Influence functions of trimmed likelihood estimators for lifetime experiments

Christine H. Müller
TU Dortmund University

Sebastian Szugat *
TU Dortmund University

Nuri Celik †
Bartin University

Brenton Clarke
Murdoch University

May 1, 2015

Abstract

We provide a general approach for deriving the influence function for trimmed likelihood estimators using the implicit function theorem. The approach is applied to lifetime models with exponential or lognormal distributions possessing a linear or nonlinear link function. A side result is that the functional form of the trimmed estimator for location and linear regression used by Bednarski and Clarke (1993, 2002) and Bednarski et al. (2010) is not generally always the correct functional form of a trimmed likelihood estimator. However, it is a version for which the influence function has a treatable form. A real data example shows the effect of trimming using a nonlinear link function for either the exponential or lognormal distribution.

2000 Mathematics Subject Classification. Primary 62F35, 62J02, 62J12; secondary 62N05, 62F10, 62G35.
Keywords: Outlier robustness, robust estimation, generalized linear model, lifetime distribution.

*Research supported by DFG via the Collaborative Research Centre SFB 823 'Statistical modelling of nonlinear dynamic processes' (Project B5)

†Research supported by TÜBITAK

1 Introduction

In this article, we consider simple lifetime experiments, where the observations of the lifetimes are independent and identically distributed, and accelerated lifetime experiments, where the lifetimes are observed at different stress levels, typically at stress levels at higher values than in practice to reduce the observation time. Usually maximum likelihood estimators are considered for these experiments where typical lifetime distributions as exponential distributions, Weibull distributions or lognormal distributions are used. However, maximum likelihood estimators are very sensitive to outliers. Therefore, we consider here the trimmed likelihood estimators (TLE) proposed by Müller and Neykov (2003) which are outlier robust modifications of maximum likelihood estimators. The trimmed likelihood estimators extend the least median squares estimators and the least trimmed squares estimators of Rousseeuw (1984) and Rousseeuw and Leroy (1987) by replacing the likelihood functions of the normal distribution by likelihood functions of other distributions. Although the TLEs are constructed for a specific likelihood function, for example like the likelihood given by the exponential distribution, the TLEs can be applied to data from any other distribution.

Many exponents of trimmed likelihood estimators use breakdown point as a principal robustness measure, see Vandev (1993), Müller (1995), Mili, L. and Coakley, C.W. (1996), Müller (1997), Vandev and Neykov (1998), Müller and Neykov (2003), Müller (2013a). Additionally, Ahmed et al. (2005) provide a relative bias and a quadratic risk as robustness measures for trimmed likelihood estimators for the exponential distribution in the case of one stress level. They show in particular that their estimator is asymptotically equivalent with the simple one-sided α-trimmed mean.

Another important robustness measure is the influence function introduced by Hampel (1974). It is well known that the influence function is not only an important robustness measure but also a useful tool for obtaining the asymptotic distribution of the estimator, see in particular Hampel et al. (1986) and Rieder (1994).

However, even the influence function for the one-sided α-trimmed mean is not easy to derive. In Staudte and Sheather (1990), it is given via the influence function of quantiles. But this derivation has some drawbacks. There is a vast literature on influence functions for many other robust estimators, also for robust methods for lifetime distributions as that of Boudt et al. (2011). But the influence function of trimmed likelihood estimators is not treated. The only exception is that asymptotic expansions for trimmed likelihood estimators are derived by Bednarski and Clarke (1993, 2002) for the location and scale case and by Bednarski et al. (2010) for the regression case. From these asymptotic expansions, the influence function can be derived. However they consider in the main only trimmed likelihood estimators for the normal distribution which leads to the least trimmed squares estimators. Moreover, they allow only symmetric distributions for the asymptotic expansions and work with a modified version of the trimmed likelihood estimator.

This modified version of the trimmed likelihood estimator is not easy to calculate but its corresponding functionals, the modified trimmed likelihood functional (MTLF) $\widetilde{\theta}_M$, is given by a rather simple equation. Since the influence function is defined for the corresponding functionals of the estimators, a simple form of the functional is advantageous. However, this MTLF is not the functional $\widetilde{\theta}_O$ corresponding to the original trimmed likelihood estimator, which is called here the original trimmed likelihood functional (OTLF). The definitions of the modified and the original trimmed likelihood functional are given in Section 2 together with the definition of the trimmed likelihood estimator and the influence function.

It is not obvious that the modified functional $\widetilde{\theta}_M$ should coincide with the original functional $\widetilde{\theta}_O$. In Section 3, we compare the both versions for two likelihood functions, namely the likelihood given by the exponential distribution (Section 3.1) and the likelihood given by the (log)normal distribution (Section 3.2). We show that $\widetilde{\theta}_M$ and $\widetilde{\theta}_O$ coincide if only one stress level is used, i.e. in the one-sample case. However, for regression, where several stress levels are used, this is not satisfied in general and we quantify the difference between the defining equations of $\widetilde{\theta}_M$ and $\widetilde{\theta}_O$. This is done by using the implicit function theorem. Although we base the trimmed likelihood functionals on the exponential distribution and the (log)normal distribution, we allow quite general distributions P to which the functionals are applied. In particular, P can be the empirical distribution P_N and then $\widetilde{\theta}_M(P_N)$ and $\widetilde{\theta}_O(P_N)$ coincide under some assumptions. This is important since the original trimmed likelihood functionals $\widetilde{\theta}_O$ are much easier to calculate at empirical distributions while the influence functions of the modified trimmed likelihood functionals $\widetilde{\theta}_M$ show a much simpler form.

The influence functions of the modified trimmed likelihood functionals are derived in Section 4. Here again the implicit function theorem is an important tool since the functionals are given implicitly. The influence function of the exponential regression TLF is treated in Section 4.1 and correspondingly the one for the (log)normal regression TLF in Section 4.2. In both cases, the influence functions are derived at quite general central distributions P. In particular, we do not assume symmetry as Bednarski and Clarke (1993, 2002) and Bednarski et al. (2010) did for the MTLE based on the normal distribution. However, their results appear as special cases.

We believe that the present approach can be used also for trimmed likelihood functionals where the likelihood function is based on other distributions and for censored data.

Section 5 provides an application of the exponential and (log)normal regression TLF for a real data set using a nonlinear link function. Finally, we provide in Section 6 a discussion of the results.

2 Definitions

Let $z_{1N}, \ldots, z_{NN}$ be realizations of independent random variables $Z_{1N}, \ldots, Z_{NN}$, $z_N = (z_{1N}, \ldots, z_{NN})$, and $\widehat{\theta}(z_N)$ an estimate of a parameter $\theta \in \Theta$ of the underlying distribution. Typically it is difficult to measure the influence of an outlier z_* on the estimate $\widehat{\theta}(z_N)$. Therefore Hampel (1974) proposed to consider the influence of an outlier z_* on the asymptotic value of $\widehat{\theta}(Z_N)$. Usually, the estimator $\widehat{\theta}(Z_N)$ converges for $N \to \infty$ in probability or almost surely to a value $\widetilde{\theta}(P)$, where P is the underlying distribution. If $\widetilde{\theta}(P)$ is defined for a class $\mathcal{P}$ of distributions then $\widetilde{\theta} : \mathcal{P} \to \Theta$ is called a statistical functional. Usually, also the contaminated distribution $(1-\epsilon)P + \epsilon\delta_{z_*}$, where δ_{z_*} is the one-point (Dirac) measure on z_*, lies in $\mathcal{P}$ so that $\widetilde{\theta}((1-\epsilon)P + \epsilon\delta_{z_*})$ is defined. Then the influence function at z_* is the directional derivative of $\widetilde{\theta}$ in the direction of $(1-\epsilon)P + \epsilon\delta_{z_*}$ and measures the influence of an outlier z_* on the asymptotic value of the estimator.

Definition 1 (See Hampel et al. 1986)**.** *The influence function $IF(\widehat{\theta}, P, z_*)$ of a statistical functional $\widehat{\theta}$ at a probability measure P and an observation z_* is defined as*

$$IF(\widetilde{\theta}, P, z_*) = \lim_{\epsilon \downarrow 0} \frac{\widetilde{\theta}((1-\epsilon)P + \epsilon\delta_{z_*}) - \widetilde{\theta}(P)}{\epsilon}.$$

To take into account accelerated lifetime experiments, set $z_{1N} = (t_{1N}, s_{1N}), \ldots, z_{NN} = (t_{NN}, s_{NN})$, where t_{nN} is the observed lifetime at stress level s_{nN}. Let $f_{\theta,s}$ be the density of the lifetime distribution at stress s, then l given by $l(\theta, t, s) = \log(f_{\theta,s}(t))$ denotes the loglikelihood function.

Definition 2 (See Müller and Neykov 2003)**.** *A h-trimmed likehood estimator (TLE) $\widehat{\theta}(z_n)$ at z_N is defined as*

$$\widehat{\theta}(z_n) = \arg\max_{\theta \in \Theta} \sum\nolimits_{n=h+1}^{N} l_{(n)}(\theta, z_N),$$

where $l_n(\theta, z_N) = l(\theta, t_{nN}, s_{nN})$ and $l_{(1)}(\theta, z_N) < l_{(2)}(\theta, z_N) < \ldots < l_{(N)}(\theta, z_N)$.

In a h-trimmed likelihood estimator the observations with the h smallest likelihood values are not used.

The functional form of this estimator is given in Definition 3. Thereby, we use $\alpha = \lim_{N\to\infty} \frac{h_N}{N}$ with $h_N = h$. Moreover, to model stress levels s given by an experimenter, the distribution is given by $P = P^{T|S} \otimes P^S$, where T is the random variable for the lifetime and S the random variable for the stress. If fixed designs for the stress variables are used, then P^S is the asymptotic distribution of the stress variables and can be interpreted as a generalized design, see e.g. Müller (1997).

Definition 3. *The original α-trimmed likelihood functional (OTLF) $\widetilde{\theta}_O(P)$ at $P = P^{T|S} \otimes P^S$ is given by*

$$\widetilde{\theta}_O(P) = \arg\max_{\theta\in\Theta} \int\int 1\!\!\mathrm{I}\{l(\theta,t,s) \geq b(\theta)\}\, l(\theta,t,s)\, P^{T|S=s}(dt)\, P^S(ds)$$

where $b(\theta)$ satisfies

$$b(\theta) = \arg\max \left\{ b\,; \int\int 1\!\!\mathrm{I}\{l(\theta,t,s) \geq b\}\, P^{T|S=s}(dt)\, P^S(ds) \geq 1-\alpha \right\} \tag{1}$$

and $1\!\!\mathrm{I}\{x \in A\} = 1\!\!\mathrm{I}_A(x)$ denotes the indicator function.

The functional form of the modified version used by Bednarski and Clarke (1993, 2002) and Bednarski et al. (2010) is given in Definition 4.

Definition 4. *The modified α-trimmed likelihood functional (MTLF) $\widetilde{\theta}_M = \widetilde{\theta}_M(P)$ at $P = P^{T|S} \otimes P^S$ is a solution of*

$$0 = \int\int 1\!\!\mathrm{I}\{l(\widetilde{\theta}_M,t,s) \geq b(\widetilde{\theta}_M)\}\, \dot{l}(\widetilde{\theta}_M,t,s)\, P^{T|S=s}(dt)\, P^S(ds)$$

where $b(\theta)$ is defined by (1) and $\dot{l}(\theta,t,s) = \frac{\partial}{\partial\theta} l(\theta,t,s)$.

3 Comparison of MTLF and OTLF

To check whether the MTLF given by Definition 4 and the OTLF given by Definition 3 coincide, we have to check the equality of

$$\int\int 1\!\!\mathrm{I}\{l(\theta,t,s) \geq b(\theta)\}\, \dot{l}(\theta,t,s)\, P^{T|S=s}(dt)\, P^S(ds) \tag{2}$$

and

$$\frac{\partial}{\partial\theta}\int\int 1\!\!\mathrm{I}\{l(\theta,t,s) \geq b(\theta)\}\, l(\theta,t,s)\, P^{T|S=s}(dt)\, P^S(ds). \tag{3}$$

We will consider here only the case where $b(\theta)$ given by (1) satisfies

$$1-\alpha = \int 1\!\!\mathrm{I}\{l(\theta,t,s) \geq b(\theta)\}\, P^{T|S=s}(dt)\, P^S(ds) \tag{4}$$

for all θ in a neighborhood of $\widetilde{\theta}_M(P)$ and $\widetilde{\theta}_O(P)$, respectively. This is in particular the case for continuous distributions $P^{T|S=s}$ but not restricted to them. Hence $b(\theta)$ is implicitly defined by (4).

3.1 Exponential regression TLFs

If the lifetimes at different stress levels have exponential distributions then the loglikeli-function is given by

$$l(\theta, t, s) = \log(\lambda_s(\theta)) - \lambda_s(\theta) t \tag{5}$$

where $\lambda_s(\theta)$ is the link function between the stress levels and the parameter of the exponential distribution. Typical link functions in accelerated lifetime experiments are $\lambda_s(\theta) = \theta s$ with $\theta \in (0, \infty)$, $\lambda_s(\theta) = \exp(\vartheta_0 + \vartheta_1 s)$ with $\theta = (\vartheta_0, \vartheta_1)^\top \in \mathbb{R} \times (0, \infty)$, or $\lambda_s(\theta) = \exp(-\vartheta_0 + \vartheta_1 s - \vartheta_2 s^{-\vartheta_3})$ with $\theta = (\vartheta_0, \vartheta_1, \vartheta_2, \vartheta_3)^\top \in [0, \infty)^4$, see e.g. Müller (2013b) or the real data example in Section 5.

Definition 5. *The exponential regression OTLF and MTLF are the OTLF and the MTLF where $l(\theta, t, s)$ is given by (5).*

Then we have

$$l(\theta, t, s) \geq b(\theta) \Longleftrightarrow t \leq \frac{\log(\lambda_s(\theta)) - b(\theta)}{\lambda_s(\theta)}.$$

Set

$$\eta_s(\theta, b) := \frac{\log(\lambda_s(\theta)) - b}{\lambda_s(\theta)}$$

and use $\dot{l}(\theta, t, s) = \left(\frac{1}{\lambda_s(\theta)} - t\right) \dot{\lambda}_s(\theta)$ with $\dot{\lambda}_s(\theta) := \frac{\partial}{\partial \theta} \lambda_s(\theta)$. Then we obtain for expression (2) with partial integration of $\int t \, P^{T|S=s}(dt)$

$$\begin{aligned}
& \int \int \mathbb{1}\{l(\theta, t, s) \geq b(\theta)\} \, \dot{l}(\theta, t, s) \, P^{T|S=s}(dt) \, P^S(ds) \qquad (6) \\
= \; & \int \int_0^{\eta_s(\theta, b(\theta))} \left(\frac{1}{\lambda_s(\theta)} - t\right) \dot{\lambda}_s(\theta) \, P^{T|S=s}(dt) \, P^S(ds) \\
= \; & \int \left[F_s(\eta_s(\theta, b(\theta))) \left(\frac{1}{\lambda_s(\theta)} - \eta_s(\theta, b(\theta))\right) + \mathcal{F}_s(\eta_s(\theta, b(\theta)))\right] \dot{\lambda}_s(\theta) \, P^S(ds) \\
=: \; & U_P^M(\theta),
\end{aligned}$$

where F_s is the cumulative distribution function of an arbitrary lifetime distribution $P^{T|S=s}$ on $[0, \infty)$ and $\mathcal{F}_s$ is the antiderivative of F_s, i.e. $\frac{\partial}{\partial t} \mathcal{F}_s(t) = F_s(t)$. In particular, it is not necessary to assume an exponential distribution for $P^{T|S=s}$. Hence we arrive at the following lemma.

Lemma 1. *The exponential regression MTLE $\widetilde{\theta}_M$ at P is given as a solution of* $0 = U_P^M(\theta)$.

Similarly, the integral in (3) is given by

$$\int\int 1\!\!\mathrm{I}\{l(\theta,t,s)\geq b(\theta)\}\, l(\theta,t,s)\, P^{T|S=s}(dt)\, P^S(ds) \tag{7}$$
$$= \int\int_0^{\eta_s(\theta,b(\theta))} (\log(\lambda_s(\theta))-\lambda_s(\theta)t)\; P^{T|S=s}(dt)\, P^S(ds)$$
$$= \int \Big[F_s(\eta_s(\theta,b(\theta)))\,[\log(\lambda_s(\theta))-\lambda_s(\theta)\eta_s(\theta,b(\theta))]+\lambda_s(\theta)\mathcal{F}_s(\eta_s(\theta,b(\theta)))\Big]\, P^S(ds).$$

To calculate the derivative of (7), we need the derivative of $b(\theta)$ which is implicitly given by $W_1(\theta,b(\theta))=0$, where

$$W_1(\theta,b):=\int\int_0^{\eta_s(\theta,b)} P^{T|S=s}(dt)\, P^S(ds)-(1-\alpha)=\int F_s(\eta_s(\theta,b))\, P^S(ds)-(1-\alpha).$$

We assume here that

$$F_s \text{ is differentiable in a neighborhood of } \eta_s(\theta,b(\theta)) \text{ with derivative } f_s \text{ for all } s \text{ in the support of } P^S. \tag{8}$$

Since

$$\frac{\partial}{\partial\theta}\eta_s(\theta,b)=\frac{1+b-\log(\lambda_s(\theta))}{\lambda_s(\theta)^2}\,\dot\lambda_s(\theta)$$

and

$$\frac{\partial}{\partial b}\eta_s(\theta,b)=-\frac{1}{\lambda_s(\theta)}$$

we have

$$\frac{\partial}{\partial\theta}W_1(\theta,b)=\int f_s(\eta_s(\theta,b))\,\frac{1+b-\log(\lambda_s(\theta))}{\lambda_s(\theta)^2}\,\dot\lambda_s(\theta)\, P^S(ds)$$

and

$$\frac{\partial}{\partial b}W_1(\theta,b)=-\int f_s(\eta_s(\theta,b))\,\frac{1}{\lambda_s(\theta)}\, P^S(ds).$$

If $\frac{\partial}{\partial b}W_1(\theta,b)|_{b=b(\theta)}\neq 0$, then the implicit function theorem provides

$$\dot b(\theta):=\frac{\partial}{\partial\theta}b(\theta)=\frac{\int f_s(\eta_s(\theta,b(\theta)))\;\; [1+b(\theta)-\log(\lambda_s(\theta))]\;\; \lambda_s(\theta)^{-2}\;\; \dot\lambda_s(\theta)\, P^S(ds)}{\int f_s(\eta_s(\theta,b(\theta)))\;\; \lambda_s(\theta)^{-1}\;\; P^S(ds)}.$$

If $\frac{\partial}{\partial b}W_1(\theta,b)|_{b=b(\theta)}=0$, then $f_s(\eta_s(\theta,b(\theta)))=0$ for all s of the support of P^S, so that $\frac{\partial}{\partial\theta}W_1(\theta,b)|_{b=b(\theta)}=0$ holds as well. Hence we can use $\dot b(\theta)=0$ in this case.

Setting

$$\dot\eta_s(\theta):=\frac{\partial}{\partial\theta}\eta_s(\theta,b(\theta))=\frac{1+b(\theta)-\log(\lambda_s(\theta))}{\lambda_s(\theta)^2}\,\dot\lambda_s(\theta)-\frac{\dot b(\theta)}{\lambda_s(\theta)}$$

and using $b(\theta) = \log(\lambda_s(\theta)) - \lambda_s(\theta)\,\eta_s(\theta, b(\theta))$, the derivative of (7) is

$$
\begin{aligned}
U_P^O(\theta) := \int \Big[& f_s(\eta_s(\theta, b(\theta)))\, \dot{\eta}_s(\theta)\, b(\theta) \\
& + F_s(\eta_s(\theta, b(\theta))) \left[\frac{1}{\lambda_s(\theta)} - \eta_s(\theta, b(\theta)) \right] \dot{\lambda}_s(\theta) + \dot{\lambda}_s(\theta)\, \mathcal{F}_s(\eta_s(\theta, b(\theta))) \Big] P^S(ds).
\end{aligned}
$$

Hence we obtain the following lemma.

Lemma 2. *Under the assumption (8), the exponential regression OTLE $\widetilde{\theta}_O$ at P is given as a solution of $0 = U_P^O(\theta)$.*

Corollary 1. *The difference between (2) and (3) for exponential regression trimmed likelihood functionals is given by*

$$
\int f_s(\eta_s(\theta, b(\theta)))\, \dot{\eta}_s(\theta)\, b(\theta)\, P^S(ds). \tag{9}
$$

The difference (9) is zero if $\dot{\eta}_s(\theta) = 0$ or $f_s(\eta_s(\theta, b(\theta))) = 0$ holds for all stress levels s in the support of P^S. In general, this will not be the case. However, if only one stress level s_0 is used, i.e. we have the one-sample case, and $f_{s_0}(\eta_s(\theta, b(\theta))) \neq 0$ is satisfied, then we obtain

$$
\dot{b}(\theta) = \frac{1 + b(\theta) - \log(\lambda_{s_0}(\theta))}{\lambda_{s_0}(\theta)}\, \dot{\lambda}_{s_0}(\theta)
$$

and thus $\dot{\eta}_{s_0}(\theta) = 0$ which is also clear from the definition of $\eta_s(\theta)$.

For empirical distributions satisfying (8), the difference (9) is zero since then the derivative of F_s is zero. Hence if we prefer to work with the modified trimmed likelihood functional, then frequently the h-trimmed likelihood estimator will provide the correct corresponding estimator in the finite sample case.

Example 1 (One-sample case).

If P^S is given by a one-point measure at s_0, then the TLF $\widetilde{\theta} := \widetilde{\theta}_M(P) = \widetilde{\theta}_O(P)$ can be given more explicitly. Setting $\eta := \eta_{s_0}(\widetilde{\theta}, b(\widetilde{\theta}))$ and using $\mathcal{F}_s(\eta) = \eta F_s(\eta) - \int_0^\eta t dP^{T|S=s}(dt)$, we have $F_{s_0}(\eta) = 1 - \alpha$ and the TLF at P satisfies

$$
\begin{aligned}
0 &= \left[F_{s_0}(\eta) \left(\frac{1}{\lambda_{s_0}(\widetilde{\theta})} - \eta \right) + \mathcal{F}_{s_0}(\eta) \right] \dot{\lambda}_{s_0}(\widetilde{\theta}) \\
&= \left[(1-\alpha) \frac{1}{\lambda_{s_0}(\widetilde{\theta})} - \int_0^\eta t\, dP^{T|S=s_0}(dt) \right] \dot{\lambda}_{s_0}(\widetilde{\theta}) \\
\iff \quad \frac{1}{\lambda_{s_0}(\widetilde{\theta})} &= \frac{1}{1-\alpha} \int_0^\eta t\, dP^{T|S=s_0}(dt).
\end{aligned}
$$

Since $\frac{1}{1-\alpha} \int_0^\eta t\, dP^{T|S=s_0}(dt)$ is the functional of the one-sided trimmed mean, we see that the TLF is given by the one-sided trimmed mean in the one-sample case. This corresponds to a result of Ahmed et al. (2005) who showed that the TLE for the exponential distribution behaves asymptotically like a one-sided trimmed mean.

3.2 TLFs for (log)normal distribution

Another often used lifetime distribution is the lognormal distribution, where the logarithm of the lifetime T has a normal distribution. For simplicity, we use here directly the normal distribution, i.e. we work with $Y = \log(T)$. A typical link between the mean of the normal distribution and the stress level, is a linear link given by $m_s(\theta) = x(s)^\top\theta$ with e.g. $x(s) = \frac{1}{s}$ with $\theta \in (0, \infty)$ or $x(s) = (1, -s)^\top$ with $\theta = (\vartheta_0, \vartheta_1)^\top \in [0, \infty)^2$. But also nonlinear links like $m_s(\theta) = \vartheta_0 + \vartheta_1(\frac{1}{s})^{\vartheta_2}$ with $\theta = (\vartheta_0, \vartheta_1, \vartheta_2)^\top \in [0, \infty)^3$ or $m_s(\theta) = \vartheta_0 - \vartheta_1 s + \vartheta_2 s^{-\vartheta_3}$ with $\theta = (\vartheta_0, \vartheta_1, \vartheta_2, \vartheta_3)^\top \in [0, \infty)^4$ are used in accelerated lifetime experiments, see e.g. the real data example in Section 5. The loglikelihood function is then given up to a constant by

$$l(\theta, y, s) = -(y - m_s(\theta))^2. \tag{10}$$

Definition 6. *The (log)normal regression OTLF and MTLF are the OTLF and the MTLF where $l(\theta, t, s)$ is given by (10).*

Then we have

$$\begin{aligned} &l(\theta, y, s) \geq b(\theta) =: -a(\theta)^2 \\ &\quad \Leftrightarrow \quad |y - m_s(\theta)| \leq a(\theta) \Leftrightarrow m_s(\theta) - a(\theta) \leq y \leq m_s(\theta) + a(\theta). \end{aligned}$$

With partial integration of $\int y\, P^{Y|S=s}(dy)$ and $\dot{l}(\theta, y, s) = 2(y - m_s(\theta))\dot{m}_s(\theta)$ where $\dot{m}_s(\theta) = \frac{\partial}{\partial\theta} m_s(\theta)$, we obtain for expression (2)

$$\begin{aligned} &\int\int \mathbb{1}\{l(\theta, y, s) \geq b(\theta)\}\, \dot{l}(\theta, y, s)\, P^{Y|S=s}(dt)\, P^S(ds) \qquad (11) \\ &= \quad 2\int\int_{m_s(\theta)-a(\theta)}^{m_s(\theta)+a(\theta)} (y - m_s(\theta))\, \dot{m}_s(\theta)\, P^{Y|S=s}(dy)\, P^S(ds) \\ &= \quad 2\int \Big[a(\theta)\ \left[F_s(m_s(\theta) + a(\theta)) + F_s(m_s(\theta) - a(\theta))\right] \\ &\qquad - \ \mathcal{F}_s(m_s(\theta) + a(\theta)) + \mathcal{F}_s(m_s(\theta) - a(\theta))\Big]\ \dot{m}_s(\theta)\, P^S(ds) \\ &=: \quad V_P^M(\theta), \end{aligned}$$

where F_s is again the cumulative distribution function of $P^{Y|S=s}$ and $\mathcal{F}_s$ is the antiderivative of F_s. Thereby F_s can be any distribution function on $\mathbb{R}$. Hence the following lemma is shown.

Lemma 3. *The (log)normal regression MTLE $\widetilde{\theta}_M$ at P is given as a solution of $0 = V_P^M(\theta)$.*

For the integral in (3), we get using partial integration of $\int y\, P^{Y|S=s}(dy)$ and $\int y^2\, P^{Y|S=s}(dy)$

$$\begin{aligned}
&\int\int 1\!\mathrm{I}\{l(\theta,t,s) \geq b(\theta)\}\; l(\theta,t,s)\, P^{Y|S=s}(dy)\, P^S(ds) \qquad (12)\\
&= -\int\int_{m_s(\theta)-a(\theta)}^{m_s(\theta)+a(\theta)} (y - m_s(\theta))^2\, P^{Y|S=s}(dy)\, P^S(ds)\\
&= -\int \Big\{a(\theta)^2\, \left[F_s(m_s(\theta)+a(\theta)) - F_s(m_s(\theta)-a(\theta))\right]\\
&\quad - 2\,\left[\mathcal{H}_s(m_s(\theta)+a(\theta)) - \mathcal{H}_s(m_s(\theta)-a(\theta))\right]\\
&\quad + 2\, m_s(\theta)\, \left[\mathcal{F}_s(m_s(\theta)+a(\theta)) - \mathcal{F}_s(m_s(\theta)-a(\theta))\right]\Big\}\; P^S(ds),
\end{aligned}$$

where $\mathcal{H}_s$ is the antiderivative of H_s given by $H_s(y) = y\, F_s(y)$. To obtain the derivative of (12), we have to calculate the derivative of $a(\theta)$ which is implicitly given by $W_1(\theta, a(\theta)) = 0$, where

$$\begin{aligned}
W_1(\theta,a) &:= \int\int_{m_s(\theta)-a}^{m_s(\theta)+a} P^{Y|S=s}(dy)\, P^S(ds) - (1-\alpha)\\
&= \int \left[F_s(m_s(\theta)+a) - F_s(m_s(\theta)-a)\right]\; P^S(ds) - (1-\alpha).
\end{aligned}$$

We assume here that

F_s is differentiable in a neighborhood of $m_s(\theta)+a(\theta)$ and $m_s(\theta)-a(\theta)$ (13)
with derivative f_s for all s in the support of P^S.

Since

$$\frac{\partial}{\partial a} W_1(\theta,a) = \int \left[f_s(m_s(\theta)+a) + f_s(m_s(\theta)-a)\right]\; P^S(ds)$$

and

$$\frac{\partial}{\partial \theta} W_1(\theta,a) = \int \left[f_s(m_s(\theta)+a) - f_s(m_s(\theta)-a)\right]\; \dot m_s(\theta)\, P^S(ds)$$

the implicit function theorem provides

$$\begin{aligned}
\dot a(\theta) &:= \frac{\partial}{\partial\theta}\, a(\theta)\\
&= -\left(\int \left[f_s(m_s(\theta)+a(\theta)) + f_s(m_s(\theta)-a(\theta))\right]\; P^S(ds)\right)^{-1}\\
&\qquad \cdot \int \left[f_s(m_s(\theta)+a(\theta)) - f_s(m_s(\theta)-a(\theta))\right]\; \dot m_s(\theta)\, P^S(ds)
\end{aligned}$$

if $\frac{\partial}{\partial a} W_1(\theta,a)|_{a=a(\theta)} \neq 0$. If $\frac{\partial}{\partial a} W_1(\theta,a)|_{a=a(\theta)} = 0$ then also $\frac{\partial}{\partial \theta} W_1(\theta,a)|_{a=a(\theta)} = 0$ so that we can set $\dot a(\theta) = 0$ in this case. Hence the derivative of (12) is

$$\begin{aligned}
V_P^O(\theta) \;:=\; & -\int \Big\{a(\theta)^2\; \left[f_s(m_s(\theta)+a(\theta)) - f_s(m_s(\theta)-a(\theta))\right]\, \dot m_s(\theta)\\
& + a(\theta)^2\; \left[f_s(m_s(\theta)+a(\theta)) + f_s(m_s(\theta)-a(\theta))\right]\, \dot a(\theta)\\
& - 2\, a(\theta)\, \left[F_s(m_s(\theta)+a(\theta)) + F_s(m_s(\theta)-a(\theta))\right]\; \dot m_s(\theta)\\
& + 2\; \left[\mathcal{F}_s(m_s(\theta)+a(\theta)) - \mathcal{F}_s(m_s(\theta)-a(\theta))\right]\; \dot m_s(\theta)\Big\}\; P^S(ds).
\end{aligned}$$

Lemma 4. *Under the assumption (13), the (log)normal regression OTLE $\widetilde{\theta}_O$ at P is given as a solution of* $0 = V_P^O(\theta)$.

Corollary 2. *The difference between (2) and (3) for (log)normal regression trimmed likelihood functionals is given by*

$$\int \Big\{ a(\theta)^2 \; [f_s(m_s(\theta)+a(\theta)) - f_s(m_s(\theta)-a(\theta))] \; \dot{m}_s(\theta) \\ + \;\; a(\theta)^2 \; [f_s(m_s(\theta)+a(\theta)) + f_s(m_s(\theta)-a(\theta))] \; \dot{a}(\theta) \Big\} \; P^S(ds). \tag{14}$$

The difference (14) is zero if only one stress level s_0 is used or if f_s is symmetric around $m_s(\theta)$ for P^S-almost all s. Bednarski et al. (2010) considered symmetric distributions for the central distribution. However, any neighborhood around a central symmetric distribution contains also asymmetric distributions. But again we can argue that the difference (14) is zero for empirical distributions satisfying (13) since the derivative of F_s is then zero. Hence frequently the h-trimmed likelihood estimator will provide the correct corresponding estimator of the modified trimmed likelihood functional in the finite sample case.

4 The influence function of TLFs

Although we have seen that the modified trimmed likelihood functional of Definition 4 does not coincide in general with the original trimmed likelihood functional of Definition 3, we will derive the influence function for the MTLF since its form is simpler. That the treatment of the MTLF instead of the OTLF makes sense is due to the fact that the MTLF applied on the empirical distribution coincides with the h-trimmed likelihood estimator of Definition 2 in many cases. This is important since only the h-trimmed likelihood estimator can be calculated efficiently. I.e. for small sample sizes, the TLEs can be obtained by calculating the maximum likelihood estimator for all subsamples with $N-h$ elements. For larger samples sizes, special methods for the TLE have been developed, see e.g. Neykov and Müller (2003) or Rousseeuw and Driessen (2006).

Since the stress levels are given by the experimenter, only contamination with respect to $P^{T|S}$ is considered. Set $P_\epsilon = P_\epsilon^{T|S} \otimes P^S$ with

$$P_\epsilon^{T|S=s} = (1-\epsilon)P^{T|S=s} + \epsilon Q^{T|S=s}$$

and corresponding distribution function

$$F_{s,\epsilon} = (1-\epsilon)F_s + \epsilon G_s.$$

We will derive

$$\lim_{\epsilon \downarrow 0} \frac{\widetilde{\theta}_M(P_\epsilon) - \widetilde{\theta}_M(P)}{\epsilon} = \frac{\partial}{\partial \epsilon} \widetilde{\theta}_M(P_\epsilon)\Big|_{\epsilon=0},$$

which provides for the special case of $Q^{T|S=s_*} = \delta_{t_*}$ and $Q^{T|S=s} = P^{T|S=s}$ for $s \neq s_*$ the influence function at P and $z_* = (t_*, s_*)$ of Definition 1. Thereby, $\widetilde{\theta}_M(P_\epsilon)$ is implicitly given by

$$W_2(\epsilon, \widetilde{\theta}_M(P_\epsilon)) = 0$$

where

$$W_2(\epsilon, \theta) = \int \int 1\!\!1\{l(\theta, t, s) \geq b(\epsilon, \theta)\}\ \dot{l}(\theta, t, s)\, P_\epsilon^{T|S=s}(dt)\, P^S(ds) \tag{15}$$

and $b(\epsilon, \theta)$ is implicitly given by

$$W_1(\epsilon, \theta, b(\epsilon, \theta)) = 0$$

with

$$W_1(\epsilon, \theta, b) = \int \int 1\!\!1\{l(\theta, t, s) \geq b\}\ P_\epsilon^{T|S=s}(dt)\, P^S(ds) - (1 - \alpha). \tag{16}$$

4.1 The influence function of the exponential regression MTLF

Here, (15) becomes according to (6)

$$\begin{aligned} W_2(\epsilon, \theta) = \int \Bigg[& F_{s,\epsilon}(\eta_s(\theta, b(\epsilon, \theta))) \left(\frac{1}{\lambda_s(\theta)} - \eta_s(\theta, b(\epsilon, \theta)) \right) \\ & + \mathcal{F}_{s,\epsilon}(\eta_s(\theta, b(\epsilon, \theta)))\Bigg]\ \dot{\lambda}_s(\theta)\, P^S(ds), \end{aligned}$$

and W_1 of (16) is given by, see Section 3.1,

$$W_1(\epsilon, \theta, b) = \int F_{s,\epsilon}(\eta_s(\theta, b))\, P^S(ds) - (1 - \alpha).$$

As in Section 3.1, we have

$$\frac{\partial}{\partial \theta} W_1(\epsilon, \theta, b) = \int f_{s,\epsilon}(\eta_s(\theta, b))\, \frac{1 + b - \log(\lambda_s(\theta))}{\lambda_s(\theta)^2}\, \dot{\lambda}_s(\theta)\, P^S(ds)$$

and

$$\frac{\partial}{\partial b} W_1(\epsilon, \theta, b) = - \int f_{s,\epsilon}(\eta_s(\theta, b))\, \frac{1}{\lambda_s(\theta)}\, P^S(ds).$$

Additionally, we use here

$$\left. \frac{\partial}{\partial \epsilon} W_1(\epsilon, \theta, b) \right|_{\epsilon = 0} = \int (G_s - F_s)(\eta_s(\theta, b))\, P^S(ds).$$

Setting $\theta_0 = \widetilde{\theta}_M(P_0) = \widetilde{\theta}_M(P)$, $b_0 = b(0, \theta_0)$, we make the following assumption:

$$\begin{gathered} F_s \text{ and } G_s \text{ are differentiable in a neighborhood of } \eta_s(\theta_0, b_0) \\ \text{for all } s \text{ in the support of } P^S. \end{gathered} \tag{17}$$

Clearly this is satisfied for F_s since the central distribution P_s should be a continuous distribution. However, G_s could be also the distribution function of a one-point measure so that the differentiability is not everywhere satisfied. However, we consider here only the cases where the differentiability is satisfied. Then the implicit function theorem provides

$$\begin{aligned}\dot{b}_\theta(0) &:= \frac{\partial}{\partial\theta} b(\epsilon,\theta)\Big|_{(\epsilon,\theta)=(0,\theta_0)} \\ &= \frac{\int f_s(\eta_s(\theta_0,b_0)) \;\; [1+b_0-\log(\lambda_s(\theta_0))] \;\; \lambda_s(\theta_0)^{-2} \;\; \dot{\lambda}_s(\theta_0)\, P^S(ds)}{\int f_s(\eta_s(\theta_0,b_0)) \;\; \lambda_s(\theta_0)^{-1} \;\; P^S(ds)}\end{aligned}$$

and

$$\dot{b}_\epsilon(0) := \frac{\partial}{\partial\epsilon} b(\epsilon,\theta)\Big|_{(\epsilon,\theta)=(0,\theta_0)} = \frac{\int (G_s-F_s)(\eta_s(\theta_0,b_0))\, P^S(ds)}{\int f_s(\eta_s(\theta_0,b_0)) \;\; \lambda_s(\theta_0)^{-1} \;\; P^S(ds)}.$$

Using this notation, we obtain for the derivatives of $\eta_s(\theta,b(\epsilon,\theta)) = \frac{\log(\lambda_s(\theta))-b(\epsilon,\theta)}{\lambda_s(\theta)}$

$$\dot{\eta}_{s,\theta}(0) := \frac{\partial}{\partial\theta}\eta_s(\theta,b(\epsilon,\theta))\Big|_{(\epsilon,\theta)=(0,\theta_0)} = \frac{1+b_0-\log(\lambda_s(\theta_0))}{\lambda_s(\theta_0)^2}\,\dot{\lambda}_s(\theta_0) - \frac{\dot{b}_\theta(0)}{\lambda_s(\theta_0)}$$

and

$$\dot{\eta}_{s,\epsilon}(0) := \frac{\partial}{\partial\epsilon}\eta_s(\theta,b(\epsilon,\theta))\Big|_{(\epsilon,\theta)=(0,\theta_0)} = -\frac{\dot{b}_\epsilon(0)}{\lambda_s(\theta_0)}.$$

Let here $\mathcal{G}_s$ be the antiderivative of G_s so that $\mathcal{F}_{s,\epsilon} = (1-\epsilon)\mathcal{F}_s + \epsilon\mathcal{G}_s = \mathcal{F}_s + \epsilon(\mathcal{G}_s - \mathcal{F}_s)$ and set $\ddot{\lambda}_s(\theta) = \frac{\partial}{\partial\theta}\dot{\lambda}_s(\theta)^\top$. Now, we can calculate the derivatives of $W_2(\epsilon,\theta)^\top$:

$$\begin{aligned}&\frac{\partial}{\partial\theta} W_2(\epsilon,\theta)^\top\Big|_{(\epsilon,\theta)=(0,\theta_0)} \\ &= \int \Bigg[f_s(\eta_s(\theta_0,b_0))\, \dot{\eta}_{s,\theta}(0) \;\left(\frac{1}{\lambda_s(\theta_0)} - \eta_s(\theta_0,b_0)\right) \\ &\quad - \; F_s(\eta_s(\theta_0,b_0))\, \frac{\dot{\lambda}_s(\theta_0)}{\lambda_s(\theta_0)^2}\Bigg]\, \dot{\lambda}_s(\theta_0)^\top\, P^S(ds) \\ &\quad + \int \left[F_s(\eta_s(\theta_0,b_0)) \left(\frac{1}{\lambda_s(\theta_0)} - \eta_s(\theta_0,b_0)\right) + \mathcal{F}_s(\eta_s(\theta_0,b_0))\right] \; \ddot{\lambda}_s(\theta_0)\, P^S(ds) \\ &=: \; A(P)\end{aligned}$$

and

$$\begin{aligned}&\frac{\partial}{\partial\epsilon} W_2(\epsilon,\theta)^\top\Big|_{(\epsilon,\theta)=(0,\theta_0)} \\ &= \; -B(P) \;\; \int (G_s-F_s)(\eta_s(\theta_0,b_0))\, P^S(ds) \\ &\quad + \int \left[(G_s-F_s)(\eta_s(\theta_0,b_0))\, a_s(P) + \int_0^{\eta_s(\theta_0,b_0)} (G_s-F_s)(t)\, dt \right] \; \dot{\lambda}_s(\theta_0)^\top\, P^S(ds),\end{aligned}$$

where

$$a_s(P) := \left(\frac{1}{\lambda_s(\theta_0)} - \eta_s(\theta_0, b_0)\right)$$

and

$$B(P) := \int \frac{f_s(\eta_s(\theta_0, b_0))\ \lambda_s(\theta_0)^{-1}}{\int f_{\tilde{s}}(\eta_{\tilde{s}}(\theta_0, b_0))\ \ \lambda_{\tilde{s}}(\theta_0)^{-1}\ \ P^S(d\tilde{s})}\ a_s(P)\ \dot{\lambda}_s(\theta_0)^\top\ P^S(ds).$$

Hence with the implicit function theorem, we obtain the following theorem.

Theorem 1. *Under the assumption (17), the exponential regression MTLF $\widetilde{\theta}_M$ of Definition 5 satisfies*

$$\lim_{\epsilon \downarrow 0} \frac{\widetilde{\theta}_M(P_\epsilon) - \widetilde{\theta}_M(P)}{\epsilon} = -A(P)^{-1} \left[-B(P) \ \ \int (G_s - F_s)(\eta_s(\theta_0, b_0))\ P^S(ds) \right.$$
$$\left. + \int \left[(G_s - F_s)(\eta_s(\theta_0, b_0))\ a_s(P) + \int_0^{\eta_s(\theta_0, b_0)} (G_s - F_s)(t)\ dt \right] \dot{\lambda}_s(\theta_0)^\top\ P^S(ds) \right].$$

Using $G_{s_*}(t) = 1\!\!\mathrm{I}_{[t_*, \infty)}(t)$ if $Q^{T|S=s_*} = \delta_{t_*}$ and $G_s(t) = F_s(t)$ for $s \neq s_*$, assumption (17) is satisfied if $t_* \neq \eta_{s_*}(\theta_0, b_0)$. Hence we get at once the following influence function.

Corollary 3. *If P^S has finite support and $t_* \neq \eta_* := \eta_{s_*}(\theta_0, b_0)$, then the influence function of the exponential regression MTLF $\widetilde{\theta}_M$ at $z_* = (t_*, s_*) \in [0, \infty)^2$ is given by*

$$\begin{aligned} & IF(\widetilde{\theta}_M, P, z_*) \\ = \ & -A(P)^{-1} \left[\left\{ \left(\eta_* - t_* \right) 1\!\!\mathrm{I}_{[0, \eta_*]}(t_*) - \int_0^{\eta_*} F_{s_*}(t)\ dt \right\} \dot{\lambda}_{s_*}(\theta_0)^\top\ P^S(\{s_*\}) \right. \\ & \left. + \left(1\!\!\mathrm{I}_{[0, \eta_*]}(t_*) - F_{s_*}(\eta_*) \right) \left(a_{s_*}(P)\ \dot{\lambda}_{s_*}(\theta_0)^\top - B(P) \right)\ P^S(\{s_*\}) \right]. \end{aligned}$$

Obviously, this influence function is a bounded function in t_* so that outliers t_* at s_* have a bounded influence on the trimmed likelihood estimator.

Example 2 (One-sample case)**.**

If P^S is given by a one-point measure at s_0, then the results of Theorem 1 and Corollary 3 concern also the original trimmed likelihood function as shown in Section 3.1. In this case, we have as in Section 3.1 $\dot{\eta}_{s,\theta}(0) = 0$. Moreover, θ should be one-dimensional and reasonable choices for $\lambda_s(\theta)$ are $\lambda_s(\theta) = \theta s$ or $\lambda_s(\theta) = \theta$. Then it holds $\ddot{\lambda}_s(\theta) = 0$ so that $A(P)$ becomes

$$A(P) = -F_{s_0}(\eta_{s_0}(\theta_0, b_0))\, \frac{\dot{\lambda}_{s_0}(\theta_0)^2}{\lambda_{s_0}(\theta_0)^2} = -(1 - \alpha)\ \theta_0^{-2}.$$

With $B(P) = a_{s_0}(P)\ \dot{\lambda}_{s_0}(\theta_0)$ and setting $\eta_* = \eta_{s_0}(\theta_0, b_0)$, $\widetilde{\theta} = \theta_0 = \widetilde{\theta}_M(P) = \widetilde{\theta}_O(P)$ we obtain

$$\lim_{\epsilon \downarrow 0} \frac{\widetilde{\theta}_M(P_\epsilon) - \widetilde{\theta}_M(P)}{\epsilon} = \frac{\widetilde{\theta}^2}{1 - \alpha} \int_0^{\eta_*} (G_{s_0} - F_{s_0})(t)\ dt\ \dot{\lambda}_{s_0}(\widetilde{\theta}).$$

Using partial integration of $\int_0^{\eta_*} F_{s_0}(t)\,dt$, the influence function is given by

$$IF(\widetilde{\theta}_M, P, z_*) = IF(\widetilde{\theta}_O, P, z_*) \tag{18}$$
$$= \begin{cases} \widetilde{\theta}^2 \left(\frac{\eta_*\,\alpha - t_*}{1-\alpha} + \frac{1}{\lambda_{s_0}(\widetilde{\theta})} \right) \dot{\lambda}_{s_0}(\widetilde{\theta}) & \text{if } t_* < \eta_*, \\ \widetilde{\theta}^2 \left(-\eta_* + \frac{1}{\lambda_{s_0}(\widetilde{\theta})} \right) \dot{\lambda}_{s_0}(\widetilde{\theta}) & \text{if } t_* > \eta_*, \end{cases}$$

since according to Example 1 $F_{s_0}(\eta_*) = 1 - \alpha$ and

$$\frac{1}{\lambda_{s_0}(\widetilde{\theta})} = \frac{1}{1-\alpha} \int_0^{\eta_*} t\, dP^{T|S=s_0}(dt). \tag{19}$$

Note that the influence function of the one-sided trimmed mean $\widetilde{\mu} = \widetilde{\mu}(P) = \frac{1}{1-\alpha}\int_0^{\eta_*} t\, dP(dt)$ is given by, see e.g. Staudte and Sheather (1990) pp.55,

$$IF(\widetilde{\mu}, P, t_*) = \begin{cases} \frac{t_* - \eta_*\,\alpha}{1-\alpha} - \widetilde{\mu} & \text{if } t_* < \eta_*, \\ \eta_* - \widetilde{\mu} & \text{if } t_* > \eta_*. \end{cases}$$

This coincides with (18) using (19) and $\lambda_s(\mu) = \frac{1}{\mu}$. Note that in contrast to the result of Staudte and Sheather (1990), it holds $IF(\widetilde{\mu}, P, t_*) = \frac{t_*}{1-\alpha} - \widetilde{\mu}$ for $t_* = \eta_*$ so that $IF(\widetilde{\mu}, P, t_*)$ is not continuous in t_*. But this cannot be shown with the implicit function theorem since differentiability at t_* is not given for $G = \mathbb{I}_{[t_*,\infty)}$. This can only be obtained by studying the influence function of quantiles.

4.2 The influence function of the (log)normal regression MTLF

Here, (15) becomes according to (11)

$$W_2(\epsilon,\theta) = 2 \int [a(\epsilon,\theta)\ \ [F_{s,\epsilon}(m_s(\theta) + a(\epsilon,\theta)) + F_{s,\epsilon}(m_s(\theta) - a(\epsilon,\theta))]$$
$$-\ \ \mathcal{F}_{s,\epsilon}(m_s(\theta) + a(\epsilon,\theta)) + \mathcal{F}_{s,\epsilon}(m_s(\theta) - a(\epsilon,\theta))]\ \dot{m}_s(\theta)\, P^S(ds)$$

and W_1 of (16) is given by, see Section 3.2,

$$W_1(\epsilon,\theta,a) = \int [F_{s,\epsilon}(m_s(\theta) + a) - F_{s,\epsilon}(m_s(\theta) - a)]\ P^S(ds) - (1-\alpha).$$

As in Section 3.2, we have

$$\frac{\partial}{\partial\theta} W_1(\epsilon,\theta,a) = \int [f_{s,\epsilon}(m_s(\theta) + a) - f_{s,\epsilon}(m_s(\theta) - a)]\ \dot{m}_s(\theta)\, P^S(ds)$$

and

$$\frac{\partial}{\partial a} W_1(\epsilon,\theta,a) = \int [f_{s,\epsilon}(m_s(\theta) + a) + f_{s,\epsilon}(m_s(\theta) - a)]\ P^S(ds).$$

Additionally, we use here

$$\frac{\partial}{\partial \epsilon} W_1(\epsilon, \theta, a)\bigg|_{\epsilon=0} = \int \left[(G_s - F_s)(m_s(\theta) + a) - (G_s - F_s)(m_s(\theta) - a) \right] \, P^S(ds).$$

Setting $\theta_0 = \widetilde{\theta}_M(P_0) = \widetilde{\theta}_M(P)$, $a_0 = a(0, \theta_0)$, we make the following assumption:

$$\begin{array}{c} F_s \text{ and } G_s \text{ are differentiable in a neighborhood of} \\ m_s(\theta_0) + a_0 \text{ and } m_s(\theta_0) - a_0 \text{ for all } s \text{ in the support of } P^S. \end{array} \tag{20}$$

Under this assumption, the implicit function theorem provides

$$\begin{aligned} \dot{a}_\theta(0) :=& \frac{\partial}{\partial \theta} a(\epsilon, \theta)\bigg|_{(\epsilon,\theta)=(0,\theta_0)} \\ =& -\left(\int \left[f_s(m_s(\theta_0) + a_0) + f_s(m_s(\theta_0) - a_0) \right] \, P^S(ds) \right)^{-1} \\ & \cdot \int \left[f_s(m_s(\theta_0) + a_0) - f_s(m_s(\theta_0) - a_0) \right] \, \dot{m}_s(\theta_0) \, P^S(ds) \end{aligned}$$

and

$$\dot{a}_\epsilon(0) := \frac{\partial}{\partial \epsilon} a(\epsilon, \theta)\bigg|_{(\epsilon,\theta)=(0,\theta_0)} = -\left(\int \left[f_s(m_s(\theta_0) + a_0) + f_s(m_s(\theta_0) - a_0) \right] \, P^S(ds) \right)^{-1}$$

$$\int \left[(G_s - F_s)(m_s(\theta_0) + a_0) - (G_s - F_s)(m_s(\theta_0) - a_0) \right] \, P^S(ds).$$

Let here $\mathcal{G}_s$ be the antiderivative of G_s so that $\mathcal{F}_{s,\epsilon} = (1-\epsilon)\mathcal{F}_s + \epsilon \mathcal{G}_s = \mathcal{F}_s + \epsilon(\mathcal{G}_s - \mathcal{F}_s)$ and set $\ddot{m}_s(\theta) = \frac{\partial}{\partial \theta} \dot{m}_s(\theta)^\top$. Now, we can calculate the derivatives of $W_2(\epsilon, \theta)^\top$:

$$\begin{aligned} & \frac{\partial}{\partial \theta} W_2(\epsilon, \theta)^\top \bigg|_{(\epsilon,\theta)=(0,\theta_0)} \\ =& \; 2 \int \Big\{ a_0 \Big[f_s(m_s(\theta_0) + a_0)(\dot{m}_s(\theta_0) + \dot{a}_\theta(0)) + f_s(m_s(\theta_0) - a_0)(\dot{m}_s(\theta_0) - \dot{a}_\theta(0)) \Big] \\ & - \; \Big[F_s(m_s(\theta_0) + a_0) - F_s(m_s(\theta_0) - a_0) \Big] \dot{m}_s(\theta_0) \Big\} \dot{m}_s(\theta_0)^\top P^S(ds) \\ & + \; 2 \int \Big\{ a_0 \Big[F_s(m_s(\theta_0) + a_0) + F_s(m_s(\theta_0) - a_0)) \Big] - \mathcal{F}_s(m_s(\theta_0) + a_0) \\ & + \; \mathcal{F}_s(m_s(\theta_0) - a_0) \Big\} \ddot{m}_s(\theta_0) P^S(ds) \\ =:& \; C(P) \end{aligned}$$

and

$$
\begin{aligned}
&\frac{\partial}{\partial \epsilon} W_2(\epsilon,\theta)^\top\Big|_{(\epsilon,\theta)=(0,\theta_0)} \\
&= \; D(P) \\
&\quad + \; 2\,a_0 \int \Big[(G_s - F_s)(m_s(\theta_0)+a_0) + (G_s - F_s)(m_s(\theta_0)-a_0)\Big]\dot{m}_s(\theta_0)^\top P^S(ds) \\
&\quad - \; 2 \int \int_{m_s(\theta_0)-a_0}^{m_s(\theta_0)+a_0} (G_s - F_s)\,(y)\,dy\,\dot{m}_s(\theta_0)^\top P^S(ds)
\end{aligned}
$$

where

$$
D(P) := 2 \int \Big\{a_0\,\dot{a}_\epsilon(0)\Big[f_s(m_s(\theta_0)+a_0) - f_s(m_s(\theta_0)-a_0)\Big]\Big\}\dot{m}_s(\theta)^\top\, P^S(ds).
$$

As before, the implicit function theorem provides the following theorem.

Theorem 2. *Under the assumption (20), the (log)normal regression MTLF $\widetilde{\theta}_M$ of Definition 6 satisfies*

$$
\begin{aligned}
&\lim_{\epsilon\downarrow 0} \frac{\widetilde{\theta}_M(P_\epsilon) - \widetilde{\theta}_M(P)}{\epsilon} = -C(P)^{-1}\Big\{D(P) \\
&\quad +2\,a_0 \int \Big[(G_s - F_s)(m_s(\theta_0)+a_0) + (G_s - F_s)(m_s(\theta_0)-a_0)\Big]\,\dot{m}_s(\theta_0)\,P^S(ds) \\
&\quad -2 \int \int_{m_s(\theta_0)-a_0}^{m_s(\theta_0)+a_0} (G_s - F_s)(y)\,dy\,\dot{m}_s(\theta_0)^\top\,P^S(ds)\Big\}.
\end{aligned}
$$

Using here $G_{s_*}(y) = 1\!\!1_{[y_*,\infty)}(y)$ if $Q^{Y|S=s_*} = \delta_{y_*}$ and $G_s(y) = F_s(y)$ for $s \neq s_*$, assumption (20) is satisfied if $y_* \neq m_{s_*}(\theta_0)+a_0$ and $y_* \neq m_{s_*}(\theta_0)-a_0$. Hence we get at once the following influence function.

Corollary 4. *If P^S has finite support and $m_{s_*}(\theta_0)-a_0 \neq y_* \neq m_{s_*}(\theta_0)+a_0$, then the influence function of the (log)normal regression MTLF $\widetilde{\theta}_M$ at $z_* = (y_*, s_*)$ is given by*

$$
\begin{aligned}
&IF(\widetilde{\theta}_M, P, z_*) \\
&\quad = \; -C(P)^{-1}\Big\{D(P) + 2\Big(\,(y_* - m_{s_*}(\theta_0))1\!\!1_{(m_{s_*}(\theta_0)-a_0,m_{s_*}(\theta_0)+a_0]}(y_*) \\
&\qquad - \; a_0\Big[F_{s_*}(m_{s_*}(\theta_0)+a_0) + F_{s_*}(m_{s_*}(\theta_0)-a_0)\Big] \\
&\qquad + \; \int_{m_{s_*}(\theta_0)-a_0}^{m_{s_*}(\theta_0)+a_0} F_{s_*}(y)\,dy\Big)\,\dot{m}_{s_*}(\theta_0)^\top\;P^S(\{s_*\})\;\Big\}.
\end{aligned}
$$

Obviously, this influence function is again a bounded function in y_*.

Example 3 (Symmetric case)**.**

If f_s is symmetric about $m_s(\theta_0)$ for all s of the support of P^S, then the results of Theorem 2 and Corollary 4 concern also the original trimmed likelihood functional $\widetilde{\theta}_O$ as shown in Section 3.2. In this case, $f_s(m_s(\theta_0)+a) = f_s(m_s(\theta_0)-a)$ and $F_s(m_s(\theta_0)-a) = 1 - F_s(m_s(\theta_0)+a)$ for any $a > 0$ so that with partial integration

$$\mathcal{F}_s(m_s(\theta_0)+a) - \mathcal{F}_s(m_s(\theta_0)-a) = a\,F_s(m_s(\theta_0)+a) + a(1-F_s(m_s(\theta_0)+a)) = a$$

implying

$$a\Big[F_s(m_s(\theta_0)+a) + F_s(m_s(\theta_0)-a)\Big] - \mathcal{F}_s(m_s(\theta_0)+a) + \mathcal{F}_s(m_s(\theta_0)-a) = 0 \tag{21}$$

for any $a > 0$. The equality (21) provides at once $W_2(0,\theta_0) = 0$ for any $a_0 := a(0,\theta_0)$ given by

$$W_1(0,\theta_0,a_0) = 2\int F_s(m_s(\theta_0)+a_0)\,P^S(ds) - 2 + \alpha = 0. \tag{22}$$

Hence we obtain $D(P) = 0$, $\dot{a}_\theta(0) = 0$, and

$$\begin{aligned} &C(P) \\ &= \;2\int \Big\{a_0\,2\,f_s(m_s(\theta_0)+a_0) - \Big[2\,F_s(m_s(\theta_0)+a_0) - 1\Big]\Big\}\dot{m}_s(\theta_0)\dot{m}_s(\theta_0)^\top P^S(ds), \end{aligned}$$

so that the influence function at y_* with $m_{s_*}(\theta_0) - a_0 \neq y_* \neq m_{s_*}(\theta_0) + a_0$ is given by

$$\begin{aligned} &IF(\widetilde{\theta}_M, P, z_*) = IF(\widetilde{\theta}_O, P, z_*) \\ &= \;-C(P)^{-1}\Big\{2\,(y_* - m_{s_*}(\theta_0))\;\mathbb{1}_{(m_{s_*}(\theta_0)-a_0,\, m_{s_*}(\theta_0)+a_0]}(y_*)\;\;\dot{m}_{s_*}(\theta_0)\;\;P^S(\{s_*\})\;\Big\}. \end{aligned}$$

If we additionally assume $f_s(y) = f_*(y - m_s(\theta_0))$ for all s of the support of P^S, where f_* is symmetric about 0, then equality (22) is equivalent to

$$F_*(a_0) = 1 - \frac{\alpha}{2},$$

so that $a_0 = F_*^{-1}\left(1 - \frac{\alpha}{2}\right)$ is the $1-\frac{\alpha}{2}$-quantile of the distribution given by F_*. In this case we get

$$C(P) = 2\Big\{2\;a_0\,f_*(a_0) - (1-\alpha)\Big\}\int \dot{m}_s(\theta_0)\dot{m}_s(\theta_0)^\top P^S(ds)$$

so that the influence function is given by

$$\begin{aligned} &IF(\widetilde{\theta}_M, P, z_*) = IF(\widetilde{\theta}_O, P, z_*) \\ &= \begin{cases} \Big(\int \dot{m}_s(\theta_0)\dot{m}_s(\theta_0)^\top P^S(ds)\Big)^{-1}\dot{m}_{s_*}(\theta_0)\;\frac{y_* - m_{s_*}(\theta_0)}{1-\alpha-2\,a_0\,f_*(a_0)}\;\;P^S(\{s_*\}) & \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{if } y_* \in (m_{s_*}(\theta_0) - a_0, m_{s_*}(\theta_0)+a_0), & \\ 0 \qquad\qquad\qquad\qquad\qquad\qquad \text{if } y_* \notin [m_{s_*}(\theta_0) - a_0, m_{s_*}(\theta_0)+a_0]. & \end{cases} \end{aligned}$$

For the special case of $m_s(\theta) = x(s)^\top \theta$, this influence function appears in the expansion which Bednarski et al. (2010), p. 212, derived for the least trimmed squares estimator for linear regression. The matrix $\int IF(\widetilde{\theta}_M, P, z)\, IF(\widetilde{\theta}_M, P, z)^\top\, P^{Y|S} \otimes P^S(dz)$ is also the asymptotic covariance matrix of the least trimmed squares estimator which Víšek (1999) derived for linear regression and Čížek (2005) for nonlinear regression.

5 Application to a real data set

As an example of the TLE we consider accelerated lifetime experiments carried out on freerunning pre-stressed steel within the SFB 823 at TU Dortmund University. In these experiments 25 steel samples were exposed to cyclic loads with stress s from an interval of $[300, 1050]\ N/mm^2$. The recorded times $t_1, \ldots, t_{25}$ describe the number of applied load cycles until the first tension wire in the material broke. A parametric approach used within the SFB to model the influence of s on t is given by the relation

$$\mathrm{g}(\theta, s) := \vartheta_0 - \vartheta_1 s + \vartheta_2 s^{-\vartheta_3} \quad (\theta = (\vartheta_0, \vartheta_1, \vartheta_2, \vartheta_3)^\top \in [0, \infty)^4). \tag{23}$$

It combines a nonlinear influence of stress levels with a linear one and models the expectation of the random variable T given $S = s$ on the log scale, i.e. $\log(\mathrm{E}_\theta(T|S = s)) = g(\theta, s)$.

At first we assume $T|S = s \sim \mathrm{Exp}(\lambda_s(\theta))$. As equation (23) is used to model the expectation of T given $S = s$ on the log scale, it must hold that

$$\log(\mathrm{E}(T|S = s)) = \log\left(\frac{1}{\lambda_s(\theta)}\right) = \mathrm{g}(\theta, s) = \vartheta_0 - \vartheta_1 s + \vartheta_2 s^{-\vartheta_3}.$$

Therefore, the link function $\lambda_s(\theta)$ is given by

$$\lambda_s(\theta) = \exp(-\vartheta_0 + \vartheta_1 s - \vartheta_2 s^{-\vartheta_3}) = \exp(-\mathrm{g}(\theta, s)).$$

When we assume a lognormal distribution for T given $S = s$, we have $\log(T)|S = s \sim \mathcal{N}(m_s(\theta), \sigma^2)$. Hence, the link function is directly given by $m_s(\theta) = \mathrm{g}(\theta, s)$.

The resulting loglikelihood functions can be maximized numerically for both distributions for the $N = 25$ observations. For the computation of the h-trimmed likelihood estimator the Fast TLE algorithm from Neykov and Müller (2003) was used. In Figure 1 we compare the results for the untrimmed likelihood estimator to the TLE with $h = 5$ for both distributions. In the case of the exponential distribution the shape of the fitted curve changes noticeably. When $h = 5$ observations are trimmed the fitted curve describes the remaining observations very well, whereas the untrimmed fit is a straight line which does not fit well to the data at all. For the (log)normal distribution the effect of the trimming is not that large but the fit is also better. Moreover, using the trimmed estimators, the fitted curves do not differ much if the likelihood is based on the exponential or the lognormal distribution.

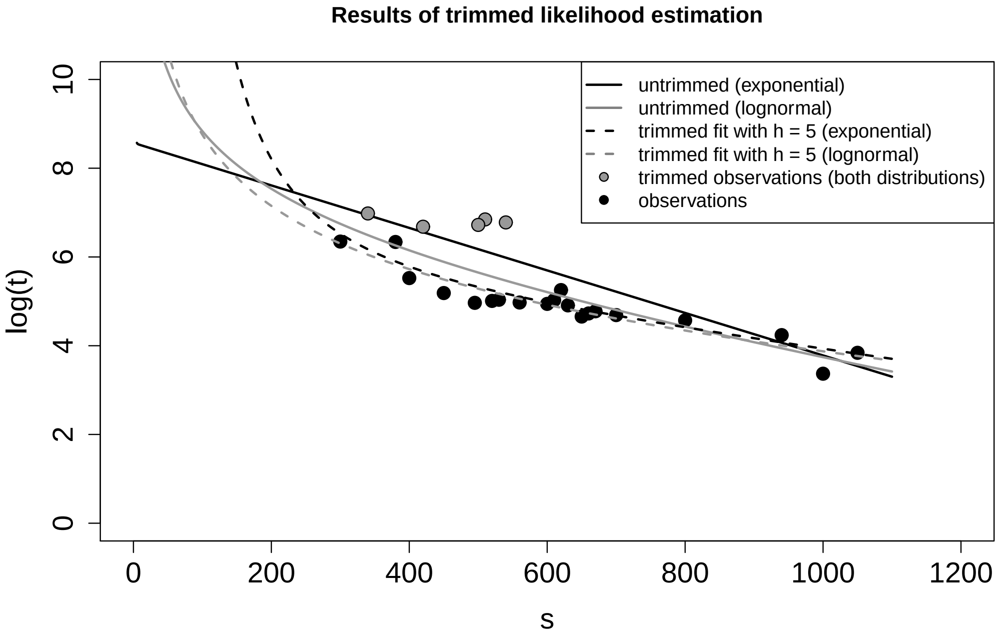

Figure 1: Estimated regression function assuming exponential or lognormal distribution with 5 trimmed observations and without trimming

6 Discussion

Since the influence function is defined for the functional defining an estimator, we considered at first two versions of the functional of a trimmed likelihood estimator, one, called the original trimmed likelihood functional (OTLF), which corresponds to the original trimmed likelihood estimator, and a modified version, called modified trimmed likelihood functional (MTLF), used by Bednarski and Clarke (1993, 2002) and by Bednarski et al. (2010). We showed that these two versions do not coincide in general and indicated situations for coincidence. Since we used the implicit function theorem, we could not show the coincidence at any empirical distribution. Moreover, often no parameter will satisfy the defining equations of the MTLF for empirical distributions. Nevertheless, we derived the influence function only for the modified version using again the implicit function theorem. However, the influence function could be derived similarly for the original version. On the other hand the results will be more complicated since then additionally derivatives of the densities of the central distribution are necessary. The approach was only demonstrated for trimmed likelihood functionals based on the exponential and the (log)normal distribution in regression models with linear and nonlinear link function. It is possible that it can be used also for other distributions. We expect that censoring, an important issue in lifetime experiments, can be treated with this approach as well. Another extension of the presented work will be to derive tests, confidence intervals and prediction intervals using the asymptotic distribution. In this context it would be important to know whether the trimmed estimators are asymptotically linear in the derived influence functions. For that it is useful to note that the presented results show Gâteaux differentiability of the modified trimmed likelihood functionals. A question is whether stronger differentiability notions like Hadamard differentiability can be shown.

Acknowledgments

This work has been supported by the Collaborative Research Center ”Statistical modeling of nonlinear dynamic processes” (SFB 823) of the German Research Foundation (DFG) in project B5 ”Statistical methods for damage processes under cyclic load”. We want to thank Professor R. Maurer and Guido Heeke for their experiments and providing us the data.

References

[1] Ahmed, E.S., Volodin, A.I. and Hussein, A.A. (2005). Robust weighted likelihood estimation of exponential parameters. *IEEE Transactions on Reliability* **54**, 389-395.

[2] Bednarski, T. and Clarke, B.R. (1993). Trimmed likelihood estimation of location and scale of the normal distribution. *Austral. J. Statist.* **35**, 141-153.

[3] Bednarski, T. and Clarke, B.R. (2002). Asymptotics for an adaptive trimmed likelihood location estimator. *Statistics* **36**, 1-8.

[4] Bednarski, T., Clarke, B.R. and Schubert, D. (2010). Adaptive trimmed likelihood estimation in regression. *Discussiones Mathematicae. Probability and Statistics* **30**, 203-219.

[5] Boudt, K., Caliskan, D. and Croux, C. (2011). Robust explicit estimators of Weibull parameters. *Metrika* **73**, 187-209.

[6] Čížek, P. (2005). Least trimmed squares in nonlinear regression under dependence. *J. Statist. Plann. Inference.* **136**, 3967-3988.

[7] Hampel, F.R. (1974). The influence curve and its role in robust estimation. *J. Amer. Statist. Assoc.* **69**, 383-393.

[8] Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. and Stahel, W.A. (1986). *Robust Statistics - The Approach Based on Influence Functions.* John Wiley, New York.

[9] Heuser, H. (1981). *Lehrbuch der Analysis. Teil 2.* Teubner, Stuttgart.

[10] Mili, L. and Coakley, C.W. (1996). Robust estimation in structured linear regression. *Ann. Statist.* **24**, 2297-2778.

[11] Müller, Ch.H. (1995). Breakdown points for designed experiments. *J. Statist. Plann. Inference.* **45**, 413-427.

[12] Müller, Ch.H. (1997). *Robust Planning and Analysis of Experiments.* Lecture Notes in Statistics 124, Springer, New York.

[13] Müller, Ch.H. (2013a). Upper and lower bounds for breakdown points. In: *Robustness and Complex Data Structures. Festschrift in Honour of Ursula Gather.* Eds. C. Becker, R. Fried and S. Kuhnt, Springer, Berlin, Heidelberg, 67-84.

[14] Müller, Ch.H. (2013b). D-optimal designs for lifetime experiments with exponential distribution and censoring. In: *mODa 10 - Advances in Model-Oriented Design and Analysis.* Eds. D. Ucinski, A.C. Atkinson and M. Patan, Physica-Verlag, Heidelberg, 179-186.

[15] Müller, Ch.H., Neykov, N. (2003). Breakdown points of trimmed likelihood estimators and related estimators in generalized linear models. *J. Statist. Plann. Inference* **116**, 503–519.

[16] Neykov, N. and Müller, Ch.H. (2003). Breakdown point and computation of trimmed likelihood estimators in generalized linear models. In: *Developments in Robust Statistics*, eds. R. Dutter, P. Filzmoser, U. Gather, P.J. Rousseeuw, Physica-Verlag, Heidelberg, 277-286.

[17] Rieder, H. (1994). *Robust Asymptotic Statistics.* Springer, New York.

[18] Rousseeuw, P.J. (1984). Least median of squares regression. *J. Amer. Statist. Assoc.* **79**, 871-880.

[19] Rousseeuw, P.J. and Leroy, A.M. (1987). *Robust Regression and Outlier Detection.* John Wiley, New York.

[20] Rousseeuw, P. J., Driessen, K. (2006) Computing LTS regression for large data sets. *Data Min. Knowl. Disc.* **12**, 29–45.

[21] Staudte, R.G. and Sheather, S.J. (1990). *Robust Estimation and Testing.* Wiley, New York.

[22] Vandev, D.L. (1993). A note on breakdown point of the least median squares and least trimmed squares. *Statist. Probab. Lett.* **16**, 117–119.

[23] Vandev, D., Neykov, N. (1998). About regression estimators with high breakdown point. *Statistics* **32**, 111–129.

[24] Víšek, J.Á. (1999). The least trimmed squares - random carriers. *Bull. Czech Econometric Soc.* **10**, 1-30.