Statistics
A Journal of Theoretical and Applied Statistics

ISSN: 0233-1888 (Print) 1029-4910 (Online) Journal homepage: https://www.tandfonline.com/loi/gsta20

Trimmed likelihood estimators for lifetime experiments and their influence functions

Christine H. Müller, Sebastian Szugat, Nuri Celik & Brenton R. Clarke

To cite this article: Christine H. Müller, Sebastian Szugat, Nuri Celik & Brenton R. Clarke (2016) Trimmed likelihood estimators for lifetime experiments and their influence functions, Statistics, 50:3, 505-524, DOI: 10.1080/02331888.2015.1104313

To link to this article: https://doi.org/10.1080/02331888.2015.1104313

Published online: 18 Nov 2015.

Statistics, 2016
Vol. 50, No. 3, 505–524, http://dx.doi.org/10.1080/02331888.2015.1104313

Taylor & Francis
Taylor & Francis Group

Trimmed likelihood estimators for lifetime experiments and their influence functions

Christine H. Müller[a], Sebastian Szugat[a*], Nuri Celik[b] and Brenton R. Clarke[c]

[a]*Faculty of Statistics, TU Dortmund University, Dortmund, Germany;* [b]*Department of Statistics, Bartin University, Turkey;* [c]*School of Engineering and Information Technology, Murdoch University, Perth, Australia*

(*Received 6 May 2015; accepted 30 September 2015*)

We study the behaviour of trimmed likelihood estimators (TLEs) for lifetime models with exponential or lognormal distributions possessing a linear or nonlinear link function. In particular, we investigate the difference between two possible definitions for the TLE, one called original trimmed likelihood estimator (OTLE) and one called modified trimmed likelihood estimator (MTLE) which is the finite sample version of a form for location and linear regression used by Bednarski and Clarke [Trimmed likelihood estimation of location and scale of the normal distribution. Aust J Statist. 1993;35:141–153, Asymptotics for an adaptive trimmed likelihood location estimator. Statistics. 2002;36:1–8] and Bednarski et al. [Adaptive trimmed likelihood estimation in regression. Discuss Math Probab Stat. 2010;30:203–219]. The OTLE is always an MTLE but the MTLE may not be unique even in cases where the OLTE is unique. We compare especially the functional forms of both types of estimators, characterize the difference with the implicit function theorem and indicate situations where they coincide and where they do not coincide. Since the functional form of the MTLE has a simpler form, we use it then for deriving the influence function, again with the help of the implicit function theorem. The derivation of the influence function for the functional form of the OTLE is similar but more complicated.

Keywords: outlier robustness; robust estimation; generalized linear model; lifetime distribution

AMS Subject Classification: Primary: 62F35; 62J02; 62J12; Secondary: 62N05; 62F10; 62G35

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 (TLEs) proposed by Müller and Neykov [1] which are outlier robust modifications of maximum likelihood estimators. The TLEs extend the least median of squares estimators and the least trimmed squares estimators of Rousseeuw [2] and Rousseeuw and Leroy [3] by replacing the likelihood functions of the normal distribution by likelihood functions of other distributions. Although the

*Corresponding author. Email: szugat@statistik.tu-dortmund.de

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.

The principal robustness measure used for TLEs is the breakdown point, see [1,4–9]. Additionally, Ahmed et al. [10] provide a relative bias and a quadratic risk as robustness measures for TLEs 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.[11] 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. [12] and Rieder.[13]

However, even the influence function for the one-sided α-trimmed mean is not easy to derive. In [14], it is given via the influence function of quantiles and is not completely correct, see p. 16. 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.[15] However, the influence function of TLEs is not treated. The only exception is that asymptotic expansions for TLEs are derived by Bednarski and Clarke [16,17] for the location and scale case and by Bednarski et al. [18] for the regression case. From these asymptotic expansions, the influence function can be derived. However, mostly they consider only TLEs 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 TLE.

This modified version of the TLE is not easy to calculate but its corresponding functionals, the modified trimmed likelihood functional (MTLF) $\tilde{\theta}_M$, 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 $\tilde{\theta}_O$ corresponding to the original trimmed likelihood estimator (OTLE), which is called here the original trimmed likelihood functional (OTLF). The definitions of the modified and the OTLF are given in Section 2 together with the definition of the TLE and the influence function.

It is not obvious that the modified functional $\tilde{\theta}_M$ should coincide with the original functional $\tilde{\theta}_O$. In Section 3, we compare the two 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 $\tilde{\theta}_M$ and $\tilde{\theta}_O$ 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 $\tilde{\theta}_M$ and $\tilde{\theta}_O$. 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 $\tilde{\theta}_M(P_N)$ (MTLE) and $\tilde{\theta}_O(P_N)$ (OTLE) coincide under some assumptions. This holds in particular when $\tilde{\theta}_M(P_N)$ is uniquely defined. While a simple example in Section 2 demonstrates that this is not satisfied in general, we show in Section 4 using a real data set that this holds quite often for more realistic situations. This is important since the OTLFs $\tilde{\theta}_O$ are much easier to calculate at empirical distributions while the influence functions of the MTLFs $\tilde{\theta}_M$ show a much simpler form.

The influence functions of the MTLFs are derived in Section 5. 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 5.1 and correspondingly the one for the (log)normal regression TLF in Section 5.2. 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 [16,17] and Bednarski et al. [18] did for the modified trimmed likelihood estimator (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.

Finally, we provide in Section 6 a discussion of the results.

2. Definitions

Let $z_{1N}, \dots, z_{NN}$ be realizations of independent random variables $Z_{1N}, \dots, Z_{NN}$, $z_N = (z_{1N}, \dots, z_{NN})$, and $\hat{\theta}(z_N)$ 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 $\hat{\theta}(z_N)$. Therefore, Hampel [11] proposed to consider the influence of an outlier z_* on the asymptotic value of $\hat{\theta}(Z_N)$. Usually, the estimator $\hat{\theta}(Z_N)$ converges for $N \to \infty$ in probability or almost surely to a value $\tilde{\theta}(P)$, where P is the underlying distribution. If $\tilde{\theta}(P)$ is defined for a class $\mathcal{P}$ of distributions, then $\tilde{\theta} : \mathcal{P} \to \Theta$ 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 $\tilde{\theta}((1-\epsilon)P + \epsilon\delta_{z_*})$ is defined. Then the influence function at z_* is the directional derivative of $\tilde{\theta}$ 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 2.1 (See [12]) *The influence function* $\mathrm{IF}(\hat{\theta}, P, z_*)$ *of a statistical functional* $\hat{\theta}$ *at a probability measure* P *and an observation* z_* *is defined as*

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

To take into account accelerated lifetime experiments, set $z_{1N} = (t_{1N}, s_{1N}), \dots, z_{NN} = (t_{NN}, s_{NN})$, 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 2.2 (See [1]) *The original h-trimmed likelihood estimator (OTLE)* $\hat{\theta}(z_N)$ *at* z_N *is defined as*

$$\hat{\theta}(z_N) = \arg\max_{\theta \in \Theta} \sum_{n=h+1}^{N} l_{(n)}(\theta, z_N),$$

where

$$l_n(\theta, z_N) = l(\theta, t_{nN}, s_{nN}) \quad \textit{and} \quad l_{(1)}(\theta, z_N) \le l_{(2)}(\theta, z_N) \le \dots \le l_{(N)}(\theta, z_N). \tag{1}$$

In an h-TLE the observations with the h smallest likelihood values are not used.

The functional form of this estimator is given in Definition 2.3. Thereby, we use $\alpha = \lim_{N\to\infty}(h_N/N)$ 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. It should be noted that S could be any explanatory variable. However, since our main applications are accelerated lifetime experiments, we call S here a stress variable. 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. [6].

Definition 2.3 *The original* α*-trimmed likelihood functional (OTLF)* $\tilde{\theta}_O(P)$ *at* $P = P^{T|S} \otimes P^S$ *is given by*

$$\tilde{\theta}_O(P) = \arg\max_{\theta \in \Theta} \int\int \mathbb{1}\{l(\theta, t, s) \ge b(\theta)\} l(\theta, t, s) P^{T|S=s}(\mathrm{d}t) P^S(\mathrm{d}s),$$

where $b(\theta)$ satisfies

$$b(\theta) = \arg\max \left\{ b; \int \int \mathrm{II}\{l(\theta, t, s) \geq b\} P^{T|S=s}(\mathrm{d}t) P^S(\mathrm{d}s) \geq 1 - \alpha \right\} \tag{2}$$

and $\mathrm{II}\{x \in A\} = \mathrm{II}_A(x)$ *denotes the indicator function.*

The functional form of the modified version used by Bednarski and Clarke [16,17] and Bednarski et al. [18] is given in Definition 2.4.

DEFINITION 2.4 *The modified α-trimmed likelihood functional (MTLF) $\tilde{\theta}_M = \tilde{\theta}_M(P)$ at $P = P^{T|S} \otimes P^S$ is a solution of*

$$0 = \int \int \mathrm{II}\{l(\tilde{\theta}_M, t, s) \geq b(\tilde{\theta}_M)\} \dot{l}(\tilde{\theta}_M, t, s) P^{T|S=s}(\mathrm{d}t) P^S(\mathrm{d}s).$$

where $b(\theta)$ is defined by Equation (2) and $\dot{l}(\theta, t, s) = (\partial/\partial\theta) l(\theta, t, s)$.

Note that the empirical version of this definition, which is also called here the modified h-trimmed likelihood estimator (MTLE), is a parameter estimate $\hat{\theta}_M(z_N)$ satisfying

$$0 = \sum_{n=h+1}^{N} \dot{l}_{(n)}(\hat{\theta}_M(z_N), z_N), \tag{3}$$

where $\dot{l}_{(n)}(\theta, z_N) = (\partial/\partial\theta) l_{(n)}(\theta, z_N)$, $l_{(n)}(\theta, z_N)$ is given by Equation (1).

The original h-TLE can be obtained by calculating the maximum likelihood estimator $\hat{\theta}_{\mathcal{Z}}$ of each subset $\mathcal{Z}$ of data points with $N - h$ elements and defining $\mathcal{Z}_0$ as $\mathcal{Z}_0 = \arg\max_{\mathcal{Z}} \sum_{n=h+1}^{N} l_{(n)}(\hat{\theta}_{\mathcal{Z}}, z_N)$. Then it holds that $\hat{\theta} = \hat{\theta}_{\mathcal{Z}_0}$. This implies that an h-TLE $\hat{\theta}$ satisfies Equation (3). However, there could be many more solutions of Equation (3). This is in particular the case when $l(\cdot, t, s)$ is not concave. However, it can also happen when $l(\cdot, t, s)$ is concave as the following simple example shows. Thereby note that for calculating all solutions of Equation (3), all solutions $\theta_{\mathcal{Z}}$ of $0 = \sum_{n \in \mathcal{Z}} \dot{l}_n(\theta_{\mathcal{Z}}, t_{nN}, s_{nN})$ must be found for all subsets $\mathcal{Z}$ of data points with $N - h$ elements. If the $N - h$ largest loglikelihood functions with respect to this solution $\theta_{\mathcal{Z}}$ provide the same subset $\mathcal{Z}$ from which the solution was obtained, i.e. the ordering given by Equation (1) reproduces $\mathcal{Z}$, then $\theta_{\mathcal{Z}}$ is a solution of Equation (3).

Example 2.5 Let $l(\theta, t, s)$ be the loglikelihood function of the normal distribution in a simple linear regression model with $\theta \in I\!R^2$, i.e. $l(\theta, t, s) = -(t - (1, s)\theta)^2$ up to a constant, so that the OTLE is the least trimmed squares estimator of Rousseeuw [2] and Rousseeuw and Leroy.[3] Consider five data points given by $(2, 0)^\top$, $(1, 1)^\top$, $(3, 1)^\top$, $(3.5, 1.5)^\top$, $(12, 10)^\top$ and use $h = 1$. Then $\hat{\theta}_1 = (-2, 1)^\top$ obtained by the subset $\mathcal{Z} = \{(2, 0)^\top, (3, 1)^\top, (3.5, 1.5)^\top, (12, 10)^\top\}$ is the OTLE and thus also an MTLE. The maximum likelihood estimator for the subsets $\mathcal{Z} = \{(2, 0)^\top, (1, 1)^\top, (3, 1)^\top, (12, 10)^\top\}$, $\mathcal{Z} = \{(2, 0)^\top, (1, 1)^\top, (3.5, 1.5)^\top, (12, 10)^\top\}$, and $\mathcal{Z} = \{(1, 1)^\top, (3, 1)^\top, (3.5, 1.5)^\top, (12, 10)^\top\}$, do not reproduce these sets when regarding the $N - h$ largest loglikelihood functions. However this is done by the maximum likelihood estimator given by $\hat{\theta}_2 = (0.2712274, 0.2542400)^\top$ for $\mathcal{Z} = \{(2, 0)^\top, (1, 1)^\top, (3, 1)^\top, (3.5, 1.5)^\top\}$ so that $\hat{\theta}_2$ is also a solution of Equation (3) and thus another MTLE. Hence in this example, two MTLE exist, see also Figure 1.

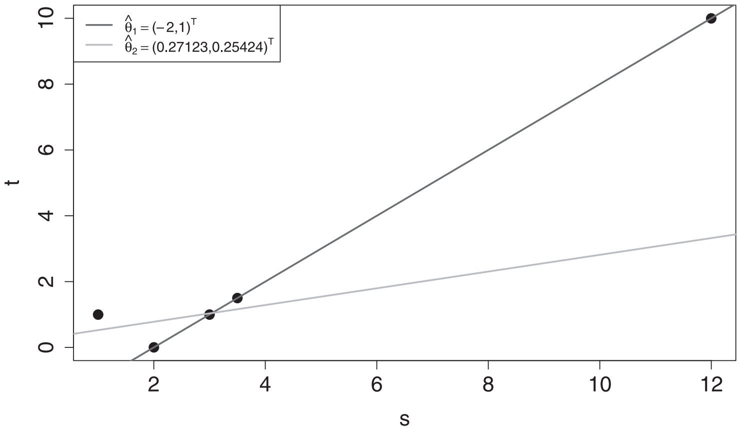

Figure 1. Visualization of the two different MTLE in Example 2.5. $\hat{\theta}_1$ is the maximum likelihood estimator, when the point $(1,1)^\top$ is trimmed and $\hat{\theta}_2$ when $(12,10)^\top$ is trimmed.

3. Comparison of MTLF and OTLF

To check under which conditions the MTLF given by Definition 2.4 and the OTLF given by Definition 2.3 coincide, we have to check the equality of

$$\int\int \mathbb{I}\{l(\theta,t,s) \geq b(\theta)\}\dot{l}(\theta,t,s)P^{T|S=s}(\mathrm{d}t)P^S(\mathrm{d}s) \tag{4}$$

and

$$\frac{\partial}{\partial\theta}\int\int \mathbb{I}\{l(\theta,t,s) \geq b(\theta)\}l(\theta,t,s)P^{T|S=s}(\mathrm{d}t)P^S(\mathrm{d}s). \tag{5}$$

We will consider here only the case where $b(\theta)$ given by Equation (2) satisfies

$$1-\alpha = \int \mathbb{I}\{l(\theta,t,s) \geq b(\theta)\}P^{T|S=s}(\mathrm{d}t)P^S(\mathrm{d}s) \tag{6}$$

for all θ in a neighbourhood of $\tilde{\theta}_M(P)$ and $\tilde{\theta}_O(P)$, 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 Equation (6).

3.1. *Exponential regression TLFs*

If the lifetimes at different stress levels have exponential distributions, then the loglikelihood function is given by

$$l(\theta,t,s) = \log(\lambda_s(\theta)) - \lambda_s(\theta)t, \tag{7}$$

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 $\theta = (\vartheta_0, \vartheta_1)^\top \in I\!R \times (0,\infty)$, 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. [19] or the real data example in Section 4.

Definition 3.1 *The exponential regression OTLF and MTLF are the OTLF and the MTLF where $l(\theta,t,s)$ is given by Equation* (7).

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 $\dot{l}(\theta,t,s) = (1/\lambda_s(\theta) - t)\dot{\lambda}_s(\theta)$ with $\dot{\lambda}_s(\theta) := (\partial/\partial\theta)\lambda_s(\theta)$. Then we obtain for expression (4) with partial integration of $\int tP^{T|S=s}(\mathrm{d}t)$

$$\begin{aligned}
&\int\int \mathbb{I}\{l(\theta,t,s) \geq b(\theta)\}\dot{l}(\theta,t,s)P^{T|S=s}(\mathrm{d}t)P^S(\mathrm{d}s)\\
&\quad = \int\int_0^{\eta_s(\theta,b(\theta))} \left(\frac{1}{\lambda_s(\theta)} - t\right)\dot{\lambda}_s(\theta)P^{T|S=s}(\mathrm{d}t)P^S(\mathrm{d}s)\\
&\quad = \int \left[F_s(\eta_s(\theta,b(\theta)))\left(\frac{1}{\lambda_s(\theta)} - \eta_s(\theta,b(\theta))\right) + \mathcal{F}_s(\eta_s(\theta,b(\theta)))\right]\dot{\lambda}_s(\theta)P^S(\mathrm{d}s)\\
&\quad =: U_P^M(\theta), \qquad (8)
\end{aligned}$$

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. $(\partial/\partial t)\mathcal{F}_s(t) = F_s(t)$. In particular, it is not necessary to assume an exponential distribution for $P^{T|S=s}$. Hence we arrive at the following lemma.

LEMMA 3.2 *The exponential regression MTLE $\tilde{\theta}_M$ at P is given as a solution of* $0 = U_P^M(\theta)$.

Similarly, the integral in Equation (5) is given by

$$\begin{aligned}
&\int\int \mathbb{I}\{l(\theta,t,s) \geq b(\theta)\}l(\theta,t,s)P^{T|S=s}(\mathrm{d}t)P^S(\mathrm{d}s)\\
&\quad = \int\int_0^{\eta_s(\theta,b(\theta))} (\log(\lambda_s(\theta)) - \lambda_s(\theta)t)P^{T|S=s}(\mathrm{d}t)P^S(\mathrm{d}s)\\
&\quad = \int [F_s(\eta_s(\theta,b(\theta)))[\log(\lambda_s(\theta)) - \lambda_s(\theta)\eta_s(\theta,b(\theta))] + \lambda_s(\theta)\mathcal{F}_s(\eta_s(\theta,b(\theta)))]P^S(\mathrm{d}s). \qquad (9)
\end{aligned}$$

To calculate the derivative of Equation (9), we need the derivative of $b(\theta)$ which is implicitly given by $W_1(\theta,b(\theta)) = 0$, where

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

We assume here that

F_s is differentiable in a neighbourhood of $\eta_s(\theta,b(\theta))$ with derivative f_s for all s in the support of P^S. (10)

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

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

and

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

If $(\partial/\partial b)W_1(\theta,b)|_{b=b(\theta)} \neq 0$, then the implicit function theorem (see [20,pp.210–211]) provides

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

If $(\partial/\partial b)W_1(\theta,b)|_{b=b(\theta)} = 0$, then $f_s(\eta_s(\theta,b(\theta))) = 0$ for all s of the support of P^S, so that $(\partial/\partial\theta)W_1(\theta,b)|_{b=b(\theta)} = 0$ 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 Equation (9) is

$$U_P^O(\theta) := \int\Bigg[f_s(\eta_s(\theta,b(\theta)))\dot{\eta}_s(\theta)b(\theta) + F_s(\eta_s(\theta,b(\theta)))\left[\frac{1}{\lambda_s(\theta)} - \eta_s(\theta,b(\theta))\right]\dot{\lambda}_s(\theta)$$
$$+\dot{\lambda}_s(\theta)\mathcal{F}_s(\eta_s(\theta,b(\theta)))\Bigg]P^S(\mathrm{d}s).$$

Hence we obtain the following lemma.

Lemma 3.3 *Under the assumption* (10), *the exponential regression OTLE* $\tilde{\theta}_O$ *at P is given as a solution of* $0 = U_P^O(\theta)$.

Corollary 3.4 *The difference between Equations* (4) *and* (5) *for exponential regression trimmed likelihood functionals is given by*

$$\int f_s(\eta_s(\theta,b(\theta)))\dot{\eta}_s(\theta)b(\theta)P^S(\mathrm{d}s). \tag{11}$$

The difference (11) 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 $\eta_s(\theta)$.

Example 3.5 (One-sample case) If P^S is given by a one-point measure at s_0, then the TLF $\tilde{\theta} := \tilde{\theta}_M(P) = \tilde{\theta}_O(P)$ can be given more explicitly. Setting $\eta := \eta_{s_0}(\tilde{\theta}, b(\tilde{\theta}))$ and using $\mathcal{F}_s(\eta) = \eta F_s(\eta) - \int_0^\eta t\, \mathrm{d}P^{T|S=s}(\mathrm{d}t)$, we have $F_{s_0}(\eta) = 1 - \alpha$ and the TLF at P satisfies

$$\begin{aligned} 0 &= \left[F_{s_0}(\eta) \left(\frac{1}{\lambda_{s_0}(\tilde{\theta})} - \eta \right) + \mathcal{F}_{s_0}(\eta) \right] \dot{\lambda}_{s_0}(\tilde{\theta}) \\ &= \left[(1-\alpha) \frac{1}{\lambda_{s_0}(\tilde{\theta})} - \int_0^\eta t\, \mathrm{d}P^{T|S=s_0}(\mathrm{d}t) \right] \dot{\lambda}_{s_0}(\tilde{\theta}) \\ \Longleftrightarrow \frac{1}{\lambda_{s_0}(\tilde{\theta})} &= \frac{1}{1-\alpha} \int_0^\eta t\, \mathrm{d}P^{T|S=s_0}(\mathrm{d}t). \end{aligned}$$

Since $(1/(1-\alpha)) \int_0^\eta t\, \mathrm{d}P^{T|S=s_0}(\mathrm{d}t)$ 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. [10] 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. $x(s) = 1/s$ with $\theta \in (0, \infty)$ or $x(s) = (1, -s)^\top$ with $\theta = (\vartheta_0, \vartheta_1)^\top \in [0, \infty)^2$. However, also nonlinear links like $m_s(\theta) = \vartheta_0 + \vartheta_1 (1/s)^{\vartheta_2}$ 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 4. The loglikelihood function is then given up to a constant by

$$l(\theta, y, s) = -(y - m_s(\theta))^2. \tag{12}$$

Definition 3.6 *The (log)normal regression OTLF and MTLF are the OTLF and the MTLF where $l(\theta, t, s)$ is given by Equation (12).*

Then we have

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

With partial integration of $\int y P^{Y|S=s}(\mathrm{d}y)$ and $\dot{l}(\theta, y, s) = 2(y - m_s(\theta)) \dot{m}_s(\theta)$ where $\dot{m}_s(\theta) = (\partial/\partial\theta) m_s(\theta)$, we obtain for expression (4)

$$\begin{aligned} &\int \int \mathbb{I}\{l(\theta, y, s) \geq b(\theta)\} \dot{l}(\theta, y, s) P^{Y|S=s}(\mathrm{d}t) P^S(\mathrm{d}s) \\ &= 2 \int \int_{m_s(\theta)-a(\theta)}^{m_s(\theta)+a(\theta)} (y - m_s(\theta)) \dot{m}_s(\theta) P^{Y|S=s}(\mathrm{d}y) P^S(\mathrm{d}s) \\ &= 2 \int \big[a(\theta) [F_s(m_s(\theta) + a(\theta)) + F_s(m_s(\theta) - a(\theta))] \\ &\qquad - \mathcal{F}_s(m_s(\theta) + a(\theta)) + \mathcal{F}_s(m_s(\theta) - a(\theta)) \big] \dot{m}_s(\theta) P^S(\mathrm{d}s) \\ &=: V_P^M(\theta), \end{aligned} \tag{13}$$

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 $I\!R$. Hence the following lemma is shown.

LEMMA 3.7 *The (log)normal regression MTLE $\tilde{\theta}_M$ at P is given as a solution of* $0 = V_P^M(\theta)$.

For the integral in Equation (5), we get using partial integration of $\int yP^{Y|S=s}(\mathrm{d}y)$ and $\int y^2 P^{Y|S=s}(\mathrm{d}y)$

$$\begin{aligned}
&\int\int \mathrm{1\!I}\{l(\theta,t,s) \geq b(\theta)\} l(\theta,t,s) P^{Y|S=s}(\mathrm{d}y)P^S(\mathrm{d}s)\\
&\quad = -\int\int_{m_s(\theta)-a(\theta)}^{m_s(\theta)+a(\theta)} (y-m_s(\theta))^2 P^{Y|S=s}(\mathrm{d}y)P^S(\mathrm{d}s)\\
&\quad = -\int \big\{a(\theta)^2[F_s(m_s(\theta)+a(\theta)) - F_s(m_s(\theta)-a(\theta))]\\
&\qquad - 2[\mathcal{H}_s(m_s(\theta)+a(\theta)) - \mathcal{H}_s(m_s(\theta)-a(\theta))]\\
&\qquad + 2m_s(\theta)[\mathcal{F}_s(m_s(\theta)+a(\theta)) - \mathcal{F}_s(m_s(\theta)-a(\theta))]\big\} P^S(\mathrm{d}s), \qquad (14)
\end{aligned}$$

where $\mathcal{H}_s$ is the antiderivative of H_s given by $H_s(y) = yF_s(y)$. To obtain the derivative of Equation (14), we have to calculate the derivative of $a(\theta)$ which is implicitly given by $W_1(\theta, a(\theta)) = 0$, where

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

We assume here that

$$\begin{aligned}
&F_s \text{ is differentiable in a neighbourhood of } m_s(\theta)+a(\theta) \text{ and } m_s(\theta)-a(\theta)\\
&\quad \text{with derivative } f_s \text{ for all } s \text{ in the support of } P^S. \qquad (15)
\end{aligned}$$

Since

$$\frac{\partial}{\partial a} W_1(\theta,a) = \int [f_s(m_s(\theta)+a) + f_s(m_s(\theta)-a)]P^S(\mathrm{d}s)$$

and

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

the implicit function theorem provides

$$\begin{aligned}
\dot{a}(\theta) &:= \frac{\partial}{\partial\theta} a(\theta)\\
&= -\left(\int [f_s(m_s(\theta)+a(\theta)) + f_s(m_s(\theta)-a(\theta))]P^S(\mathrm{d}s)\right)^{-1}\\
&\quad \cdot \int [f_s(m_s(\theta)+a(\theta)) - f_s(m_s(\theta)-a(\theta))]\dot{m}_s(\theta)P^S(\mathrm{d}s)
\end{aligned}$$

if $(\partial/\partial a)W_1(\theta,a)|_{a=a(\theta)} \neq 0$. If $(\partial/\partial a)W_1(\theta,a)|_{a=a(\theta)} = 0$, then also $(\partial/\partial\theta)W_1(\theta,a)|_{a=a(\theta)} = 0$ so that we can set $\dot{a}(\theta) = 0$ in this case. Hence the derivative of Equation (14) is

$$\begin{aligned} V_P^O(\theta) := -\int \{ & a(\theta)^2[f_s(m_s(\theta)+a(\theta)) - f_s(m_s(\theta)-a(\theta))]\dot{m}_s(\theta) \\ & + a(\theta)^2[f_s(m_s(\theta)+a(\theta)) + f_s(m_s(\theta)-a(\theta))]\dot{a}(\theta) \\ & - 2a(\theta)[F_s(m_s(\theta)+a(\theta)) + F_s(m_s(\theta)-a(\theta))]\dot{m}_s(\theta) \\ & + 2[\mathcal{F}_s(m_s(\theta)+a(\theta)) - \mathcal{F}_s(m_s(\theta)-a(\theta))]\dot{m}_s(\theta)\}P^S(\mathrm{d}s). \end{aligned}$$

Lemma 3.8 *Under the assumption* (15), *the* (*log*)*normal regression OTLE* $\tilde{\theta}_O$ *at P is given as a solution of* $0 = V_P^O(\theta)$.

Corollary 3.9 *The difference between Equations* (4) *and* (5) *for* (*log*)*normal regression trimmed likelihood functionals is given by*

$$\begin{aligned} \int \{ & a(\theta)^2[f_s(m_s(\theta)+a(\theta)) - f_s(m_s(\theta)-a(\theta))]\dot{m}_s(\theta) \\ & + a(\theta)^2[f_s(m_s(\theta)+a(\theta)) + f_s(m_s(\theta)-a(\theta))]\dot{a}(\theta)\}P^S(\mathrm{d}s). \end{aligned} \tag{16}$$

The difference (16) 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. [18] considered symmetric distributions for the central distribution. However, any neighbourhood around a central symmetric distribution contains also asymmetric distributions.

4. Comparison of the OTLE and the MTLE using 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 $t_1,\dots,t_{25}$ 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

$$\mathrm{g}(\theta,s) := \vartheta_0 - \vartheta_1 s + \vartheta_2 s^{-\vartheta_3} \quad (\theta = (\vartheta_0,\vartheta_1,\vartheta_2,\vartheta_3)^\top \in [0,\infty)^4). \tag{17}$$

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 (17) 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 OTLE, the Fast TLE algorithm from Neykov and Müller [21] was used. In Figure 2, we compare the results for the untrimmed likelihood estimator to the OTLE 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 lognormal distribution, the effect of 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.

To determine the MTLE, the maximum likelihood estimator must be calculated for $\binom{25}{5}$ subsets which is not practicable. Therefore we checked only for the subsets used in the Fast TLE algorithm whether the subsets reproduce the ordering of the loglikelihood function. It turned out for the exponential distribution that there is no other subset satisfying this property so that no other MTLE besides the OTLE were found. For the lognormal distribution, three subsets which preserve the ordering of the likelihood were found via the Fast TLE algorithm. These solutions are given in Table 1. The corresponding index sets of the trimmed observations are $\{4, 5, 13, 16, 18\}$ for Solution 1, which coincides with the index set of the solution for the exponential distribution, $\{4, 5, 13, 16, 19\}$ for Solution 2, and $\{4, 5, 8, 13, 16\}$ for Solution 3.

Since the three different MTLE are quite different from each other, the corresponding estimated regression functions are depicted in Figure 3. All three functions are plausible solutions

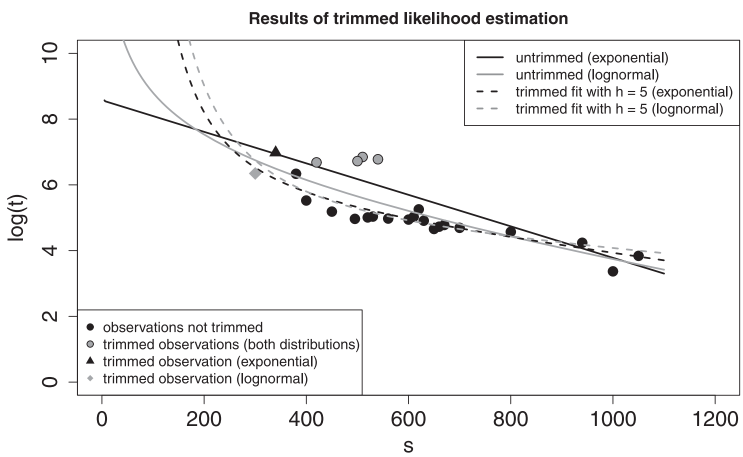

Figure 2. Estimated regression function assuming exponential or lognormal distribution with 5 trimmed observations and without trimming.

Table 1. Three MTLE for the (log)normal distribution with $h = 5$ trimmed observations. The solution with the highest value of the loglikelihood is also the OTLE.

	Solution 1	Solution 2	Solution 3
$\hat{\vartheta}_0$	3.568	6.343	5.003
$\hat{\vartheta}_1$	0.001	0.002	0.001
$\hat{\vartheta}_2$	59.644	1.958×10^{20}	44883.810
$\hat{\vartheta}_3$	0.516	7.948	1.748
Loglikelihood	0.29	3.06	1.63

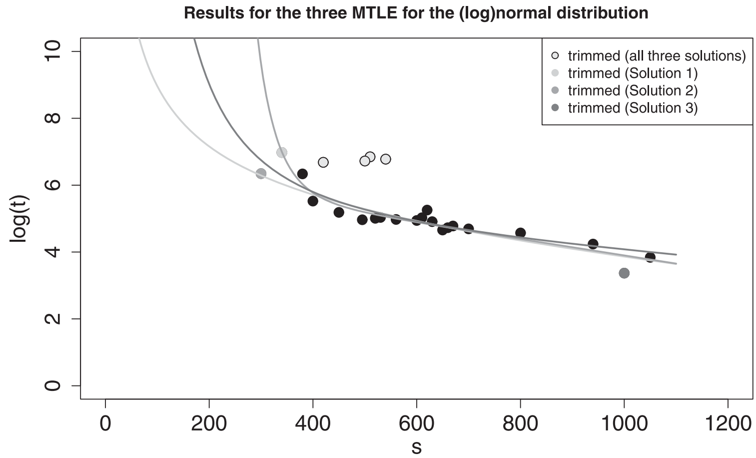

Figure 3. Estimated regression function using the three different MTLE for the (log)normal distribution.

for the given data set although the OTLE which coincides with Solution 2 leads to the highest value of the loglikelihood.

For $h = 1$, all $\binom{25}{1}$ subsets can be checked. In this case, there is no other MTLE besides the OTLE for the exponential distribution as well as for the lognormal distribution.

5. The influence function of TLFs

Although we have seen that the MTLF of Definition 2.4 does not coincide in general with the OTLF of Definition 2.3, 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 any OTLE is also an MTLE. Although Example 2.5 provides a simple example where the MTLE is not unique, the application in Section 4 also indicates that in many more realistic situations the MTLE is unique. Moreover, if the MTLE is not unique, one can restrict oneself to the solution which coincides with the OTLE. Hence it makes sense to derive the influence function only for the MTLF although the estimator which is used is the OTLE. This is important since only the OTLE can be calculated efficiently. While for very small sample sizes, the OTLEs and the MTLEs can be obtained by calculating the maximum likelihood estimator for all subsamples with $N - h$ elements, special methods for larger sample sizes have been developed only for the OTLE, see, e.g. [21] or [22].

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

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

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 2.1. Thereby, $\tilde{\theta}_M(P_\epsilon)$ is implicitly given by

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

where

$$W_2(\epsilon,\theta) = \int\int \mathrm{II}\{l(\theta,t,s) \geq b(\epsilon,\theta)\}\dot{l}(\theta,t,s)P_\epsilon^{T|S=s}(\mathrm{d}t)P^S(\mathrm{d}s) \tag{18}$$

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

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

with

$$W_1(\epsilon,\theta,b) = \int\int \mathrm{II}\{l(\theta,t,s) \geq b\}P_\epsilon^{T|S=s}(\mathrm{d}t)P^S(\mathrm{d}s) - (1-\alpha). \tag{19}$$

5.1. *The influence function of the exponential regression MTLF*

Here, Equation (18) becomes according to Equation (8)

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

and W_1 of Equation (19) is given by, see Section 3.1,

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

As in Section 3.1, we have

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

and

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

Additionally, we use here

$$\frac{\partial}{\partial\epsilon}W_1(\epsilon,\theta,b)\bigg|_{\epsilon=0} = \int (G_s - F_s)(\eta_s(\theta,b))P^S(\mathrm{d}s).$$

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

$$F_s \text{ and } G_s \text{ are differentiable in a neighbourhood of } \eta_s(\theta_0,b_0) \text{ for all } s \text{ in the support of } P^S. \tag{20}$$

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. We consider here only the cases where the differentiability is satisfied though. Then the implicit function theorem provides

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

and

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

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

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

and

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

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 $\ddot{\lambda}_s(\theta) = (\partial/\partial\theta)\dot{\lambda}_s(\theta)^\top$. Now, we can calculate the derivatives of $W_2(\epsilon,\theta)^\top$:

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

and

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

where

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

and

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

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

THEOREM 5.1 *Under the assumption* (20), *the exponential regression MTLF* $\tilde{\theta}_M$ *of Definition* 7 *satisfies*

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

Using $G_{s_*}(t)=\mathrm{I\!I}_{[t_*,\infty)}(t)$ if $Q^{T|S=s_*}=\delta_{t_*}$ and $G_s(t)=F_s(t)$ for $s\neq s_*$, assumption (20) is satisfied if $t_*\neq\eta_{s_*}(\theta_0,b_0)$. Hence we get at once the following influence function.

COROLLARY 5.2 *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* $\tilde{\theta}_M$ *at* $z_*=(t_*,s_*)\in[0,\infty)^2$ *is given by*

$$\mathrm{IF}(\tilde{\theta}_M,P,z_*) = -A(P)^{-1}\Bigg[\Bigg\{(\eta_*-t_*)\,\mathrm{I\!I}_{[0,\eta_*]}(t_*)-\int_0^{\eta_*}F_{s_*}(t)\,\mathrm{d}t\Bigg\}\dot{\lambda}_{s_*}(\theta_0)^\top P^S(\{s_*\})$$
$$+\ (\mathrm{I\!I}_{[0,\eta_*]}(t_*)-F_{s_*}(\eta_*))(a_{s_*}(P)\dot{\lambda}_{s_*}(\theta_0)^\top-B(P))P^S(\{s_*\})\Bigg].$$

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

Example 5.3 (One-sample case) If P^S is given by a one-point measure at s_0, then the results of Theorem 5.1 and Corollary 5.2 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 $\eta_*=\eta_{s_0}(\theta_0,b_0)$, $\tilde{\theta}=\theta_0=\tilde{\theta}_M(P)=\tilde{\theta}_O(P)$ we obtain

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

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

$$\begin{aligned}\mathrm{IF}(\tilde{\theta}_M,P,z_*)&=\mathrm{IF}(\tilde{\theta}_O,P,z_*)\\&=\begin{cases}\tilde{\theta}^2\left(\dfrac{\eta_*\alpha-t_*}{1-\alpha}+\dfrac{1}{\lambda_{s_0}(\tilde{\theta})}\right)\dot{\lambda}_{s_0}(\tilde{\theta}) & \text{if } t_*<\eta_*,\\ \tilde{\theta}^2\left(-\eta_*+\dfrac{1}{\lambda_{s_0}(\tilde{\theta})}\right)\dot{\lambda}_{s_0}(\tilde{\theta}) & \text{if } t_*>\eta_*,\end{cases}\end{aligned}\tag{21}$$

since according to Example 3.5 $F_{s_0}(\eta_*)=1-\alpha$ and

$$\frac{1}{\lambda_{s_0}(\tilde{\theta})}=\frac{1}{1-\alpha}\int_0^{\eta_*}t\,\mathrm{d}P^{T|S=s_0}(\mathrm{d}t).\tag{22}$$

Note that the influence function of the one-sided trimmed mean $\tilde{\mu} = \tilde{\mu}(P) = (1/(1-\alpha)) \int_0^{\eta_*} t\,\mathrm{d}P(\mathrm{d}t)$ is given by, see, e.g.,[14,p.55]

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

This coincides with Equation (21) using Equation (22) and $\lambda_s(\mu) = 1/\mu$. Note that in contrast to the result of Staudte and Sheather,[14] it holds $\mathrm{IF}(\tilde{\mu}, P, t_*) = t_*/(1-\alpha) - \tilde{\mu}$ for $t_* = \eta_*$ so that $\mathrm{IF}(\tilde{\mu}, P, t_*)$ is not continuous in t_*, see [23,pp.43–45]. However, this cannot be shown with the implicit function theorem since differentiability at t_* is not given for $G = \mathbb{1}_{[t_*,\infty)}$. This can only be obtained by studying the influence function of quantiles.

5.2. *The influence function of the (log)normal regression MTLF*

Here, Equation (18) becomes according to Equation (13)

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

and W_1 of Equation (19) is given by, see Section 3.2,

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

As in Section 3.2, we have

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

and

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

Additionally, we use here

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

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

F_s and G_s are differentiable in a neighbourhood of $m_s(\theta_0) + a_0$ and $m_s(\theta_0) - a_0$ for all s in the support of P^S. (23)

Under this assumption, the implicit function theorem provides

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

and

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

$$\int [(G_s-F_s)(m_s(\theta_0)+a_0)-(G_s-F_s)(m_s(\theta_0)-a_0)]P^S(\mathrm{d}s).$$

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 $\ddot{m}_s(\theta)=(\partial/\partial\theta)\dot{m}_s(\theta)^\top$. Now, we can calculate the derivatives of $W_2(\epsilon,\theta)^\top$:

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

and

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

where

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

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

THEOREM 5.4 *Under the assumption* (23), *the* (*log*)*normal regression MTLF* $\tilde{\theta}_M$ *of Definition* 3.6 *satisfies*

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

Using here $G_{s_*}(y) = 1\!\!\mathrm{I}_{[y_*,\infty)}(y)$ if $Q^{Y|S=s_*} = \delta_{y_*}$ and $G_s(y) = F_s(y)$ for $s \neq s_*$, assumption (23) 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 5.5 *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 $\tilde{\theta}_M$ at $z_* = (y_*, s_*)$ is given by*

$$\begin{aligned}\mathrm{IF}(\tilde{\theta}_M, P, z_*) = -C(P)^{-1}\Bigg\{D(P) + 2\Bigg(&(y_* - m_{s_*}(\theta_0))\, 1\!\!\mathrm{I}_{(m_{s_*}(\theta_0)-a_0, m_{s_*}(\theta_0)+a_0]}(y_*)\\ &- a_0[F_{s_*}(m_{s_*}(\theta_0) + a_0) + F_{s_*}(m_{s_*}(\theta_0) - a_0)]\\ &+ \int_{m_{s_*}(\theta_0)-a_0}^{m_{s_*}(\theta_0)+a_0} F_{s_*}(y)\,\mathrm{d}y\Bigg)\dot{m}_{s_*}(\theta_0)^\top P^S(\{s_*\})\Bigg\}.\end{aligned}$$

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

Example 5.6 (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 5.4 and Corollary 5.5 concern also the OTLF $\tilde{\theta}_O$ 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

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

implying

$$a[F_s(m_s(\theta_0) + a) + F_s(m_s(\theta_0) - a)] - \mathcal{F}_s(m_s(\theta_0) + a) + \mathcal{F}_s(m_s(\theta_0) - a) = 0 \tag{24}$$

for any $a > 0$. The equality (24) provides at once $W_2(0, \theta_0) = 0$ for any $a_0 := a(0, \theta_0)$ given by

$$W_1(0, \theta_0, a_0) = 2\int F_s(m_s(\theta_0) + a_0) P^S(\mathrm{d}s) - 2 + \alpha = 0. \tag{25}$$

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

$$C(P) = 2\int \{a_0 2f_s(m_s(\theta_0) + a_0) - [2F_s(m_s(\theta_0) + a_0) - 1]\}\dot{m}_s(\theta_0)\dot{m}_s(\theta_0)^\top P^S(\mathrm{d}s),$$

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

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

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 (25) is equivalent to

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

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

$$C(P) = 2\{2a_0 f_*(a_0) - (1 - \alpha)\}\int \dot{m}_s(\theta_0)\dot{m}_s(\theta_0)^\top P^S(\mathrm{d}s)$$

so that the influence function is given by

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

For the special case of $m_s(\theta) = x(s)^\top \theta$, this influence function appears in the expansion which Bednarski et al.,[18,p.212] derived for the least trimmed squares estimator for linear regression. The matrix $\int \mathrm{IF}(\tilde{\theta}_M, P, z)\mathrm{IF}(\tilde{\theta}_M, P, z)^\top P^{Y|S} \otimes P^S(\mathrm{d}z)$ is also the asymptotic covariance matrix of the least trimmed squares estimator which Víšek [24] derived for linear regression and Čížek [25] for nonlinear regression.

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 TLE, one, called the OTLF, which corresponds to the OTLE, and a modified version, called MTLF, used by Bednarski and Clarke [16,17] and by Bednarski et al.[18] 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. For empirical distributions, the OTLF is always an MTLF but a simple example demonstrated that the MTLF may not be unique while the OTLF is unique. However, the application to a real data set indicates that for realistic situations the MTLF for finite samples is often unique. Therefore, 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. However, it can similarly be used also for other distributions. In particular a unified derivation can be used for some parts of the derivation. 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 MTLFs. A question is whether stronger differentiability notions like Hadamard differentiability can be shown. A problem in this context is that the exponential regression MTLF is not Fisher consistent at the exponential distribution so that a bias correction would be necessary.

Acknowledgements

We want to thank Professor R. Maurer and Guido Heeke from TU Dortmund University for their experiments and providing us the data. We are also grateful for the valuable comments and suggestions of the referees.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The work of the second author was supported by the German Research Foundation (DFG) under the Collaborative Research Center 'Statistical modeling of nonlinear dynamic processes' (SFB 823) in project B5 'Statistical methods for damage processes under cyclic load', and TÜBITAK supported the visit of the third author in Dortmund.

References

[1] Müller ChH, Neykov NM. Breakdown points of trimmed likelihood estimators and related estimators in generalized linear models. J Statist Plan Inference. 2003;116:503–519.
[2] Rousseeuw PJ. Least median of squares regression. J Amer Statist Assoc. 1984;79:871–880.
[3] Rousseeuw PJ, Leroy AM. Robust regression and outlier detection. New York: Wiley; 1987.
[4] Mili L, Coakley CW. Robust estimation in structured linear regression. Ann Statist. 1996;24:2297–2778.
[5] Müller ChH. Breakdown points for designed experiments. J Statist Plan Inference. 1995;45:413–427.
[6] Müller ChH, Robust planning and analysis of experiments, Lecture Notes in Statistics. Vol. 124. New York: Springer, 1997.
[7] Müller ChH, Upper and lower bounds for breakdown points. In: Becker C, Fried R, Kuhnt S, editors. Robustness and complex data structures. Festschrift in honour of Ursula Gather. Berlin: Springer; 2013. p. 67–84.
[8] Vandev DL. A note on the breakdown point of the least median of squares and least trimmed squares estimators. Statist Probab Lett. 1993;16:117–119.
[9] Vandev DL, Neykov NM, About regression estimators with high breakdown point. Statistics. 1998;32:111–129.
[10] Ahmed ES, Volodin AI, Hussein AA. Robust weighted likelihood estimation of exponential parameters. IEEE Trans Reliab. 2005;54:389–395.
[11] Hampel FR. The influence curve and its role in robust estimation. J Amer Statist Assoc. 1974;69:383–393.
[12] Hampel FR, Ronchetti EM, Rousseeuw PJ, Stahel WA. Robust statistics – the approach based on influence functions. New York: Wiley; 1986.
[13] Rieder H. Robust asymptotic statistics. New York: Springer; 1994.
[14] Staudte RG, Sheather SJ. Robust estimation and testing. New York: Wiley; 1990.
[15] Boudt K, Caliskan D, Croux C. Robust explicit estimators of Weibull parameters. Metrika. 2011;73:187–209.
[16] Bednarski T, Clarke BR. Trimmed likelihood estimation of location and scale of the normal distribution. Aust J Statist. 1993;35:141–153.
[17] Bednarski T, Clarke BR. Asymptotics for an adaptive trimmed likelihood location estimator. Statistics. 2002;36: 1–8.
[18] Bednarski T, Clarke BR, Schubert D. Adaptive trimmed likelihood estimation in regression. Discuss Math Probab Stat. 2010;30:203–219.
[19] Müller ChH, D-optimal designs for lifetime experiments with exponential distribution and censoring. In: Ucinski D, Atkinson AC, Patan M, editors. mODa 10 – advances in model-oriented design and analysis. Heidelberg: Physica-Verlag; 2013. p. 179–186.
[20] Marsden JE. Elementary classical analysis. New York: W. H. Freeman and Company; 1974.
[21] Neykov NM, Müller ChH, Breakdown point and computation of trimmed likelihood estimators in generalized linear models. In: Dutter R, Filzmoser P, Gather U, Rousseeuw PJ, editors. Developments in robust statistics. Heidelberg: Physica-Verlag; 2003. p. 277–286.
[22] Rousseeuw PJ, Van Driessen K. Computing LTS regression for large data sets. Data Min Knowl Discov. 2006;12:29–45.
[23] Keppler J, Maximum-Likelihood-Schätzung für zensierte und getrimmte Daten [Diploma thesis]. University of Kassel; 2012. Available from: http://www.statistik.tu-dortmund.de/1960.html.
[24] Víšek JÁ. The least trimmed squares - random carriers. Bull Czech Econom Soc. 1999;10:1–30.
[25] Čížek P. Least trimmed squares in nonlinear regression under dependence. J Statist Plan Inference. 2005;136:3967–3988.