Parameter estimation for a generalized Gaussian distribution
Beelen, T.G.J.; Dohmen, J.J.

Published: 01/12/2015

Document Version
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the author’s version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.
Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 11. Dec. 2018
Parameter estimation for a generalized Gaussian distribution

by

T.G.J. Beelen, J.J. Dohmen
1 General Introduction

In 2014 several companies, institutes and universities in Europe started a joined project named ‘CORTIF’ (Coexistence Of RF Transmissions In the Future). See the website http://cortif.xlim.fr for more details.

The CORTIF partners in the Netherlands are NXP Semiconductors (NXP-NL), IMEC, Technolution, and Eindhoven University of Technology, departments EE and Mathematics & Computer Science (CASA group). In 2015 CASA started a close co-operation with NXP-NL on the subject of so-called ‘trimmed distributions’. The underlying background is as follows.

As might be known NXP is a semiconductor company designing and producing a large variety of integrated circuits (ICs) for a wide range of products. Of course, NXP aims at a high yield as well as good IC-performance. Therefore, most of the produced ICs are measured and checked against the specifications. For some RF applications a specific output signal of produced ICs can be measured. In case of out-of-spec that IC can be ‘tuned’ by adapting a built-in resistor such that the IC does meet the specs eventually.

Clearly, when looking at the statistical distribution of a large amount of such ‘tuned’ ICs, this distribution will not be standard Gaussian anymore. In fact, the corresponding probability density function has a more flat shape than in case of standard Gaussian. In order to optimize the yield it would be beneficial to have a statistical model for the observed distribution. One of the promising approaches is to use the so-called generalized Gaussian distribution function and to estimate its 3 defining parameters using the maximum likelihood technique.

In this report we will investigate this approach in detail and propose a numerical fast and reliable method for computing these parameters.
2 Abstract

In this report we analyze and solve the problem of computing a reliable estimation of the parameters of a generalized gaussian distribution function. It is well-known that this problem can be formulated as a maximization problem of the maximum likelihood function associated with the distribution function. Next this can easily be transformed to finding the zero of a function \( g(\beta) \) where \( \beta \) is the unknown exponent parameter of the distribution. The evaluation of \( g(\beta) \) requires a few summations of a (large) number of randomly generated values \( x_i \) from a generalized Gaussian distribution. This will be indicated by \( g(\beta) = g(\beta; x_i) \).

Among other approaches, this zero can be computed numerically using Newton’s method. However, we show that this approach is far from trivial since \( g(\beta) \) behaves in an extreme way near its zero (if existing). In fact, \( g(\beta) \) is rather steep at one side of its zero, but is almost completely flat and often (very) close to zero at the other side. Hence, we include an extensive and detailed analysis of all formulas involved for a full understanding of the behaviour of \( g(\beta) \).

We propose a new procedure consisting of several runs where per run \((k)\) a different set of random values \( x_i^{(k)} \) is taken for computing the zero \( \beta^{(k)} \) of \( g(\beta; x_i^{(k)}) \). Then the average of all these zeros \( \beta^{(k)} \) is considered as a good estimate of the exponent parameter \( \beta \).

Finally, we show that this procedure indeed gives reliable results when applied to empirical data.
3 Mathematical problem description

The starting point is the paper "Generalized normal distribution" in [1]. The problem is to find a numerical reliable and accurate method for estimating the parameters of the following probability density function

\[
f(x) = \frac{\beta}{2\alpha \Gamma(1/\beta)} \exp\left(-\left(\frac{|x-\mu|}{\alpha}\right)^\beta\right)
\]

(1)

In the figures 1 and 2 the probability density function (1) and the cumulative density function \(cdf\) are shown for \(\alpha = 1\) and a few values of the parameter \(\beta\). It is easily seen that in case \(\alpha = 1\) the major part of the x-values where \(tol < cdf(x) < 1 - to\) is in the range \((-2, 2)\) for \(2 \leq \beta \leq 8\). Here \(tol\) is a relatively small positive number, say \(tol \leq 10^{-3}\).

![Figure 1: Generalized Gaussian density functions with \(\mu = 0\) and \(\alpha = 1\).](image)

We assume \(N\) points \(x_i\) be given in the interval \([U, V]\). The points \(x_i\) can be randomly distributed corresponding to (1).

The likelihood function corresponding to \(f(x)\) is given by

\[
\mathcal{L} = \prod_{i=1}^N f(x_i) = \left(\frac{\beta}{2\alpha \Gamma(1/\beta)}\right)^N \exp\left(-\sum_{i=1}^N \left(\frac{|x_i-\mu|}{\alpha}\right)^\beta\right) = (2\alpha)^{-N} \beta^N (\Gamma(1/\beta))^{-N} \exp\left(-\sum_{i=1}^N \left(\frac{|x_i-\mu|}{\alpha}\right)^\beta\right)
\]

(2)
The log maximum likelihood function $L$ is given by

$$L = \log(\mathcal{L}) = -N \log(2\alpha) + N \log(\beta) - N \log((\Gamma(1/\beta)) - \sum_{i=1}^{N} \left(\frac{|x_i - \mu|}{\alpha}\right)^{\beta}$$  \hspace{1cm} (3)

In figure 3 the function $L$ is shown for $\mu = 0$ and a few values of $\alpha$ and $\beta$.

The necessary conditions to maximize $L$ are

$$\frac{\partial L}{\partial \alpha} = 0, \quad \frac{\partial L}{\partial \beta} = 0 \quad \text{and} \quad \frac{\partial L}{\partial \mu} = 0. \hspace{1cm} (4)$$

In the rest of this document we assume that $\mu$ is known. Hence, we can ignore the condition $\frac{\partial L}{\partial \mu} = 0$.

However, for completeness, the formula for $\frac{\partial L}{\partial \mu}$ are given in appendix ??.

The condition $\frac{\partial L}{\partial \alpha} = 0$ gives

$$\alpha = \hat{\alpha} = \left(\frac{\beta}{N} \sum_{i=1}^{N} |x_i - \mu|^{\beta}\right)^{1/\beta} \hspace{1cm} (5)$$

The condition $\frac{\partial L}{\partial \beta} = 0$ gives

$$\frac{\partial L}{\partial \beta} = \frac{1}{\beta} + \frac{\Psi(1/\beta)}{\beta^2} - \frac{1}{N} \sum_{i=1}^{N} \left|\frac{x_i - \mu}{\alpha}\right|^{\beta} \log \left|\frac{x_i - \mu}{\alpha}\right| = 0 \hspace{1cm} (6)$$

The typical behaviour of $\partial L/\partial \beta$ is shown in the figures (4) and (5) below. We see

---

1 Throughout this document we will indicate the ‘ln’-function by ‘log’.
from figure (5) that $\partial L / \partial \beta$ for $\beta \in [6.3, 6.4]$ which corresponds to $\alpha \approx 4.55$ using formula (5).

It is easily seen from these figures that solving $\partial L / \partial \beta = 0$ is not straightforward, especially in case $\alpha > 2$. This suggests that further investigation of $\partial L / \partial \beta$ is needed. We will proceed with this in the next sections. First we mention the approach in [1] for computing the maximum likelihood function $L$. Their main idea is taking expression (6) for $\partial L / \partial \beta = 0$ and next substituting $\hat{\alpha}$ given in (5). Then, after multiplying with the factor $\beta / N$, we get

$$g(\beta) = 0 \quad \text{where}$$

$$g(\beta) = 1 + \frac{\Psi(1/\beta)}{\beta} - \frac{\sum_{i=1}^{N} |x_i - \mu|^\beta \log |x_i - \mu|}{\sum_{i=1}^{N} |x_i - \mu|^\beta} + \frac{\log \left( \frac{\beta}{N} \sum_{i=1}^{N} |x_i - \mu|^\beta \right)}{\beta}, \quad (8)$$
Equation (7) is solved using Newton-Raphson’s method with

\[
  g'(\beta) = -\frac{\Psi'(1/\beta)}{\beta^2} - \frac{\Psi(1/\beta)}{\beta^3} + \frac{1}{\beta^2} \\
  - \frac{\sum_{i=1}^{N} |x_i - \mu|^{\beta} (\log|x_i - \mu|)^2}{\sum_{i=1}^{N} |x_i - \mu|^{\beta}} - \left( \frac{\sum_{i=1}^{N} |x_i - \mu|^{\beta} \log|x_i - \mu|}{\sum_{i=1}^{N} |x_i - \mu|^{\beta}} \right)^2 \\
  + \frac{\sum_{i=1}^{N} |x_i - \mu|^{\beta} \log|x_i - \mu|}{\beta \sum_{i=1}^{N} |x_i - \mu|^{\beta}} - \frac{\log \left( \frac{\sum_{i=1}^{N} |x_i - \mu|^{\beta}}{\sum_{i=1}^{N} |x_i - \mu|^{\beta}} \right)}{\beta^2}.
\]

(9)

The typical behaviour of the functions \( g(\beta) \) and \( g'(\beta) \) is shown in figure (6).

Remark

The zero \( \beta_0 \) in figure 6 is computed by applying Newton-Raphson’s method using \( g(\beta) \) and \( g'(\beta) \) to the interval \( x = [-3 : 0.01 : 5] \). As usual, \( \mu = 0 \). So, we have

\[
  \beta_{k+1} = \beta_k - \frac{g(\beta_k)}{g'(\beta_k)}, \quad k \geq 1 \text{ and } \beta_1 = 3.
\]

(10)

We used as stopping criterion:

\[
  |\beta_{k+1} - \beta_k| < e_{\text{abs}} + e_{\text{rel}} \cdot |\beta_k| \quad \text{with } e_{\text{abs}} = 1.0 \times 15, \quad e_{\text{rel}} = 1.0 \times 6.
\]

(11)
The numerical results are shown in the table in figure 5. The function \( \partial L / \partial \beta \) where \( \partial L / \partial \beta = 0 \) for \( \beta_0 \approx 6.3 \) and \( \hat{a}(\beta_0) \approx 4.55 \).

It should be noticed that this method is sensitive to the choice of the interval for the x-values. This is shown in figure 8. Notice that this figure suggests that the function \( g(\beta) \) has no zeros at all when the x-values are in the interval \([-3 : 0.01 : 3]\). If so, then clearly the Newton-Raphson method will fail.

In Appendix C it is shown that

\[
\lim_{\beta \to \infty} g(\beta) = 0 \quad (12)
\]

4 Behaviour of the function \( \hat{\alpha}(\beta) \)

For convenience, we recall formula (5) for the function \( \hat{\alpha}(\beta) \)

\[
\hat{\alpha}(\beta) = \hat{\alpha}_{\text{sum}}(\beta) = \left( \frac{\beta}{N} \sum_{i=1}^{N} |x_i - \mu|^{\beta} \right)^{1/\beta} \quad (13)
\]
The functions $g(\beta)$ and $g'(\beta)$. The marker denotes $\beta_0 \approx 6.4205$ with $g(\beta_0) = 0$ and $\hat{a}(\beta_0) \approx 4.5637$. Recall also figure 5.

The typical behaviour of $\hat{\alpha}(\beta)$ is depicted in figure 9. We see that the function $\hat{\alpha}(\beta)$ is monotonically increasing for $\beta > 0$.

It is shown in (40) that the function $\hat{\alpha}_{\text{sum}}(\beta)$ can be approximated by

$$\hat{\alpha}_{\text{analyt}}(\beta) \approx M \left( \frac{\beta}{\beta + 1} \right)^{1/\beta}, \quad N \to \infty$$

where $M = \max_{1 \leq i \leq N} |x_i - \mu|$ \hfill (14)

When evaluating formulas (13) and (14) numerically (with $\mu = 0$, $|x_{i+1} - x_i| = 0.001, -4 \leq x_i \leq 4$) we found $|\hat{\alpha}_{\text{sum}}(\beta) - \hat{\alpha}_{\text{analyt}}(\beta)| < 5e^{-4}$ for all $0 < \beta \leq 10$. See also figure 10. Furthermore, it is shown in Appendix D.1 that

$$\lim_{\beta \to 0} \hat{\alpha}(\beta) = 0 \quad \text{and} \quad \lim_{\beta \to \infty} \hat{\alpha}(\beta) = M \quad \text{where} \quad M = \max_{1 \leq i \leq N} |x_i - \mu|$$

(15)

Notice that $\hat{\alpha}(\beta) < M$ for all $\beta > 0$ and that $\hat{\alpha}(\beta)$ is quite sensitive to the choice of the endpoint of the x-interval.

### 4.1 Asymptotic behaviour of $g(\beta)$

We split the function $g(\beta)$ given by (8) into 3 parts as follows:

$$g(\beta) = g_1(\beta) - P(\beta | \mu) + Q(\beta | \mu)$$

where

$$(16)$$
Table: Newton-Raphson results for $g(\beta)$ for the x-interval $[-3 : 0.01 : 5]$.

<table>
<thead>
<tr>
<th>k</th>
<th>$g(\beta)$</th>
<th>$g'(\beta)$</th>
<th>$\beta_k$</th>
<th>$\beta_{i+1} - \beta_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.053152806467897</td>
<td>0.036172506943895</td>
<td>3.000000000000000</td>
<td>0.000000E+00</td>
</tr>
<tr>
<td>1</td>
<td>0.017539793552050</td>
<td>0.015289945849182</td>
<td>4.469425565398000</td>
<td>1.46943E+00</td>
</tr>
<tr>
<td>2</td>
<td>0.004836768693093</td>
<td>-0.007665901007709</td>
<td>5.616571129225360</td>
<td>1.14715E+00</td>
</tr>
<tr>
<td>3</td>
<td>0.000841939332000</td>
<td>-0.005146591060444</td>
<td>6.247519322820430</td>
<td>6.30948E-01</td>
</tr>
<tr>
<td>4</td>
<td>0.000043191215275</td>
<td>-0.004625770562097</td>
<td>6.411139587253290</td>
<td>1.63620E-01</td>
</tr>
<tr>
<td>5</td>
<td>0.0000000131961955</td>
<td>-0.004597529286314</td>
<td>6.420476672760490</td>
<td>9.33709E-03</td>
</tr>
<tr>
<td>6</td>
<td>0.000000000001241</td>
<td>-0.004597442701170</td>
<td>6.420505375558600</td>
<td>2.87028E-05</td>
</tr>
<tr>
<td>7</td>
<td>0.000000000000000</td>
<td>0.000000000000000</td>
<td>6.420505375828630</td>
<td>2.70031E-10</td>
</tr>
</tbody>
</table>

Figure 7: Table: Newton-Raphson results for $g(\beta)$ for the x-interval $[-3 : 0.01 : 5]$.

\begin{equation}
g_1(\beta) = 1 + \frac{\Psi(1/\beta)}{\beta}, \quad (17)
\end{equation}

\begin{equation}
P(\beta | \mu) = \frac{\sum_{i=1}^{N} |x_i - \mu|^\beta \log |x_i - \mu|}{\sum_{i=1}^{N} |x_i - \mu|^\beta} \quad \text{and} \quad (18)
\end{equation}

\begin{equation}
Q(\beta | \mu) = \frac{\log(\frac{\beta}{\pi} \sum_{i=1}^{N} |x_i - \mu|^\beta)}{\beta} \quad (19)
\end{equation}

Recalling (100) and (5) we notice that $g_1(\beta) = \beta \ast F(\beta)$ and $Q(\beta | \mu) = \log(\hat{a})$.

It is shown in Appendix C that

\begin{equation}
\lim_{\beta \to \infty} g_1(\beta) = 0, \quad \lim_{\beta \to \infty} P(\beta | \mu) = \log(M), \quad \text{and} \quad \lim_{\beta \to \infty} Q(\beta | \mu) = \log(M),
\end{equation}

where $M = \max_{1 \leq i \leq N} |x_i - \mu| \quad (20)$

Thus,

\begin{equation}
\lim_{\beta \to \infty} g(\beta) = 0 \quad (21)
\end{equation}

5 Numerical experiments

5.1 Procedure for parameter estimation of an empirical pdf

We will now consider a pdf obtained from numerical simulations or from measurements in practice. Such a pdf will be called an empirical pdf. We assume that this pdf can be modelled by a generalized Gaussian pdf where its parameters $\alpha$ and $\beta$ are still unknown, but that $\mu$ is known. We will determine these parameters using
the formulas given earlier. This is globally done using the procedure described below. We assume the data of the pdf to be modelled are given by the set

\[
\{(x_{0i}, y_{0i})| 1 \leq i \leq N\} \quad \text{where}
\]
\[
x_{0i} \in [X_{0 \text{min}}, X_{0 \text{max}}], \quad y_{0i} \in [Y_{0 \text{min}}, Y_{0 \text{max}}],
\]
\[
\text{mean } \mu_0 = \Sigma x_{0i} / N
\]

1. Transform the given data such that their mean is equal to 0:
   (set \( x_i = x_{0i} - \mu_0, \ 1 \leq i \leq N \), and set \( \mu = 0 \))

2. Determine the empirical pdf \( Epdf \) based on the given data:
   (use, for example, the Matlab function \( ksdensity.m \))

3. Compute the cumulative density function \( Ecdf \) corresponding to \( Epdf \)
   (for example, by numerical integration of \( Epdf \) using the trapezoid rule)

4. Generate a set of random variables \( x_{rnd,i} \) using \( Ecdf \)
   (see Appendix I)

5. Compute the parameter \( \beta \) as the zero \( \beta_0 \) of the function \( g(x) \) in (8), using \( x_{rnd,i} \)
Figure 9: The function $\hat{\alpha}(\beta)$ for $\beta > 0$ with $\lim_{\beta \to \infty} \hat{\alpha}(\beta) = V = \max_{1 \leq i \leq N} |x_i - \mu|$.

6. Compute the parameter $\alpha = \hat{\alpha}(\beta_0)$ with formula (5)

Let us now comment on this global procedure.

- Since the $x_{\text{rnd},i}$ are random values, they are not sorted in general. Therefore, the magnitudes of the successive values $|x_{\text{rnd},i} - \mu|^\beta$ may vary considerably, and might cause numerical inaccuracy when computing the sum $\Sigma |x_{\text{rnd},i} - \mu|^\beta$, as required in the computation of $g(x)$.

So, we suggest to first sort the values $x_{\text{rnd},i}$ to increasing order and then to use them in further calculations.

- As said before (see section 3) the computation of $g(x)$ is sensitive to the endpoint $M = \max_{1 \leq i \leq N} |x_{\text{rnd},i}|$ of the x-interval. So, the computed zero $\beta_0$ of $g(x)$ and next also $\hat{\alpha}(\beta_0)$ strongly depend on $M$.

Therefore, we propose to modify the procedure as follows

- First, carry out the steps 1 up to 3 as before
- Secondly, repeat the steps 4 and 5 a number of times $k = 1, \ldots, K$, and saving the values in run $k$, step 4 :
  
  $* \min_{\text{rnd}}^{(k)} = \min_{1 \leq i \leq N} \{x_{\text{rnd},i}^{(k)}\} \quad \text{and} \quad \max_{\text{rnd}}^{(k)} = \max_{1 \leq i \leq N} \{x_{\text{rnd},i}^{(k)}\}$,
in run \( k \), step 5:

\[ \beta^{(k)}_0 \text{ and } \hat{\alpha}^{(k)} = \alpha(\beta^{(k)}_0) \]

Finally, compute

the mean \( \hat{\alpha} = \frac{1}{K} \sum_0^{(k)} \alpha \), the mean \( \hat{\beta} = \frac{1}{K} \sum_0^{(k)} \beta \) and

\[ \text{rnd}_{\text{min}} = \min_{1 \leq k \leq K} \text{min}_{\text{rnd}}^{(k)}, \quad \text{rnd}_{\text{max}} = \max_{1 \leq k \leq K} \text{max}_{\text{rnd}}^{(k)} \]

Remarks

- Recall that the computation of \( \hat{\alpha} \) involves the computation of \( \Sigma|\gamma_k - \mu| \hat{\beta} \). This should be done as accurate as possible and preferably not be sensitive to the choice of random interval points.

- Notice that the interval \( I_{\text{rnd}} = [\text{rnd}_{\text{min}}, \text{rnd}_{\text{max}}] \) is the interval where all generated random values \( x_{\text{rnd},i}^k \) obtained in all runs \( k \) are located.
• Therefore, an alternative approach for $\hat{\alpha}$ might be calculating $\hat{\alpha} = \alpha(\hat{\beta})$
on the interval $[\text{rnd}_{min}, \text{rnd}_{max}]$, equidistantly subdivided by $R (\gg N)$ points. However, it turned out that this $\hat{\alpha}$-value was much larger than that obtained above, and gave a worse approximation of the empirical pdf (see also next section).

• An alternative for the equidistant subdivision of $I_{\text{rnd}}$ might be using all the random points $x_{\text{rnd},i}^{(k)}$, after having sorted them in increasing order. Clearly, this requires saving the generated values in each run $k$.

5.2 Application of the estimation procedure

We want to apply the described procedure to an empirical pdf obtained when simulating measurements of a voltage output signal of an industrial IC design. The pdf is calculated using the Matlab function $k$sdensity.m ansd is shown in figure 11.

![Figure 11: An empirical pdf obtained from industrial data.](image)

By numerical integration of this pdf (using the trapezoidal method) we find the empirical cdf as in figure 12. Notice that in this figure we hardly can recognize the curvatures due to the local peaks in the corresponding pdf. See Appendix G showing a similar figure in a standard Gaussian case where the peaks in the pdf are even more pronounced than in fig.11. Once having the empirical cdf ($Ecdf$) available, we start the procedure described above. In the figures below a few typical examples are given. Per example we we have chosen 20 runs, and per run we generated 1000 random numbers distributed conform the Ecdf. The final values $\hat{\beta}$ and $\hat{\alpha}$ for $\beta$ and $\alpha$ are
found by averaging all computed $\beta_k^k$ and $\alpha_k^k = \phi_{\alpha}(\beta_k^k)$, respectively. Similarly for $\hat{\mu}$.

In order to check the fit of the generalized Gaussian pdf $GG(x|\mu, \alpha, \beta)$ defined by the computed $\hat{\alpha}, \hat{\beta}, \hat{\mu}$ we show this pdf and the original empirical pdf simultaneously in figure 14. Moreover, in figure 15 we also add the plots of $GG(x|\hat{\mu}, \hat{\alpha} - 10\%, \hat{\beta})$ and $GG(x|\hat{\mu}, \hat{\alpha} + 10\%, \hat{\beta})$ in order to see the influence of a $10\%$ change in $\alpha$ to the deviation of generalized Gaussian pdf from the $Epdf$.

Clearly, the curve corresponding to $\hat{\alpha}$ fits best the $Epdf$.

This statement can also be checked quantitatively, using the MSE-criterion:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} |Epdf(x_i) - GG(x_i|\mu, \alpha, \beta)|^2 \quad (23)$$

We chose $N=100$ equidistant points $x_i$ in the given $x$-interval and computed the MSE-values, see table in figure 16.
Figure 13: Computation of 50 $\beta$'s using random values from the Ecdf.

Figure 14: The original Epdf and the estimated generalized Gaussian pdf $GG(x | \hat{\mu}, \hat{\beta})$. 

15
Figure 15: The original Epdf and 3 estimated generalized Gaussian pdf’s.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\alpha}$</td>
<td>0.00454</td>
</tr>
<tr>
<td>$\hat{\alpha} + 10%$</td>
<td>0.00499</td>
</tr>
<tr>
<td>$\hat{\alpha} - 10%$</td>
<td>0.00409</td>
</tr>
</tbody>
</table>

Figure 16: The MSE-values for the 3 estimated generalized Gaussian pdf’s.
A Approximation of a sum $\frac{1}{N}\sum_{i=1}^{N} f(x_i)$ by an integral

Problem
Consider the interval $\mathcal{I} = [U, V]$ with $0 \leq U < V$. Let $f(x)$ be an integrable function on this interval and let $\{x_i | 1 \leq i \leq N\}$ be a finite set of points in $\mathcal{I}$. Our goal is to find an approximation of the sum $\frac{1}{N}\sum_{i=1}^{N} f(x_i)$ in terms of an analytical integral, in case $N \to \infty$.

Approach to problem solving
We have $x_i \in [U, V]$ with $U \leq 0 < V$, for all $1 \leq i \leq N$. We can assume that these $x_i$ are ordered such that $x_1 = U, \cdots, x_N = V$. Furthermore, we assume that the $x_i$ are equidistant on $[U, V]$. In this case we can derive some results concerning the approximation of certain summations involving $x_i$ by an integral. A more detailed and more accurate analysis is given in a next section. In the next appendix we consider the situation where the $x_i$ are not equidistantly but randomly distributed.

Let $m = \min_{1 \leq i \leq N} |x_i|$, $M = \max_{1 \leq i \leq N} |x_i|$.
Clearly, $m < M$ (indeed, if $m = M$ then all $x_i$ are equal to $M$ which is not the case).
Let $v_i = \frac{x_i}{M}$. Then $x_i = M v_i$ and $0 \leq \frac{m}{M} \leq |v_i| \leq 1$.
Since the $x_i$ are assumed to be equidistant we see that $|x_{i+1} - x_i|$ is independent of $i$. Hence, we can define $\Delta x = |x_{i+1} - x_i|$ and $\Delta v = |v_{i+1} - v_i| = \frac{M}{M}$.
Let $x = 0$ be in the interval $[x_k, x_{k+1}]$, so $x_k \leq 0 < x_{k+1}$ or $x_k < 0 \leq x_{k+1}$. Then
$$\Delta x = |x_{k+1} - x_k| = |x_k| + |x_{k+1}| \geq 2m \quad (24)$$
Let $L = V - U$. Then we have $N - 1$ intervals $[x_i, x_{i+1}]$, so
$$\Delta x = \frac{L}{N - 1} \quad (25)$$
Combining (24) and (25) we find $0 \leq \frac{m}{M} \leq \frac{L}{2M(N-1)}$. Thus,
$$\frac{m}{M} = O\left(\frac{1}{N}\right), \quad N \to \infty \quad (26)$$
Since the $x_i$ are equidistant on $[U, V]$, we see that the $v_i$ are equidistant on $[\frac{m}{M}, 1] \approx [0, 1]$. Hence,
$$\Delta v \approx \frac{1}{(N - 1)} \quad \text{and} \quad \frac{1}{N\Delta v} \approx 1, \quad N \to \infty \quad (27)$$
Let $f(|v_i|)$ be an integrable function of $|v_i|$ defined on the interval $[0, 1]$ and let $N$
be sufficiently large.

Then, using (27) we have

\[
\frac{1}{N} \sum f(|v_i|) = \frac{1}{N} \sum \left( f(|v_i|) \Delta v \right) \approx \sum \left( f(|v_i|) \Delta v \right)
\]

\[
\approx \int_{m/M}^{1} f(v) \, dv = \left( \int_{0}^{1} f(v) \, dv \right) \approx \int_{0}^{1} f(v) \, dv, \quad \text{if } N \to \infty
\]

(28)

Recalling (26) and provided that

\[
\int_{0}^{m/M} f(v) \, dv = O\left(\frac{1}{N}\right),
\]

formula (28) simplifies to

\[
\frac{1}{N} \sum f(|v_i|) \approx \int_{0}^{1} f(v) \, dv + O\left(\frac{1}{N}\right), \quad N \to \infty
\]

(29)

**Note:**

For use later, we also mention that in case $N \to \infty$

\[
\left( \frac{m}{M} \right)^{\beta + 1} = O\left(\frac{1}{N}\right), \quad 1 - \left( \frac{m}{M} \right)^{\beta + 1} = 1 + O\left(\frac{1}{N}\right),
\]

\[
\log \left( 1 - \left( \frac{m}{M} \right)^{\beta + 1} \right) = O\left(\frac{1}{N}\right),
\]

\[
\left( \frac{m}{M} \right)^{\beta + 1} \log \left( \frac{m}{M} \right) = O\left(\frac{1}{N}\right), \quad \text{for all } \beta > 0
\]

(30)

Outline of the proof for the last formula in (30):

Consider the function $H(x) = x^{\beta + 1} \log(x) = xh(x)$ with $h(x) = x^{\beta} \log(x) \to 0$ as $x \to 0$, using $x = e^{-\beta y}$. Hence, for each $\beta > 0$ there exists a constant $C = C(\beta)$ such that $|h(x)| \leq 1$ for all $x$ and $|H(x)| \leq x$ if $-\log(x) > C(\beta)$.

Consequently, $H(x) = O\left( x \right), \quad x \to 0$

**End of Note**
A.1 A few examples

Example 1: \( f(x) = x^\beta, \ x \geq 0. \)
We have \( \int_0^{m/M} f(x)dx = \frac{1}{\beta+1} \left( \frac{m}{M} \right)^{\beta+1} = O\left( \frac{1}{N} \right). \) see (26).
Applying (29) we have (in case \( N \to \infty \))
\[
\frac{1}{N} \sum |v_i|^{\beta} \approx \int_0^1 f(v)dv = \frac{1}{\beta + 1}\]
(31)
Hence,
\[
\frac{1}{N} \sum |x_i|^{\beta} = M^{\beta} \frac{1}{N} \sum |v_i|^{\beta} \approx M^{\beta} \frac{1}{\beta + 1}\]
(32)
Hence, using (30),
\[
\frac{1}{\beta} \log \left( \frac{1}{N} \sum |x_i|^{\beta} \right) = \frac{1}{\beta} \log \left( \frac{M^{\beta}}{N} \sum |v_i|^{\beta} \right) = \log(M) + \frac{1}{\beta} \log \left( \frac{1}{N} \sum |v_i|^{\beta} \right) \\
\quad \approx \log(M) + \frac{1}{\beta} \log \left( \frac{1}{\beta + 1} + O\left( \frac{1}{N} \right) \right) \quad \text{(33)}
\]

Example 2: \( f(x) = x^\beta \log(x), \ x \geq 0. \)
Straightforward integration gives
\[
\int_0^{m/M} f(v)dv = \int_0^{m/M} v^{\beta} \log(v)dv \\
= \frac{1}{(\beta + 1)} \left( \frac{m}{M} \right)^{\beta+1} \left\{ \log \left( \frac{m}{M} \right) - \frac{1}{\beta + 1} \right\} = O\left( \frac{1}{N} \right)
\]
(34)
and
\[
\int_0^1 f(v)dv = \int_0^1 v^{\beta} \log(v)dv = -\frac{1}{(\beta + 1)^2}
\]
(35)
Hence,
\[
\frac{1}{N} \sum_{i=1}^N |v_i|^{\beta} \log |v_i| \approx -\frac{1}{(\beta + 1)^2}, \quad N \to \infty
\]

and
\[
\frac{1}{N} \sum_{i=1}^N |x_i|^{\beta} \log |x_i| = M^{\beta} \left( \log(M) \frac{1}{N} \sum_{i=1}^N |v_i|^{\beta} + \frac{1}{N} \sum_{i=1}^N |v_i|^{\beta} \log |v_i| \right)
\]
\[
\approx M^{\beta} \left( \log(M) \frac{1}{\beta + 1} - \frac{1}{(\beta + 1)^2} \right), \quad N \to \infty
\]

**Example 3:** \( \log(\hat{\alpha}(\beta)) \).

We recall the formula for \( \hat{\alpha}(\beta) \):
\[
\hat{\alpha}(\beta) = \left( \frac{\beta}{N} \sum_{i=1}^N |x_i - \mu|^{\beta} \right)^{1/\beta}
\]  
(37)

As before, we can assume \( \mu = 0 \) without loss of generality. Then
\[
\log(\hat{\alpha}(\beta)) = \frac{1}{\beta} \log(\beta) + \frac{1}{\beta} \log \left( \frac{1}{N} \sum_{i=1}^N |x_i|^{\beta} \right)
\]  
(38)

Substitution of (33) into (38), we find
\[
\log(\hat{\alpha}(\beta)) = \log(M) + \frac{1}{\beta} \log \left( \frac{\beta}{\beta + 1} \right) + O \left( \frac{1}{N} \right)
\]
\[
\approx \log(M) + \frac{1}{\beta} \log \left( \frac{\beta}{\beta + 1} \right), \quad N \to \infty
\]  
(39)

Hence,
\[
\hat{\alpha}(\beta) \approx M \left( \frac{\beta}{\beta + 1} \right)^{1/\beta}, \quad N \to \infty
\]  
(40)

**A.2 An accurate approximation of the function \( g(\beta) \)**

In this section we include a more accurate integral approximation of the sum, given by Dr. A. J.E.M. Janssen, Eindhoven University of Technology, see [5].

We recall the formula for the function \( g(\beta) \):
\[
g(\beta) = g_1(\beta) - P(\beta) + Q(\beta) \quad \text{where}
\]  
(41)
\[ g_1(\beta) = 1 + \frac{\Psi(1/\beta)}{\beta}, \quad (42) \]

\[ P(\beta) = \frac{\sum_{i=1}^{N}|x_i|^\beta \log |x_i|}{\sum_{i=1}^{N}|x_i|^\beta}, \quad (43) \]

and

\[ Q(\beta) = \frac{\log \left( \frac{\beta}{N} \sum_{i=1}^{N}|x_i|^\beta \right)}{\beta} \quad (44) \]

Here \( \Psi \) is the digamma function, \( \Psi(z) = \Gamma'(z)/\Gamma(z) \) as given in [4], Ch. 6.

The \( x_i, i = 1, ..., N \) are in an interval \([m, M]\) with \( 0 < m < M \) and are defined by

\[ x_i = m + (i - 1/2) \frac{M - m}{N}, \quad i = 1, ..., N \quad (45) \]

Using the midpoint integration rule we approximate

\[ \sum_{i=1}^{N}|x_i|^\beta \approx \frac{N}{M - m} \int_{m}^{M} u^\beta du = \frac{N}{M - m} \left( \frac{M^{\beta+1} - m^{\beta+1}}{\beta + 1} \right) \quad (46) \]

Notice that it follows from (46) that

\[ \frac{1}{N} \sum_{i=1}^{N} |x_i|^\beta \approx \frac{M^\beta}{\beta + 1} \left( \frac{1 - (m/M)^{\beta+1}}{1 - (m/M)} \right) \to \frac{M^\beta}{\beta + 1} \quad \text{as } m \to 0 \quad (47) \]

Similarly,

\[ \sum_{i=1}^{N} |x_i|^\beta \log(x_i) \approx \frac{N}{M - m} \int_{m}^{M} u^\beta \log(u) du \]

\[ = \frac{N}{M - m} \frac{1}{\beta + 1} \left( \left[ u^{\beta+1} \log(u) \right]_m^M - \int_{m}^{M} u^\beta du \right) \]

\[ = \frac{N}{M - m} \frac{1}{\beta + 1} \left( M^{\beta+1} \log(M) - m^{\beta+1} \log(m) - \frac{M^{\beta+1} - m^{\beta+1}}{\beta + 1} \right) \quad (48) \]
We find from (46) and (48) that
\[
P(\beta) \approx \frac{N_{M-m}^{\beta+1}(M^{\beta+1}\log(M) - m^{\beta+1}\log(m) - \frac{M^{\beta+1} - m^{\beta+1}}{\beta+1})}{M^{\beta+1} - m^{\beta+1}}
\]
and
\[
Q(\beta) \approx \frac{1}{\beta} \log \left( \frac{M^{\beta+1} - m^{\beta+1}}{M - m} \right)
\]
Notice that
\[
P(\beta) \approx \log(M) - \frac{1}{\beta + 1} \text{ as } m \to 0
\]
and
\[
Q(\beta) \approx \log(M) + \frac{1}{\beta} \log \left( \frac{M^{\beta+1} - m^{\beta+1}}{M - m} \right) \text{ as } m \to 0
\]
By combining (41), (49) and (50) we get
\[
g(\beta) \approx 1 + \frac{\Psi(1/\beta)}{\beta} - \left( \frac{M^{\beta+1}\log(M) - m^{\beta+1}\log(m)}{M^{\beta+1} - m^{\beta+1}} - \frac{1}{\beta+1} \right) + \frac{1}{\beta} \log \left( \frac{M^{\beta+1} - m^{\beta+1}}{M - m} \right)
\]
Finally, by combining (51), (52), (53) and letting \( m \to 0 \) we find
\[
g(\beta) \approx \left( 1 + \frac{\Psi(1/\beta)}{\beta} \right) + \left( \frac{1}{\beta + 1} + \frac{1}{\beta} \log \left( \frac{M^{\beta+1} - m^{\beta+1}}{M - m} \right) \right)
\]
As \( m \to 0 \)

A.3 Behaviour of \( g(\beta) \) as \( \beta \searrow 0 \)

Let us now investigate the behaviour of \( g(\beta) \) as \( \beta \searrow 0 \).
To this end, we first consider formula (42) for \( g_1(\beta) \).
We have (see [4], Eq. 6.3.18)
\[
\Psi(z) = \log(z) - \frac{1}{2z} - \frac{1}{12z^2} + O(z^{-4}), \quad z \to \infty
\]
Hence,

\[ \Psi(1/\beta) = \log(1/\beta) - \frac{1}{2} \beta + \mathcal{O}(\beta^2), \quad \beta \searrow 0 \quad (56) \]

and

\[ 1 + \frac{1}{\beta} \Psi(1/\beta) = \frac{1}{2} + \frac{1}{\beta} \log(1/\beta) + \mathcal{O}(\beta), \quad \beta \searrow 0 \quad (57) \]

Combining (53) and (57) we obtain

\[
g(\beta) \approx \frac{1}{2} + \frac{1}{\beta} \log(1/\beta) \\
- \left( \frac{M^\beta + \log(M) - m^\beta + \log(m)}{M^\beta + 1 - m^\beta + 1} \right) - \frac{1}{\beta + 1} \\
+ \frac{1}{\beta} \log \left( \frac{\beta}{\beta + 1} \right) \frac{M^\beta + 1 - m^\beta + 1}{M - m} \quad (58) \]

or equivalently,

\[
g(\beta) \approx \frac{1}{2} - \frac{1}{\beta} \log(1 + \beta) \\
- \left( \frac{M^\beta + \log(M) - m^\beta + \log(m)}{M^\beta + 1 - m^\beta + 1} \right) - \frac{1}{\beta + 1} \\
+ \frac{1}{\beta} \log \left( \frac{M^\beta + 1 - m^\beta + 1}{M - m} \right) \quad (59) \]

Let us now return to expression (49) for \( P(\beta) \) in the case \( \beta \searrow 0 \).

\[
P(\beta) \approx \frac{M^\beta + \log(M) - m^\beta + \log(m)}{M^\beta + 1 - m^\beta + 1} - \frac{1}{\beta + 1} \\
\rightarrow \frac{M \log(M) - m \log(m)}{M - m} - 1 \quad (60) \]
Furthermore, we have

\[
\frac{M^{\beta+1} - m^{\beta+1}}{M - m} = M^{\beta} \left( \frac{1 - (m/M)^{\beta+1}}{1 - (m/M)} \right)
\]

\[
= M^{\beta} \left( 1 - (m/M) \right) + \left( (m/M) - (m/M)^{\beta+1} \right) \frac{1}{1 - (m/M)}
\]

\[
= M^{\beta} \left( 1 + \left( \frac{m}{M} \right) 1 - (m/M)^{\beta} \right)
\]

\[
= M^{\beta} \left( 1 + \left( \frac{m}{M} \right) 1 - e^{-\beta \log(m/M)} \right)
\]

\[
= M^{\beta} \left( 1 + \left( \frac{m}{M} \right) 1 - \left( 1 + \beta \log(m/M) + O(\beta^2) \right) \right)
\]

\[
= M^{\beta} \left( 1 - \left( \frac{m}{M} \right) \frac{\beta \log(m/M)}{1 - (m/M)} + O(\beta^2) \right)
\]

\[
= M^{\beta} \left( 1 - \beta \frac{m \log(m/M)}{M - m} + O(\beta^2) \right)
\]

Hence,

\[
\frac{1}{\beta} \log \left( \frac{M^{\beta+1} - m^{\beta+1}}{M - m} \right) = \frac{1}{\beta} \log \left[ M^{\beta} \left( 1 - \beta \frac{m \log(m/M)}{M - m} + O(\beta^2) \right) \right]
\]

\[
= \log(M) - \frac{m \log(m/M)}{M - m} + O(\beta)
\]

\[
= M \log(M) - m \log(m) \left( \frac{M - m}{M - m} \right) + O(\beta)
\]

(62)

Combining (50) and (62) we find

\[
Q(\beta) \approx \frac{1}{\beta} \log \left( \frac{\beta}{\beta + 1} \right) + \frac{1}{\beta} \log \left( \frac{M^{\beta+1} - m^{\beta+1}}{M - m} \right)
\]

\[
= \frac{1}{\beta} \log \left( \frac{\beta}{\beta + 1} \right) + \frac{M \log(M) - m \log(m)}{M - m} + O(\beta)
\]

(63)

Notice that in case \( m \to 0 \) formula (63) reduces to

\[
Q(\beta) \approx \log(M) + \frac{1}{\beta} \log \left( \frac{\beta}{\beta + 1} \right) + O(\beta)
\]

(64)
We conclude from (59), (60), (61), and (62) that
\[
\frac{1}{2} - \left[ \frac{1}{\beta} \log(1 + \beta) \right]_{\beta = 0} - \left( \frac{M \log(M) - m \log(m)}{M - m} - 1 \right) + \frac{M \log(M) - m \log(m)}{M - m} = \\
\frac{1}{2} - \left[ \frac{1}{\beta} (\beta + O(\beta^2)) \right]_{\beta = 0} + 1 = \frac{1}{2} \quad \text{as} \quad \beta \searrow 0
\] (65)

**B Approximation of a sum $\frac{1}{N} \sum_{i=1}^{N} f(x_i)$ for random $x_i$**

In this section we will discuss the approximation of a sum $\frac{1}{N} \sum_{i=1}^{N} f(x_i)$ in case the $x_i$ are random variables. We refer to [6], Ch. 5.4.

Here we only consider the case that the $x_i$ correspond to the generalized Gaussian density function
\[
p(x|\alpha,\beta,\mu) = \frac{\beta}{2\alpha \Gamma(1/\beta)} \exp\left(-\left|\frac{x-\mu}{\alpha}\right|^\beta\right),
\] (66)
with $\alpha > 0, \beta > 0$

We assume $\mu = 0$ and we will denote $p(x|\alpha,\beta,\mu = 0)$ simply by $p_{\alpha,\beta}(x)$. Thus,
\[
p_{\alpha,\beta}(x) = c(\alpha,\beta) \exp\left(-\left|\frac{x}{\alpha}\right|^\beta\right), \quad \text{with} \quad c(\alpha,\beta) = \frac{\beta}{2\alpha \Gamma(1/\beta)}
\] (67)

Let $\{x_i | 1 \leq i \leq N\}$ be a set of random variables, where each $x_i$ corresponds to the density function (67).

Then we have
\[
\frac{1}{N} \sum_{i=1}^{N} x_i \approx \mathcal{E}\{x\} = \int_{-\infty}^{\infty} x p_{\alpha,\beta}(x) dx
\] (68)

Notice that (68) gives $\mathcal{E}\{x\} = 0$ since $\mu = 0$ was assumed.

Let $h(x)$ be an integrable function over $\mathbb{R}$.

As stated in [6], Ch. 5.4, Eq. (5-28) we have
\[
\frac{1}{N} \sum_{i=1}^{N} h(x_i) \approx \mathcal{E}\{h(x)\} = \int_{-\infty}^{\infty} h(x) p_{\alpha,\beta}(x) dx
\] (69)
We now consider two special cases. Below we will use the notation: ‘formula(x) \to [x = y] \to = formula(y)’ when the variable x will be replaced by y.

**Case 1:**  
\[ h(x) = h_1(x) = |x|^{\beta} \]

\[ \frac{1}{N} \sum_{i=1}^{N} |x_i|^{\beta} \approx \mathcal{E}\{h_1(x)\} \]

\[ = c(\alpha, \beta) \int_{-\infty}^{\infty} |x|^\beta \exp\left(-\frac{|x|^{\beta}}{\alpha}\right) dx = 2c(\alpha, \beta) \int_{0}^{\infty} x^\beta \exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right) dx \]

\[ \to [x = \alpha t, \ dx = \alpha \ dt] \to \]

\[ = 2c(\alpha, \beta) \alpha^{\beta+1} \int_{0}^{\infty} t^\beta \exp(-t^\beta) dt \]

\[ \to [t = u^{1/\beta}, \ dt = (1/\beta)u^{1/\beta-1} du] \to \]

\[ = 2c(\alpha, \beta) \frac{\alpha^{\beta+1}}{\beta} \int_{0}^{\infty} u^{1/\beta} e^{-u} du = 2c(\alpha, \beta) \frac{\alpha^{\beta+1}}{\beta} \Gamma(1/\beta + 1) \]

\[ = 2 \cdot \frac{\beta}{2\alpha \Gamma(1/\beta)} \cdot \frac{\alpha^{\beta+1}}{\beta} \left( \frac{1}{\beta} \Gamma(1/\beta) \right) = \frac{\alpha^\beta}{\beta} \]

(70)

**Remark 1**

We used the identity \( \Gamma(z + 1) = z \Gamma(z) \) for \( z > 0 \) with \( z = \frac{1}{\beta} \), see [4], Eq.(6.1.15).

**Remark 2** By carefully inspecting the formulas in (70) we see

\[ 2c(\alpha, \beta) \alpha^{\beta+1} \int_{0}^{\infty} t^\beta \exp(-t^\beta) dt = \frac{\alpha^\beta}{\beta} \]

(71)

**Remark 3**

Consider the case \( \beta = 2 \). Then we have from (70) and the fact that \( \mathcal{E}\{x\} = 0 \)

\[ \frac{1}{N} \sum_{i=1}^{N} |x_i|^2 \approx \mathcal{E}\{x^2\} = \mathcal{E}\{x^2\} - \left( \mathcal{E}\{x\} \right)^2 = \frac{1}{2} \alpha^2 \]

(72)

Notice that this formula is consistent with the fact that for \( \beta = 2 \) the generalized pdf (67) equals the standard Gaussian pdf with \( \mu = 0 \) and \( \sigma^2 = \alpha^2/2 \).

**End of Case 1**
Case 2: \( h(x) = h_2(x) = |x|^\beta \log |x| \)

\[
\frac{1}{N} \sum_{i=1}^{N} |x_i|^\beta \log |x_i| \approx \mathcal{E}\{h_2(x)\} = c(\alpha, \beta) \int_{-\infty}^{\infty} |x|^\beta \log |x| \exp\left(-\frac{|x|}{\alpha}\right) dx \\
= 2c(\alpha, \beta) \int_{0}^{\infty} x^\beta \log(x) \exp\left(-\frac{x}{\alpha}\right) dx \\
\rightarrow [x = \alpha t, \ dx = \alpha dt] \rightarrow \\
= 2c(\alpha, \beta) \alpha^{\beta+1} \int_{0}^{\infty} t^\beta \log(\alpha t) \exp(-t^\beta) dt \\
= 2c(\alpha, \beta) \alpha^{\beta+1} \left\{ \log(\alpha) \int_{0}^{\infty} t^\beta \exp(-t^\beta) dt + K_\beta \right\} \\
\text{with } K_\beta = \int_{0}^{\infty} t^\beta \log(t) \exp(-t^\beta) dt
\]

We have

\[
K_\beta \rightarrow [t^\beta = u, \ t = u^{1/\beta}, \ dt = (1/\beta) u^{1/\beta-1} du] \rightarrow \\
= (1/\beta^2) \int_{0}^{\infty} u^{1/\beta} \log(u) e^{-u} du \\
\text{Combining (73), (74), (71), and (67) we find}
\]

\[
\mathcal{E}\{h_2(x)\} = \frac{\alpha^\beta}{\beta} \log(\alpha) + 2c(\alpha, \beta) \alpha^{\beta+1} K_\beta \\
= \frac{\alpha^\beta}{\beta} \log(\alpha) + \frac{\beta}{\Gamma(1/\beta)} \alpha^\beta K_\beta
\]

End of Case 2
Recalling formula (43) for \( P(\beta) \) and using the results (70), (75), (74) we have

\[
P(\beta) = \frac{1}{N} \sum |x_i|\beta \log |x_i| \approx \frac{\varepsilon\{h_2(x)\}}{\varepsilon\{h_1(x)\}}
\]

\[
= \frac{\alpha^\beta}{\beta} \log(\alpha) + \frac{\beta}{\Gamma(1/\beta)} \alpha^\beta K_\beta
\]

\[
= \log(\alpha) + \frac{1}{\Gamma(1/\beta)} \int_0^\infty u^{1/\beta} \log(u) e^{-u} du
\]

Recall that \( Q(\beta) = (1/\beta) \log\left( (\beta/N) \sum |x_i|\beta \right) \), see (44).

Then by using (70) we find

\[
Q(\beta) \approx (1/\beta) \log\left( \beta \varepsilon\{h_1(x)\} \right)
\]

\[
= (1/\beta) \log\left( \beta \frac{\alpha^\beta}{\beta} \right) = \log(\alpha)
\]

We can conclude from (76) and (77) that

\[
P(\beta) - Q(\beta) \approx \frac{1}{\Gamma(1/\beta)} \int_0^\infty u^{1/\beta} \log(u) e^{-u} du
\]

Hence, we have the following approximation for the function \( g(\beta) \) defined in (41)

\[
g(\beta) = 1 + \frac{\Psi(1/\beta)}{\beta} - \left( P(\beta) - Q(\beta) \right)
\]

\[
\approx 1 + \frac{\Psi(1/\beta)}{\beta} - \frac{1}{\Gamma(1/\beta)} \int_0^\infty u^{1/\beta} \log(u) e^{-u} du
\]

**Remark**

Clearly, this equation merely depends on \( \beta \) (as should be), despite the fact that in the above derivation of the formulas for \( P(\beta) \) and \( Q(\beta) \) we have used the density function \( p_{\alpha\beta} \) that depends on \( \beta \) as well as on \( \alpha \).

**End of Remark**
C Analysis of the behaviour of $g(\beta)$

For convenience, we recall the formula (41) for $g(\beta)$.

$$g(\beta) = g_1(\beta) - P(\beta) + Q(\beta)$$

where

$$g_1(\beta) = 1 + \frac{\Psi(1/\beta)}{\beta}, \quad \text{and}$$

$$P(\beta) = \frac{\sum_{i=1}^{N} |x_i|^\beta \log |x_i|}{\sum_{i=1}^{N} |x_i|^\beta}, \quad Q(\beta) = \frac{\log(\frac{\beta}{N} \sum_{i=1}^{N} |x_i|^\beta)}{\beta}$$

C.1 Behaviour of $g_1(\beta)$

We have $g_1(\beta) = 1 + \frac{\Psi(1/\beta)}{\beta}$, see figure 17.

\[
\begin{align*}
  g_1(\beta) &= 1 + \frac{1}{\beta} \Psi\left(\frac{1}{\beta}\right) \\
  \frac{dg_1(\beta)}{d\beta} &= -\frac{1}{\beta^2} \left\{ \frac{\Psi(1/\beta)}{\beta} + \frac{1}{\beta} \Psi'(1/\beta) \right\}, \\
  g_1'(\beta^*) &= 0, \quad \text{with } \beta^* \approx 4.628, \\
  g_1'(\beta^*) < 0, \quad \text{for } 0 < \beta < \beta^*, \\
  g_1'(\beta^*) > 0, \quad \text{for } \beta > \beta^*
\end{align*}
\]
As shown in Appendix F, eqs. (131), (132) we have

\[
g'_1(1/\beta) = -\beta \left\{ \beta \Psi(\beta) + \beta^2 \Psi'(\beta) \right\} \to 0 \text{ as } \beta \to 0 \tag{82}
\]

Thus,

\[
\lim_{\beta \to \infty} g'_1(\beta) = 0. \tag{83}
\]

The function \(g'_1(\beta)\) is shown in figure 18.

We can conclude from (81) that

- \(g_1(\beta)\) is monotonically decreasing for \(0 < \beta < \beta^*\),
- \(g_1(\beta)\) is monotonically increasing for \(\beta > \beta^*\),
- \(g_1(\beta)\) has a minimum \(g'_1(\beta^*) \approx -0.05\) with \(\beta^* \approx 4.628\),
- \(g_1(\beta) = 0\) for \(\beta = \beta_0 \approx 2.166\) (by inspecting Fig.17)
- \(g_1(\beta) > 0\) for \(0 < \beta < \beta_0\),
- \(g_1(\beta) \to \infty\) as \(\beta \to 0\) (since \(g'_1(\beta) < 0\) for \(0 < \beta < \beta_0 < \beta^*\))
As shown in Appendix F, eq. (132) we have \( \lim_{x \to 0} x \Psi(x) = -1 \), or \( \lim_{\beta \to \infty} \frac{\Psi(1/\beta)}{\beta} = -1 \). Hence,

\[
\lim_{\beta \to \infty} g_1(\beta) = 0. \quad (85)
\]

### C.2 Behaviour of \( Q(\beta) \)

First notice that by comparing (44) and (38) we have

\[
Q(\beta) = \log(\hat{a}). \quad (86)
\]

Recalling (44) and (52) we have

\[
Q(\beta) = \frac{1}{\beta} \log\left(\frac{\beta}{N} \sum_{i=1}^{N} |x_i|^\beta\right) \approx \log(M) + \frac{1}{\beta} \log\left(\frac{\beta}{\beta + 1}\right) \quad (87)
\]

with \( M = \max_{1 \leq i \leq N} |x_i| \)

Hence,

\[
\lim_{\beta \to \infty} Q(\beta) = \log(M) \quad \text{and} \quad \lim_{\beta \to 0} Q(\beta) = -\infty \quad (88)
\]

### C.3 Behaviour of \( P(\beta) \)

Recalling (43) and (51) we have

\[
P(\beta) = \frac{\sum_{i=1}^{N} |x_i|^\beta \log|x_i|}{\sum_{i=1}^{N} |x_i|^\beta} \approx \log(M) - \frac{1}{\beta + 1} \quad (89)
\]

Hence,

\[
\lim_{\beta \to \infty} P(\beta) = \log(M) \quad \text{and} \quad \lim_{\beta \to 0} P(\beta) = \log(M) - 1 \quad (90)
\]

### C.4 Behaviour of \( P(\beta) - Q(\beta) \)

It is easily seen from (87) and (89) that

\[
P(\beta) - Q(\beta) \approx -\left(\frac{1}{\beta + 1} + \frac{1}{\beta} \log\left(\frac{\beta}{\beta + 1}\right)\right) \rightarrow 0 \quad \text{as} \quad \beta \to \infty \quad (91)
\]
C.5 Behaviour of $g(\beta)$

Finally, by combining (41), (42) and (91) we have

$$g(\beta) = g_1(\beta) - \left( P(\beta) - Q(\beta) \right)$$

$$\approx \left( 1 + \frac{1}{\beta}\Psi(1/\beta) \right) + \left( \frac{1}{\beta+1} + \frac{1}{\beta} \log\left( \frac{\beta}{\beta+1} \right) \right)$$

(92)

Hence, using (85), (91) and (65),

$$\lim_{\beta \to \infty} g(\beta) = 0 \quad \text{and} \quad \lim_{\beta \to 0} g(\beta) = 1/2$$

(93)

The behaviour of $g(\beta)$ is shown in figure 19.

Figure 19: The behaviour of function $g(\beta)$ including $\beta \to 0$.
Here $g(\beta_0) = 0$ with $\beta_0 \approx 9.68$
D Asymptotics

D.1 Asymptotic behaviour of \( \hat{a}(\beta) \)

We will prove that

1) \( \lim_{\beta \to \infty} \hat{a}(\beta) \approx M \) where \( M = \max_{1 \leq i \leq N} |x_i| \) and

2) \( \lim_{\beta \to 0} \hat{a}(\beta) = 0 \)

Proof of Part 1

It is easily seen from (38) that

\[
\lim_{\beta \to \infty} \log(\hat{a}(\beta)) \approx \log(M), \quad \text{or equivalently} \quad \lim_{\beta \to \infty} \hat{a}(\beta) \approx M. \tag{94}
\]

End of Part 1

Proof of Part 2

We introduce

\[
T(\beta) = \log\left( \frac{1}{N} \sum |x_i|^\beta \right) \tag{95}
\]

Using l’Hôpital’s rule and the fact that \( \lim_{\beta \to 0} T(\beta) = 0 \) we find

\[
\lim_{\beta \to 0} \frac{1}{\beta} \left( \log\left( \frac{1}{N} \sum |x_i|^\beta \right) \right) = \lim_{\beta \to 0} \frac{T(\beta)}{\beta} = \lim_{\beta \to 0} T'(\beta) = \frac{1}{N} \sum |x_i| \log |x_i| \tag{96}
\]

Recalling (38) , using (96) and \( \lim_{y \to \infty} \frac{1}{y} \log(y) = \lim_{y \to \infty} (-ye^y) = -\infty \) we find

\[
\log(\hat{a}(\beta)) = \frac{1}{\beta} \log(\beta) + \frac{1}{\beta} \log \left( \frac{1}{N} \sum_{i=1}^{N} |x_i|^\beta \right) \to -\infty \quad \text{as} \quad \beta \to 0 \tag{97}
\]

Hence,

\[
\lim_{\beta \to 0} \hat{a}(\beta) = \lim_{\beta \to 0} \left( \frac{\beta}{N} \sum |x_i|^\beta \right)^{1/\beta} = 0 \tag{98}
\]

End of Part 2
E  Analysis of $\partial L/\partial \beta$

In this section we want to analyze the behaviour of $\partial L/\partial \beta$ as given (6), i.e.,

$$\frac{\partial L}{\partial \beta} = \frac{1}{\beta} + \frac{\Psi(1/\beta)}{\beta^2} + \frac{1}{N} \sum_{i=1}^{N} \left| \frac{x_i - \mu}{\alpha} \right|^{\beta \log \left| \frac{x_i - \mu}{\alpha} \right|} = F(\beta) + S(\beta)$$

where

$$F(\beta) = \frac{1}{\beta} + \frac{\Psi(1/\beta)}{\beta^2}$$

and

$$S(\beta | \mu, \alpha) = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{x_i - \mu}{\alpha} \right|^{\beta \log \left| \frac{x_i - \mu}{\alpha} \right|}$$

Notice that $F(\beta)$ depends on $\beta$ only. Recall that $\Psi(x)$ is the logarithmic derivative of the well-known $\Gamma$-function.

We want to answer the

**Question:**

In which situations has the equation $\partial L/\partial \beta = 0$ a solution for $\beta > 0, \alpha \geq 1$?

Therefore, we will analyze the functions $F(\beta)$ and $S(\beta)$ given in (100), (101) separately.

We will make the following assumptions:

- $\beta > 0, \alpha \geq 1$
- The real points $x_i (i = 1, N)$ are located in the interval $[U, V]$
- In general, the $x_i$ are random values corresponding to the probability function (1)
- For analysis purposes we choose $x_i$ to be nonrandom, i.e., equidistantly in the interval $[U, V]$
- The values $x_i$ has mean $\mu = \sum_{i=1}^{N} x_i$
- The mean $\mu$ is in the interval $(U, V)$, so $U < \mu = 0 < V$
- If $x_i = \mu$ then these $x_i$ will be excluded (otherwise $\partial L/\partial \beta$ is not defined).
- Consequently, $N$ denotes the number of points $x_i \neq \mu$
- Without loss of generality, we can take $\mu = 0$
E.1 Analysis of $F(\beta)$

The first part of $\partial L/\partial \beta$ is denoted by

$$F(\beta) = \frac{1}{\beta} + \frac{\Psi(1/\beta)}{\beta^2}$$  \hfill (102)

We recall (see [4], Ch.6)

- $\Psi(x)$ is the digamma function: $\Psi(x) = \frac{d}{dx} \ln(\Gamma(x)) = \frac{\Gamma'(x)}{\Gamma(x)}$ where
- $\Gamma(x)$ is the well-known gamma function

**Behaviour of the functions $F(\beta)$ and $F'(\beta)$**

The functions $F(\beta)$ and $F'(\beta)$ are shown in figures 20 and 21.

![Figure 20: $F(\beta)$ and $F'(\beta)$ for $1 \leq \beta \leq 10$, with $\lim_{\beta \to \infty} F(\beta) = \lim_{\beta \to \infty} F'(\beta) = 0$.](image)

For convenience, we mention a few properties of these functions. For a proof, see Appendix F.

- $F(\beta) = 0$ for $\beta = \beta_0 \approx 2.13$
- $F(\beta) > 0$ for $\beta < \beta_0$ and $F(\beta) < 0$ for $\beta > \beta_0$ (see figure 21).
In particular, $F(\beta = 2) \approx 0.0091 \approx 0.01$

- $F(\beta) \rightarrow +\infty$ for $\beta \nearrow 0$

- $\lim_{\beta \rightarrow \infty} F(\beta) = 0$

- $\lim_{\beta \rightarrow \infty} F'(\beta) = 0$

- As shown in figure 21 and Appendix F we have

- $F'(\beta) = 0$ for $\beta = \beta_1 \approx 3.38$

- $F'(\beta) < 0$ for $\beta < \beta_1$ and $F'(\beta) > 0$ for $\beta > \beta_1$

- So, $F(\beta)$ is decreasing for $0 < \beta < \beta_1$, and increasing for $\beta > \beta_1$

- and $\min_{\beta > 0} F(\beta) = F(\beta_1) \approx -0.015$

- Thus,

$$0 \leq F(\beta) \lesssim 0.01 \text{ for } 2 \leq \beta \leq \beta_0,$$

$$-0.015 \lesssim F(\beta) < 0 \text{ for } \beta > \beta_0$$

(103)
E.2 Analysis of $S(\beta)$

We will now start with analyzing the second part of $\partial L / \partial \beta$, i.e.,

$$S(\beta \mid \mu, \alpha) = \frac{1}{N} \sum_{i=1}^{N} \frac{x_i - \mu}{\alpha} \left| \beta \log \left( \frac{x_i - \mu}{\alpha} \right) \right|, \quad \text{for } \beta > 0. \quad (104)$$

As mentioned before we may assume that $\mu = 0$. Where needed we will indicate the dependence of $\mu$ explicitly in the formula for $S$. The dependence on $\alpha$ will be omitted. So, in most cases we will write $S(\beta)$ or $S(\beta \mid \mu = 0)$.

We introduce the function $f(x)$ for $\beta \geq 0, \alpha \geq 1$ as follows:

$$f(x) = \begin{cases} \left( \frac{x}{\alpha} \right)^\beta \log \left( \frac{x}{\alpha} \right) & \text{for } x > 0 \\ 0 & \text{for } x = 0 \end{cases} \quad (105)$$

Notice that $\lim_{x \to 0} f(x) = 0$. See Appendix H.

Then we can write $S(\beta) = S(\beta \mid \mu = 0)$ as

$$S(\beta) = \frac{1}{N} \sum_{|x_i| > 0} f(|x_i|) = S^-(\beta) + S^+(\beta) \quad \text{where}$$

$$S^-(\beta) = \frac{1}{N} \sum_{x_i < 0} f(|x_i|) = \frac{1}{N} \sum_{x_i < 0} f(-x_i) = \frac{1}{N} \sum_{y_i > 0} f(y_i)$$

$$S^+(\beta) = \frac{1}{N} \sum_{x_i > 0} f(x_i)$$

To simplify the analysis of $S(\beta)$ we assume that the points $x_i \neq 0$, $1 \leq i \leq N$ are distributed equidistantly over the interval $[U, V]$. By assumption, $\mu = 0$. Moreover, we assume that $\mu$ is in $[U, V]$. Hence, $U < 0$ and $V > 0$.

Let $L = V - U$ be the length of the interval $[U, V]$ and let $\Delta x = L / N$ be the distance between two subsequent points $x_i$ and $x_{i+1}$. By assumption $x_i \neq 0$. So, $\mu \in (-\Delta x, \Delta x)$ and $x_i \in [U, -\Delta x]$ or $x_i \in [\Delta x, V]$.

If $N$ is sufficiently large, then we have

$$\sum_{|x_i| > 0} f(|x_i|) \Delta x = \int_{U}^{V} f(x) dx + O(\Delta x) \quad (107)$$

Hence,

$$S^-(\beta) = \frac{1}{N} \sum_{x_i < 0} f(-x_i) = \frac{1}{N \Delta x} \sum_{x_i < 0} f(-x_i) \Delta x = \frac{1}{L} \int_{U}^{U} f(-x) dx + O(\Delta x)$$

$$= \frac{1}{L} \int_{U}^{0} f(-x) dx - \frac{1}{L} \int_{-\Delta x}^{0} f(-x) dx + O(\Delta x)$$

$$= \frac{1}{L} \int_{0}^{U} f(x) dx - \frac{1}{L} \int_{0}^{\Delta x} f(x) dx + O(\Delta x) \quad (108)$$
and

\[ S^+ (\beta) = \frac{1}{N} \sum_{x_i > 0} f(x_i) = \frac{1}{N} \sum_{x_i > 0} f(x_i) \Delta x = \frac{1}{L} \sum_{x_i > 0} f(x_i) \Delta x \]
\[ = \frac{1}{L} \int_0^V f(x) \, dx + O(\Delta x) \]
\[ = \frac{1}{L} \int_0^V f(x) \, dx - \frac{1}{L} \int_0^\Delta x f(x) \, dx + O(\Delta x) \]  

(109)

It is shown in Appendix H that

\[ \int_0^\Delta x f(x) \, dx = O(\Delta x), \quad \Delta x \to 0 \]  

(110)

Thus,

\[ S^-(\beta) = \frac{1}{L} \int_0^|U| f(x) \, dx + O(\Delta x), \quad \Delta x \to 0 \]
\[ S^+ (\beta) = \frac{1}{L} \int_0^V f(x) \, dx + O(\Delta x), \quad \Delta x \to 0 \]
\[ S(\beta) = \frac{1}{N} \sum_{|x_i| > 0} f(|x_i|) = S^- (\beta) + S^+ (\beta) \]
\[ = \frac{1}{L} \left( \int_0^|U| + \int_0^V \right) f(x) \, dx + O(\Delta x), \quad \Delta x \to 0 \]  

(111)

Notice that if \( \int_0^V f(x) \, dx < 0 \) for some \( \alpha \) and \( \beta \) then \( S^- (\beta) \) might be nonnegative due to the term \( O(\Delta x) \) in (111).

We will use the notation \( S^- (\beta) \leq 0 \) to refer to its formula (111) with \( \int_0^V f(x) \, dx < 0 \), and similarly for \( S^+ (\beta) \) and \( S(\beta) \). Using standard calculus we find we find

\[ S^+ (\beta) = \frac{1}{L} \int_0^V f(x) \, dx + O(\Delta x) \]
\[ = \frac{1}{L} \frac{V}{|\beta + 1|} \left( \frac{V}{\alpha} \right)^\beta \left( \log(\frac{V}{\alpha}) - \frac{1}{\beta + 1} \right) + O(\Delta x) \]  

(112)

For simplicity, we introduce the function

\[ H(\beta \mid \alpha, V) = \frac{1}{L} \frac{V}{|\beta + 1|} \left( \frac{V}{\alpha} \right)^\beta \left( \log(\frac{V}{\alpha}) - \frac{1}{\beta + 1} \right) \]  

(113)

Thus,

\[ S^+ (\beta) = H(\beta \mid \alpha, V) + O(\Delta x), \quad \Delta x \to 0 \]  

(114)
A similar expression can be given for $S^-(\beta)$ by replacing $V$ by $|U|$ in (113), i.e.,

$$S^-(\beta) = H(\beta | \alpha, |U|) + \mathcal{O}(\Delta x), \quad \Delta x \to 0 \tag{115}$$

Finally, combination of (111), (112), (115) gives

$$S(\beta) = H(\beta | \alpha, V) + H(\beta | \alpha, |U|) + \mathcal{O}(\Delta x), \quad \Delta x \to 0 \tag{116}$$

Let us return to formula (113) for $H(\beta | \alpha, V)$ to find out when $S(\beta) < 0$.

For fixed $\alpha$ and $V$ we introduce $H(\beta) = H(\beta | \alpha, V)$.

Furthermore, let $K = V / \alpha$ and $c = \log(K)$, then

$$H(\beta) = \frac{V}{L} \frac{K^\beta}{\beta + 1} \left( \log(K) - \frac{1}{\beta + 1} \right)$$

$$H'(\beta) = \frac{V}{L} \frac{K^\beta}{(\beta + 1)^3} \left( c^2(\beta + 1)^2 - 2c(\beta + 1) + 2 \right)$$

$$= \frac{V}{L} \frac{K^\beta}{(\beta + 1)^3} \left( (c(\beta + 1) - 1)^2 + 1 \right) > 0 \text{ for all } \beta \geq 0 \tag{117}$$

Thus, $H(\beta)$ is a monotonically increasing function for $\beta \geq 0$ and $H(\beta) > H(0) = \frac{V}{L} \left( \log(V / \alpha) - 1 \right)$. Clearly, $\text{sign}(H(\beta)) = \text{sign}\left( \log(V / \alpha) - \frac{1}{\beta + 1} \right)$ and

$$\log \left( \frac{V}{\alpha} \right) - \frac{1}{\beta + 1} > 0 \quad \text{if} \quad \alpha < Ve^{-\frac{1}{\beta + 1}} \tag{118}$$

$$< 0 \quad \text{if} \quad \alpha > Ve^{-\frac{1}{\beta + 1}}$$

and

$$\frac{V}{e} \leq Ve^{-\frac{1}{\beta + 1}} < V \quad \text{for all } \beta \geq 0 \tag{119}$$

See also figure 22.

We can distinguish three cases:

**Case 1**

If $\alpha \geq V$ then $\alpha \geq V > Ve^{-\frac{1}{\beta + 1}}$ for all $\beta \geq 0$. Moreover, in this case $0 < V / \alpha \leq 1$. So, we can conclude

$$\text{if } \alpha \geq V \text{ then } S^+(\beta) < 0 \text{ for all } \beta \geq 0 \text{ and } S^+(\beta) \to 0 \text{ as } \beta \to \infty \tag{120}$$

**Note:** if $\alpha$ is given by (5) then this case can not occur as indicated in section 4.
Figure 22: The function $f(\beta) = Ve^{-\frac{1}{\beta+1}}$ with $f(0) = V/e$ and $\lim_{\beta \to \infty} f(\beta) = V$.

Case 2
If $\alpha \leq V/e$ then $\alpha \leq V/e < Ve^{-\frac{1}{\beta+1}}$ for all $\beta > 0$, and thus $log(\frac{V}{\alpha}) - \frac{1}{\beta+1} > 0$ for all $\beta > 0$. Then it follows from (112)

$$\text{if } \alpha \leq V/e \text{ then } S^+(\beta) > 0 \text{ for all } \beta > 0$$

(121)

Case 3
If $V/e < \alpha < V$ then (see figure 23) there exists a $\beta_V$ such that $Ve^{-\frac{1}{\beta_V+1}} = \alpha$ and $Ve^{-\frac{1}{\beta+1}} < \alpha$ for all $0 \leq \beta < \beta_V$, with

$$\beta_V = \frac{1}{log(V/\alpha)} - 1$$

(122)

Recalling (112) and (118) we can now conclude

$$\text{if } V/e < \alpha < V \text{ then }$$

$$H(\beta | \alpha, V) < 0 \quad \text{and}$$

$$S^+(\beta) < 0 \quad \text{for all } 0 \leq \beta < \beta_V$$

(123)

In figure 23 the behaviour of $H(\beta | \alpha, V)$ is shown for a few values of $\alpha$. Clearly, by replacing $V$ by $|U|$ in the formulas above, similar results can be given for $S^-(\beta)$:

$$\text{if } \alpha \geq |U| \quad \text{then } S^-(\beta) < 0 \text{ for all } \beta > 0;$$

$$\text{if } \alpha \leq |U|/e \quad \text{then } S^-(\beta) > 0 \text{ for all } \beta > 0;$$

$$\text{if } V/e < \alpha < |U| \text{ then } S^-(\beta) < 0 \text{ for all } 0 \leq \beta < \beta_V$$

(124)

40
The function $H(\beta | \alpha, V)$ approximating $S^+(\beta)$. We have $H(\beta | \alpha, V) = 0$ for $\beta = \beta_V = (\log(V/\alpha)^{-1} - 1)$, indicated by a ■.

The value $\beta_{|U|}$ is given by

$$\beta_{|U|} = \frac{1}{\log(|U|/\alpha)} - 1$$

(125)

Now we can combine the results above for $S(\beta) = S^-(\beta) + S^+(\beta)$. To this end, we refer to figure 24 and we find

**Case 1**

If $V = |U|$ then $S(\beta) \leq 0$ for $0 \leq \beta < \beta_V$, in case $V/e < \alpha < V$.

**Case 2**

If $V > |U|$ (as shown in figure 24) then

if $V/e < \alpha < |U|$ then $S(\beta) \leq 0$, $0 < \beta < \beta_V$

if $\alpha \leq V/e$ or $\alpha \geq |U|$ then $S(\beta) > 0$ for all $\beta > 0$

(126)

**Case 3**
If $V < |U|$ then interchange $V$ and $|U|$ in the previous formulas, i.e.,

$$
\begin{align*}
\text{if } |U|/e < \alpha < V & \quad \text{then } S(\beta) < 0, \ 0 < \beta < \beta_{|U|} \\
\text{if } \alpha \leq |U|/e \text{ or } \alpha \geq V & \quad \text{then } S(\beta) > 0 \text{ for all } \beta > 0
\end{align*}
$$

(127)

The formulas (126) and (127) can be combined to

$$
\begin{align*}
\text{if } \max(|U|/e, V/e) < \alpha < \min(|U|, V) & \quad \text{then } \\
S(\beta) < 0, \quad 0 < \beta < \min(\beta_{|U|}, \beta_V)
\end{align*}
$$

(128)

$$
\begin{align*}
\text{if } \alpha \leq \max(|U|/e, V/e) \text{ or } \alpha \geq \min(|U|, V) & \quad \text{then } \\
S(\beta) > 0 \text{ for all } \beta > 0
\end{align*}
$$

(129)
F Properties of $\Gamma(x)$, $\Psi(x)$, $\Psi'(x)$, $F(\beta)$ and $F'(\beta)$

In figure 25 the behaviour of $\Gamma(x)$, $\Psi(x)$ and $\Psi'(x)$ is shown.

![The functions $\Gamma(x)$, $\Psi(x)$ and $\Psi'(x)$](image)

Figure 25: The functions $\Gamma(x)$, $\Psi(x)$ and $\Psi'(x)$

We need the following properties of $\Psi$, $\Psi'$ and $\Gamma$ for a detailed analysis of $F(\beta)$.

F.1 Properties of $\Psi'(x)$

- $\Psi^{(n)}(x) = (-1)^{n+1}n!\sum_{k=0}^{\infty} \frac{1}{(x+k)^{n+1}}$ (see [4], Eq. 6.4.10)
- In particular, $\Psi'(x) = \sum_{k=0}^{\infty} \frac{1}{(x+k)^2} > 0$ for $x > 0$
- $\Psi''(x) = \frac{1}{x^2} + \sum_{k=1}^{\infty} \frac{1}{(x+k)^2} > 0$, thus $\Psi''(x) \sim \frac{1}{x^2}$ for $x \to 0$
- Hence,

$$\lim_{x \to 0} x^2 \Psi'(x) = 1 \quad (130)$$
F.2 Properties of $\Psi(x)$

- $\Psi(x)$ is monotonically increasing for $x > 0$ (since $\Psi'(x) > 0$)
- $\Psi(3/2) = -\gamma - 2\ln(2) + 2 \approx 0$ (see [4], Eq. 6.3.4)
- $\Psi(x_0) = 0$ for $x_0 \approx 1.460 \approx 3/2$ (see [4], Table 6.1)
- $\Psi(x) < 0$ for $0 < x < x_0$ and $\Psi(x) > 0$ for $x > x_0$
- $\Psi(1/2) \approx -2$, so $\Psi(x) \leq -2$ for $0 < x \leq 1/2$ (see [4], Eq. 6.3.3)
- $\Psi(1 - x) = \Psi(x) + \pi \cot(\pi x)$ (see [4], Eq. 6.3.7)
- So, $\lim_{x \to 0} x^2 \Psi(x) = \lim_{x \to 0} (x^2[\Psi(1 - x) - \pi \cot(\pi x)]) = -\lim_{x \to 0} \pi x^2 \cot(\pi x)$
- and $\pi x^2 \cot(\pi x) = \pi x^2 \frac{\cos(\pi x)}{\sin(\pi x)} = x \cos(\pi x) \frac{\pi x}{\sin(\pi x)}$
- Then $\lim_{x \to 0} \pi x^2 \cot(\pi x) = \lim_{x \to 0} x \cos(\pi x) \frac{\pi x}{\sin(\pi x)} = \lim_{y \to 0} \frac{y \cos(y)}{\sin(y)} = 0$
- Thus,
  \[\lim_{x \to 0} x^2 \Psi(x) = 0\] (131)
- Similarly, $\lim_{x \to 0} x \Psi(x) = -\lim_{x \to 0} \pi x \cot(\pi x)$
- and $\pi x \cot(\pi x) = \cos(\pi x) \frac{\pi x}{\sin(\pi x)} = \cos(y) \frac{y}{\sin(y)}$, $(y = \pi x)$
- Thus,
  \[\lim_{x \to 0} x \Psi(x) = \lim_{y \to 0} \cos(y) \frac{y}{\sin(y)} = -1\] (132)

F.3 Properties of $\Gamma(x)$

- $\Gamma(x) > 0$ for $x > 0$
- $\Psi(x) = \frac{d}{dx} \ln(\Gamma(x))$
- $\Psi(x) < 0$, $x < x_0$, $\Psi(x_0) = 0$ and $\Psi(x) > 0$, $x > x_0$
- Hence, $\ln(\Gamma(x))$ has a unique minimum for $x = x_0 \approx 3/2$
- Since the ln-function is increasing, also $\Gamma(x)$ has a unique minimum for $x = x_0$
F.4 Behaviour of $F(\beta)$ and $F'(\beta)$

For convenience, we duplicate figure 21 with $F(\beta)$ and $F'(\beta)$ for $2 \leq \beta \leq 5$. We have the following properties.

Figure 26: $F(\beta)$ and $F'(\beta)$ for $2 \leq \beta \leq 5$ and $F'(\beta) = 0$ for $\beta \approx 3.38$.

- $\lim_{\beta \to \infty} F(\beta) = \lim_{\beta \to \infty} \left( \frac{1}{\beta} + \frac{\Psi(1/\beta)}{\beta^2} \right) = \lim_{\beta \to \infty} \frac{\Psi(1/\beta)}{\beta^2} = \lim_{x \to 0} x^2 \Psi(x) = 0$ (see (132))
- $F'(\beta) = -\frac{1}{\beta^2} \left( 1 + 2\Psi(1/\beta) + \Psi'(1/\beta)/\beta^2 \right)$
- Let $H(\beta) = 1 + 2\Psi(1/\beta)/\beta + \Psi'(1/\beta)/\beta^2$ or
- $H(x) = 1 + 2x \Psi(x) + x^2 \Psi'(x)$ where $x = 1/\beta$
- Using (132) and (130), we see that $\lim_{x \to 0} H(x) = 1 + 2 \times (-1) + 1 = 0$
- So, $\lim_{\beta \to \infty} F'(\beta) = \lim_{x \to 0} F'(x) = -\lim_{x \to 0} x^2 H(x) = 0$ (see also figure 27)
- Numerical experiments show the following properties
- $F'(\beta) = 0$ for $\beta = \beta_1 \approx 3.38$
• $F'(\beta) < 0$ for $\beta < \beta_1$ and $F'(\beta) > 0$ for $x > \beta_1$
• Thus, $F(\beta)$ is decreasing for $0 < \beta < \beta_1$, increasing for $\beta > \beta_1$ and
• $\min_{\beta > 0} F(\beta) = F(\beta_1) \approx -0.015$
• $F(\beta) \to +\infty$ for $\beta \nearrow 0$
• $F(\beta) = 0$ for $\beta = \beta_0 \approx 2.13$
• $F(\beta) > 0$ for $\beta < \beta_0$ and $F(\beta) < 0$ for $\beta > \beta_0$
• $F(\beta = 2) \approx 0.0091 \approx 0.01$
• Thus, $-0.015 \lesssim F(\beta) \lesssim 0.01$ for $\beta \geq 2$

![Figure 27: $F'(x)$ where $x = 1/\beta$ and $\lim_{x \to 0} F'(x) = \lim_{\beta \to \infty} F'(1/\beta) = 0.$](image)
The sum of two Gaussian pdf’s and its cdf

The standard Gaussian probability density function (pdf) \( f(x) \) and its corresponding cumulative density function \( F(x) \) are given by the well-known formulas

\[
f(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left( -\frac{|x - \mu|^2}{2\sigma^2} \right) \tag{133}
\]

and

\[
F(x) = \int_{-\infty}^{x} f(t) \, dt = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{x - \mu}{\sigma \sqrt{2}} \right) \right] \tag{134}
\]

Both functions are well-known in literature, and are easily available via tables or by computer packages. However, when a pdf is given as a sum of a few pdf’s in analytical form, then the computation of its corresponding cdf in analytical form will cost more effort. In such cases a numerical approach can be helpful. As an example, we give the result of a numerical computation of the cdf corresponding to the sum of two standard Gaussian pdf’s, each with different mean \( \mu \) and standard deviation \( \sigma \). See figures 28 and 29. Notice that when comparing both figures we see that the pdf curve of \((y_1 + y_2)/2\) is more pronounced than the corresponding pdf.

**Remark:** In the figure above we have used the notation \( \mathcal{N}(\mu, \sigma) \) instead of the more commonly used notation \( \mathcal{N}(\mu, \sigma^2) \).
Figure 29: The cdf of the sum of 2 standard Gaussian density functions.
Consider the function

\[ f(x) = \left( \frac{x}{\alpha} \right)^{\beta} \log \left( \frac{x}{\alpha} \right) \quad \text{for } x > 0 \text{ and } \beta > 0, \alpha \geq 1 \]  

(135)

\[ f(x) = 0 \quad \text{for } x = 0 \]

Then we have (see also figure 30)

- \[ f'(x) = \frac{1}{\alpha} \left( \frac{x}{\alpha} \right)^{\beta-1} \left( 1 + \beta \log \left( \frac{x}{\alpha} \right) \right) \]
- \[ f'(x) = 0 \quad \text{for } x = \hat{x} = \alpha e^{-1/\beta} \text{ and } f(\hat{x}) = -1/(e\beta) \]
- \[ f'(x) > 0 \quad \text{for } x > \alpha \text{ and } f'(x) < 0 \quad \text{for } 0 < x < \hat{x} \]
- \[ f(x) \text{ has a minimum for } x = \hat{x} \]

Figure 30: The function \( f(x) = (x/\alpha)^\beta \log(x/\alpha) \). Its minimum is marked by an *.
• $f(x)$ is monotonically increasing for $x > \hat{x}$

• $f(x)$ is monotonically decreasing for $0 < x < \hat{x}$

• $f(\alpha) = 0$ with $f'(\alpha) = \frac{1}{\alpha}$

• Substituting $x/\alpha = \exp(-y)$ in $f(x)$ gives $f(x(y)) = -ye^{-\beta y}$. Hence,

• $\lim_{x \to 0} f(x) = 0$

\[
\int_0^P f(x)dx = \frac{P}{\beta + 1}f(P) - \frac{P}{(\beta + 1)^2}\left(\frac{P}{\alpha}\right)^\beta
\]

(136)

and

\[
\int_0^{\Delta x} f(x)dx = \frac{\Delta x}{\beta + 1}f(\Delta x) - \frac{\Delta x}{(\beta + 1)^2}\left(\frac{\Delta x}{\alpha}\right)^\beta
\]

(137)

It can easily be seen from (137) and Appendix H that

\[
\int_0^{\Delta x} f(x)dx = O(\Delta x), \Delta x \to 0.
\]

(138)
I Generating a random number from an empirical pdf

Let $E_{cdf}(x)$ denote an empirical cumulative density function computed by numerical integration of a given empirical pdf denoted by $E_{pdf}(x)$. We assume that $E_{cdf}(x)$ is a monotonically increasing function. Let the coordinates \(\{(x_i, y_i) | i = 1, \ldots, N\}\) be located on the curve $E_{cdf}(x)$. Thus, $y_i = E_{cdf}(x_i)$ or equivalently, $x_i = E_{cdf}^{-1}(y_i)$ for all $i$.

Let \(\{\text{rnd}(i) | i = 1, \ldots, K\}\) be a given set of uniformly on $[0, 1]$ distributed random values. We want to determine the set \(\{x_i | i = 1, \ldots, K\}\) such that each $x_i = E_{cdf}^{-1}(\text{rnd}(i))$ as accurate as possible. This is done by linear interpolation as follows. See figure 31.

Figure 31: Generating a random value corresponding to an empirical cdf, using linear interpolation.

Let $r$ be an arbitrarily chosen value from the set $\{\text{rnd}(i) | i = 1, \ldots, K\}$. We introduce

$$y^+ = \min_{1 \leq i \leq N} \{y_i | y_i \geq r\} \quad \text{and}$$

$$y^- = \max_{1 \leq i \leq N} \{y_i | y_i \leq r\}$$

(139)

Notice that $y^+ > y^-$ due to the assumed monotonicity of $E_{cdf}(x)$.

Let $x^+$ and $x^-$ correspond to $y^+$ and $y^-$, respectively. Thus,

$$x^+ = E_{cdf}^{-1}(y^+) \quad \text{and} \quad x^- = E_{cdf}^{-1}(y^-)$$

(140)
Clearly, $x^+ > x^-$. Moreover, we define

\[
\begin{align*}
    dy^+ &= |y^+ - r|, \quad dy^- = |y^- - r|, \\
    dx^+ &= |x^+ - x_r|, \quad dx^- = |x^- - x_r|, \\
    dy &= dy^+ + dy^- , \quad dx = dx^+ + dx^- 
\end{align*}
\]

(141)

Notice that $dy^+ \geq 0$ and $dy^- \geq 0$, but not both of them can be equal to zero.

If $dy^+ = 0$ then $r = y^+$ and $dx^+ = 0$.

If $dy^- = 0$ then $r = y^-$ and $dx^- = 0$

Recalling fig(31) and using linear interpolation, we have

\[
\begin{align*}
    \frac{x_r - x^-}{dx} &= \frac{dy^-}{dy}, \quad \frac{x^+ - x_r}{dx} = \frac{dy^+}{dy} \quad \text{and} \\
    \tan(\alpha) &= \frac{dy}{dx} = \frac{dy^+}{dx^+} = \frac{dy^-}{dx^-} 
\end{align*}
\]

(142)

It is easily seen from (142) that

\[
x_r = x^- + dx^- = x^+ - dx^+ 
\]

(143)

Hence, we can define the value $x_r = E_{cdf}^{-1}(r)$ as follows:

\[
x_r = \begin{cases} 
    x^+ - dx^+ & \text{if } 0 \leq dy^+ < dy^- \\
    x^- + dx^- & \text{if } 0 \leq dy^- < dy^+ 
\end{cases}
\]

(144)

This value $x_r$ can now be interpreted as a random value corresponding to the probability function $E_{pdf}(x)$. By taking successively $r = rnd(i), i = 1, \ldots, K$ and computing the corresponding $x_r$-values as indicated above, we obtain a set of K values being randomly distributed conform the pdf $E_{pdf}(x)$.
J  Conclusions

In this report we addressed the problem of estimating the parameters of a Generalized Gaussian distribution. We studied in detail a numerical method known from literature. We proposed a new numerical reliable and fast procedure to compute the desired parameters. This procedure is based on a detailed analysis of all functions involved in order to have fast and safe convergence of the used Newton-Raphson method.

K  Acknowledgements

The authors wish to thank Dr. E.J.W. ter Maten, Dr. M. Hochstenbach, Dr. A. Di Bucchianico and Dr. A.J.E.M. Janssen for their stimulating discussions and useful suggestions.

References


## Contents

1 General Introduction .......................... 1

2 Abstract ................................... 2

3 Mathematical problem description ............ 3

4 Behaviour of the function $\hat{\alpha}(\beta)$ .... 7
   4.1 Asymptotic behaviour of $g(\beta)$ .............. 8

5 Numerical experiments ........................ 9
   5.1 Procedure for parameter estimation of an empirical pdf .... 9
   5.2 Application of the estimation procedure .......... 13

A Approximation of a sum $\frac{1}{N}\sum_{i=1}^{N} f(x_i)$ by an integral .... 17
   A.1 A few examples ............................. 19
   A.2 An accurate approximation of the function $g(\beta)$ ......... 20
   A.3 Behaviour of $g(\beta)$ as $\beta \searrow 0$ .............. 22

B Approximation of a sum $\frac{1}{N}\sum_{i=1}^{N} f(x_i)$ for random $x_i$ .... 25

C Analysis of the behaviour of $g(\beta)$ ............. 29
   C.1 Behaviour of $g_1(\beta)$ ........................ 29
   C.2 Behaviour of $Q(\beta)$ .......................... 31
   C.3 Behaviour of $P(\beta)$ .......................... 31
   C.4 Behaviour of $P(\beta) - Q(\beta)$ ................. 31
   C.5 Behaviour of $g(\beta)$ .......................... 32

D Asymptotics .................................. 33
   D.1 Asymptotic behaviour of $\hat{a}(\beta)$ .............. 33

E Analysis of $\frac{\partial L}{\partial \beta}$ .................... 34
   E.1 Analysis of $F(\beta)$ ............................ 35
   E.2 Analysis of $S(\beta)$ ............................. 37

F Properties of $\Gamma(x)$, $\Psi(x)$, $\Psi'(x)$, $F(\beta)$ and $F'(\beta)$ .... 43
   F.1 Properties of $\Psi'(x)$ ............................ 43
   F.2 Properties of $\Psi(x)$ .............................. 44
   F.3 Properties of $\Gamma(x)$ ............................. 44
   F.4 Behaviour of $F(\beta)$ and $F'(\beta)$ .............. 45
<table>
<thead>
<tr>
<th></th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>The sum of two Gaussian pdf’s and its cdf</td>
<td>47</td>
</tr>
<tr>
<td>H</td>
<td>Behaviour of f(x)</td>
<td>49</td>
</tr>
<tr>
<td>I</td>
<td>Generating a random number from an empirical pdf</td>
<td>51</td>
</tr>
<tr>
<td>J</td>
<td>Conclusions</td>
<td>53</td>
</tr>
<tr>
<td>K</td>
<td>Acknowledgements</td>
<td>53</td>
</tr>
</tbody>
</table>
**PREVIOUS PUBLICATIONS IN THIS SERIES:**

<table>
<thead>
<tr>
<th>Number</th>
<th>Author(s)</th>
<th>Title</th>
<th>Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>15-37</td>
<td>N. Kumar J.H.M. ten Thije Boonkamp B. Koren</td>
<td>Flux approximation scheme for the incompressible Navier-Stokes equations using local boundary value problems</td>
<td>Dec. ‘15</td>
</tr>
<tr>
<td>15-38</td>
<td>C. Filosa J.H.M. ten Thije Boonkamp W.J. IJzerman</td>
<td>A new ray tracing method in phase space using $\alpha$-shapes</td>
<td>Dec. ‘15</td>
</tr>
<tr>
<td>15-39</td>
<td>F.A. Radu K. Kumar J.M. Nordbotten I.S. Pop</td>
<td>A convergent mass conservative numerical scheme based on mixed finite elements for two-phase flow in porous media</td>
<td>Dec. ‘15</td>
</tr>
<tr>
<td>15-40</td>
<td>T.G.J. Beelen J.J. Dohmen</td>
<td>Parameter estimation for a generalized Gaussian distribution</td>
<td>Dec. ‘15</td>
</tr>
</tbody>
</table>

Ontwerp: de Tantes, Tobias Ranstorp, CWI