ML-Estimation in the Location-Scale-Shape Model of the Generalized Logistic Distribution

Klaus Abberger, University of Konstanz, 78457 Konstanz, Germany

Abstract: A three parameter (location, scale, shape) generalization of the logistic distribution is fitted to data. Local maximum likelihood estimators of the parameters are derived. Although the likelihood function is unbounded, the likelihood equations have a consistent root. ML-estimation combined with the ECM algorithm allows the distribution to be easily fitted to data.

Keywords: ECM algorithm, generalized logistic distribution, location-scale-shape model, maximum likelihood estimation

1 Introduction

The generalized logistic distribution with location (θ), scale (σ) and shape (b) parameters has the density:

$$f(x) \;=\; \frac{\frac{b}{\sigma}e^{-\frac{(x-\theta)}{\sigma}}}{(1+e^{-\frac{(x-\theta)}{\sigma}})^{b+1}},\ b>0,\ \sigma>0,\ \theta\in R,\ x\in R. \tag{1}$$

Additional generalizations of the logistic distribution are discussed by Johnson, Kotz and Balakrishnan (1995). For $b=1$ the distribution is symmetric, for $b<1$ it is skewed to the left and for $b>1$ it is skewed to the right. The moment generating function is

$$m(t) = e^{\theta t} \frac{\Gamma(1-\sigma t)\Gamma(b+\sigma t)}{\Gamma(b)}. \tag{2}$$

From that follows

$$E[X] = m'(0) = \theta + \sigma(\psi(b) - \psi(1)) \tag{3}$$

$$E\left[X^2\right] - E^2[X] = m''(0) - m'(0)^2 = \sigma^2(\psi'(1) + \psi'(b)), \tag{4}$$

with $\psi(b) = \Gamma'(b)/\Gamma(b)$ the digamma function (see e.g. Abramowitz, Stegun (1972)).

Fisher´s coefficient of skewness is

$$\gamma_1 = \frac{E[(X - E[X])^3]}{E[(X - E[X])^2]^{3/2}} = \frac{\psi''(b) - \psi''(1)}{(\psi'(b) + \psi'(1))^{3/2}}. \tag{5}$$

Since γ_1 is location and scale invariant the skewness of the distribution depends only on parameter b. Invariance of the measure is requested because two random variables Y and $\theta + \sigma Y$ should have the same degree of skewness. That means a random variable X with generalized logistic distribution has a variance depending on the parameters b and σ, with σ a part only affecting scale and a part b affecting scale and furthermore shape of the distribution.

The logistic distribution has been one of the most important statistical distributions because of its simplicity and also its historical importance as growth curve. Some applications of logistic distributions in the analysis of quantal response data, probit analysis, and dosage response studies, and in several other situations have been mentioned by Johnson, Kotz and Balakrishnan (1995). The generalized logistic distributions are very useful classes of densities as they posses a wide range of indices of skewness and kurtosis. Therefore an important application of the generalized logistic distribution is its use in studying robustness of estimators and tests.

Balakrishnan and Leung (1988) present two real data examples for the usefulness of the generalized logistic distribution. One about oxygen concentration and another about resistance of automobile paint. In both examples the authors choose $b = 2$ by eye and verify the validity of this assumption by Q-Q plots.

Zelterman (1987) describes method of moments and Bayesian estimation of the parameters θ, σ and b. The method of moments estimators rely on the first three sample moments and have large variances. So Zeltermann concludes the the method of moments estimators are useful only as starting values for other iterative estimation techniques. He further shows that maximum likelihood estimators do not exists (a point which we demonstrate graphically in Sec. 2 of this paper) and presents as conclusion Bayesian estimators.

Gerstenkorn (1992) discusses ML-estimators of b alone, because there is no problem with existency in this case.

In this paper the three parameter generalized logistic distribution is fitted to data. It is shown that, although the likelihood function is known to be unbounded at some points on the edge of the parameter space, the likelihood equations have a root which is consistent and asymptotically normally distributed.

The following section derives ML-estimators of θ, σ and b which are consistent. The ML-estimation is combined with the ECM algorithm in Section 3 to produce an iterative procedure for fitting the distribution. The final section in this paper presents numerical examples.

2 ML-estimation

Let $X_1, ..., X_n$ be an i.i.d. sample from the generalized logistic distribution. Then the log-likelihood is given by

$$
\begin{aligned}
l(b, \theta, \sigma; x_1, ..., x_n) &= n \ln(b) - n \ln(\sigma) - \frac{1}{\sigma} \sum_i (x_i - \theta) \\
&\quad -(b+1) \sum_i \ln \left(1 + e^{-\frac{1}{\sigma}(x_i - \theta)}\right).
\end{aligned} \tag{6}
$$

A closed-form expression for estimating b is as follows

$$
\hat{b} = \frac{n}{\sum_i \ln \left(1 + e^{-\frac{1}{\sigma}(x_i - \theta)}\right)}. \tag{7}
$$

Plugging in this estimator into the log-likelihood gives the concentrated log-likelihood

$$
l_c(\theta, \sigma; x_1, ..., x_n) = \frac{n\theta}{\sigma} - \sum_i \ln \left(1 + e^{-\frac{1}{\sigma}(x_i - \theta)}\right) \tag{8}
$$

$$
-n \ln \sum_i \ln \left(1 + e^{-\frac{1}{\sigma}(x_i - \theta)}\right) + H(\sigma, x), \tag{9}
$$

with H a function not depending on θ or b. The concentrated log-likelihood function is maximized for $\theta \to -\infty$ (Zelterman (1987) p.180). That means the concentrated log-likelihood diverges to infinity and the global maximum is not a consistent estimator of the parameters under consideration. This is a very interesting behaviour which results from introducing a scale parameter and a shape parameter. There is for example no equivalent problem in the location-scale model of the usual logistic distribution.

2.1 Graphical representation of the likelihood functions

The following figures show the concentrated log-likelihood and the conditional log-likelihood for a simulated sample of size $n = 200$ and parameters $b = 1, \theta = 0, \sigma = 1$. Figure 1 presents the concentrated log-likelihood. As stated above, the likelihood diverges as θ approaches minus infinity.

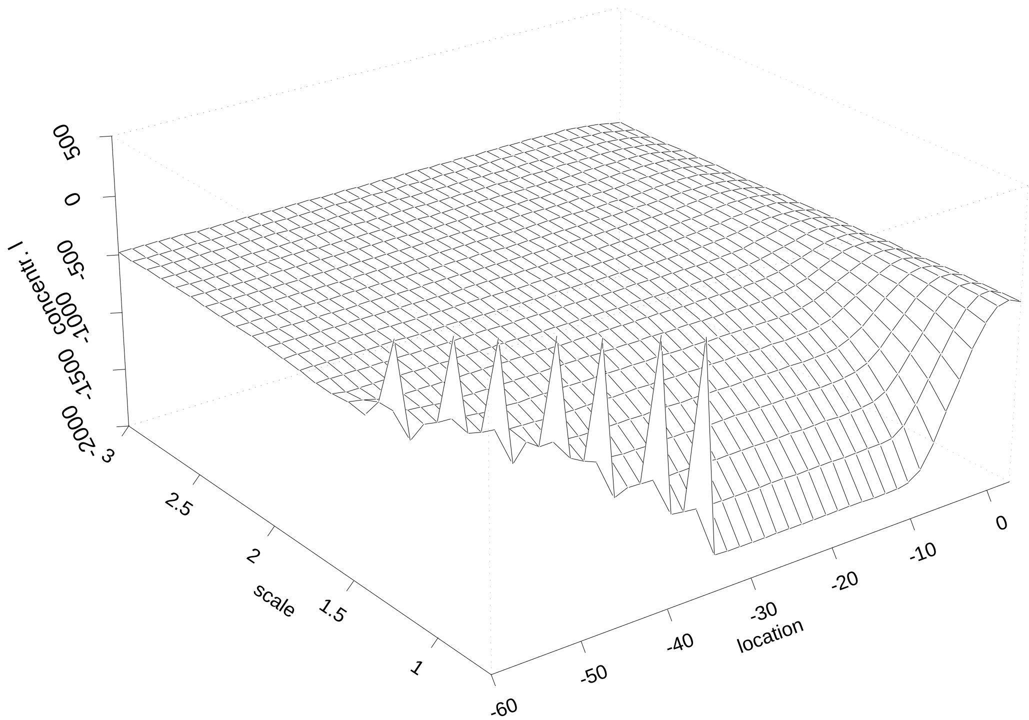

Figure 1: Concentrated log-likelihood ($b = 1, \theta = 0, \sigma = 1$)

Figure 2 contains the concentrated log-likelihood as in Figure 1 but for a smaller domain, revealing a local optimum at the true parameter values.

Figures 3 and 4 show the conditional likelihood given $b = 1$. In these graphs the problem of divergence vanishes, and the maximum of the function at the true parameters becomes clearly apparent. This behaviour motivates the use of the so called Expectation-Conditional Maximization Algorithm for the calculation of the ML-estimators.

2.2 Consistency of the ML-estimates

The global maximum of the log-likelihood is not a consistent estimator. But this does not imply that ML-estimation is impossible. There exists a sequence of local maxima which forms consistent estimators. To prove this we use a theorem of

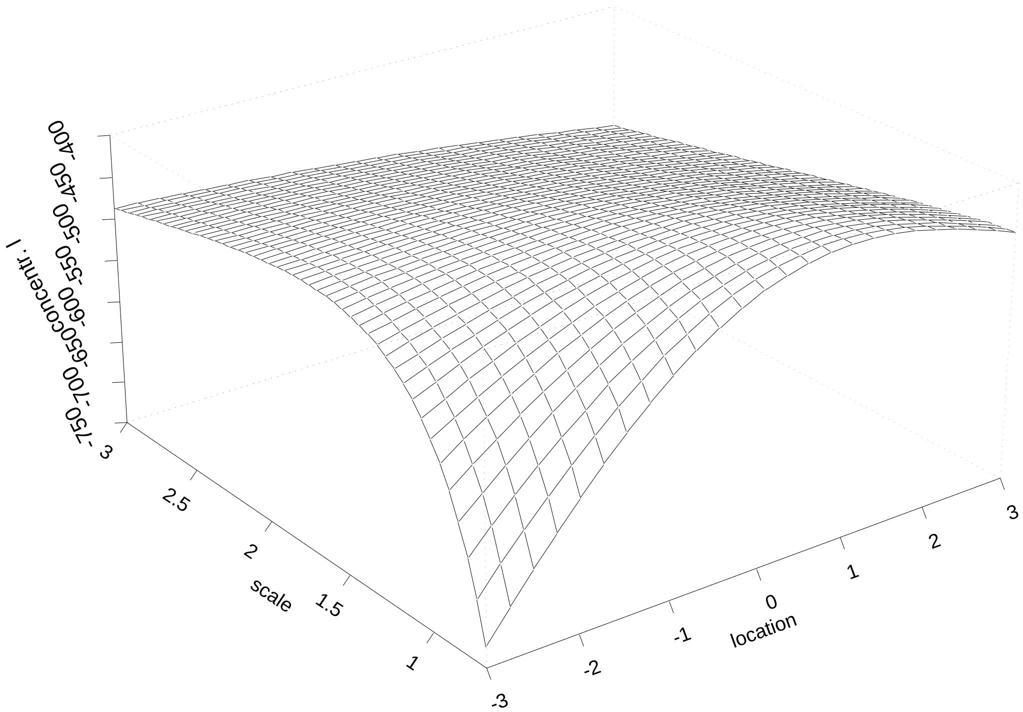

Figure 2: Concentrated log-likelihood ($b = 1, \theta = 0, \sigma = 1$)

Chanda (1954), which is stated here as a lemma.

Lemma: Let $f(x, \omega)$ be a density with $\omega = (\omega_1, \omega_2, ..., \omega_k)$ a vector of unknown parameters and let $x_1, ..., x_n$ be independent observations of X. The *likelihood equations* or *score functions* are

$$\frac{\delta \ln L}{\delta \omega} = 0, \tag{10}$$

with $\ln L = \sum_{i=1}^{n} \ln f(x_i, \omega)$.

Let ω_0 be the true value of the parameter vector ω, included in a closed region $\bar{\Omega} \subset \Omega$. If Conditions 1-3, given below, are fulfilled, then there exists a consistent estimator ω_n corresponding to a solution of the likelihood equations. Furthermore, $\sqrt{n}(\omega_n - \omega_0)$ is asymptotically normally distributed with mean zero and variance-covariance $I(\omega_0)^{-1}$, the Fisher information.

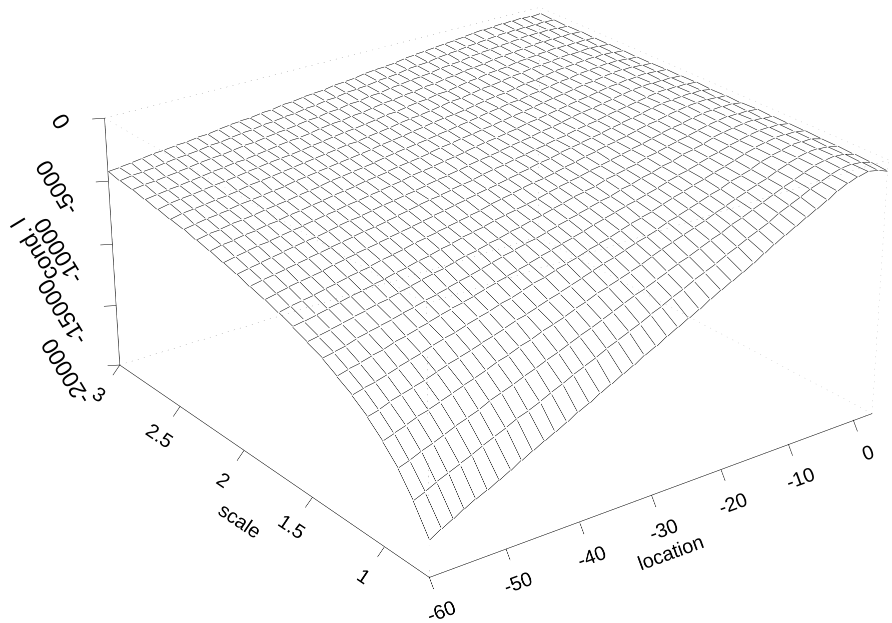

Figure 3: Conditional log-likelihood given $b = 1$ ($\theta = 0, \sigma = 1$)

Condition 1: for almost all x and for all $\omega \in \bar{\Omega}$

$$\left|\frac{\delta \ln f}{\delta \omega_r}\right|, \qquad \left|\frac{\delta^2 \ln f}{\delta \omega_r \delta \omega_s}\right| \qquad \text{and} \qquad \left|\frac{\delta^3 \ln f}{\delta \omega_r \delta \omega_s \delta \omega_t}\right|,$$

exist for all $r, s, t = 1, ..., k$.

Condition 2: for almost all x and for all $\omega \in \bar{\Omega}$

$$\left|\frac{\delta f}{\delta \omega_r}\right| < F_r(x), \qquad \left|\frac{\delta^2 f}{\delta \omega_r \delta \omega_s}\right| < F_{r,s}(x) \qquad \text{and} \qquad \left|\frac{\delta^3 \ln f}{\delta \omega_r \delta \omega_s \delta \omega_t}\right| < H_{r,s,t}(x),$$

with H such that $\int_{-\infty}^{\infty} H_{r,s,t}(x) f \, dx \leq M < \infty$ and $F_r(x)$ and $F_{r,s}(x)$ bounded for all $r, s, t = 1, ..., k$.

Condition 3: for all $\omega \in \bar{\Omega}$

$$I(\omega) = \int_{-\infty}^{\infty} \left(\frac{\delta \ln f}{\delta \omega}\right) \left(\frac{\delta \ln f}{\delta \omega}\right)^T f \, dx$$

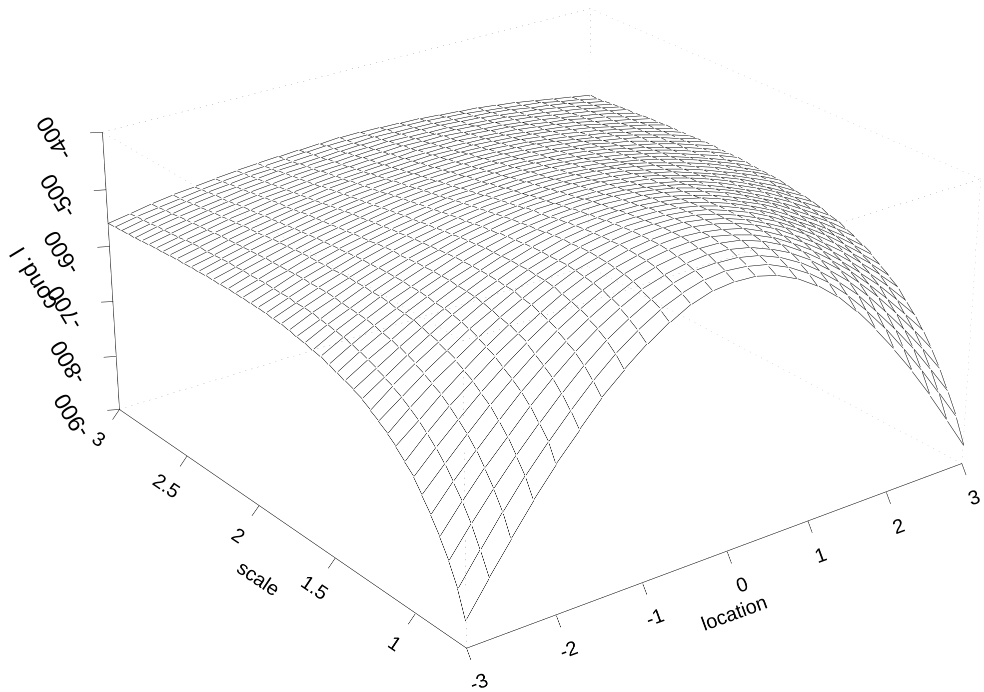

Figure 4: Conditional log-likelihood given $b = 1$ ($\theta = 0, \sigma = 1$)

is positive definite.

With the help of this lemma we can show the following theorem about the consistency of the local ML-estimates.

THEOREM 1: Let $\omega = (b, \theta, \sigma)$ *and* Ω *the parameter space be given by*

$$0 < b < \infty, \qquad -\infty < \theta < \infty, \qquad 0 < \sigma < \infty.$$

Let ω_0*, the true parameter value, be contained in a closed set* $\bar{\Omega}$ *which is a subset of* Ω*. If the random variable* X_i*, is i.i.d. with the generalized logistic distribution, then there exists a consistent root* ω_n *of the likelihood equations and* $\sqrt{n}(\omega_n - \omega_0)$ *is asymptotically normally distributed with mean zero and variance* $I(\omega_0)^{-1}$.

Proof:

The proof consists of the verification of Conditions 1-3 and is given in the appendix.

The theorem does not provide any information about which root, in the event there is more than one, is consistent. So there is a necessity for a good algorithm which helps us to find the ML-estimators. Also an initial consistent estimate is needed as starting point for the maximization. Therefore the above mentioned method of moments estimators derived by Zeltermann (1987) can be used.

3 ECM algorithm

The specific behaviour of the likelihood under consideration suggests the use of the ECM-Algorithm (Expectation-Conditional Maximation-Algorithm) discussed by Meng and Rubin (1993). This algorithm consists of several conditional maximizations. Since we know that the problem of divergence of the likelihood vanishes when the function is maximized for θ and σ conditional on b and for b conditional on θ and σ the estimator is in closed form it is reasonable to use an algorithm which separates these steps. Exactly this does the ECM-Algorithm. Also other optimization algorithms could be used but the use of the ECM-Algorithm should protect from running into divergence problems.

ECM is a generalization of the EM-Algorithm. These algorithms have been developed for estimation problems with missing data. But they can also be applied without missing data. Let $l(\omega; x)$ denote the log-likelihood under consideration, than the E-step, generating the function which has to be maximized, is without missing data an identity operation of the form

$$Q(\omega) \quad := \quad l(\omega; x), \tag{11}$$

with $\omega = (b \quad \theta \quad \sigma)' \in \Omega \subset R^3$ and $x = (x_1 \quad \ldots \quad x_n)'$ and $Q(\omega)$ the function which has to be maximized. The CM-iteration takes advantage of the simplicity of conditional maximum likelihood estimation by replacing the maximum-step with several conditional maximum-steps. More precisely, CM replaces each M-step by a sequence of S conditional maximum steps, each of which maximizes the Q function over ω but with some vector function of ω, $g_s(\omega)$ $(s = 1, ..., S)$, fixed at its previous value. The procedure is then used iteratively until it converges.

If the log-likelihood of the generalized logistic distribution is maximized in two steps, then the $S = 2$ constraints are

$$\begin{aligned} g_1(\omega) &= (\theta^{(t)} \qquad \sigma^{(t)})' && (12) \\ g_2(\omega) &= (b^{(t+1/2)}). && (13) \end{aligned}$$

The t-th iteration step finds $\omega^{(t+s/2)}$, $s = 1, 2$, so that

$$Q(\theta^{(t+s/2)}|\omega^{(t)}) \geq Q(\omega|\omega^{(t)}), \qquad \{\omega \in \Omega | g_s(\omega) = g_s(\omega^{(t+(s-1)/2)}\}. \tag{14}$$

Then the value of ω for starting the next iteration, $\omega^{(t+1)}$, is defined as the output of the final step of the previous iteration, that is $\omega^{(t+s/S)}$.

For estimating the location-scale model of the generalized distribution the first CM-step is calculating

$$b^{(t+1/2)} = \frac{n}{\sum_i \ln\left(1 + e^{-\frac{1}{\sigma^{(t)}}(x_i - \theta^{(t)})}\right)}. \tag{15}$$

The second step consists of maximizing

$$n \ln(b^{(t+1/2)}) - n\ln(\sigma) - \frac{1}{\sigma}\sum_i (x_i - \theta) - (b^{(t+1/2)} + 1)\sum_i \ln\left(1 + e^{-\frac{1}{\sigma}(x_i - \theta)}\right) \tag{16}$$

with a usual optimization algorithm, finding the values of $\theta^{(t+1)}$ and $\sigma^{(t+1)}$.

Meng and Rubin (1993) show the following theorem about the ECM-Algorithm:

THEOREM 2:

Suppose that all the conditional maximizations of ECM are unique. Then all limit points of any ECM sequence $\{\omega^{(t)}, t \geq 0\}$ belong to the set

$$\Gamma \equiv \left\{ \omega \in \Omega \mid \frac{\delta l(\omega; x)}{\delta \omega} \in J(\omega) \right\}, \tag{17}$$

with

$$J(\omega) \equiv \bigcap_{s=1}^{S} J_s(\omega) \tag{18}$$

and $J_s(\omega)$ the column space of the gradient of $g_s(\omega)$, that is

$$J_s(\omega) = \{\Delta g_s(\omega)\lambda \mid \quad \lambda \in R^{d_s}\}, \tag{19}$$

and d_s is the dimensionality of the vector function $g_s(\omega)$. It is assumed that $g_s(\omega)$ is differentiable and the corresponding gradient, $\Delta g_s(\omega)$, is of full rank at $\omega \in \Omega_0$, the interior of Ω.

Meng und Rubin (1993) state that the assumption that all conditional maximizations are unique can be eliminated, if we force the conditions: a) continuity of $D^{10}Q(\omega|\omega')$ in both ω and ω', D^{10} denoting the first derivative in the first argument of Q, and b) continuity of $\Delta g_s(\omega)$ for all s.

In the present application $g_s(\omega)$ is a partition of the parameter space. From that follows, that $J_1(\omega)$ and $J_2(\omega)$ are orthogonal and because of that $J(\omega) = \{0\}$. So the algorithm converges to a local optimum of the likelihood of the generalized logistic distribution. In the event there is more than one root, it is not guaranteed that the consistent root is found. Therefore initial consistent estimators like Zeltermann´s (1987) method of moments estimators should be used as starting values.

4 Numerical examples

The above mentioned Theorem 2 states that the iterative ECM procedure converges to a root of the likelihood. To demonstrate that the procedure really converges, two numerical examples are worked out. In both examples 1000 repetitions for every three sample sizes $n_1 = 100$, $n_2 = 200$ and $n_3 = 500$ are carried out. In the first example the true parameter values are $b = 2$, $\theta = 1$, $\sigma = 2$. The estimation results are presented in Figure 5. The second example, shown in figure 6, uses the true parameters $b = 0.5$, $\theta = 1$, $\sigma = 2$.

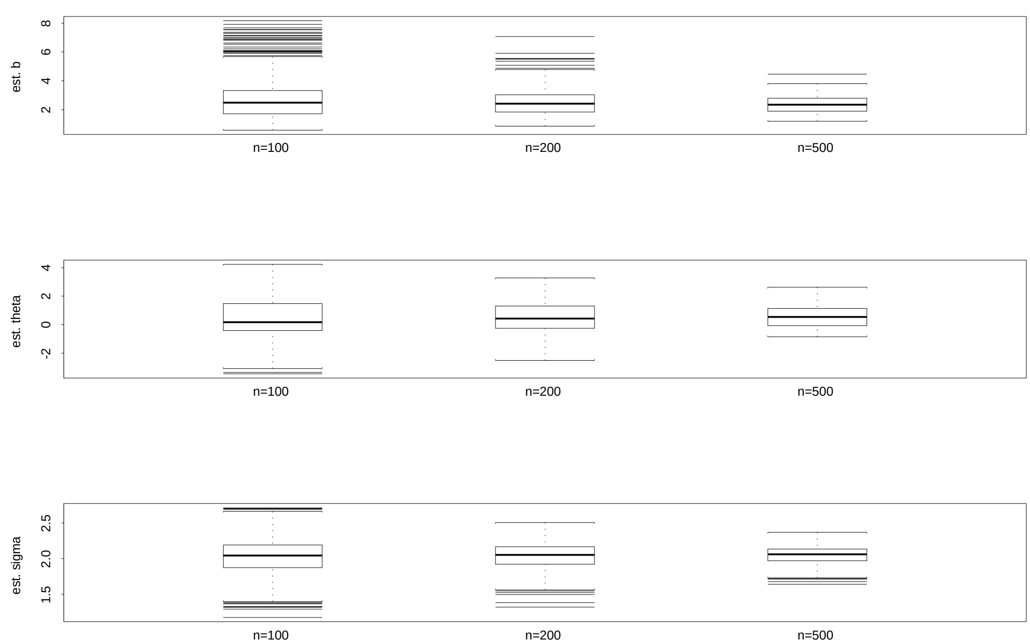

Figure 5: Simulation example for $b = 2$, $\theta = 1$, $\sigma = 2$ (1000 repetitions)

In the first example the distribution is right skewed, while in the second one the distribution is left skewed. In both examples the ECM algorithm works very well for all three parameters. That means the procedure does not run into divergence

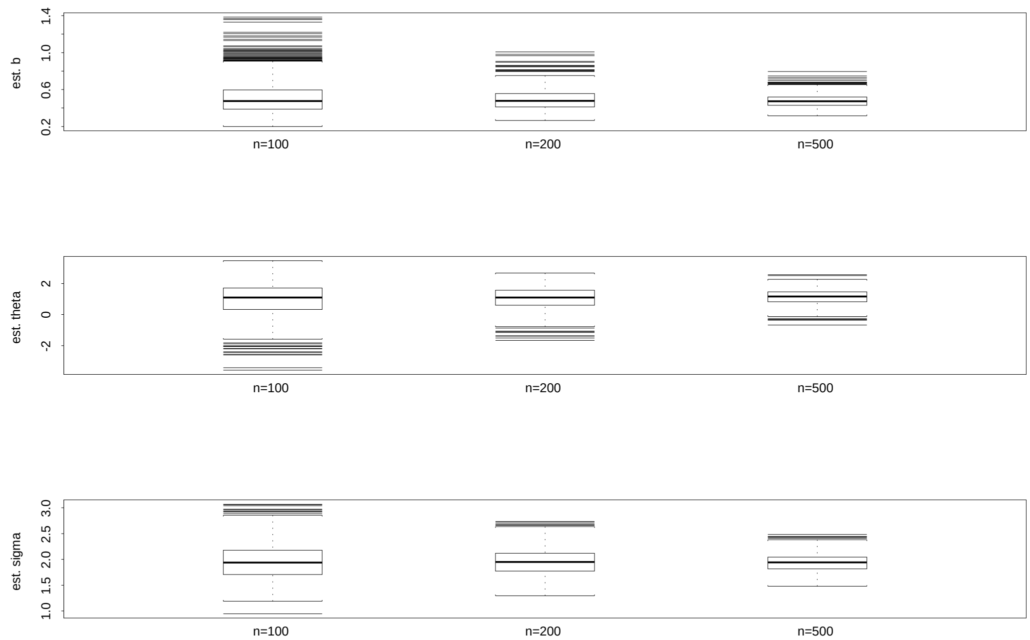

Figure 6: Simulation example for $b = 0.5$, $\theta = 1$, $\sigma = 2$ (1000 repetitions)

problems. All repetitions converged to reasonable results. None of them leads to unreliable estimations.

5 Appendix

Proof of Theorem 1:

The proof consists of verification of conditions 1-3.

Condition 1:

Let $u = e^{-\frac{1}{\sigma}(x-\theta)}$. Then the derivatives of u are as follows:

$$\begin{aligned}
u_\theta &= \frac{1}{\sigma}u \\
u_{\theta\theta} &= \frac{1}{\sigma^2}u \\
u_{\theta\theta\theta} &= \frac{1}{\sigma^3}u \\
u_\sigma &= \frac{1}{\sigma^2}(x-\theta)u \\
u_{\sigma\sigma} &= -\frac{2}{\sigma^3}(x-\theta)u + \frac{1}{\sigma^4}(x-\theta)^2u \\
u_{\sigma\sigma\sigma} &= \frac{6}{\sigma^4}(x-\theta)u - \frac{6}{\sigma^5}(x-\theta)^2u + \frac{1}{\sigma^6}(x-\theta)^3u \\
u_{\theta\sigma} &= -\frac{1}{\sigma^2}u + \frac{1}{\sigma^3}(x-\theta)u \\
u_{\theta\theta\sigma} &= -\frac{2}{\sigma^3}u + \frac{1}{\sigma^4}(x-\theta)u \\
u_{\theta\sigma\sigma} &= \frac{2}{\sigma^3}u - \frac{4}{\sigma^4}(x-\theta)u + \frac{1}{\sigma^5}(x-\theta)^2u.
\end{aligned}$$

The first and third derivatives of the likelihood are listed below, because they are also required later. They are as follows

$$\begin{aligned}
\frac{\delta l(b,\theta,\sigma;x)}{\delta b} &= \frac{1}{b} - \ln(1+u) \\
\frac{\delta l}{\delta\theta} &= \frac{1}{\sigma} - (b+1)\frac{u_\theta}{1+u} \\
\frac{\delta l}{\delta\sigma} &= -\frac{1}{\sigma} + \frac{1}{\sigma^2}(x-\theta) - (b+1)\frac{u_\sigma}{1+u} \\
\frac{\delta^3 l}{\delta b^3} &= -\frac{1}{b^2} \\
\frac{\delta^3 l}{\delta b^2\delta\theta} &= 0 \\
\frac{\delta^3 l}{\delta b^2\delta\sigma} &= 0 \\
\frac{\delta^3 l}{\delta\theta^2\delta b} &= -\left\{\frac{u_{\theta\theta}(1+u) - u_\theta^2}{(1+u)^2}\right\} \\
\frac{\delta^3 l}{\delta\sigma^2\delta b} &= -(b+1)\left\{\frac{u_{\theta\theta\sigma}(1+u) + u_{\theta\theta}u_\sigma - 2u_\theta u_{\sigma\theta} - 2(u_{\theta\theta}(1+u) - u_\theta^2)(1+u)u_\sigma}{(1+u)^4}\right\} \\
\frac{\delta^3 l}{\delta\sigma^2\delta\theta} &= \frac{2}{\sigma^3} - (b+1)\quad \frac{u_{\sigma\sigma\theta}(1+u) - u_{\sigma\sigma}u_\theta}{(1+u)^2} - \frac{2u_\sigma u_{\sigma\theta}(1+u)^2 - 2u_\sigma^2(1+u)u_\theta}{(1+u)^4} \\
\frac{\delta^3 l}{\delta\theta^3} &= -(b+1)\left\{\frac{u_{\theta\theta\theta}(1+u) + u_{\theta\theta}u_\theta - 2u_\theta u_{\theta\theta} - 2(u_{\theta\theta}(1+u) - u_\theta^2)(1+u)u_\theta}{(1+u)^4}\right\} \\
\frac{\delta^3 l}{\delta\sigma^3} &= -\frac{2}{\sigma^3} + \frac{6}{\sigma^4}(x-\theta) - (b+1)\quad \frac{u_{\sigma\sigma\sigma}(1+u) - u_{\sigma\sigma}u_\sigma}{(1+u)^2} - \frac{2u_\sigma u_{\sigma\sigma}(1+u)^2 - u_\sigma^3 2(1+u)}{(1+u)^4} \quad .
\end{aligned}$$

All of the third derivatives exist for all y and all ω in $\bar{\Omega}$ with $\bar{\Omega}$ any closed set in Ω. So Condition 1 is satisfied.

Condition 2:

The first derivatives of the density are

$$\begin{aligned}
\frac{\delta f}{\delta b} &= \frac{u}{\sigma(1+u)^{b+1}} - \frac{bu\ln(1+u)}{\sigma(1+u)^{b+1}} \\
\frac{\delta f}{\delta \theta} &= \frac{bu_\theta}{\sigma(1+u)^{b+1}} - \frac{b(b+1)uu_\theta}{\sigma(1+u)^{b+2}} \\
\frac{\delta f}{\delta \sigma} &= -\frac{bu}{\sigma^2(1+u)^{b+1}} + \frac{bu_\sigma}{\sigma(1+u)^{b+1}} - \frac{b(b+1)uu_\sigma}{\sigma(1+u)^{b+2}}.
\end{aligned}$$

These derivatives of f are continuous in x and in ω, so they are certainly bounded for $\omega \in \bar{\Omega}$ and x in any closed interval of the real line. It is only necessary to consider their behavior for extreme values of x. Since the exponential function grows faster than any power, the terms xe^{-x} are bounded. From that follows, that the derivatives of f are bounded for $x \to \infty$. For the case $x \to -\infty$ we use that $e^{-x}/(1+e^{-x})^{b+1} = O(e^{bx}) = e^{-2x}/(1+e^{-x})^{b+2}$. So the derivatives are also bounded for the case $x \to -\infty$. The second derivatives are also continuous, and for $x \to \infty$ the relevant expressions are of the form xe^{-x}, thus the second derivatives are bounded in this case. Looking at $x \to -\infty$ there are in addition terms of the form $e^{-3x}/(1+e^{-x})^{b+3}$. Since these terms are of order $O(e^{bx})$ also the second derivatives are bounded.

The third derivatives of the logarithmized density are listed under Condition 1. Similar to the considerations, the second derivatives of the density holds that for $x \to \infty$ the terms in the curved brackets tend to zero. So in this case the relevant value is x in the derivative $\delta^3 l/\delta^3\sigma$. Since the moment generating function of the generalized logistic distribution exists for t in a region around zero, all of the moments of f exist and are finite. A constant M can be choosen so that Condition 2 is satisfied. Remains the case $x \to -\infty$. The term with the largest order is contained in the derivative $(\delta^3 l/\delta\sigma^3)$. It holds $u_{\sigma\sigma\sigma}/(1+u) = x^3 O(1) = u_\sigma^3/(1+u)^3$. Since all moments of X exist, a constant M can be chosen, so that Condition 2 is fulfilled.

Condition 3:

Since $I(\omega)$ is a dispersion matrix it is positive semidefinite. The matrix is positive definite, if the statistics $\delta l/\delta b$, $\delta l/\delta\theta$ and $\delta l/\delta\sigma$ are affine independent. So $I(\omega)$ is positive definite unless

$$\sum_{i=1}^{3} \frac{\delta l}{\delta \omega_i} \cdot \lambda_i = 0,$$

for all x and with $\lambda_1, \lambda_2, \lambda_3$ not all equal to zero. An examination of the first derivatives of f, given in Condition 1, shows that Condition 3 is satisfied.

6 Literature

Abramowitz M., Stegun I.A. (1972): Handbook of Mathematical Functions. Dover Publications, New York.

Balakrishnan N., Leung M.Y. (1988): Means, variances and covariances of order statistics, BLUE´s for the type I generalized logistic distribution, and some applications. Commun. Statist. - Simula. 17, 51-88.

Chanda K.C. (1954): A note on the consistency and maxima of the roots of likelihood equations. Biometrika 41, 56-61.

Gerstenkorn T. (1992): Estimation of a parameter of the logistic distribution. Transactions of the Eleventh Prague Conference on Information Theory, Statistical Decision Functions, Random Processes, Prague: Academia Publishing House of the Czechoslovak Academy of Sciences, 441-448.

Johnson, N.L., Kotz S., Balakrishnan N. (1995): Continuous Univariate Distributions Volume 2. Wiley, New York.

Kiefer N.M. (1978): Discrete parameter variation: efficient estimation of a switching regression model. Econometrica 46, 427-434.

Lehmann E.L. (1998): Theorie of Point Estimation. Springer, New York.

Meng X.L., Rubin D.B. (1993): Maximum likelihood estimation via the ECM algorithm: A general framework. Biometrika 80, 267-278.

Zelterman D. (1987): Parameter estimation in the generalized logistic distribution. Computational Statistics & Data Analysis 5, 177-184.