A new robust method of estimation with application to the normal distribution of order $p$

Benito V. Frosini
Università Cattolica del Sacro Cuore di Milano

1. An optimal summary of a unidimensional distribution

In a paper devoted to the application of optimal criteria for the determination of descriptive statistics, Frosini (1977) found out a remarkable convergence of the most common fitting criteria in producing the same summary distribution, given by a rectangular discrete distribution which assumes $k$ equiprobable values. Formally, given a c.d.f. (cumulative distribution function) $\Phi$ of a r.v. (random variable) $X$, the non-parametric problem consists in finding the best summary of $\Phi$ by means of a rectangular discrete c.d.f. $\Psi$ which assumes $k$ equiprobable values $\theta_i$ ($k$ integer $\geq 1$), such that

$$\Psi(x) = i / k \quad \text{for} \quad \theta_i \leq x < \theta_{i+1} \quad i = 0, 1, \ldots, k$$

Conventionally, $\theta_0 = -\infty$ and $\theta_{k+1} = +\infty$. In order to simplify the derivation of $\Psi$, it is convenient to assume that $\Phi$ be absolutely continuous with density $\varphi$, and that the support of $X$ be an interval $(a,b)$. The goodness-of-fit criteria considered in the sequel are actually among the most common in the literature about goodness-of-fit tests (see e.g. Frosini, 1977, 1978; D’Agostino and Stephens, 1986).

By assuming $\Phi$ completely known, and $\Psi$ defined as above, the minimizing criteria to be examined are the following (with the usual conventional definition of $\Psi^{-1}$: $\Psi^{-1}(u) = \inf \{ x : \Psi(x) \geq u \}$ for $0 < u < 1$):
\[ D_{1A} = \int |\Phi(x) - \Psi(x, \theta)| \, dx \]

\[ D_{1B} = \int \{\Phi(x) - \Psi(x, \theta)\}^2 \, dx \]

\[ D_{2A} = \int |\Phi(x) - \Psi(x, \theta)| \, d\Phi(x) \]

\[ D_{2B} = \int \{\Phi(x) - \Psi(x, \theta)\}^2 \, d\Phi(x) \]

\[ D_{3A} = \int |x - \Psi^{-1}(\Phi(x), \theta)| \, dx = \int_0^1 \Phi^{-1}(y) - \Psi^{-1}(y, \theta) \frac{1}{\varphi(x)} \, dy \]

\[ D_{3B} = \int \{x - \Psi^{-1}(\Phi(x), \theta)\}^2 \, dx = \int_0^1 \{\Phi^{-1}(y) - \Psi^{-1}(y, \theta)\}^2 \frac{1}{\varphi(x)} \, dy \]

\[ D_{4A} = \int |x - \Psi^{-1}(\Phi(x), \theta)| \, d\Phi(x) = \int_0^1 \Phi^{-1}(y) - \Psi^{-1}(y, \theta) \, dy = D_{1A} \]

(Gini’s simple index of dissimilarity; see Gini, 1918, 1965)

\[ D_{4B} = \int \{x - \Psi^{-1}(\Phi(x), \theta)\}^2 \, d\Phi(x) = \int_0^1 \{\Phi^{-1}(y) - \Psi^{-1}(y, \theta)\}^2 \, dy \neq D_{1B} \]

(square of Gini’s squared index of dissimilarity; see Gini, 1918, 1965)

\[ D_{5A} = \sup |\Phi(x) - \Psi(x, \theta)| \quad a < x < b \]

\[ D_{5B} = \sup \{\Phi(x) - \Psi(x, \theta)\}^2 \quad a < x < b \]

(minimizing \(D_{5B}\) yields the same distribution \(\Psi\) obtained by minimizing \(D_{5A}\)).

The remarkable result is that all the above criteria are minimized for the same \(\theta = (\theta_1, \ldots, \theta_k)\), namely for the equiprobable values \(\theta_i\) such that:

\[ \Phi(\theta_i) = \frac{i - 0.5}{k} \quad i = 1, \ldots, k, \text{ or} \]

\[ \theta_i = \Phi^{-1} \left( \frac{i - 0.5}{k} \right) \quad i = 1, \ldots, k. \]

For example, when \(k = 1\), from \(\Phi(\theta_1) = 1/2\) one obtains \(\theta_1 = \text{Me}(X)\) (median of \(X\)).

When \(k = 2\) the two values assumed by \(\Psi\) are: \(\theta_1 = x(1/4) = Q_1\) (first quartile) and \(\theta_2 = x(3/4) = Q_3\) (third quartile). When \(k = 3\): \(\theta_1 = x(1/6)\), \(\theta_2 = x(1/2)\) (median), \(\theta_3 = x(5/6)\)
etc. Thus formula (1) identifies the quantiles of a distribution $\Phi$, which are maximally informative about $\Phi$, as they are non-parametric descriptive statistics (for $k = 1, 2, ...$) satisfying all the most common fitting criteria (for $k = 1, 2, ...$).

2. A robust method of estimation based on the optimal summary of distributions

A non-parametric method of estimation, that maintains the properties of optimal summary, listed above, consists in equalizing the summary distribution $\Psi$ of a distribution $\Phi$ depending on $k$ parameters $\eta_1, ... , \eta_k$, and the summary distribution $G$ of an empirical (or sampling) distribution $F$, being $G$ defined on $k$ equiprobable values $\theta_1, ... , \theta_k$. Under identifiability assumptions, the following system in $\eta = (\eta_1, ... , \eta_k)$ is established:

$$\Phi^{-1}\left(\frac{i-0.5}{k}\right); \eta) = F^{-1}\left(\frac{i-0.5}{k}\right) \quad i = 1, ... , k \quad (2)$$

Putting $\theta_i = F^{-1}\left(\frac{i-0.5}{k}\right)$, the system can be rewritten

$$\Phi(\theta_i; \eta) = \left(\frac{i-0.5}{k}\right) \quad i = 1, ... , k, \quad (3)$$

and also, if direct reference to the density $\varphi$ turns out convenient:

$$\int_{\theta_i}^{\theta_{i+1}} \varphi(x; \eta) dx = \frac{1}{k} \quad i = 1, ... , k \quad (4)$$

or other equivalent systems.

The estimation method (2) – or an equivalent one – is robust, as it is based on the equalization of quantiles, theoretical (functions of $\eta = (\eta_1, ... , \eta_k)$) and observed (for robust L-estimators see Huber, 1981, pp. 55-61). The estimator $\hat{\eta} = (\hat{\eta}_1, ... , \hat{\eta}_k)$ is consistent. In fact, let us observe that:

(a) under identifiability assumptions about the parametric distribution $\Phi$, the estimator $\hat{\eta}$ satisfies the condition of Fisher-consistency (implying simple consistency under regularity conditions); namely, if $\theta_i^*$ is the quantile of order $((i-0.5)/k)$ of the c.d.f. $\Phi$ indexed by the parameter $\eta^* = (\eta_1^*, ... , \eta_k^*)$, the system
\[ \Phi^{-1}\left(\frac{i - 0.5}{k}\right) \eta = \theta_i^* \quad i = 1, \ldots, k \]  

(5)

yields the solution \( \eta = \eta^* \);

(b) the sampling quantile \( \theta_p \) (0 < \( p < 1 \)) is a consistent estimator (strong consistency) of the corresponding population quantile, with variance of order \( n^{-1} \) (Cramér, 1946, pp. 367-370; Serfling, 1980, pp. 74-85).

In the case of \( \Phi \) normal \( N(\mu, \sigma^2) \) (special case of \( k = 2 \)), as a simple application of the above system the parameters \( \mu \) and \( \sigma \) come out as the solutions of the system (with \( Q_1 \) and \( Q_3 \) first and third sampling quartiles, respectively):

\[
\begin{align*}
\mu - 0.6745 \sigma &= Q_1 \\
\mu + 0.6745 \sigma &= Q_3
\end{align*}
\]

hence the classical quantile-based estimates

\[
\begin{align*}
\tilde{\mu} &= \frac{(Q_1 + Q_3)}{2} \\
\tilde{\sigma} &= \frac{(Q_3 - Q_1)}{(2 \times 0.6745)}
\end{align*}
\]  

(6)

being 0.6745 the third quantile of the standard normal \( N(0,1) \).

When \( k \geq 3 \) odd, and \( \Phi \) symmetric for any admissible \( \eta \) (as in the case of the r.v. “normal of order \( p \)” treated in the sequel), the system (2) – or an equivalent system – cannot be maintained. The ensuing problem is already fully characterized for \( k = 3 \). Then, let us assume that the generic quantile of the r.v. \( X \) with c.d.f. \( \Phi \) is \( x(p ; \eta) \), whereas the generic quantile of the empirical c.d.f. \( F \) is \( y(p) \). When \( k = 3 \) the system (2) consists of the three equations

\[
\begin{align*}
x(1/6 ; \eta) &= y(1/6) \\
x(1/2 ; \eta) &= y(1/2) \\
x(5/6 ; \eta) &= y(5/6)
\end{align*}
\]

The differences between second and third equations, and third and second equations, give rise to the equations

\[
\begin{align*}
x(1/2 ; \eta) - x(1/6 ; \eta) &= y(1/2) - y(1/6) \\
x(5/6 ; \eta) - x(1/2 ; \eta) &= y(5/6) - y(1/2)
\end{align*}
\]

If \( \Phi \) is symmetric for every \( \eta \), the left hand sides of these equations coincide for every \( \eta \); if the right hand sides coincide too (an exceptional fact, owing to the sampling variability of the empirical \( F \)), the two equations are equivalent, thus reducing to a
single equation; otherwise, if the right hand sides do not coincide (which is the rule), the two equations are not compatible, hence no solution exists.

Therefore, in the case of $\Phi$ symmetric, and for $k \geq 3$ odd, the above fitting criteria must be applied as follows:

(I) As the summary distribution $\Psi$ is constituted by $k$ equiprobable values $\theta_1, \ldots, \theta_k$, and couples of these values share the same distance from the central value $\text{Me}(X) = \theta_{(k+1)/2}$, if $k \geq 3$ odd is the number of parameters $\eta_1, \ldots, \eta_k$ indexing the distribution $\Phi$, the number of values $\theta_i$ of the summary distribution $\Psi$ which allows $k$ functionally independent equations is $(2^k - 1)$ (hence $2 \times 3 - 1 = 5$ when $k = 3$; $2 \times 5 - 1 = 9$ when $k = 5$ etc.).

(II) In order that the ensuing systems of kind (2) – or equivalent ones – be compatible, also the empirical distribution $F$ must be symmetric around its median; however, being practically impossible that $F$ too be symmetric (in the presence of a symmetric $\Phi$ for the population), a constrained symmetrized $F^*$ must be preliminarily constructed. As we are only interested in a few quantiles of $F^*$, the simplest way of doing so consists in determining the quantiles $y^*(1-p)$ and $y^*(p)$ of $F^*$ such that, for $p > 1/2$:

$$y^*(p) - y^*(1-p) = y(p) - y(1-p)$$

allowing for the same distance of $y^*(p)$ and $y^*(1-p)$ from the median of $F$; this distance is given by

$$\Delta(p) = \frac{|y(p) - y(1-p)|}{2}.$$

The “corrected” quantiles, to be employed in the fitting procedure, will be therefore, for $p > 1/2$:

$$y^*(p) = \text{Me} + \Delta(p) ; \quad y^*(1-p) = \text{Me} - \Delta(p).$$

Although not really necessary, the same kind of demonstrations utilized by Frosini (1977) could be resumed and directly applied to the case of a symmetric distribution; however, the development of the minimization procedures come out in a more complex writing. Just as an example of these procedures, a detailed proof concerning the criterium $D_{1A}$ will be given in the case of three parameters, hence of five values for the fitted $\Psi$ (it is just the case of the normal distribution of order $p$). In the following formulae $\theta_1 = \eta - \varepsilon_1, \theta_2 = \eta - \varepsilon_2, \theta_3 = \eta$ etc.

$$D_{1A} = \int |\Phi(x) - \Psi(x,\theta)| \, dx$$
The simplest derivative concerns \( \eta = \theta_3 \):
\[
\frac{\partial D_{1A}}{\partial \eta} = |\Phi(\eta) - 2/5| - |\Phi(\eta) - 3/5|;
\]
by equating to zero, one obtains solutions if
\[
\Phi(\eta) - 2/5 = 3/5 - \Phi(\eta) \quad \text{hence} \quad \Phi(\eta) = 1/2
\]
namely \( \eta \) is the median of \( \Phi \). The derivative with respect to \( \varepsilon_1 \) is
\[
\frac{\partial D_{1A}}{\partial \varepsilon_1} = - \Phi(\eta - \varepsilon_1) + |\Phi(\eta - \varepsilon_1) - 1/5| + |\Phi(\eta + \varepsilon_1) - 4/5| - 1 + \Phi(\eta + \varepsilon_1).
\]
By equating to zero and simplifying, a solution exists if
\[
\Phi(\eta + \varepsilon_1) - \Phi(\eta - \varepsilon_1) = 0.8;
\]
owing to the symmetry of \( \Phi \) we get also
\[
\Phi(\eta - \varepsilon_1) = 0.1 \quad ; \quad \Phi(\eta + \varepsilon_1) = 0.9.
\]
Analogously, by equating to zero the derivative of \( D_{1A} \) with respect to \( \varepsilon_2 \) one obtains
\[
\Phi(\eta + \varepsilon_2) - \Phi(\eta - \varepsilon_2) = 0.4;
\]
and also, for the symmetry of \( \Phi \),
\[
\Phi(\eta - \varepsilon_2) = 0.3 \quad ; \quad \Phi(\eta + \varepsilon_2) = 0.7.
\]
The mixed second derivatives are equal to zero, while the second direct derivatives, evaluated at the solution point, are greater than zero – thus ensuring that the same solution is a minimum.

3. An application to the estimation of the parameters of a normal distribution of order \( p \)

3.1 Mixed ML-M (Maximum Likelihood – Moments) method of estimation

The normal distribution of order \( p \), first introduced by Subbotin (1923), has been resumed and extensively studied by distinguished statisticians of Palermo and Catania universities, starting from the fundamental contributions of Lunetta (1963) and Vianelli (1963). The most relevant studies of Antonino Mineo, concerning this family of
distributions, are the papers of 1980, 1983, 1989 and 2002, and the volume of 1978 (with Vianelli); the specific proposal of Mineo about the estimators of the three parameters of the distribution will be treated in the sequel. Other contributions on the estimation problem are those of Burgio (1978 a, b), Chiodi (1994, 2000a), Agrò (1992, 1995, 1999), and Angelo A. Mineo (1994, 1995, 2003). Chiodi has contributed to this research with a programme (in R) for the generation of random values from normal distributions of order \( p \) (reproduced in Appendix A). Some preceding contributions by Chiodi, on this and some related subjects, are contained in the papers of 1986 and 1995, and in the volume of 2000.

Given severe mathematical difficulties to work directly with the c.d.f. of this r.v., it is advisable the direct study of its density:

\[
\phi(x; \mu, \sigma, p) = \frac{1}{2 p^{1/p} \Gamma(1+1/ p)\sigma} \exp \left\{ \frac{-|x - \mu|^p}{p \sigma^p} \right\} \quad x \in R; \mu \in R; \sigma, p > 0. \tag{7}
\]

If \( X \) is a r.v. having this density defined on the real axis, \( X \) is called a normal distribution of order \( p \), and we can write, by definition, \( X \sim N_p(\mu, \sigma, p) \). In particular, for \( p = 1 \) we obtain the Laplace (or double exponential) distribution, while for \( p = 2 \) we get the normal distribution \( N(\mu, \sigma^2) \).

From A. Mineo (1983, p. 467), it is important to note that “the parameter \( p \) can be interpreted as a structural parameter ... This interpretation is confirmed by relevant kurtosis indices”; in fact, the ratio \( \beta_p \) between the central moment \( \mu_{2p} \) of order \( 2p \) and the square of the central moment \( \mu_p \) of order \( p \) depends only on \( p \), according to the simple relation

\[
\beta_p = \frac{\mu_{2p}}{\mu_p^2} = p + 1. \tag{8}
\]

For a useful comparison, the results achieved by means of the above robust criterion will be confronted with two other similar criteria, which exploit both Maximum Likelihood (ML) (for the parameters \( \mu \) and \( \sigma \)) and the method of Moments (M) (for the parameter \( p \)); the first of these criteria was proposed and implemented by A. Mineo in 1983. By equating to zero the derivatives of the log-likelihood function with respect to \( \mu \) and \( \sigma \) one obtains the estimation equation for \( \mu \)

\[
\sum_{i=1}^{n} |x_i - \hat{\mu}|^{p-1} \text{sgn}(x_i - \hat{\mu}) = 0 \tag{9}
\]

and the estimator of \( \sigma \):
\[ \hat{\sigma} = \left[ \frac{1}{n} \sum_{i=1}^{n} |x_i - \hat{\mu}|^p \right]^{1/p} \] (10)

A. Mineo (1983, p. 469) observes that “the maximum likelihood estimation of \( p \) appears rather difficult”; hence he suggests the simple criterion (8) applied to the corresponding sampling statistics (this same proposal was already present in the preceding Mineo paper of 1980, p. 569). Therefore the estimator of \( p \) becomes implicitly defined by the solution of the equation

\[ n \sum_{i=1}^{n} |x_i - \hat{\mu}|^{2p} - (p + 1) \left[ \sum_{i=1}^{n} |x_i - \hat{\mu}|^p \right]^2 = 0 \] (11)

which is based on the sampling central moments of order \( p \) and \( 2p \).

Another criterion of this kind has been suggested by A.M. Mineo (1994, 2003), who maintains the two ML equations (9) and (10), but resorts to the kurtosis index (cf Vianelli, 1963, p. 463)

\[ V = \frac{\sigma}{\mu_1} = \left[ \frac{\Gamma(1/p) \Gamma(3/p)}{\Gamma(2/p)} \right]^{1/2} \] (12)

where \( \sigma \) is the standard deviation and \( \mu_1 \) is the mean absolute deviation (from the mean); thus he proposes the estimation equation for \( p \)

\[ \hat{\mu} = \left[ \frac{n \sum_{i=1}^{n} (x_i - \hat{\mu})^2}{\sum_{i=1}^{n} |x_i - \hat{\mu}|} \right]^{1/2} = \left[ \frac{\Gamma(1/p) \Gamma(3/p)}{\Gamma(2/p)} \right]^{1/2} \] (13)

The estimation method based on formulae (9)-(10)-(11) will be called LM(I) (Maximum Likelihood + Moments of kind I), while the method which replaces (11) with (13) will be called LM(II).

### 3.2 Implementation of the general robust criterion

The application of the robust criterion, exposed at section 2, starts from the estimation of \( \mu \) by means of the sampling median \( \text{Me}(F) = \text{Me}(x_1, \ldots, x_n) \):

\[ \hat{\mu} = \text{Me}(F) \] (14)

The calculation of the estimates \( \hat{\sigma} \) and \( \hat{p} \) comes out from the system in \( \sigma \) and \( p \), consisting of the two equations
It is quite obvious that, in order to solve the equation (11) (method of moments) and the system (15)-(16), a suitable mathematical software, or a suitable programme, is needed (the one employed for carrying out the ensuing calculations has been initially Matlab; afterwards, practically all the calculations have been programmed in R).

While the estimation of $\mu$, given by the sampling median (formula (14)), is immediately obtained and is independent of the other parameters, the estimation of $\sigma$ and $p$ requires the numerical solution of the system (15)-(16). Although postponing detailed comments on calculations to sections 4 and 5, we can anticipate that the numerical convergence to the solutions $\tilde{\sigma}$ and $\tilde{p}$ of the system is comfortably fast; however, the dispersion of the estimator $\tilde{p}$ appears unpleasantly high for $p \geq 2$, and especially for $p \geq 3$; thus, some kind of intervention is required. A first tentative in this direction has been carried out, in order to disentangle the estimation of $p$ from the estimation of $\sigma$ (with some kind of analogy with respect to the procedures suggested by A. Mineo and A.M. Mineo, who employ kurtosis indices only dependent on $p$), although maintaining the same formulation of equating theoretical and empirical quantiles. This (apparently) new procedure directly exploits the property of $\sigma$ of being a scale parameter of the distribution; this implies that, passing from the r.v. $X \sim N(\mu,\sigma = 1, p)$ to the r.v. $Y \sim N_p(\mu,\sigma, p)$, with $\sigma > 0$, the generic quantile of order $q$ for $Y$ is simply derivable from the corresponding quantile of $X$ multiplied by $\sigma$:

$$y(q ; \mu,\sigma, p) = \sigma x(q ; \mu,1,p).$$

For fixed $q$-values $q_1, q_2 > 0.5$ ($q_1 = 0.7$ and $q_2 = 0.9$ with $\mu = 0$, according to the system (15)-(16)), the ratio between the quantiles of $Y$, when $\mu = 0$, reduces to

$$\frac{y(q_2 ; 0,\sigma, p)}{y(q_1 ; 0,\sigma, p)} = \frac{x(q_2 ; 0,1,p)}{x(q_1 ; 0,1,p)}$$

thus showing a dependence only on $p$.

By computing the quantiles $x_1 = x(q_1 ; 0,1,p)$ and $x_2 = x(q_2 ; 0,1,p)$ with the equations

\begin{align*}
&y^*(0.7) \\
&\int \varphi(x; \tilde{\mu},\sigma, p) \, dx = 0.4 \quad (15) \\
&y^*(0.9) \\
&\int \varphi(x; \tilde{\mu},\sigma, p) \, dx = 0.8 \quad (16)
\end{align*}
\[ \int_{-\infty}^{x_1} \phi(x;0,1,p) \, dx = q_1 \quad \text{and} \quad \int_{-\infty}^{x_2} \phi(x;0,1,p) \, dx = q_2 \]

and using a fine scanning for \( 0.5 \leq p \leq 5 \) (the range of \( p \)-values concretely interesting in applications), it is possible to obtain a numerical relation between \( p \) and the ratio

\[ R(p; q_1, q_2) = \frac{x(q_2;0,1,p)}{x(q_1;0,1,p)} \]  

(19)

Therefore, it is possible to derive an estimate of \( p \) by means of the inverse relation, which relates the ratio between sampling quantiles (expressed as differences with respect to the sampling median)

\[ R^*(q_1, q_2) = \frac{y^*(q_2) - Me(F)}{y^*(q_1) - Me(F)} \]  

(20)

to the corresponding \( p \)-value. The numerical relation between \( p \) and \( R(p; 0.7, 0.9) \) is shown in tabular form in Table 1, and in graphical form in Figure 1. It is thus possible to determine (estimate) the abscissa \( \hat{p} \) corresponding to an ordinate equal to \( R^*(q_1 = 0.7 ; q_2 = 0.9) \).

Table 1 – Theoretical relationship between \( p \) and \( R(p; 0.7, 0.9) \) (formula (19))

<table>
<thead>
<tr>
<th>( p )</th>
<th>( R^*(0.7,0.9) )</th>
<th>( p )</th>
<th>( R^*(0.7,0.9) )</th>
<th>( p )</th>
<th>( R^*(0.7,0.9) )</th>
<th>( p )</th>
<th>( R^*(0.7,0.9) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>4.7325</td>
<td>1.75</td>
<td>2.5405</td>
<td>3.00</td>
<td>2.2295</td>
<td>4.25</td>
<td>2.1180</td>
</tr>
<tr>
<td>0.55</td>
<td>4.4276</td>
<td>1.80</td>
<td>2.5188</td>
<td>3.05</td>
<td>2.2229</td>
<td>4.30</td>
<td>2.1152</td>
</tr>
<tr>
<td>0.60</td>
<td>4.1797</td>
<td>1.85</td>
<td>2.4984</td>
<td>3.10</td>
<td>2.2166</td>
<td>4.35</td>
<td>2.1124</td>
</tr>
<tr>
<td>0.65</td>
<td>3.9741</td>
<td>1.90</td>
<td>2.4792</td>
<td>3.15</td>
<td>2.2104</td>
<td>4.40</td>
<td>2.1097</td>
</tr>
<tr>
<td>0.70</td>
<td>3.8009</td>
<td>1.95</td>
<td>2.4610</td>
<td>3.20</td>
<td>2.2045</td>
<td>4.45</td>
<td>2.1072</td>
</tr>
<tr>
<td>0.75</td>
<td>3.6530</td>
<td>2.00</td>
<td>2.4438</td>
<td>3.25</td>
<td>2.1988</td>
<td>4.50</td>
<td>2.1046</td>
</tr>
<tr>
<td>0.80</td>
<td>3.5252</td>
<td>2.05</td>
<td>2.4276</td>
<td>3.30</td>
<td>2.1934</td>
<td>4.55</td>
<td>2.1022</td>
</tr>
<tr>
<td>0.85</td>
<td>3.4137</td>
<td>2.10</td>
<td>2.4121</td>
<td>3.35</td>
<td>2.1881</td>
<td>4.60</td>
<td>2.0998</td>
</tr>
<tr>
<td>0.90</td>
<td>3.3155</td>
<td>2.15</td>
<td>2.3975</td>
<td>3.40</td>
<td>2.1830</td>
<td>4.65</td>
<td>2.0975</td>
</tr>
<tr>
<td>0.95</td>
<td>3.2284</td>
<td>2.20</td>
<td>2.3835</td>
<td>3.45</td>
<td>2.1780</td>
<td>4.70</td>
<td>2.0953</td>
</tr>
<tr>
<td>1.00</td>
<td>3.1507</td>
<td>2.25</td>
<td>2.3703</td>
<td>3.50</td>
<td>2.1733</td>
<td>4.75</td>
<td>2.0931</td>
</tr>
<tr>
<td>1.05</td>
<td>3.0808</td>
<td>2.30</td>
<td>2.3577</td>
<td>3.55</td>
<td>2.1687</td>
<td>4.80</td>
<td>2.0910</td>
</tr>
<tr>
<td>1.10</td>
<td>3.0177</td>
<td>2.35</td>
<td>2.3456</td>
<td>3.60</td>
<td>2.1642</td>
<td>4.85</td>
<td>2.0889</td>
</tr>
<tr>
<td>1.15</td>
<td>2.9605</td>
<td>2.40</td>
<td>2.3342</td>
<td>3.65</td>
<td>2.1599</td>
<td>4.90</td>
<td>2.0869</td>
</tr>
<tr>
<td>1.20</td>
<td>2.9083</td>
<td>2.45</td>
<td>2.3232</td>
<td>3.70</td>
<td>2.1558</td>
<td>4.95</td>
<td>2.0850</td>
</tr>
<tr>
<td>1.25</td>
<td>2.8605</td>
<td>2.50</td>
<td>2.3128</td>
<td>3.75</td>
<td>2.1518</td>
<td>5.00</td>
<td>2.0831</td>
</tr>
<tr>
<td>1.30</td>
<td>2.8166</td>
<td>2.55</td>
<td>2.3028</td>
<td>3.80</td>
<td>2.1479</td>
<td>5.05</td>
<td>2.0807</td>
</tr>
<tr>
<td>1.35</td>
<td>2.7762</td>
<td>2.60</td>
<td>2.2932</td>
<td>3.85</td>
<td>2.1441</td>
<td>5.10</td>
<td>2.0785</td>
</tr>
<tr>
<td>1.40</td>
<td>2.7389</td>
<td>2.65</td>
<td>2.2840</td>
<td>3.90</td>
<td>2.1405</td>
<td>5.15</td>
<td>2.0744</td>
</tr>
<tr>
<td>1.45</td>
<td>2.7043</td>
<td>2.70</td>
<td>2.2753</td>
<td>3.95</td>
<td>2.1370</td>
<td>5.20</td>
<td>2.0704</td>
</tr>
<tr>
<td>1.50</td>
<td>2.6721</td>
<td>2.75</td>
<td>2.2668</td>
<td>4.00</td>
<td>2.1336</td>
<td>5.25</td>
<td>2.0663</td>
</tr>
<tr>
<td>1.55</td>
<td>2.6421</td>
<td>2.80</td>
<td>2.2588</td>
<td>4.05</td>
<td>2.1302</td>
<td>5.30</td>
<td>2.0602</td>
</tr>
<tr>
<td>1.60</td>
<td>2.6142</td>
<td>2.85</td>
<td>2.2510</td>
<td>4.10</td>
<td>2.1270</td>
<td>5.35</td>
<td>2.0502</td>
</tr>
<tr>
<td>1.65</td>
<td>2.5880</td>
<td>2.90</td>
<td>2.2436</td>
<td>4.15</td>
<td>2.1239</td>
<td>5.40</td>
<td>2.0402</td>
</tr>
<tr>
<td>1.70</td>
<td>2.5635</td>
<td>2.95</td>
<td>2.2364</td>
<td>4.20</td>
<td>2.1209</td>
<td>5.45</td>
<td>2.0302</td>
</tr>
</tbody>
</table>
The same quantiles utilized in formulae (18) and (19) lend themselves to an easy estimation of the scale parameter $\sigma$. It is enough to observe, from the simple relation (17), that the ratio between the difference of these quantiles and the difference of the corresponding quantiles for $\mu = 0$ and $\sigma = 1$ is equal to $\sigma$:

$$S = \frac{x(q_2;0,\sigma,p) - x(q_1;0,\sigma,p)}{x(q_2;0,1,p) - x(q_1;0,1,p)} = \sigma$$

(21)

An analogic estimate for $\sigma$ can thus be obtained from the ratio between the difference of empirical quantiles and the difference between theoretical quantiles, for $\mu = 0$ and $\sigma = 1$:

$$\bar{\sigma} = \frac{y^*(q_2) - y^*(q_1)}{x(q_2;0,1,\bar{p}) - x(q_1;0,1,\bar{p})}$$

(22)

In order to facilitate the computation of this estimate of $\sigma$, Table 2 reports the relation between $p$ and the quantile difference in the denominator of formula (22).

Just as an easy numerical example, let us assume that the calculations on the sample values lead to the following figures for the median and the relevant quantiles of order 0.7 and 0.9: $\text{Me}(F) = 10$, $y^*(0.7) = 14$, $y^*(0.9) = 19.24$. As $R^*(q_1,q_2) = 9.24/4 = 2.31$, a linear interpolation on the points (2.50,2.3128) and (2.55,2.3028) in Table 1 leads to an estimate $\hat{p} = 2.514$. In order to use this estimate for the difference $\Delta(0.7,0.9)$, we can make another linear interpolation on the points (2.50,0.6874) and (2.55,0.6819) in Table 2, and obtain – as the ordinate corresponding to the abscissa $p = 2.514$, $\Delta = 0.6859$; by applying formula (22) we obtain the estimate $\bar{\sigma} = 7.64$. Of
course, if a (slightly) better approximation would be required, a four-point interpolation on the relevant points in Tables 1 and 2 could be performed.

Table 2 – Theoretical relation between \( p \) and the quantile difference \( \Delta(0.7,0.9) = [x(0.9;0,1,p) - x(0.7;0,1,p)] \)

<table>
<thead>
<tr>
<th>( p )</th>
<th>( \Delta(0.7,0.9) )</th>
<th>( p )</th>
<th>( \Delta(0.7,0.9) )</th>
<th>( p )</th>
<th>( \Delta(0.7,0.9) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>1.7678</td>
<td>1.75</td>
<td>0.8067</td>
<td>3.00</td>
<td>0.6409</td>
</tr>
<tr>
<td>0.55</td>
<td>1.6455</td>
<td>1.80</td>
<td>0.7957</td>
<td>3.05</td>
<td>0.6371</td>
</tr>
<tr>
<td>0.60</td>
<td>1.5440</td>
<td>1.85</td>
<td>0.7853</td>
<td>3.10</td>
<td>0.6334</td>
</tr>
<tr>
<td>0.65</td>
<td>1.4584</td>
<td>1.90</td>
<td>0.7754</td>
<td>3.15</td>
<td>0.6298</td>
</tr>
<tr>
<td>0.70</td>
<td>1.3851</td>
<td>1.95</td>
<td>0.7661</td>
<td>3.20</td>
<td>0.6264</td>
</tr>
<tr>
<td>0.75</td>
<td>1.3216</td>
<td>2.00</td>
<td>0.7572</td>
<td>3.25</td>
<td>0.6230</td>
</tr>
<tr>
<td>0.80</td>
<td>1.2660</td>
<td>2.05</td>
<td>0.7487</td>
<td>3.30</td>
<td>0.6198</td>
</tr>
<tr>
<td>0.85</td>
<td>1.2169</td>
<td>2.10</td>
<td>0.7406</td>
<td>3.35</td>
<td>0.6167</td>
</tr>
<tr>
<td>0.90</td>
<td>1.1731</td>
<td>2.15</td>
<td>0.7329</td>
<td>3.40</td>
<td>0.6136</td>
</tr>
<tr>
<td>0.95</td>
<td>1.1340</td>
<td>2.20</td>
<td>0.7255</td>
<td>3.45</td>
<td>0.6107</td>
</tr>
<tr>
<td>1.00</td>
<td>1.0986</td>
<td>2.25</td>
<td>0.7184</td>
<td>3.50</td>
<td>0.6078</td>
</tr>
<tr>
<td>1.05</td>
<td>1.0666</td>
<td>2.30</td>
<td>0.7117</td>
<td>3.55</td>
<td>0.6050</td>
</tr>
<tr>
<td>1.10</td>
<td>1.0374</td>
<td>2.35</td>
<td>0.7052</td>
<td>3.60</td>
<td>0.6023</td>
</tr>
<tr>
<td>1.15</td>
<td>1.0106</td>
<td>2.40</td>
<td>0.6990</td>
<td>3.65</td>
<td>0.5997</td>
</tr>
<tr>
<td>1.20</td>
<td>0.9861</td>
<td>2.45</td>
<td>0.6931</td>
<td>3.70</td>
<td>0.5972</td>
</tr>
<tr>
<td>1.25</td>
<td>0.9635</td>
<td>2.50</td>
<td>0.6874</td>
<td>3.75</td>
<td>0.5947</td>
</tr>
<tr>
<td>1.30</td>
<td>0.9425</td>
<td>2.55</td>
<td>0.6819</td>
<td>3.80</td>
<td>0.5923</td>
</tr>
<tr>
<td>1.35</td>
<td>0.9231</td>
<td>2.60</td>
<td>0.6767</td>
<td>3.85</td>
<td>0.5899</td>
</tr>
<tr>
<td>1.40</td>
<td>0.9050</td>
<td>2.65</td>
<td>0.6716</td>
<td>3.90</td>
<td>0.5876</td>
</tr>
<tr>
<td>1.45</td>
<td>0.8881</td>
<td>2.70</td>
<td>0.6667</td>
<td>3.95</td>
<td>0.5854</td>
</tr>
<tr>
<td>1.50</td>
<td>0.8724</td>
<td>2.75</td>
<td>0.6620</td>
<td>4.00</td>
<td>0.5833</td>
</tr>
<tr>
<td>1.55</td>
<td>0.8576</td>
<td>2.80</td>
<td>0.6575</td>
<td>4.05</td>
<td>0.5811</td>
</tr>
<tr>
<td>1.60</td>
<td>0.8437</td>
<td>2.85</td>
<td>0.6531</td>
<td>4.10</td>
<td>0.5791</td>
</tr>
<tr>
<td>1.65</td>
<td>0.8307</td>
<td>2.90</td>
<td>0.6489</td>
<td>4.15</td>
<td>0.5771</td>
</tr>
<tr>
<td>1.70</td>
<td>0.8183</td>
<td>2.95</td>
<td>0.6448</td>
<td>4.20</td>
<td>0.5751</td>
</tr>
</tbody>
</table>

Although the numerical computation of the above two criteria, namely \( S (= \text{System, from (15)-(16), or equivalent systems}) \), and \( \text{RD (= Ratio-Difference, from (19)-(21))} \), can lead to different solutions in particular cases of unstable calculations (when \( R \) is very near to 2.0 in Figure 1), it can be shown that both estimation criteria are equivalent. First, it is easily checked, from examination of the equations (15)-(16), that multiplication of \( \sigma \) and the quantiles by a constant \( k \) yields an estimate of \( \sigma \) multiplied by \( k \) as well, but the estimated value of \( p \) (solution of the system) is unchanged; in other words, the estimate of \( p \) is independent of the estimate of \( \sigma \). Now, let us formulate the equations (15)-(16) in a different form (calling to mind that \( \tilde{\mu} \) is the sampling median \( \text{Me}(F) \)); letting for brevity \( y_1 = y^*(0.7) \) and \( y_2 = y^*(0.9) \), the two equations can be written as

\[
\int_{\tilde{\mu}}^x \phi(x;\tilde{\mu},\sigma, p)dx = 0.2 \quad ; \quad \int_{\tilde{\mu}}^x \phi(x;\tilde{\mu},\sigma, p)dx = 0.4
\]
or \( \int_{0}^{y_1 - \hat{\mu}} \phi(x; 0, \sigma, p) \, dx = 0.2 \quad ; \quad \int_{0}^{y_2 - \hat{\mu}} \phi(x; 0, \sigma, p) \, dx = 0.4 \)

and also, letting \( \Phi \) be the cumulative distribution function corresponding to the density \( \phi \):

\[
\Phi(y_1 - \hat{\mu}; 0, \sigma, p) = 0.7; \quad \Phi(y_2 - \hat{\mu}; 0, \sigma, p) = 0.9;
\]

hence, by application of the inverse \( \Phi^{-1}(q) = x(q) \), and letting \( k = 1/\sigma \), \( q_1 = 0.7 \) and \( q_2 = 0.9 \), one obtains

\[
k(y_1 - \hat{\mu}) = x(q_1; 0, 1, p) \quad ; \quad k(y_2 - \hat{\mu}) = x(q_2; 0, 1, p);
\]

by equating the ratios of the left hand sides and the right hand sides, the same equalization between the right hand sides of (19) and (20) is obtained.

### 3.3 An application to the chest measurement of children

As an interesting application, we consider the same distribution utilized by A. Mineo (1980, p. 573), concerning the chest measurement (in centimetres) of 4139 children aged 6; this same distribution is only reported here as a histogram in Figure 2 (over 27 classes of one centimetre). The estimates obtained by A. Mineo with the L-M method (Maximum Likelihood and Moments, formulae (9)-(11)) have been the following:

\[
\hat{\mu} = 60.099 \quad ; \quad \hat{\sigma} = 2.6995 \quad ; \quad \hat{p} = 1.5917
\]

showing an exceptional good fitting between theoretical and empirical distributions, as expressed by the sum of absolute differences – over the 27 classes – of relative frequencies \( f_r \) (for the empirical distribution) and probabilities \( p_r \) (of theoretical distribution):

\[
A = \sum_{r=1}^{27} |f_r - p_r| = 0.0452.
\]

The robust criterion S (derived from (14) and the System (15)-(16)) yields the estimates

\[
\tilde{\mu} = 60.067 \quad ; \quad \tilde{\sigma} = 2.8015 \quad ; \quad \tilde{p} = 1.8279
\]

showing an even better fitting, expressed by a value of the goodness-of-fit index \( A = 0.0430 \). The interpolated density curve and the original histogram are reported in Figure 2. As expected from the equivalence pointed out at Section 3.2, the computation of the estimates by means of the RD criterion (Ratio-Difference, formulae (19)-(21)) yields exactly the same estimates of \( \sigma \) and \( p \) just reported.
Figure 2 – Histogram and interpolating density (of a normal distribution of order $p$) computed by means of the method of estimation S-RD (equations (14)-(16) or (19)-(21)), concerning the chest measurement of 4.139 children aged 6. (from A. Mineo, 1980, p. 573).

It is worth mentioning that the interpolation of the same histogram by a normal distribution, carried out by the method of Maximum Likelihood (in this case, coincident with the method of Moments), yields a valid interpolation all the same, with estimates $\hat{\mu} = 60.099$ and $\hat{\sigma} = 2.9557$, and a goodness-of-fit index $A = 0.0640$. Instead, the normal interpolation carried out with the robust criterion of the estimation equations (6) leads to estimates $\hat{\mu} = 60.099$ and $\hat{\sigma} = 2.8414$, with an index $A = 0.0563$.

4. Two preliminary regularization procedures

A huge amount of simulations have been computed, in order to make useful comparisons of various estimation methods, all these computations have been programmed in R, and this implementation has been possible by the invaluable collaboration of Professor Gabriele Cantaluppi, who is also the author of the programmes reported in Appendix B of this paper. For various combinations of parameter values ($\sigma, p$), a thousand samples of size $n = 15, 45, 105, 225$ have been generated by means of the programme in R provided by Professor Marcello Chiodi, and here reported in Appendix A. On each sample a number of estimates and related measures have been computed; Section 5 will be devoted to some principal comparisons, appropriately chosen among the many available.
This Section is devoted to an explanation and a justification for the proposal of a preliminary regularization of the samples, which gives rise to parallel proposals of the same estimation methods already presented, however applied to the modified samples. Let us turn to a deeper examination of the (equivalent) estimation methods S and RD, and particularly to the behaviour of the relation between \( p \) and \( R \) in Figure 1; as already observed, from the computation of the sampling value \( R^* \) one can go back to an estimated \( p \) by the inverse relation in Figure 1 (or Table 1). From Figure 1 one can observe that \( p \) has a limited variability (say between 0.5 and 1.5) for \( R \) values relatively high (> 2.70), whereas the variation of \( p \) can easily be of dozens, or even hundreds, when \( R \) varies in a small interval near 2.0 (take note that 2.0 is the value taken by the ratio \( R \), applied to quantiles of order 0.7 and 0.9, for the limiting case – as \( p \) tends to infinity – of a uniform distribution).

What has been said refers to a theoretical relation. But, also assuming that the available sample comes from a distribution of \( N_p \) kind, e.g. with \( p = 2 \) (the normal case), we could have been unfortunate with drawing a sample far from representative of the parent population; thus the value of \( R \) could be very near to 2.0, suggesting \( p \) values of 300, or even more. Anyway, it could also happen that \( R < 2 \), thus excluding the recourse to the above relation for estimation purposes (the answer of the computer would be NA (Not Available), or infinity).

A more precise idea of the many samples giving rise to NA results for \( p \) estimation can be obtained by the examination of Table 3, which counts such cases over 1,000 samples for \( n = 15, 45, 105, 225 \), and some couples \((\sigma, p)\). These controls have been repeated also for \( \sigma = 10 \) (showing a pattern very similar to the one for \( \sigma = 5 \), as expected), and \( p = 0.5 \); this last parameter value, certainly very unusual, shows a plain reduction of cases for S-RD and a remarkable increase of cases for LM(II), especially when \( n = 105 \) and \( n = 225 \). Also excepting such “impossible” NA results, for \( n = 15 \) and \( n = 45 \) we have found, by the method S-RD, unacceptable estimates \( \tilde{p} > 50 \) for about 10% of the samples, when \( p \) is only 2 or 3.

It is true that one could choose not to apply to the S-RD method of estimation for the parameter \( p \), on a post-sample basis, namely when the sampling histogram appears scarcely representative of a distribution of the \( N_p \) kind; actually, in all cases of samples reasonably representative of a \( N_p \) distribution, the S-RD method behaves very well.
Table 3 – Cases NA (no finite solutions) for the estimation of $p$ over 1,000 samples.

<table>
<thead>
<tr>
<th>$\sigma$, $p$</th>
<th>Method S-RD</th>
<th>Method LM(I)</th>
<th>Method LM(II)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n = 15$</td>
<td>$n = 45$</td>
<td>$n = 105$</td>
</tr>
<tr>
<td>$\sigma = 1$, $p = 1$</td>
<td>144</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>$\sigma = 1$, $p = 1.5$</td>
<td>198</td>
<td>46</td>
<td>13</td>
</tr>
<tr>
<td>$\sigma = 1$, $p = 2$</td>
<td>271</td>
<td>127</td>
<td>44</td>
</tr>
<tr>
<td>$\sigma = 1$, $p = 3$</td>
<td>347</td>
<td>256</td>
<td>150</td>
</tr>
<tr>
<td>$\sigma = 1$, $p = 5$</td>
<td>447</td>
<td>413</td>
<td>348</td>
</tr>
<tr>
<td>$\sigma = 2$, $p = 1$</td>
<td>144</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>$\sigma = 2$, $p = 1.5$</td>
<td>198</td>
<td>46</td>
<td>13</td>
</tr>
<tr>
<td>$\sigma = 2$, $p = 2$</td>
<td>271</td>
<td>127</td>
<td>44</td>
</tr>
<tr>
<td>$\sigma = 2$, $p = 3$</td>
<td>347</td>
<td>256</td>
<td>150</td>
</tr>
<tr>
<td>$\sigma = 2$, $p = 5$</td>
<td>447</td>
<td>413</td>
<td>348</td>
</tr>
<tr>
<td>$\sigma = 5$, $p = 1$</td>
<td>144</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>$\sigma = 5$, $p = 1.5$</td>
<td>198</td>
<td>46</td>
<td>13</td>
</tr>
<tr>
<td>$\sigma = 5$, $p = 2$</td>
<td>271</td>
<td>127</td>
<td>44</td>
</tr>
<tr>
<td>$\sigma = 5$, $p = 3$</td>
<td>347</td>
<td>256</td>
<td>150</td>
</tr>
<tr>
<td>$\sigma = 5$, $p = 5$</td>
<td>447</td>
<td>413</td>
<td>348</td>
</tr>
</tbody>
</table>

However, one could wonder whether the above undesirable behaviour of the sampling $R^*$ could be avoided, of course at some price. A device aiming at avoiding this kind of non-regular behaviour of the sampling histogram consists just in a preliminary estimate, or regularization, of the $N_p$ density, which is known to be symmetric unimodal. Thus, a first device, named HR-S (Histogram Regularization of kind S – Single values) consists of the following stages applied to the sample values, already ordered from lowest to highest, $x_1, ..., x_n$, which transform the original sampling distribution $H$ into a regularized distribution $H_S$:

1. the original sampling median is simply transferred to the new distribution;
2. every couple of symmetric values $x_r$ and $x_{n-r+1}$ gives rise to new values $z_r$ and $z_{n-r+1}$, each with the same distance from the median, given by the average of the two original distances;
3. taking into consideration the right part of this symmetrized distribution, the successive distances are computed: of the value just right to the median with respect to the median, of the second value right to the median with respect to the first value etc.;
4. these distances between successive values are transformed, so that a given distance must be $\geq$ with respect to the preceding distance; if a distance is $<$ than the preceding distance, it is equated to such distance; then the right part of the distribution is mirrored to the left part;
5. in order to maintain a dispersion of the new distribution comparable to the one of the original distribution, all the above distances are rescaled, so that to ensure the same
difference between quantiles of order 0.95 and 0.05, of the original values and the final transformed values.

A second analogous device, named HR-Q (Histogram Regularization of kind Q – selected Quantiles) ensures the same kind of regularization (symmetrization plus regular widening of successive distances between values, starting from the median and pointing to the tails of the sampling distribution), however limited to the quantiles of order 0.45-0.55, 0.40-0.60 etc.; the final regularized distribution, thus obtained, will be called $H_Q$. The application of these regularization procedures renders it impossible a value $R^*$ (formula (20)) less than 2, and practically impossible $R^* = 2$.

For the sake of brevity, we present a limited number of comparisons in Table 4, which nonetheless give sufficient information concerning the relative merits of the three estimation methods $S-RD$, $LM(I)$ and $LM(II)$, as well gains and losses following the implementation of the regularization procedures $HR-S$ and $HR-Q$; note that this last procedure allows the application only of method $S-RD$, as the other two methods require the knowledge of all $n$ sampling values (original or transformed). Note also that, being the estimation of $p$ independent of $\sigma$, the results in Table 4 hold for any $\sigma$, (compatibly with convergence and stability of computations). As large values of $p$ are practically never encountered, the computations presented in Table 4 are referred only to the $N$ samples (see last column) for which the estimated $p$ is < 10.6675. Further clarifications about this choice will be given at the end of Section 5.

While postponing to Section 5 a comprehensive appraisal of the three estimation methods for the three parameters $\mu$, $\sigma$ and $p$, some comments about comparisons of original and regularized samples will follow; on the whole, it must be said that the results coming from the regularization procedures are rather disappointing, outside the obligatory condition of preventing impossible estimates by the criterion $S-RD$. For both $n$ values considered in Table 4, the comparison between $H$ samples (original values) and $H_S$ samples shows a satisfying strong reduction of the highest quantiles, unfortunately accompanied by a reduction also for the lowest quantiles, and even of the median, with some worsening of the bias. Clearly, when applied to method $LM(I)$, the sampling regularization displays very bad results, both for bias and dispersion; thus, $LM(I)$ must not be applied to regularized samples. Instead, the influence of regularization on $LM(II)$ is comprehensively positive, especially for $p = 2$, both for bias and dispersion; when $p = 3$ there is a conflicting behaviour of bias worsening and dispersion decrease. The final
rows of Table 4 are devoted to the quantile regularization $HR-Q$, giving rise to regularized samples $HQ$, the behaviour of these ordinates is quite like the $HS$ samples, with the following special feature: for $n = 45$ there is a smaller bias, however still rather high; for $n = 105$ the behaviour of the estimator is better than the case $HS$ concerning the bias, however accompanied by an increase of the highest quantiles.

Table 4 – Selected quantiles over $N$ samples, with $\sigma = 2$, for the estimation of $p$ by means of three methods, applied to the original samples $H$, and to regularized samples $HS$ and $HQ$, conditionally on estimated $p < 10.6675$. ($N =$ No. of samples with estimated $p < 10.6675$)

<table>
<thead>
<tr>
<th>$n$</th>
<th>$p$</th>
<th>Method</th>
<th>0.01</th>
<th>0.05</th>
<th>0.10</th>
<th>0.20</th>
<th>0.50</th>
<th>0.80</th>
<th>0.90</th>
<th>0.95</th>
<th>0.99</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>2</td>
<td>$S-RD$</td>
<td>0.589</td>
<td>0.817</td>
<td>0.963</td>
<td>1.157</td>
<td>1.800</td>
<td>3.044</td>
<td>4.054</td>
<td>5.111</td>
<td>7.832</td>
<td>873</td>
</tr>
<tr>
<td>45</td>
<td>2</td>
<td>$LM(\text{I})$</td>
<td>1.091</td>
<td>1.315</td>
<td>1.456</td>
<td>1.704</td>
<td>2.431</td>
<td>3.975</td>
<td>5.325</td>
<td>6.548</td>
<td>9.624</td>
<td>975</td>
</tr>
<tr>
<td>45</td>
<td>2</td>
<td>$LM(\text{II})$</td>
<td>0.969</td>
<td>1.103</td>
<td>1.187</td>
<td>1.333</td>
<td>1.752</td>
<td>2.411</td>
<td>3.037</td>
<td>3.835</td>
<td>5.862</td>
<td>993</td>
</tr>
<tr>
<td>105</td>
<td>2</td>
<td>$S-RD$</td>
<td>0.750</td>
<td>0.986</td>
<td>1.146</td>
<td>1.394</td>
<td>2.298</td>
<td>3.981</td>
<td>5.330</td>
<td>6.865</td>
<td>9.800</td>
<td>744</td>
</tr>
<tr>
<td>105</td>
<td>2</td>
<td>$LM(\text{I})$</td>
<td>0.881</td>
<td>1.671</td>
<td>2.030</td>
<td>2.449</td>
<td>3.683</td>
<td>5.934</td>
<td>7.694</td>
<td>8.938</td>
<td>10.467</td>
<td>939</td>
</tr>
<tr>
<td>45</td>
<td>3</td>
<td>$S-RD$</td>
<td>0.532</td>
<td>0.728</td>
<td>0.855</td>
<td>1.033</td>
<td>1.636</td>
<td>2.651</td>
<td>3.388</td>
<td>4.350</td>
<td>7.081</td>
<td>983</td>
</tr>
<tr>
<td>45</td>
<td>3</td>
<td>$LM(\text{I})$</td>
<td>0.151</td>
<td>0.432</td>
<td>1.130</td>
<td>2.041</td>
<td>3.826</td>
<td>6.807</td>
<td>8.565</td>
<td>9.409</td>
<td>10.479</td>
<td>845</td>
</tr>
<tr>
<td>45</td>
<td>3</td>
<td>$LM(\text{II})$</td>
<td>0.976</td>
<td>1.130</td>
<td>1.238</td>
<td>1.383</td>
<td>1.747</td>
<td>2.268</td>
<td>2.764</td>
<td>3.239</td>
<td>4.876</td>
<td>1000</td>
</tr>
<tr>
<td>105</td>
<td>3</td>
<td>$S-RD$</td>
<td>0.603</td>
<td>0.857</td>
<td>1.024</td>
<td>1.287</td>
<td>1.970</td>
<td>3.238</td>
<td>4.294</td>
<td>4.989</td>
<td>7.389</td>
<td>957</td>
</tr>
<tr>
<td>105</td>
<td>3</td>
<td>$LM(\text{I})$</td>
<td>0.111</td>
<td>0.213</td>
<td>0.314</td>
<td>0.497</td>
<td>4.487</td>
<td>7.802</td>
<td>9.323</td>
<td>9.951</td>
<td>10.459</td>
<td>699</td>
</tr>
<tr>
<td>105</td>
<td>3</td>
<td>$LM(\text{II})$</td>
<td>1.147</td>
<td>1.327</td>
<td>1.438</td>
<td>1.609</td>
<td>2.096</td>
<td>2.950</td>
<td>3.558</td>
<td>4.138</td>
<td>5.524</td>
<td>1000</td>
</tr>
<tr>
<td>45</td>
<td>2</td>
<td>$S-RD$</td>
<td>0.628</td>
<td>0.857</td>
<td>0.996</td>
<td>1.177</td>
<td>1.725</td>
<td>2.602</td>
<td>3.417</td>
<td>4.151</td>
<td>6.587</td>
<td>992</td>
</tr>
<tr>
<td>45</td>
<td>2</td>
<td>$LM(\text{I})$</td>
<td>1.224</td>
<td>1.976</td>
<td>2.234</td>
<td>2.591</td>
<td>3.734</td>
<td>5.822</td>
<td>7.478</td>
<td>8.637</td>
<td>10.081</td>
<td>918</td>
</tr>
<tr>
<td>45</td>
<td>2</td>
<td>$LM(\text{II})$</td>
<td>1.233</td>
<td>1.375</td>
<td>1.475</td>
<td>1.612</td>
<td>1.945</td>
<td>2.547</td>
<td>2.947</td>
<td>3.534</td>
<td>4.649</td>
<td>1000</td>
</tr>
<tr>
<td>105</td>
<td>2</td>
<td>$S-RD$</td>
<td>0.774</td>
<td>1.004</td>
<td>1.223</td>
<td>1.445</td>
<td>2.148</td>
<td>3.304</td>
<td>4.196</td>
<td>4.923</td>
<td>6.935</td>
<td>989</td>
</tr>
<tr>
<td>105</td>
<td>2</td>
<td>$LM(\text{I})$</td>
<td>0.200</td>
<td>0.374</td>
<td>1.947</td>
<td>3.445</td>
<td>5.707</td>
<td>8.267</td>
<td>9.298</td>
<td>9.903</td>
<td>10.444</td>
<td>645</td>
</tr>
<tr>
<td>45</td>
<td>3</td>
<td>$S-RD$</td>
<td>0.611</td>
<td>0.824</td>
<td>0.955</td>
<td>1.148</td>
<td>1.753</td>
<td>2.777</td>
<td>3.410</td>
<td>4.181</td>
<td>6.156</td>
<td>976</td>
</tr>
<tr>
<td>45</td>
<td>3</td>
<td>$S-RD$</td>
<td>0.715</td>
<td>0.958</td>
<td>1.145</td>
<td>1.424</td>
<td>2.137</td>
<td>3.323</td>
<td>4.126</td>
<td>5.089</td>
<td>7.557</td>
<td>957</td>
</tr>
<tr>
<td>105</td>
<td>2</td>
<td>$S-RD$</td>
<td>0.844</td>
<td>1.052</td>
<td>1.200</td>
<td>1.398</td>
<td>1.933</td>
<td>2.851</td>
<td>3.440</td>
<td>4.059</td>
<td>5.928</td>
<td>994</td>
</tr>
<tr>
<td>105</td>
<td>3</td>
<td>$S-RD$</td>
<td>1.029</td>
<td>1.313</td>
<td>1.519</td>
<td>1.813</td>
<td>2.561</td>
<td>3.743</td>
<td>4.587</td>
<td>5.364</td>
<td>7.740</td>
<td>962</td>
</tr>
</tbody>
</table>
5. A summary overview of estimation methods applied to the parameters $\mu$, $\sigma$, $p$ of a normal distribution of order $p$.

The easiest appraisal and comparison of the various estimation methods are concerned with the location parameter $\mu$ (mean and median of the parent population $N_p$), also because its estimation is independent of $p$ and of any regularization procedure in case of S-RD method, while the methods $LM(I)$ and $LM(II)$ – although depending on a preliminary estimate of $p$ – do not show any appreciable improvement when passing to regularized samples (actually, the practical unbiasedness is maintained, however with a slight enlargement of dispersion). As our principal aim is to compare the three methods, we must be content of making this comparison for a fixed value of $\sigma$ ($\sigma = 2$ in Tables 4-7), because the dispersion measures (signed distances of the quantiles $q(0.05)$ and $q(0.95)$ from the median, $a.m.d =$ absolute mean deviation from the median) can be exactly rescaled by a different $\sigma$, as well as (approximately) happens for the bias. On the whole, from Table 5 the three estimation methods for $\mu$ appear almost equivalent, the overall bias is negligible, thus all three methods can be judged practically unbiased; the sampling distributions are very near to symmetry, and have a comparable dispersion. No doubt that the median estimator (method S-RD) is preferable to the other estimators, because it does not depend on a preliminary estimate of $p$, and its approximate distribution – if needed – is easily established (see e.g. Serfling, 1980, pp. 77-80).

Passing to the estimation of $\sigma$ and $p$, we must remind, first of all, that the estimators of these parameters are strictly tied, as the estimation of $p$ must precede the estimation of $\sigma$, which depends on the estimated $p$. For the sake of brevity, concerning all 1,000 samples generated for each combination ($\sigma, p, n$), from Tables 6 and 7 we can confirm the appraisal already made about the results in Table 4; on the whole, the method most recommended, also in its $H_5$ version (regularization on single data), seems $LM(II)$, proposed by A.M. Mineo; nonetheless, the method S-RD could be adopted all the same in many instances, as it is enough competitive with $LM(II)$, and mostly for its overwhelming simplicity in computations, and practical significance of the ratio $R^*$ in formula (20).

However, some further insights are worthwhile, from the practitioner viewpoint. First, it must be observed that excessively cuspidate distributions $N_p$ ($p < 1$), or flat distributions $N_p$ in a large interval around the median ($p > 10$), are practically never
encountered, and can be safely excluded for practically every real phenomenon. On this thought, A.M. Mineo (2003, pp. 117-118) suggested to exclude estimates of \( p \) outside the range \( 1-10 \), when we are practically certain that \( 1 \leq p \leq 10 \). About the ratios \( R(p \ ; \ 0.7,0.9) \) and \( R*(0.7 \ ; \ 0.9) \) (formulae (19)-(20)) we have already observed that values of \( R \) very near to 2 (from above) feature a practically flat density in a central interval of coverage probability 0.80. On this thought, a value of \( R* < 2.01 \), with corresponding \( p > 10.6675 \) through the relation (19), can be judged as producing an unacceptably large \( p \); on this basis, we could (or should) decide that an unfortunate sample was obtained, unsuitable for the estimate of \( p \) (also, we could challenge the appropriateness of the \( N_p \) assumption).

As already observed, the production of very large \( p \)-estimates was a feature mainly of \( S-RD \) method; thus, we could restrict the possible estimators to the samples with \( R* \geq 2.01 \) (for the \( S-RD \) method) and to the samples with \( p \leq 10.6675 \) (for the other two methods).

Table 5 – Comparison of three methods for the estimation of \( \mu \), based on 1,000 samples \( H_i \) for each combination of \((n, \sigma = 2, \ p)\). M = mean, Me = median, \( q(0.05) \) and \( q(0.95) \) quantiles, a.m.d. = absolute mean deviation from the median.

<table>
<thead>
<tr>
<th></th>
<th>S-RD</th>
<th>LM(I)</th>
<th>LM(II)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n= 15 45 105</td>
<td>15 45 105</td>
<td>15 45 105</td>
</tr>
<tr>
<td>( p = 1.5 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>0.034 -0.006 -0.004</td>
<td>0.000 -0.006 -0.001</td>
<td>0.025 -0.005 -0.002</td>
</tr>
<tr>
<td>Me</td>
<td>0.001 -0.010 -0.002</td>
<td>0.038 -0.006 -0.003</td>
<td>-0.018 -0.001 -0.005</td>
</tr>
<tr>
<td>( q(0.05)-Me )</td>
<td>-0.967 -0.570 -0.374</td>
<td>-1.317 -0.594 -0.348</td>
<td>-0.962 -0.555 -0.337</td>
</tr>
<tr>
<td>( q(0.95)-Me )</td>
<td>1.090 0.588 0.384</td>
<td>1.089 0.575 0.361</td>
<td>1.080 0.560 0.366</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>0.482 0.287 0.183</td>
<td>0.577 0.289 0.175</td>
<td>0.476 0.283 0.172</td>
</tr>
<tr>
<td>( p = 2 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>-0.007 0.005 0.002</td>
<td>-0.030 0.004 -0.002</td>
<td>-0.017 0.006 -0.003</td>
</tr>
<tr>
<td>Me</td>
<td>-0.002 0.011 0.007</td>
<td>-0.056 0.000 0.001</td>
<td>-0.011 0.005 0.002</td>
</tr>
<tr>
<td>( q(0.05)-Me )</td>
<td>-0.995 -0.578 -0.392</td>
<td>-1.061 -0.508 -0.318</td>
<td>-0.940 -0.518 -0.315</td>
</tr>
<tr>
<td>( q(0.95)-Me )</td>
<td>1.063 0.627 0.367</td>
<td>1.205 0.547 0.306</td>
<td>1.020 0.517 0.306</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>0.505 0.287 0.187</td>
<td>0.542 0.252 0.151</td>
<td>0.469 0.245 0.152</td>
</tr>
<tr>
<td>( p = 3 )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>-0.003 0.005 0.007</td>
<td>0.005 0.005 0.006</td>
<td>0.001 0.013 0.002</td>
</tr>
<tr>
<td>Me</td>
<td>0.032 0.001 0.007</td>
<td>0.028 -0.001 0.006</td>
<td>-0.016 0.021 0.002</td>
</tr>
<tr>
<td>( q(0.05)-Me )</td>
<td>-1.064 -0.610 -0.389</td>
<td>-1.326 -0.433 -0.279</td>
<td>-0.855 -0.456 -0.289</td>
</tr>
<tr>
<td>( q(0.95)-Me )</td>
<td>0.977 0.601 0.407</td>
<td>1.272 0.468 0.276</td>
<td>0.911 0.432 0.291</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>0.498 0.299 0.198</td>
<td>0.571 0.222 0.136</td>
<td>0.411 0.213 0.136</td>
</tr>
</tbody>
</table>
Table 6 – Comparison of three methods for the estimation of $p$, based on 1,000 samples (of kind $H$, $H_3$), for each combination of $(n, \sigma = 2, p)$. Only the samples giving rise to a ratio $R\ast \geq 2.01$, or an estimate of $p \leq 10.6675$, have been considered.

<table>
<thead>
<tr>
<th>$p = 1.5, H$</th>
<th>S-RD</th>
<th>LM(I)</th>
<th>LM(II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>15</td>
<td>45</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.660</td>
<td>1.841</td>
<td>1.710</td>
</tr>
<tr>
<td>Me</td>
<td>1.185</td>
<td>1.487</td>
<td>1.491</td>
</tr>
<tr>
<td>$q(0.05)$-Me</td>
<td>-0.778</td>
<td>-0.819</td>
<td>-0.633</td>
</tr>
<tr>
<td>$q(0.95)$-Me</td>
<td>3.494</td>
<td>2.927</td>
<td>1.731</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>0.930</td>
<td>0.812</td>
<td>0.557</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p = 2, H$</th>
<th>S-RD</th>
<th>LM(I)</th>
<th>LM(II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>15</td>
<td>45</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.977</td>
<td>2.224</td>
<td>2.218</td>
</tr>
<tr>
<td>Me</td>
<td>1.369</td>
<td>1.800</td>
<td>1.874</td>
</tr>
<tr>
<td>$q(0.05)$-Me</td>
<td>-0.903</td>
<td>-0.982</td>
<td>-0.783</td>
</tr>
<tr>
<td>$q(0.95)$-Me</td>
<td>4.123</td>
<td>3.311</td>
<td>2.442</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>1.116</td>
<td>0.977</td>
<td>0.761</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p = 3, H$</th>
<th>S-RD</th>
<th>LM(I)</th>
<th>LM(II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>15</td>
<td>45</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2.218</td>
<td>2.870</td>
<td>3.125</td>
</tr>
<tr>
<td>Me</td>
<td>1.555</td>
<td>2.298</td>
<td>2.676</td>
</tr>
<tr>
<td>$q(0.05)$-Me</td>
<td>-1.047</td>
<td>-1.312</td>
<td>-1.312</td>
</tr>
<tr>
<td>$q(0.95)$-Me</td>
<td>4.928</td>
<td>4.567</td>
<td>3.945</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>1.243</td>
<td>1.325</td>
<td>1.193</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p = 1.5, H_3$</th>
<th>S-RD</th>
<th>LM(I)</th>
<th>LM(II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>15</td>
<td>45</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.543</td>
<td>1.628</td>
<td>1.674</td>
</tr>
<tr>
<td>Me</td>
<td>1.216</td>
<td>1.377</td>
<td>1.418</td>
</tr>
<tr>
<td>$q(0.05)$-Me</td>
<td>-0.825</td>
<td>-0.761</td>
<td>-0.650</td>
</tr>
<tr>
<td>$q(0.95)$-Me</td>
<td>2.521</td>
<td>2.041</td>
<td>1.960</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>0.796</td>
<td>0.677</td>
<td>0.606</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p = 2, H_3$</th>
<th>S-RD</th>
<th>LM(I)</th>
<th>LM(II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>15</td>
<td>45</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.893</td>
<td>1.964</td>
<td>2.019</td>
</tr>
<tr>
<td>Me</td>
<td>1.458</td>
<td>1.636</td>
<td>1.725</td>
</tr>
<tr>
<td>$q(0.05)$-Me</td>
<td>-0.979</td>
<td>-0.908</td>
<td>-0.868</td>
</tr>
<tr>
<td>$q(0.95)$-Me</td>
<td>3.285</td>
<td>2.713</td>
<td>2.426</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>0.986</td>
<td>0.837</td>
<td>0.765</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$p = 3, H_3$</th>
<th>S-RD</th>
<th>LM(I)</th>
<th>LM(II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>15</td>
<td>45</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2.056</td>
<td>2.364</td>
<td>2.468</td>
</tr>
<tr>
<td>Me</td>
<td>1.670</td>
<td>1.970</td>
<td>2.148</td>
</tr>
<tr>
<td>$q(0.05)$-Me</td>
<td>-1.140</td>
<td>-1.112</td>
<td>-1.144</td>
</tr>
<tr>
<td>$q(0.95)$-Me</td>
<td>3.310</td>
<td>3.019</td>
<td>2.775</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>1.052</td>
<td>0.995</td>
<td>0.929</td>
</tr>
</tbody>
</table>
Table 7 – Comparison of three methods for the estimation of $\sigma$, based on 1,000 samples (of kind $H, H_S$), for each combination of $(n, \sigma = 2, p)$. Only the samples giving rise to a ratio $R^* \geq 2.01$, or an estimate of $p \leq 10.6675$, have been considered. M = mean, Me = median, $q(0.05)$ and $q(0.95)$ quantiles, a.m.d. = absolute mean deviation from the median.

<table>
<thead>
<tr>
<th></th>
<th>S-RD</th>
<th>LM(I)</th>
<th>LM(II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n=$</td>
<td>15  45 105</td>
<td>15  45 105</td>
<td>15  45 105</td>
</tr>
<tr>
<td>$p = 1.5, H$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.861 1.983 2.003</td>
<td>2.158 2.112 2.059</td>
<td>1.924 1.946 1.981</td>
</tr>
<tr>
<td>Me</td>
<td>1.796 1.952 1.987</td>
<td>2.039 2.065 2.037</td>
<td>1.875 1.923 1.970</td>
</tr>
<tr>
<td>$q(0.05)$-Me</td>
<td>-0.792 -0.548 -0.387</td>
<td>-0.823 -0.589 -0.358</td>
<td>-0.644 -0.481 -0.331</td>
</tr>
<tr>
<td>$q(0.95)$-Me</td>
<td>1.187 0.673 0.484</td>
<td>1.520 0.835 0.463</td>
<td>0.940 0.607 0.372</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>0.469 0.302 0.208</td>
<td>0.560 0.339 0.198</td>
<td>0.393 0.256 0.170</td>
</tr>
<tr>
<td>$p = 2, H$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.784 1.942 1.985</td>
<td>1.977 2.115 2.055</td>
<td>1.832 1.946 1.968</td>
</tr>
<tr>
<td>Me</td>
<td>1.742 1.936 1.975</td>
<td>1.903 2.076 2.029</td>
<td>1.777 1.922 1.958</td>
</tr>
<tr>
<td>$q(0.05)$-Me</td>
<td>-0.737 -0.542 -0.378</td>
<td>-0.672 -0.541 -0.350</td>
<td>-0.588 -0.457 -0.316</td>
</tr>
<tr>
<td>$q(0.95)$-Me</td>
<td>0.978 0.582 0.426</td>
<td>1.066 0.753 0.440</td>
<td>0.863 0.595 0.371</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>0.410 0.283 0.194</td>
<td>0.429 0.309 0.191</td>
<td>0.354 0.255 0.166</td>
</tr>
<tr>
<td>$p = 3, H$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.662 1.866 1.939</td>
<td>1.700 2.058 2.064</td>
<td>1.724 1.902 1.958</td>
</tr>
<tr>
<td>Me</td>
<td>1.635 1.860 1.935</td>
<td>1.642 2.064 2.056</td>
<td>1.694 1.885 1.935</td>
</tr>
<tr>
<td>$q(0.05)$-Me</td>
<td>-0.713 -0.493 -0.371</td>
<td>-0.555 -0.607 -0.356</td>
<td>-0.606 -0.459 -0.314</td>
</tr>
<tr>
<td>$q(0.95)$-Me</td>
<td>0.802 0.536 0.392</td>
<td>0.806 0.537 0.442</td>
<td>0.790 0.541 0.428</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>0.376 0.260 0.188</td>
<td>0.330 0.273 0.193</td>
<td>0.333 0.247 0.183</td>
</tr>
<tr>
<td>$p = 1.5, H_S$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.885 1.987 2.062</td>
<td>2.383 2.380 2.409</td>
<td>1.948 1.976 2.064</td>
</tr>
<tr>
<td>Me</td>
<td>1.828 1.954 2.031</td>
<td>2.335 2.312 2.373</td>
<td>1.897 1.954 2.043</td>
</tr>
<tr>
<td>$q(0.05)$-Me</td>
<td>-0.817 -0.635 -0.515</td>
<td>-1.213 -0.937 -0.640</td>
<td>-0.653 -0.506 -0.429</td>
</tr>
<tr>
<td>$q(0.95)$-Me</td>
<td>1.109 0.821 0.662</td>
<td>1.605 1.145 0.804</td>
<td>0.893 0.652 0.541</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>0.473 0.345 0.284</td>
<td>0.700 0.483 0.338</td>
<td>0.380 0.277 0.230</td>
</tr>
<tr>
<td>$p = 2, H_S$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.820 1.926 1.989</td>
<td>2.074 2.291 2.371</td>
<td>1.820 1.890 1.972</td>
</tr>
<tr>
<td>Me</td>
<td>1.761 1.925 1.976</td>
<td>2.051 2.323 2.360</td>
<td>1.778 1.878 1.953</td>
</tr>
<tr>
<td>$q(0.05)$-Me</td>
<td>-0.735 -0.641 -0.503</td>
<td>-0.958 -1.049 -0.605</td>
<td>-0.587 -0.487 -0.377</td>
</tr>
<tr>
<td>$q(0.95)$-Me</td>
<td>1.007 0.714 0.578</td>
<td>1.191 0.823 0.659</td>
<td>0.782 0.564 0.475</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>0.421 0.317 0.268</td>
<td>0.569 0.448 0.315</td>
<td>0.331 0.251 0.213</td>
</tr>
<tr>
<td>$p = 3, H_S$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>1.652 1.816 1.879</td>
<td>1.643 1.976 2.090</td>
<td>1.630 1.757 1.837</td>
</tr>
<tr>
<td>Me</td>
<td>1.625 1.812 1.875</td>
<td>1.577 2.097 2.223</td>
<td>1.624 1.756 1.833</td>
</tr>
<tr>
<td>$q(0.05)$-Me</td>
<td>-0.683 -0.563 -0.449</td>
<td>-0.658 -1.062 -1.127</td>
<td>-0.542 -0.438 -0.342</td>
</tr>
<tr>
<td>$q(0.95)$-Me</td>
<td>0.810 0.551 0.460</td>
<td>0.972 0.677 0.464</td>
<td>0.587 0.456 0.379</td>
</tr>
<tr>
<td>a.m.d.</td>
<td>0.370 0.276 0.226</td>
<td>0.434 0.475 0.381</td>
<td>0.278 0.217 0.178</td>
</tr>
</tbody>
</table>
APPENDIX A

Programme in R for the generation of random numbers from \( N_p(\mu, \sigma^p) \) distributions.
Author: Marcello Chiodi, March 2009.

```r
# definition of the distribution "normp" a normal distribution of
# order p
# in all functions of this file:
# "mean" is the location parameter such that E[X]=mean
# "scale" is the dispersion or scale parameter such that E|X-E[X]|^p
# =scale^p
# "p" is the norm of the distribution
# the functions in this file use the property that:
# if x is distributed according to a normp distribution of parameters
# p,mean and scale, then
# abs[x-mean]^p is distributed according to a gamma distribution
# of shape 1/p and rate 1/(p scale^p)
# functions defined:

# rnormp(x,p =2,mean=0,scale=1) generates a sample of n pseudorandom
# numbers from a normal distribution of order p with parameters p,mean,scale
# dnormp(x,p =2,mean=0,scale=1) gives the density in x (vector or scalar)
# pnormp(x,p =2,mean=0,scale=1) gives the cumulative distribution
# function in x (vector or scalar)
# qnormp(q,p =2,mean=0,scale=1) gives the quantiles of the normp
# distribution given as input the probabilities q (vector or scalar)
# (inverse function of pnormp)

# rsign =function(n,prob=0.5) 1-2*(runif(n)>prob)
rsign   =function(n,prob=0.5) 1-2*(runif(n)>prob)
rnormp =function(n,p=2,mean=0,scale=1){
g     =rgamma(n,shape=1/p,rate=1/(p * scale^p))
sig     =rsign(n)
return(mean+sig*(g^(1/p)))
}
dnormp =function(x,p=2,mean=0,scale=1) { z=(abs(x-mean)/scale)^p
d=exp(-z/p-1gamma(1+1/p)-log(p)/p)/(2*scale)
return(d)
}
pnormp =function(x,p=2,mean=0,scale=1){ z=(abs(x-mean))^p
d=pgamma(z,shape=1/p,rate=1/(p * scale^p))
return(0.5+(d)*sign(x-mean)*0.5)
}
qnormp =function(q,p=2,mean=0,scale=1){ sg=sign(q-0.5)
z=gamma(sg*(2*q-1),shape=1/p,rate=1/(p * scale^p))
x=mean+sg*z^(1/p)
```

---

23
APPENDIX B

Programme in R for the estimation of the location, scale and shape parameters of a \( N_p(\mu, \sigma, p) \) distribution by the S-RD, LM(I) and LM(II) criteria

Author: Gabriele Cantaluppi, October 2009.

```r
return(x)
}

# the following packages are required;
# the functions, reported here, do not comply
# with the presence of missing data
source("random_p.R")  # by M. Chiodi, see Appendix A
library(nleqslv)  # by B. Hasselman

# the function "riordina.unimodale.simmetrica" performs the data
# transformation according to the assumptions of symmetry and
# unimodality
"riordina.unimodale.simmetrica" <- function(dati)
{
a=round(dati,12)
ndim=length(a)
perc05=a[floor(ndim*.05)+1]
perc95=a[floor(ndim*.95)+1]
dist095mediana=(perc95-perc05)/2
pos=trunc(ndim/2,0)
if (ndim/2==pos) mediana=(a[pos]+a[pos+1])/2 else mediana=a[pos+1]
deltaa=(rev(rev(a)[1:pos])-a[pos:1])/2
cc=c(deltaa[1],diff(deltaa))
cc=floor(cc*10^12)
while (min(sign(round(diff(cc),10)))==-1)
{
  pos1=min(which((1-(round(cc,10)==sort(round(cc,10))))>0))
  cctemp=cc[pos1:length(cc)]
  cctemp[cctemp<=cctemp[1]]=cctemp[1]
  cc[pos1:length(cc)]=cctemp
  deltaa=cumsum(cc)
  cc=c(deltaa[1],diff(deltaa))
}
deltaa=deltaa/10^12
if (ndim/2==pos) amod=c(-rev(deltaa),deltaa)
if (ndim/2!=pos) amod=c(-rev(deltaa),0,deltaa)
nperc05=amod[floor(ndim*.05)+1]
nperc95=amod[floor(ndim*.95)+1]
ndist095mediana=(nperc95-nperc05)/2
amod=amod/ndist095mediana*dist095mediana+mediana
amod
}

# the function "estimate_srd" performs the estimation of the
# location, scale and shape parameters of a normalp distribution
```
# by the S-RD criterion

```r
estimate_srd=function(data,regularization=c("no","HS","HQ"))
{
  regularization=match.arg(regularization)
  dataord=sort(data)
  if (regularization=="no") {regHS=0;regHQ=0}
  if (regularization=="HS") {regHS=1;regHQ=0}
  if (regularization=="HQ") {regHS=0;regHQ=1}
  # data transformation according to the HS criterion
  if (regHS==1) dataord=riordina.unimodale.simmetrica(dati=dataord)
  ndim=length(data)
  intervallo=c(.8,.4)
  ordperc=floor(ndim*ordperc)+1
  percentili=dataord[posizioni]
  matruovipercentili=matrix(c(.5,0,1,0,-.5,0,.5,0,1,0,0,0,0,0,0,0,-.5,.5,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0),byrow=T,5,5)
  nuovipercentili=matrnuovipercentili*percentili
  matrcostintegrale=matrix(c(0,0,-1,0,1,0,0,-1,1,0),byrow=T,2,5)
  # data transformation according to the HQ criterion
  # with reference to the 05, 10, ..., 95 percentiles
  if (regHQ==1)
  {
    posizioni=floor(ndim*seq(.05,.95,.05))+1
    percentilicoer=riordina.unimodale.simmetrica(dati=percentilicoer)
    nuovipercentili=percentilicoer[c(2,6,10,14,18)]
    matrcostintegrale=matrix(c(0,0,-1,0,1,0,0,-1,1,0),byrow=T,2,5)
  }
  diffperc=matrcostintegrale%*%nuovipercentili
  if (diffperc[1]/diffperc[2]<=2.01) stimap=10.6675
  if (diffperc[1]/diffperc[2]>2.01)
  {
    # estimation of \( p \), see relationships (19) and (20)
    myfun = function(x) {
      y = numeric(1)
      y = (diffperc[1]/diffperc[2]) - 
        qnormp(ordperc[5],p=x[1],scale=1)/qnormp(ordperc[4],p=x[1],scale=1)
    }
    x0 = c(1.5)                        # Starting guess
    outputaist=nleqslv(x0, myfun)
    stimap=outputaist$x
    if (regularization=="no")
    {
      if (is.nan(stimap)==TRUE)
      {
        stimap=NA
        stimas=NA
      }
    }
  }
  if (is.nan(stimap)==FALSE)
  {
    # estimation of \( s \), see relationship (22)
    stimas = (diffperc[1]-
              diffperc[2])/(qnormp(ordperc[5],p=stimap,scale=1)-
                            qnormp(ordperc[4],p=stimap,scale=1))
  }
  out=list()
}
```
```
# the function "estimate_lmI" performs the estimation of the
# location, scale and shape parameters of a normalp distribution
# by the (Maximum Likelihood + Moments of kind I) criterion
# Since, in the presence of some unfortunate samples, the system of
equations (9) and (11) may have no unique solution, several starting
values have been considered to search the estimates of the shape
parameter p and of the location parameter.
# The solution proposed consists in minimizing an objective function
linked to (9) and (11)

```
estimate_lmI=function(data,regularization=c("no","HS"))
{
    regularization=match.arg(regularization)
    if (regularization=="HS")
    {
        # data transformation according to the HS criterion
        data=riordina.unimodale.simmetrica(dati=sort(data))
    }

    startvals=seq(.5,6,.25)
    atemp=array(NA,c(length(startvals)+1,2))
    ctemp=array(NA,length(startvals)+1)
    # objective function to estimate the location and shape parameters,
    # see relationships (9) and (11)
    myfun1 = function(x) {
        y = numeric(1)
        scarti = data-x[2]
        y = mean(abs(scarti)^(x[1]-1)*sign(scarti))^2     +
        ((mean(abs(scarti)^(2*x[1]))/mean(abs(scarti)^x[1])^2)/(x[1]+1)-1)^2
        y
    }
    centro=mean(data)
    for (startv in 1:length(startvals))
    {
        x0 = c(startvals[startv], centro)      # Starting guess
        outputsist=nlminb(x0, myfun1, lower=c(0,-Inf))
        if (is.nan(outputsist$objective)) outputsist$convergence=1
        if (outputsist$spar[1]>.1&outputsist$convergence==0) # only
        estimates larger than 0.1 of p are considered as feasible
        {
            atemp[startv,]=outputsist$spar
            ctemp[startv]=outputsist$objective
        }
    }
    if (sum(is.na(atemp[1:length(startvals),1]))<length(startvals))
    {
        atemp[length(startvals)+1,]=atemp[which.min(ctemp[1:length(startvals)])
    }
```

out$location=nuovipercentili[3]    # location parameter estimate
out$p=stimap    # shape parameter estimate
out$s=stimas    # scale parameter estimate
out$flag="estimation achieved"    # termination code
if (diffperc[1]/diffperc[2]<=2.01) out$flag="difference percentile
ratio less than 2.01; the estimation of p has been performed by fixing
R*=2.01"
if (is.nan(stimap)==TRUE) out$flag="You have to apply data
regularization"
ctemp[length(startvals)+1]=ctemp[which.min(ctemp[1:length(startvals)])]
}
if (sum(is.na(atemp[1:length(startvals),1]))==length(startvals))
    atemp[length(startvals)+1,]=NA
outputsist$par=atemp[length(startvals)+1,]
flag="estimation achieved"
if (sum(is.na(outputsist$par))!=0)
    if (outputsist$par[1]>10.6675)
        outputsist$par[1]=10.6675
        flag="the estimate of p was greater than 10.6675; it has been
        constrained to 10.6675"
    }
if (sum(is.na(outputsist$par))==2)
    flag="There was no convergence; Some missing value was present
or You can try to apply data regularization"
out=list()
out$location=outputsist$par[2] # location parameter estimate
out$p=outputsist$par[1]        # shape parameter estimate
out$s=NA                       # scale parameter estimate
if (is.na(out$p)==0) {
    out$s=(sum((abs(data-
out$location))^
out$p)/length(data))^(1/out$p) # scale parameter estimate
}
out$flag=flag                  # termination code
out

# the function "estimate_lmII" performs the estimation of the
# location, scale and shape parameters of a normalp distribution
# with the criterion implemented by A.M. Mineo in his library normalp.
# The functions estimatep and paramp, available in the library
"normalp"
# by A. M. Mineo, allow only values in the set [1,10] for the
# parameter p;
# the following code contains an adaptation of the original
# estimatep and paramp functions and allows the
# parameter p to assume values in the set (0,10.6675].
estimate_lmII=function(data,regularization=c("no","HS"),mineoconstr)
{
    regularization=match.arg(regularization)
    estimatep=function (x, mu, p = 2)
    {
        if (!is.numeric(x) || !is.numeric(mu) || !is.numeric(p))
            stop(" Non-numeric argument to mathematical function")
        ssp <- sum(abs(x - mu)^p)/length(x)
        sp <- ssp^(1/p)
        xstand <- (x - mu)/sp
        sa <- sum(abs(xstand))
        sb <- sum(xstand * xstand)
    }
vi <- sqrt(length(x) * sb)/sa
vi <- vi + ((vi - 1)/length(x)) * 5
# fvi modified
fvi <- function(p) (vi - sqrt(gamma(1/p) * gamma(3/p))/gamma(2/p))
x0 = c(1.5) # Starting guess
pp=nleqslv(x0, fvi)$x
pp

# new formulation of A.M. Mineo function "paramp"
# it checks if p is enclosed in the set (0,10.6675]
paramp=function (x, p = NULL)
{
  if (!is.numeric(x) || (!is.null(p) && !is.numeric(p)))
    stop(" Non-numeric argument to mathematical function")
  ff <- function(Mp) sum(abs(x - Mp)^p)
  Mp <- mean(x)
  if (is.null(p)) {
    df <- 2
    pp <- 2
    iter <- 0
    i <- 0
    p <- estimatep(x, Mp, pp)
    #aggiunta
    if (p<0) p=1.35
    if (p>10.6675) p=10.6675
    while (abs(p - pp) > 1e-04) {
      pp <- p
      op <- optim(Mp, ff, method = "BFGS")
      Mp <- op$par
      p <- estimatep(x, Mp, pp)
      #aggiunta
      if (p<0) p=.4
      if (p>10.6675) p=10.6675
      i <- i + 1
      if (i == 100) {
        iter <- 1
        break
      }
    }
  }
  if (p == 1)
    Mp <- median(x)
  Sp <- ((sum(abs(x - Mp)^p))/(length(x) - df))^(1/p)
  if (p >= 11.5) {
    Mp <- mean(c(max(x), min(x)))
    Sp <- (max(x) - min(x))/2
  }
  sd <- sqrt((length(x) - 1) * var(x)/length(x))
  mn <- mean(x)
  ris <- list()
  ris$mean <- mn
  ris$mp <- Mp
  ris$sd <- sd
  ris$sp <- Sp
  ris$p <- p
  ris$iter <- iter
  ris$flag <- 0
  #modificata
  if (p==.4||p==1.35) ris$flag <- 1
  if (p==10.6675) ris$flag <- 2
  class(ris) <- "paramp"
if (regularization=="HS")
{
    # data transformation according to the HS criterion
    data=riordina.unimodale.simmetrica(dati=sort(data))
}

# definizione funzione per trovare stima mediana e p
outputmineo=paramp(data)
out=list()
out$location=outputmineo$mp       # location parameter estimate
out$p=outputmineo$p               # shape parameter estimate
out$s=outputmineo$s
if (outputmineo$flag==0) out$flag="estimation achieved" # termination code
if (outputmineo$flag==2) out$flag="the estimate of p was greater than 10.6675; it has been constrained to 10.6675"
if (outputmineo$flag==1)
{
    out$location=NA
    out$p=NA
    out$s=NA
    out$flag="the estimate of p was negative number"
}
out

#Example

data=rnormp(101,p=2,scale=2)

print(estimate_srd(data=data,regularization="no"))
print(estimate_srd(data=data,regularization="HS"))
print(estimate_srd(data=data,regularization="HQ"))

print(estimate_lmI(data=data,regularization="no"))
print(estimate_lmI(data=data,regularization="HS"))

print(estimate_lmII(data=data,regularization="no",mineoconstr=0))
print(estimate_lmII(data=data,regularization="HS",mineoconstr=0))

To receive the code please send an e-mail to gabriele.cantaluppi@unicatt.it.

REFERENCES


