

Robustness aspects of the generalized normal distribution

Rosa Capobianco

Facoltà di Economia - Università degli Studi del Sannio di Benevento
E-mail: capobianco@unisannio.it

Summary: The aim of the paper is to study the robustness features of the generalized normal distribution. To this end the properties of maximum likelihood estimators of the location and scale parameters are investigated and compared with those of the Huber proposal II estimator and with those of the estimators of the Student t distribution.

Key words: Generalized Normal Distribution, Huber Estimator Proposal II, Location Parameter, Robustness, Scale Parameter, t Distribution.

1. Introduction

Although the Gaussian model has a prominent role in statistical applications, the analysis of real data often leads to reject the hypothesis that data have been generated by a normal distribution (Hampel *et al.*, 1986; Hill and Dixon, 1982). In these circumstances the adoption of more flexible models, which allow to represent data generated by distributions in a neighbourhood of the Gaussian one, may be appropriate. In particular, models which embed the Gaussian distribution as a special case, are of great interest, since their use permits to deal with deviations from normality, while preserving the possibility to test the adequacy of the Gaussian distribution to the data. The first papers along this direction (Subbotin, 1923; Box, 1953;

Vianelli, 1963; Mineo and Vianelli, 1980) propose the normal distribution of order α , the so-called generalized normal distribution (g.n.d.) or exponential power distribution, which fits data with heavier tails for appropriate values of α . Actually, for $\alpha=1$ it yields the Laplace distribution, for $\alpha=2$ it yields the normal model, while the limit distribution, for $\alpha \rightarrow \infty$, is Uniform in the range $[-s, s]$ (Chiodi, 2000). Here we focus on the g.n.d. with values of α between $[1,2]$.

The use of the generalized distribution to fit heavy tailed distributions, has been suggested by several authors (D'Agostino and Lee, 1977; Hogg, 1974; among others) as an alternative to the robust approach (Staudte and Sheather, 1990). The idea underlying the latter approach is that the assumed model typically provides only an approximation to reality while the actual distribution belongs to a neighbourhood of this model. An approximated model can be adopted because of its simplicity, as it is often the case when the normal model is assumed. In some cases, this approximating model is regarded as the model which fits the majority of data. Under these circumstances, robust statistics has developed a set of inferential procedures producing reliable results even when the model is misspecified to some degree. In an estimation context, the use of robust techniques leads to some efficiency loss if the assumed model is correct. On the other hand, it avoids inconsistent results when the model is only approximately true. In particular, it yields to estimates which do not significantly differ from the actual value of the parameters under the assumed model. However, the meaning of "actual value of the parameters" is not always clear. If the model is a family of distributions, the value of the parameters might be such that the model is as close as possible to the actual distribution or, alternatively, such that it gives the best fit to most of the data. Indeed, it is not always clear whether the approximation of the model affects the value of the parameters which are to be estimated.

The circumstance that the interpretation of the estimated parameters is sometimes ambiguous may be made clear by a simple example. Let us consider the case of a location-scale model: if the data have a symmetric distribution, the location parameter is clearly identified with the centre of the distribution, whereas what is

estimated as a scale parameter is not obvious. When the true model is in a neighbourhood of the normal model, one may wonder whether the estimated scale parameter is the variance of the approximating model or should be differently interpreted. In particular, one may wonder whether the estimated parameter has to be interpreted as the scale of the entire data set or as the scale of the majority of the data.

In order to overcome these ambiguities, when the true distribution is in a neighbourhood of the normal model, inference can be carried out under the generalized normal distribution. By varying the value of α , this model fits data with heavier tails and can therefore be viewed as a flexible model able to cope with deviations from normality. In this respect, we shall refer to this model as a “robust model”. Its use, as an alternative to robust inference, has the advantage that the estimated quantities have always a clear interpretation as the parameters of a given model.

Furthermore, another important advantage that arises when the g.n.d. is used to cope with deviation from normality is that standard inferential procedures, based on the likelihood function, can be used.

As a matter of fact, the type of deviations which are taken into account when considering the g.n.d. with $\alpha \in [1,2]$ are in the specific direction of heavier tails. Although this is not the only possible deviation, it is definitely the most frequent one. A more serious concern is how relevant this type of deviation can be. As it will be clear in the next section, the g.n.d. considers exponentially decreasing tails (the limit distribution obtained when $\alpha=1$ is the Laplace distribution), thus the extent of its applicability has to be investigated. Of course, the reliability of the estimates obtained on this model depends on their influence function, which is to be compared with that of popular robust estimators.

The flexibility of the g.n.d. in a neighbourhood of the normal model is obtained by considering an additional parameter, beyond location and scale. Thus another issue that has to be considered is the following one: if the actual distribution is normal, it is desirable that the presence of an additional parameter has little effect on the efficiency of the estimators of location and scale. In other words, a question to be investigated is whether the need of estimating α leads

to an efficiency loss in the estimation of the other parameters under the normal distribution.

The inferential aspects of the extension of the Gaussian distribution have been studied by some authors. In particular Agrò (1995) obtained the maximum likelihood (ML) estimates for the three parameters and derived the covariance matrix. He also examined the regularity conditions which ensure the asymptotic normality and efficiency. However the robustness properties of the ML estimators have been only partially investigated for the location parameter by D'Agostino and Lee (1977) and the influence function has not yet been studied.

The outline of the paper is the following: in section 2 we introduce the generalized normal distribution, define the ML estimators and their asymptotic covariance matrix. In section 3 the robustness properties of the ML estimators are considered by means of the influence function. In section 4 ML estimators are compared with the Huber estimator proposal II and with the estimators of the location and scale parameters of the Student t distribution.

2. The generalized normal distribution

Let X be a generalized normal variable of order α , with location parameter μ and scale parameter σ . The probability density function of X is given by

$$f(x; \mathbf{m}, \mathbf{s}, \mathbf{a}) = \frac{1}{c(\mathbf{a})} \exp\left(-\frac{1}{\mathbf{a}} \left| \frac{x - \mathbf{m}}{\mathbf{s}} \right|^{\mathbf{a}}\right)$$

for $-\infty < x < +\infty$, where $c(\mathbf{a}) = 2\mathbf{a}^{1/\mathbf{a}-1}\Gamma(1/\mathbf{a})$, its expected value is equal to μ while its variance is

$$\text{var}(X) = \mathbf{a}^{2/a} \mathbf{s}^2 \frac{\Gamma\left(\frac{3}{\mathbf{a}}\right)}{\Gamma\left(\frac{1}{\mathbf{a}}\right)}.$$

In order to obtain the ML estimator of $\boldsymbol{\theta} = (\mathbf{m}, \mathbf{s}, \mathbf{a})'$ let $z = (x - \mathbf{m})/\mathbf{s}$; the score functions are given by (Agrò, 1995)

$$s(x, \mathbf{m}) = \frac{\partial \ln f(x)}{\partial \mathbf{m}} = -\frac{1}{\mathbf{s}} |z|^{a-1} \text{sign}(z), \quad (1)$$

$$s(x, \mathbf{s}) = \frac{\partial \ln f(x)}{\partial \mathbf{s}} = -\frac{1}{\mathbf{s}} + \frac{1}{\mathbf{s}} |z|^a, \quad (2)$$

$$s(x, \mathbf{a}) = \frac{\partial \ln f(x)}{\partial \mathbf{a}} = \frac{\ln(\mathbf{a})}{\mathbf{a}^2} + \frac{\mathbf{a}-1}{\mathbf{a}^2} + \frac{1}{\mathbf{a}^2} \Psi\left(\frac{1}{\mathbf{a}}\right) - \frac{|z|^a \ln |z|}{\mathbf{a}} + \frac{|z|^a}{\mathbf{a}^2}, \quad (3)$$

where $\Psi(z) = \partial \ln\{\Gamma(z)\}/\partial z$ is the digamma function.

Given a sample of independent and identically distributed (i.i.d.) random variables (r.v.) (X_1, \dots, X_n) , the ML estimator $T_n = (\mathbf{m}_n, \mathbf{s}_n, \mathbf{a}_n)'$ of $\boldsymbol{\theta}$ can be expressed as an M-estimator, i.e. as the solution of the equation

$$\sum_{i=1}^n \mathbf{y}(x_i, T_n) = 0,$$

where $\mathbf{y}(x, T_n) = [s(x, \mathbf{m}), s(x, \mathbf{s}), s(x, \mathbf{a})]'$. It is interesting to notice that the estimator of μ is the solution of

$$\sum_{i=1}^n \text{sign}(x_i - \mathbf{m}_n) |x_i - \mathbf{m}_n|^{a-1} = 0$$

so that \mathbf{m}_n is given by

$$\mathbf{m}_n = \sum_{i=1}^n x_i \tilde{\mathbf{w}}_i, \quad (4)$$

where

$$\tilde{\mathbf{w}}_i = \frac{\mathbf{w}(x_i - \mathbf{m}_n)}{\sum_{i=1}^n \mathbf{w}(x_i - \mathbf{m}_n)}$$

and

$$\mathbf{w}(x_i) = \frac{|x_i|^{\mathbf{a}-1}}{x_i}.$$

Given that $\mathbf{a}-1 \in [0,1]$, $\mathbf{w}(\cdot)$ decreasingly weighs the observations which are far from the remaining bulk of the data.

In the same way, the estimator of the scale parameter is given by

$$\mathbf{s}_n = \sum_{i=1}^n (x_i - \mathbf{m}_n)^2 \tilde{\mathbf{w}}_i = 0, \quad (5)$$

where

$$\mathbf{w}(z_i) = \frac{z_i^{\mathbf{a}} - 1}{z_i^2 - 1}.$$

The considerations expressed for $\mathbf{w}(\cdot)$ in \mathbf{m}_n still hold.

In order to evaluate the efficiency of the ML estimators, we need to calculate the inverse of the information matrix defined as (Agrò, 1995)

$$I = \begin{bmatrix} I_{mm} & I_{ms} & I_{ma} \\ I_{ms} & I_{ss} & I_{sa} \\ I_{ma} & I_{sa} & I_{aa} \end{bmatrix}$$

whose elements are equal to

$$I_{mm} = -E \left[\frac{\partial^2 \ln f(x)}{\partial \mathbf{m}^2} \right] = \frac{(\mathbf{a}-1)\mathbf{a}^{\frac{\mathbf{a}-2}{\mathbf{a}}}}{\mathbf{s}^2} \frac{\Gamma\left(\frac{\mathbf{a}-1}{\mathbf{a}}\right)}{\Gamma\left(\frac{1}{\mathbf{a}}\right)}, \quad (6)$$

$$I_{ms} = -E \left[\frac{\partial^2 \ln f(x)}{\partial \mathbf{m} \partial \mathbf{s}} \right] = 0, \quad (7)$$

$$I_{ma} = -E \left[\frac{\partial^2 \ln f(x)}{\partial \mathbf{m} \partial \mathbf{a}} \right] = 0, \quad (8)$$

$$I_{ss} = -E \left[\frac{\partial^2 \ln f(x)}{\partial \mathbf{s}^2} \right] = \frac{\mathbf{a}}{\mathbf{s}^2}, \quad (9)$$

$$I_{sa} = -E \left[\frac{\partial^2 \ln f(x)}{\partial \mathbf{s} \partial \mathbf{a}} \right] = -\frac{1}{\mathbf{a}\mathbf{s}} \left[\ln(\mathbf{a}) + \mathbf{a} + \Psi\left(\frac{1}{\mathbf{a}}\right) \right], \quad (10)$$

$$I_{aa} = -E \left[\frac{\partial^2 \ln f(x)}{\partial \mathbf{a}^2} \right] = -\frac{\mathbf{a} + \mathbf{a}^2 + 1}{\mathbf{a}^3} + \frac{\mathbf{a} + 1}{\mathbf{a}^4} \Psi'\left(\frac{1}{\mathbf{a}}\right) + \frac{1}{\mathbf{a}^3} \left[\ln(\mathbf{a}) + \mathbf{a} + \Psi\left(\frac{1}{\mathbf{a}}\right) \right]^2. \quad (11)$$

The inverse of the information matrix is given by

$$I^{-1} = K^{-1}(\mathbf{a}) \begin{bmatrix} K(\mathbf{a})/I_{mm} & 0 & 0 \\ 0 & I_{aa} & -I_{sa} \\ 0 & -I_{sa} & I_{ss} \end{bmatrix} \quad (12)$$

where

$$K(\mathbf{a}) = \frac{-\mathbf{a}^3 - \mathbf{a}^2 - \mathbf{a} + \Psi'\left(\frac{1}{\mathbf{a}}\right)(1 + \mathbf{a})}{\mathbf{a}^3 \mathbf{s}^2}.$$

It is interesting to evaluate the elements of I^{-1} for $\alpha=1,2$ in order to compare the asymptotic variance of the location and scale estimators of the g.n.d. with the variance of the corresponding estimators under the Laplace and the Normal distribution. For $\alpha=1$ we have that

$$I^{-1} = \begin{bmatrix} \mathbf{s}^2 & 0 & 0 \\ 0 & 1.62\mathbf{s}^2 & 1.45\mathbf{s} \\ 0 & 1.45\mathbf{s} & 3.45 \end{bmatrix}$$

While, for $\alpha=2$ we have that

$$I^{-1} = \begin{bmatrix} \mathbf{s}^2 & 0 & 0 \\ 0 & 1.19\mathbf{s}^2 & 3.58\mathbf{s} \\ 0 & 3.58\mathbf{s} & 19.90 \end{bmatrix}$$

When we use the g.n.d. to estimate the location parameter, in both cases for $\alpha=1,2$ the asymptotic variance of the estimator does not change. On the other hand, regarding the σ estimator of the g.n.d., for $\alpha=2$ its asymptotic variance is equal to $(1.19/n)\sigma^2$, that is bigger than the variance obtained for the normal r.v $(0.5/n)\sigma^2$; and, in the same way, for $\alpha=1$, the asymptotic variance of the estimator is $(1.62/n)\sigma^2$ while it is $(1/n)\sigma^2$ for the Laplace distribution. We can conclude that by using the g.n.d. as a model to estimate the location and scale parameters of the Normal and Laplace r.v.'s, the estimators μ_n and σ_n are still uncorrelated but we have a loss of efficiency caused by a

larger variance for σ_n . Furthermore, we have to estimate the parameter α , positively correlated to σ .

3. The robustness properties of maximum likelihood estimators

In order to study the behaviour of the maximum likelihood estimators of the parameters of the g.n.d., we have to derive the influence functions (IF) (Hampel *et al*, 1986: p. 230) of the estimators and study their properties.

For ML estimators the IF is given by

$$IF(x, T, F) = -I^{-1} \mathbf{y}(x, t(F)). \quad (13)$$

In the case of the g.n.d. (13) yields

$$IF(x, \mathbf{m}F) = \mathbf{s} |z|^{a-1} \text{sign}(z), \quad (14)$$

$$IF(x, \mathbf{s}, F) = -K^{-1}(\mathbf{a}) [I_{aa} \mathbf{y}(x, \mathbf{s}) - I_{sa} \mathbf{y}(x, \mathbf{a})], \quad (15)$$

$$IF(x, \mathbf{a}, F) = -K^{-1}(\mathbf{a}) [-I_{sa} \mathbf{y}(x, \mathbf{s}) + I_{ss} \mathbf{y}(x, \mathbf{a})]. \quad (16)$$

Figure 1 shows the IF of μ for discrete values of $\alpha=1, 1.2, 1.5, 1.8, 2$. For $\alpha=1$, the IF is bounded as this is the case of the Laplace distribution where the median is the estimator of the location parameter. On the other hand, for $\alpha=2$, the g.n.d. reduces to the Normal r.v., so the estimator of the location parameter is the sample mean, whose IF is unbounded. For values of α between 1 and 2 the IF's of μ are still unbounded but they go to infinity at a slower rate as α approaches 1. This is due to the fact that, as showed by (4), observations located far away from the mean have a smaller weight in the estimation process that they would get when the location parameter is estimated by the mean. As a drawback, as α approaches 1, the IF becomes steeper and steeper, so that the estimator becomes extremely sensitive to the rounding errors of the data which are located around the location parameter.

Figure 2 shows the IF of the scale parameter for discrete values of $\alpha=1,1.2,1.5,1.8,2$. The corresponding estimator appears very sensitive to outlying observations. In particular, when α is greater than 1.5, the IF diverges rather quickly.

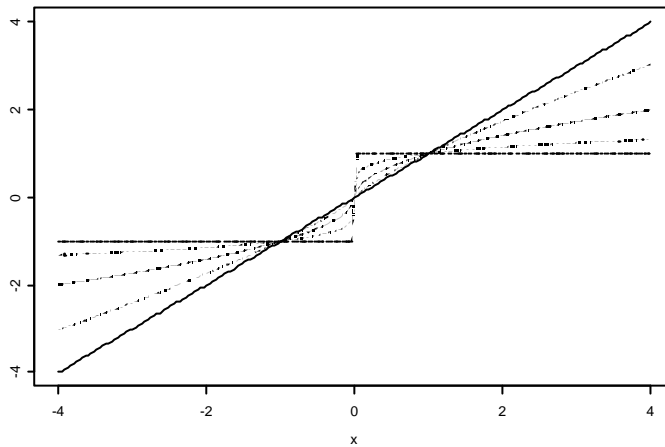


Figure.1. The IF of \mathbf{m} under the g.n.d.

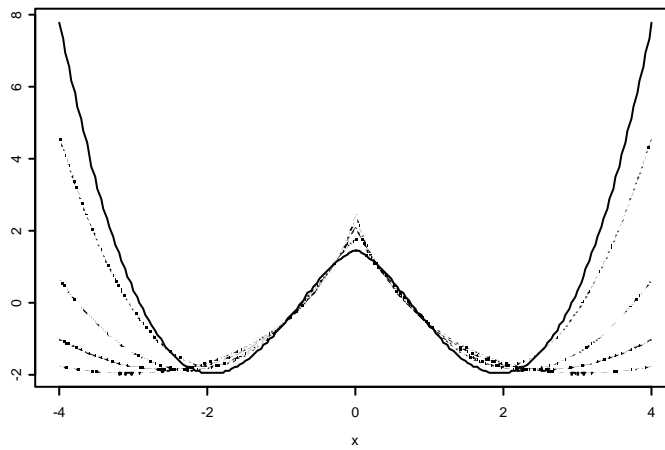


Figure 2. The IF of \mathbf{s} under the g.n.d.

Figure 3 shows the IF for the parameter α for discrete values of $\alpha=1,1.2,1.5,1.8,2$. The IF of α is close in shape to the IF of σ but, especially when α is greater than 1.5, it increases at a rather high rate.

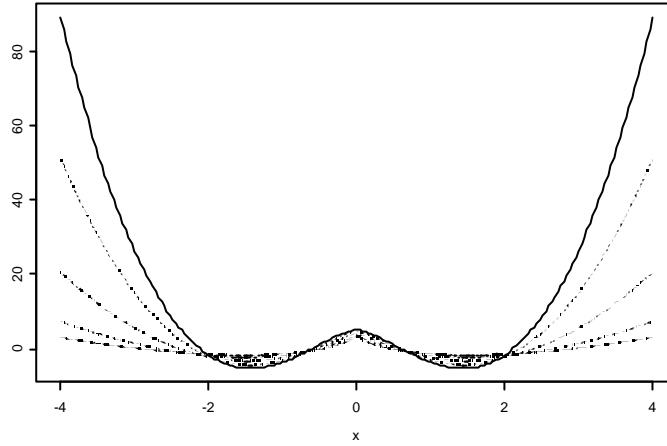


Figure 3. The IF of α under the g.n.d.

4. A comparison between the generalized normal distribution and alternative approaches

In the context of location-scale models, the natural competitor to the estimators obtained under the g.n.d. is the Huber estimator proposal II (Hampel *et al.*, 1986). The $\mathbf{y}(\cdot)$ function for the location parameter is defined as

$$\mathbf{y}_m(x) = x \cdot \min\left(1, \frac{b}{|x|}\right)$$

for $0 < b < \infty$, while for the scale parameter the $\mathbf{y}(\cdot)$ function is equal to

$$\mathbf{y}_s(x) = \mathbf{y}_m(x)^2 - \mathbf{b},$$

where

$$\mathbf{b} = \mathbf{c}_3^2(b^2) + b^2[1 - \mathbf{c}_1^2(b^2)] .$$

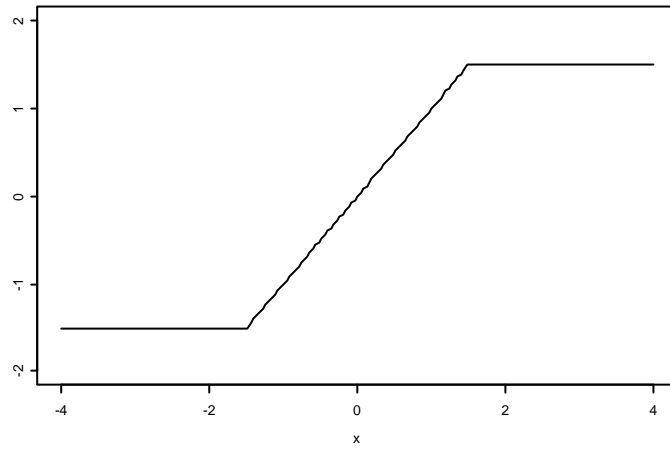


Figure 4. The $\mathbf{y}(\cdot)$ function of Huber estimator of the location parameter

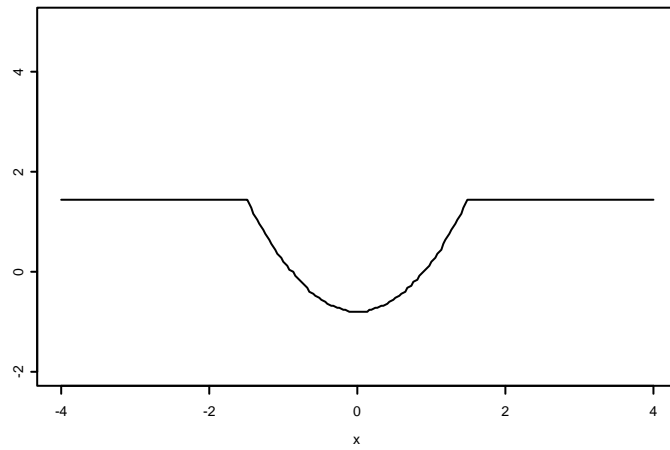


Figure 5. The $\mathbf{y}(\cdot)$ function of Huber estimator of the scale parameter

Figures 4 and 5 show the plots of the $\mathbf{y}(\cdot)$ function of the Huber estimator proposal II for the location and scale estimators, for $b=1.5$.

Although the IF of the estimators obtained under the g.n.d. and that of the Huber estimator proposal II are not comparable, as they depend on the model, an insight on the relationship between the two approaches can be obtained by comparing their $\mathbf{y}(\cdot)$ functions for values of $b=\alpha$ (Figures 6a-e and Figures 7a-e).

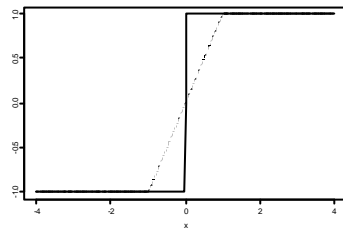


Figure 6a. Huber and ML estimators of \mathbf{m} for $b=\mathbf{a}=1$.

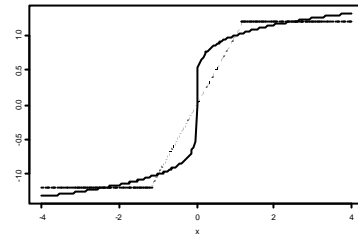


Figure 6b. Huber and ML estimators of \mathbf{m} for $b=\mathbf{a}=1.2$

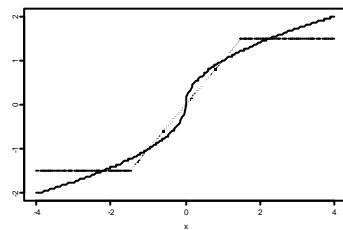


Figure 6c. Huber and ML estimators of \mathbf{m} for $b=\mathbf{a}=1.5$.

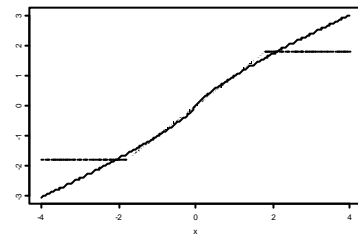


Figure 6d. Huber and ML estimators of \mathbf{m} for $b=\mathbf{a}=1.8$

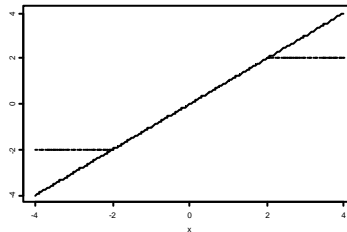


Figure 6e. Huber and ML estimators of \mathbf{m} for $b=\mathbf{a}=2$.

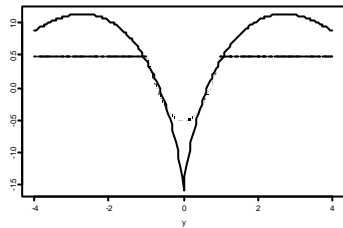


Figure 7a. Huber and ML estimators of \mathbf{s} for $b = \mathbf{a} - 1$.

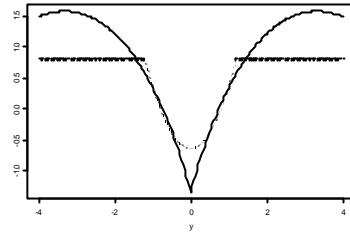


Figure 7b. Huber and ML estimators of \mathbf{s} for $b = \mathbf{a} - 1.2$.

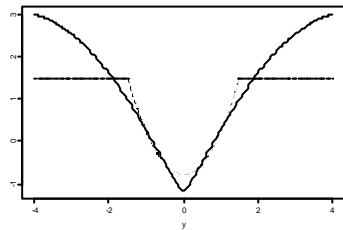


Figure 7c. Huber and ML estimators of \mathbf{s} for $b = \mathbf{a} - 1.5$.

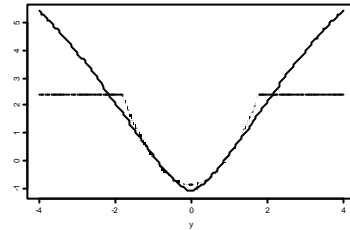


Figure 7d. Huber and ML estimators of \mathbf{s} for $b = \mathbf{a} - 1.8$.

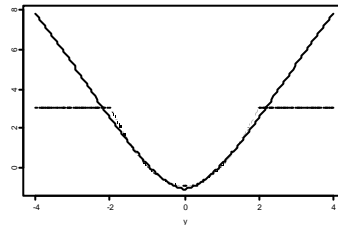


Figure 7e. Huber and ML estimators of \mathbf{s} for $b = \mathbf{a} - 2$.

For $b=\alpha=1$, the estimator of the location parameter obtained under the g.n.d. is more sensitive to rounding errors than the Huber estimator. In fact the former, being a signum function, has a discontinuity for $x=0$ while the latter varies in the range $[-1,1]$ for $x \in [-1,1]$. For $|b|>1$ the two estimators are equivalent. For $b=\alpha=2$, the $\mathbf{y}(\cdot)$ functions coincide for $x \in [-2,2]$, i.e. for the central 95% of the normal distribution, then it goes to infinity for the ML estimator while for the Huber's estimator it is bounded. For intermediate values of $1 < b = \alpha < 2$, the $\mathbf{y}(\cdot)$ functions are roughly the same within the range $x \in [-b,b]$. The overlapping improves as far as $b = \alpha$ tend to 2.

As regards the scale parameter, the $\mathbf{y}(\cdot)$ functions tend to overlap as far as $b = \alpha$ tend to 2, for values of x in the range $[-b,b]$, then the Huber estimator $\mathbf{y}(\cdot)$ is bounded while for the g.n.d. the $\mathbf{y}(\cdot)$ diverges.

In this section we also compare the robustness properties of the location and scale estimators of the g.n.d. with those of the t distribution. Many authors suggest to use the t distribution as a model for the data, instead of the normal distribution, because it is symmetric, can be parametrized by location and scale parameters (Fraser, 1976) and has the great advantage to fit distribution with heavier tails. In this respect it can be regarded as a robust model as well as the g.n.d.. A comparison between these two models has been conducted by D'Agostino and Lee (1977), who focused on the robustness of location estimators under changes of kurtosis for the two distributions.

The density function of the t distribution with ν degrees of freedom, location parameter μ and scale parameter σ , is

$$g(x; \mathbf{m}, \mathbf{s}, \mathbf{n}) = \frac{\Gamma\left(\frac{\mathbf{n}+1}{2}\right) \left(1 + \frac{(x-\mathbf{m})^2}{\mathbf{n}\mathbf{s}^2}\right)^{\left(\frac{-\mathbf{n}-1}{2}\right)}}{\mathbf{s}\sqrt{\pi}\Gamma\left(\frac{\mathbf{n}}{2}\right)}$$

for $-\infty < x < +\infty$, while the score functions are given by

$$s_1(x, \mathbf{m}) = \frac{(\mathbf{n} + 1)(x - \mathbf{m})}{\mathbf{n}\mathbf{s}^2 \left(1 + \frac{(x - \mathbf{m})^2}{\mathbf{n}\mathbf{s}^2} \right)},$$

$$s_1(x, \mathbf{s}) = \frac{(\mathbf{n} + 1)(x - \mathbf{m})^2}{\mathbf{n}\mathbf{s}^3 \left(1 + \frac{(x - \mathbf{m})^2}{\mathbf{n}\mathbf{s}^2} \right)} - \frac{1}{\mathbf{s}}$$

By remembering that the $\mathbf{y}(\cdot)$ function is equal to the score function, we plotted the two score functions for different values of $\nu=3,5,10$. Figure 8 shows the $\mathbf{y}(\cdot)$ functions of the location estimator which are re-descending. The $\mathbf{y}(\cdot)$ function for the scale parameter is depicted in Figure 9, as $|x| \rightarrow \infty$ it approaches ν , the d.f. of the referring t distribution, while the $\mathbf{y}(\cdot)$ function of the ML estimator diverges for $|x| \rightarrow \infty$. Compared to the $\mathbf{y}(\cdot)$ functions of the estimators of the g.n.d. parameters, they show a better performance given that they are bounded functions and therefore more robust in the presence of outliers.

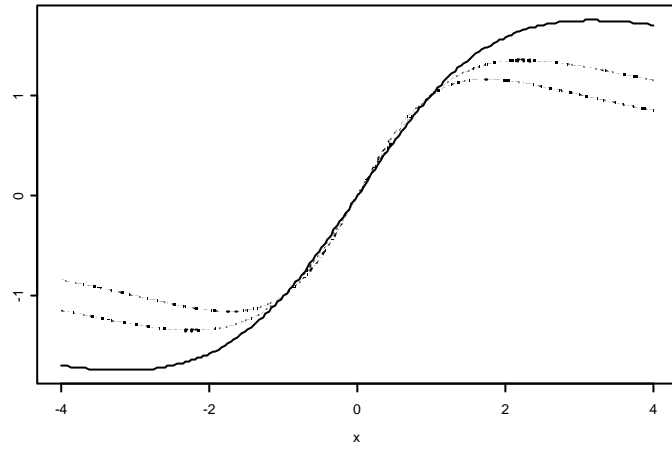


Figure 8. The $\mathbf{y}(\cdot)$ function of the location estimator of the t distribution

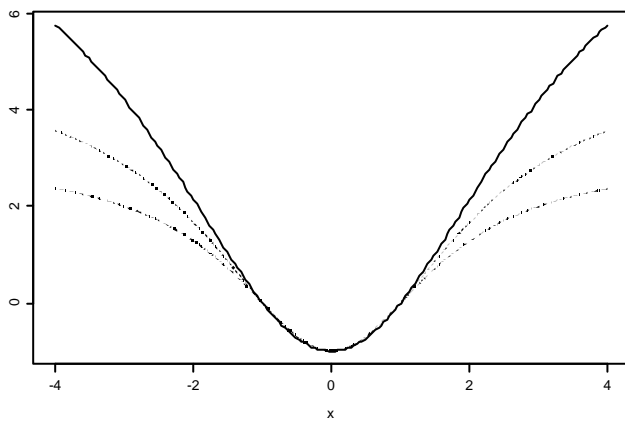


Figure 9. The $\mathbf{y}(\cdot)$ function of the scale estimator of the t distribution

5. Final remarks

In this paper we analysed some robustness properties of the generalized normal distribution, for $\alpha \in [1, 2]$. The IF's for the ML estimators of the g.n.d. were derived and plotted for different values of α . The advantage to use the g.n.d. is the possibility to refer to a well defined model, even if this leads to a loss of efficiency for the asymptotic variance of the scale estimator which increases as α increases. Furthermore the ML estimators are sensitive to outlying observations. In fact their IF's are unbounded, mostly if compared to the Huber estimator proposal II and to the location and scale estimators of the t distribution. From this point of view, it seems preferable to refer to the Student t distribution as a robust model.

Here we focused on the properties of the ML estimators of the location and scale parameters, considering α fixed. Possible further research can be carried out by estimating α from a real data set in order to test the fitness of the data to the Gaussian distribution.

Acknowledgments: Research partially supported by MURST funds.

References

Agrò G. (1995) Maximum likelihood estimation for the exponential power function parameters, *Communication in Statistics - Simulation*, 24, 523-536.

Azzalini A. (1985) A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, 12, 171-178.

Azzalini A. (1986) Further results on a class of distributions which includes the normal ones, *Statistica*, 46, 199-208.

Azzalini A. Della Valle A. (1996) The multivariate skew normal distribution, *Biometrika*, 83, 715-726.

Azzalini A. Capitanio A. (1999) Statistical applications of the multivariate skew normal distribution, *Journal of the Royal Statistical Society B*, 61, 579-602.

Box G. E. P. (1953) A note on regions of kurtosis, *Biometrika*, 40, 465-468.

Chiodi M. (2000) Le curve normali di ordine p nell'ambito delle distribuzioni di errori accidentali: una rassegna dei risultati e problemi aperti per il caso univariato e per quello multivariato, *Atti della XL Riunione Scientifica della SIS*.

D'Agostino R. B. Lee A. F. S. (1977) Robustness of location estimators under changes of population kurtosis, *Journal of the American Statistical Association*, 72, 393-396.

Hampel F. R. Ronchetti E. Rousseeuw P. J. Stahel W. A. (1986) *Robust Statistics. The Approach Based on the Influence Function*, Wiley & Sons, New York.

Hill M. A. Dixon W. J. (1982) Robustness in real life: a study of clinical laboratory data, *Biometrics*, 38, 377-396.

Hogg R. V. (1974) Adaptive robust procedure: a partial review and some suggestions for future applications and theory, *Journal of the American Statistical Association*, 69, 909-927.

Mineo A. Vianelli S. (1980) Prontuari delle probabilità integrali delle curve normali di ordine r , *Istituto di Statistica dell'Università di Palermo*.

Staudte R. G. Sheather S. J. (1990) *Robust Estimation and Testing*, Wiley & Sons, New York.

Subbotin M. T. (1923) On the law of frequency errors, *Mathematicheskii Sbornik*, 31, 296-301.

Vianelli S. (1963) La misura della variabilità condizionata in uno schema generale delle curve normali di frequenza, *Statistica*, 23, 447-474.