential and lognormal are transformed location distributions. The last three distributions are heavy-tailed, having for certain range of parameters neither mean nor variance.

Let us remark that $\gamma = M_3/M_2^{3/2}$ characterizes skewness; a real skewness for distributions with support $\mathbb{R}$ and skewness with respect to the basic form for distributions supported by $(0,\infty)$ (which is itself skewed). M_4 characterizes flatness. An analog of Pearson's measure of kurtosis,[9] M_4/M_2^2, forms a logical structure precisely contrary to the ordinary kurtosis.

Function $S^2(x)$ attains its minimum at x^*, which is the least informative point of the distribution[10], and ES^2 describes information. $S^2(x)$ can thus be thought of an information function, expressing relative information contained in observation x.

In the sample space of distribution F, a distance between $x_1, x_2 \in \mathcal{X}$ can be introduced by relation

$$d(x_1,x_2) = |S(x_2) - S(x_1)|. \tag{11}$$

Parametric forms of (11) are in fact introduced when using maximum likelihood estimators in cases of location and transformed location distributions. It is apparent that the meaning of (11) is the distance of relative influences of x_1 and x_2. For standard normal and standard exponential distributions $S(x) = x$, so that in cases of normal end exponential models, (11) is linear. However, for skewed and/or heavy-tailed distributions it can be highly nonlinear. As an analogy, the temperature scale is linear. However, the distance, say, between $-272°$C and $-272.3°$C can be expressed in millions dollars, whereas the distance, say, between $1,000,000°$C and $1,200,000°$C is not very interesting. One could imagine that a nonlinear temperature scale would be more suitable. The distance between $x_1, x_2 \in \mathcal{X} = (-273.15^oC, \infty)$ could be measured by (11) with S being a 'scalar score function' derived by

means of (4) and (8) with f describing 'a rate of incidence of temperature data in man's life'. It could resemble, perhaps, the scalar score of the inverted gamma distribution (Fig. 2).

Writing (11) in the form $d(x_1, x_2) = \int_{x_1}^{x_2} dS(x)$, we obtain another significant function of distribution F, function $w(x) = dS(x)/dx$, which could be, perhaps, explained as the weight function, giving to any $x \in \mathcal{X}$ a relative weight of the observation in the assumed model. Basic functions describing two standard probability distributions are plotted in Fig. 3.

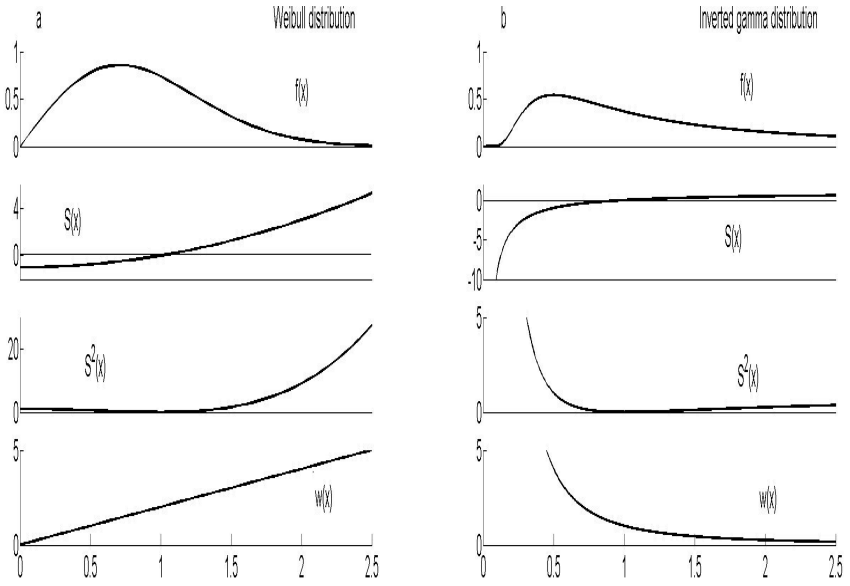


Figure 3| Functions describing distributions. Density, scalar score, information function and weight function of **a**. Weibull distribution ($c = 2$) and **b**. standard inverted gamma distribution.

4 Data characteristics

Given data $x_1, ..., x_n$ and a model family F_θ, new sample characteristics of central tendency ('center') and dispersion (square of 'radius') can be obtained as functions of estimated parameters: the sample t-mean $\hat{x}^*_{ML} = x^*(\hat{\theta}_{ML})$, the sample score variance $\hat{\omega}^2_{ML} = \omega^2(\hat{\theta}_{ML})$ (and, possibly, the sample coefficient of skewness $\hat{\gamma}_{ML} = \gamma(\hat{\theta}_{ML})$). They can be alternatively

used for the description of data samples, particularly of samples from skewed and/or heavy-tailed distributions, for which the usual characteristics cannot be used (Fig. 4, 5). By using them, it is easy to compare results of the estimation in different models.[4]

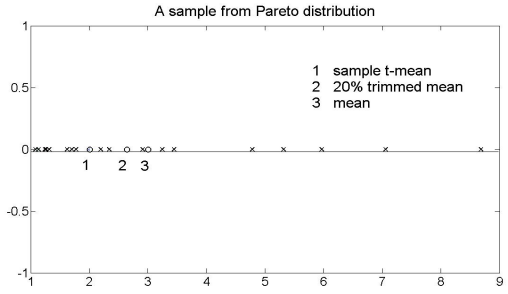


Figure 4| Typical value of a sample. A sample (the one with the small largest value) of length $n = 20$ was generated from Pareto distribution with $\alpha = 0.95$, i.e., from a distribution without mean. The sample mean (3) can be determined, as well as the standard deviation, but they are theoretically unjustified. The 20% trimmed mean (2) is known to be very robust estimator, however, it does not take into account properties of the distribution. The sample t-mean (1) characterizes the typical value of the sample.

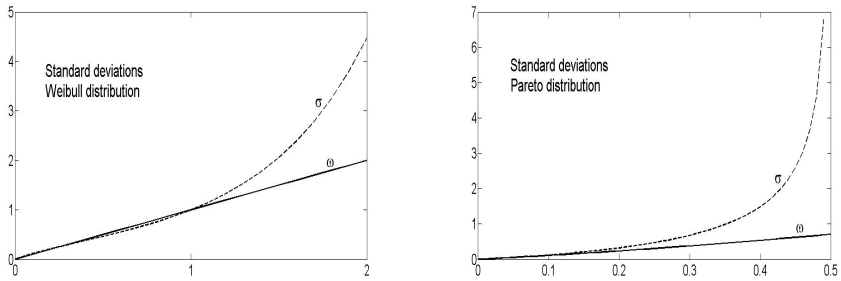


Figure 5| Standard deviations. σ and score standard deviation ω are shown as functions of the reciprocal value of parameter c of Weibull and Pareto distributions. σ of heavy-tailed Pareto distribution does not exist if $c \leq 2$. For large c, $\omega \simeq \sigma$. For c small, a description of variability of both distributions by the standard score deviation ω is the only sensible method we know.

Besides the maximum likelihood estimates, one can use the score moment estimates as the solution of equations

$$\hat{\theta}_{SM}: \qquad \frac{1}{n}\sum_{i=1}^{n} S^k(x_i;\theta) = \mathrm{E}_\theta S^k, \qquad k = 1, ..., m, \tag{12}$$

derived from (9) using the substitution principle. The score moments are often expressed by elementary functions of parameters.[11] Since the scalar scores of heavy-tailed distributions are bounded, a large observation x_i has no decisive influence on estimates since it enters into estimation equations by means of $S(x_i;\theta)$ only.

In some cases, the first equation of (12) has a form

$$\hat{x}^*_{SM}: \qquad \sum_{i=1}^{n} S(x_i;x^*) = 0.$$

Then the sample score variance of samples from distributions with support $(0,\infty)$ is given by

$$\hat{\omega}^2 = \frac{(\hat{x}^*_{SM})^2}{n^{-1}\sum_{i=1}^{n} T^2(x_i;\hat{x}^*_{SM})}.$$

Examples are given in Table 2. In a general case, however, the sample characteristics are to be determined as $\hat{x}^* = x^*(\hat{\theta})$ and $\hat{\omega}^2 = \omega^2(\hat{\theta})$, where $\hat{\theta}$ is either the maximum likelihood or the score moment estimate. The score moment estimates are generally not efficient (that is, they are asymptotically normal but their asymptotic variances are not lowest possible, contrary to the maximum likelihood estimates), but they are robust in cases of heavy-tailed models.[12]

Table 2|Sample t-mean and sample score variance of distributions from Table 1

Distribution	$\hat{x}^*$	$\hat{\omega}^2$
exponential	$\bar{x} = \frac{1}{n}\sum x_i$	$\bar{x}^2$
gamma	$\bar{x}$	$\frac{1}{n}\sum(x_i - \bar{x})^2$
lognormal	$\bar{x}_G = \frac{1}{n}\Pi x_i$	$\frac{1}{n}\sum \log^2 x_i/\bar{x}_G$
Pareto	$\bar{x}_H = n/\sum 1/x_i$	$(2\bar{x}_H - 1)(\bar{x}_H - 1)^2$
Lomax	$\bar{x}_L = \frac{\sum \frac{1}{x_i+1}}{\sum \frac{x_i}{x_i+1}}$	$\bar{x}_L^2(2\bar{x}_L + 1)$
inverted gamma	$\bar{x}_H$	$\frac{\bar{x}_H^2}{\frac{1}{n}\sum(1-\bar{x}_H/x_i)^2}$

The formulas for sample t-mean and sample score variance are determined from the score moment equations (12). $\bar{x}$ is the arithmetic mean, $\bar{x}_G$ geometric mean and $\bar{x}_H$ harmonic mean. The t-mean of the Lomax distribution (a simple member of the generalized logistic family[9]) is given by an original closed form. The sample score variance of one-parameter distributions is a function of the sample t-mean. For the sample score variance of two-parameter distributions in the table we found closed-form formulas. In a general case, both $\hat{x}^*$ and $\hat{\omega}^2$ are to be computed by an iterative way.

Confidence intervals for $\hat{x}^*_{SM}$ can be established by the modification of the Rao score test[1] or by the use of the distance $d(\hat{x}^*_{SM}, x_0)/\sqrt{ES^2}$, where d is given by (11).[5]

5 Other applications

Since the scalar score of a normal distribution is $S(x;\mu,\sigma) = (x-\mu)/\sigma^2$, the t-mean is the mean and score variance is the variance. If the model distribution is non-normal, scalar scores can be used for the solution of traditional statistical problems. We mention three of possible applications.

i) **Linear regression with non-normal residuals.**

Let $Y = \alpha_0 + \alpha_1 X + \varepsilon$, where ε is random variable with distribution F_ε. For the estimation of the coefficients α_0 and α_1, a criterion of minimal Fisher information of residual errors has been used in the form

$$\frac{1}{n}\sum_{i=1}^{n} S_\varepsilon^2(\varepsilon_i) = \min., \tag{13}$$

where S_ε is the scalar score of F_ε and $\varepsilon_i = y_i - (\alpha_0 + \alpha_1 x_i)$ are residual errors. For normally distributed residuals, (13) reduces to the least squares method. For skewed heavy-tailed F_ε we obtained results different from those provided by robust regression (Fig. 6).

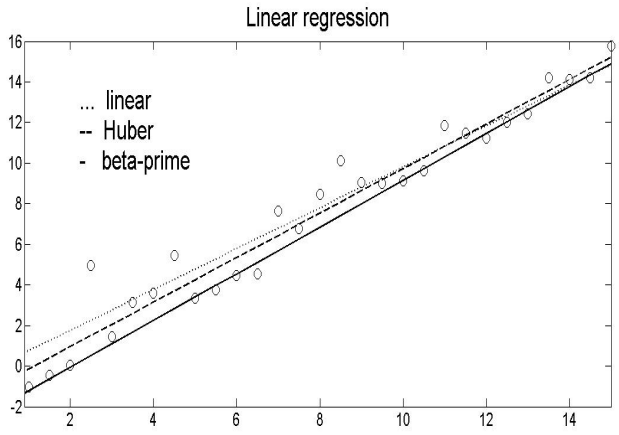


Figure 6| Linear regression. Comparison of least-squares regression, robust regression and score regression lines for data $y = -2.3 + 1.2x + \varepsilon$ where ε was generated as random variable with Lomax distribution. The score regression line is below the robust regression line since the scalar score takes into account the non-symmetry of the distribution (and, consequently, the generated data).

ii) **Distribution-dependent correlation coefficient.**

The score correlation coefficient of random variables X, Y with distributions F_X, F_Y and scalar scores S_X, S_Y, respectively, is

$$\rho_{score}(X,Y) = \frac{E(S_X, S_Y)}{(ES_X^2 ES_Y^2)^{1/2}}.$$

It holds that $-1 \leq \rho_S \leq 1$ and $\rho_S(X,Y) = 0$ for independent X and Y. Simulation experiments have shown that ρ_{score} can detect an association of random variables even if they have heavy-tailed distributions, similarly to the Spearman rank correlation coefficient.[13] This is illustrated in Fig. 7.

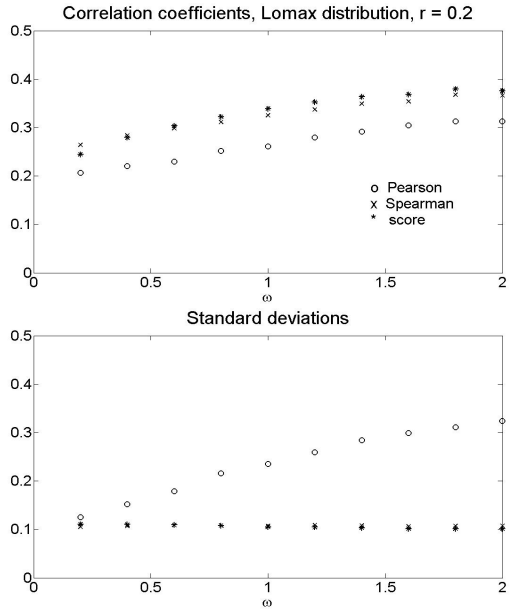


Figure 7| Correlation coefficient. Couples (X, Z) were generated as independent random samples of length 75 points from the Lomax distribution. The sample correlation coefficients between X and $Y = \alpha X + (1-\alpha)Z$ were estimated. The theoretical value of correlation is r=0.2. **a**. Average values of correlation coefficients Pearson (o), score (*) and Spearman (x) correlation coefficients, and **b**. their average standard deviations as functions of variability ω of the distribution. It is apparent from the lower plot that the usual Pearson correlation coefficient loses for heavy-tailed distribution any meaning. Estimates of the score correlation coefficient are in this particular case closed to the Spearman rank correlation coefficient and, although biased, can detect an association of random variables with heavy-tailed distributions.

iii) **Spectral properties of heavy-tailed processes.**

If the distribution of random process $\{X_t\}$ is heavy-tailed, random variable X_t does not possess finite variance and $\{X_t\}$ does not possess spectral density. Since the moments of random variable $S(X_t)$ are finite for any fixed t, one can study spectral densities of process $\{S(X_t)\}$. We have found that the score power spectrum of a process with small variability (small ω) is similar to that of the original process, whereas for processes with large variability it seems to be a sensible estimate of the spectral content of $\{X_t\}$,[14] as illustrated in Fig. 8.

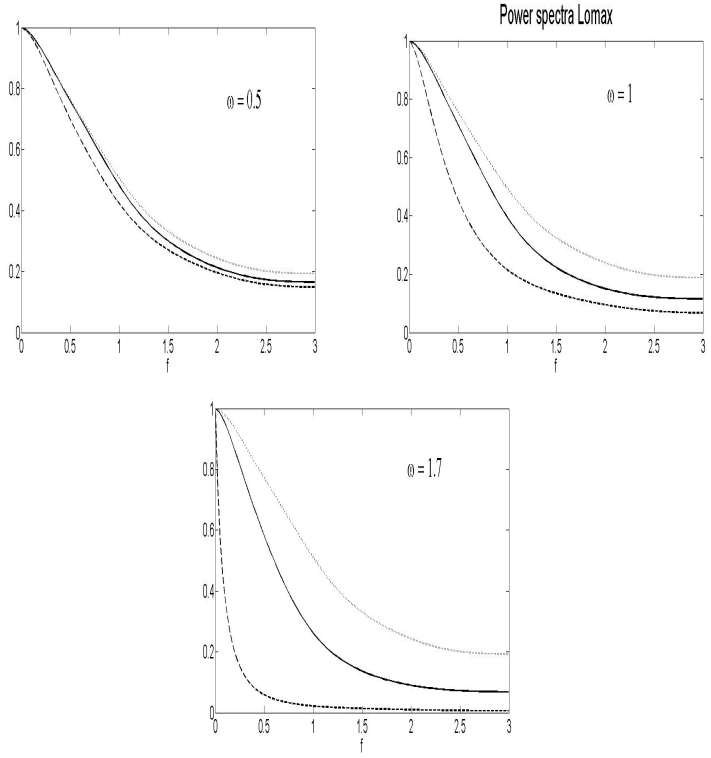


Figure 8.| Power spectra. Average power spectra of autoregressive signal $X_t = 0.4X_{t-1}+Z_t$, where Z_t is a white noise with Lomax distributions with three different ω. Dashed lines: spectra of $\log X_t$ (the used way of estimating spectra of positive signals), full lines: spectra of $S(X_t)$, where S is the scalar score of the Lomax distribution, dotted line: spectrum of X_t with standard normal Z_t. The estimated power spectra are similar for low ω. With increasing variability ω, the score spectra are still usable, whereas the log-spectra are largely distorted.

6 Conclusions

We describe the way to an introduction of a new inference function of mathematical statistics and show some of its advantages, the main of them is that it enables to introduce relevant characteristics of data samples taken from arbitrary continuous distribution. Similarly as in any parametric method, the results of processing real data are crucially dependent on the model adopted. The sample t-mean and sample score variance may now make it possible to compare the results of estimation in arbitrarily parametrized models.

Acknowledgements. I am grateful to J.Antoch, M.Stehlík and I.Vajda[†] for support and valuable comments. The research presented was supported by projects AV0Z10300504 and GACR 205/09/1079.

Bibliography.

1. Casella, G., Berger, R.L. (2002). *Statistical inference.* Duxbury.

2. Fabián, Z. (2001). Induced cores and their use in robust parametric estimation. *Comm. Statist. Theory Methods* 30, 537-556.

3. Fabián, Z. (2007). Estimation of simple characteristics of samples from skewed and heavy-tailed distribution, in Skiadas, C. (Ed.), *Recent Advances in Stochastic Modeling and Data Analysis.* World Scientific, Singapore, 43-50.

4. Fabián, Z. (2008). New measures of central tendency and variability of continuous distributions. *Comm. Statist. Theory Methods* 37, 159-174.

5. Fabián, Z. (2009). Confidence intervals for a new characteristic of central tendency of distributions. *Comm. Statist. Theory Methods* 38, 1804 - 1814.

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

7. Johnson, N.L. (1949). Systems of frequency curves generated by methods of translations. *Biometrika* 36, 149-176.

8. Marshall A. W., and Olkin I. (2007). *Life distributions. Structure of nonparametric, semiparametric and parametric families.* Springer.

9. Johnson, N. L., Kotz, S., Balakrishnan, N. (1994, 1995). *Continuous univariate distributions 1, 2.* Wiley.

10. Fabián, Z. (2009). The t-information and its use in multivariate problems and time series analysis. *J. Statist. Planning and Inference 139*, 3773-3778.

11. Fabián, Z., Stehlík, M. (2008). A note on favorable estimation when data is contaminated. *Comm. Dependability and Quality Management* 11, 36-43.

12. Fabián, Z. (2010). Score moment estimators. Proc. of COMPSTAT 2010.

13. Mori, D.D. and Kotz, S. (2001). *Correlation and dependence.* Imperial College press.

14. Fabián, Z. (2010). Spectral estimation of non-Gaussian time series. *Neural Network World*, 20, 491-499.