Quantile estimation in a generalized asymmetric distributional setting

Irène Gijbels, Rezaul Karim and Anneleen Verhasselt

Abstract Allowing for symmetry in distributions is often a necessity in statistical modelling. This paper studies a broad family of asymmetric densities, which in a regression setting shares basic philosophy with generalized (non)linear models. The main focus however for the family of densities studied here is quantile estimation instead of mean estimation. In a similar fashion a broad family of conditional densities is considered in the regression setting. We discuss estimation of the parameters in the unconditional case, and establish an asymptotic normality result, with explicit expression for the asymptotic variance-covariance matrix. In the regression setting, we allow for flexible modelling and estimate nonparametrically the location and scale functions, leading to semiparametric estimation of conditional quantiles, again in the unifying framework of the considered broad family. The practical use of the proposed methods is illustrated in a real data application on locomotor performance in small and large terrestrial mammals.

1 Introduction

Several statistical tools are developed based on the assumption that the data are symmetric about the mean. Among standard symmetric distributions are the normal and Student-t distributions. In case the data cannot be represented appropriately via

Irène Gijbels
KU Leuven, Department of Mathematics and Leuven Statistics Research Center (LStat), Celestijnenlaan 200B, B-3001 Leuven (Heverlee), Belgium, e-mail: `irene.gijbels@kuleuven.be`

Rezaul Karim
KU Leuven and UHasselt, e-mail: `mdrezaul.karim@kuleuven.be` or e-mail: `mdrezaul.karim@uhasselt.be`

Anneleen Verhasselt
Interuniversity Institute for Biostatistics and statistical Bioinformatics, Universiteit Hasselt (UHasselt), Martelarenlaan 42, 3500 Hasselt, Belgium, e-mail: `anneleen.verhasselt@uhasselt.be`

symmetric distributions, asymmetric distributions are needed. Classical asymmetric distributions include exponential, gamma, log-normal, log-Laplace, chi-squared, and Fisher distributions, which all have support on a real halfline. Many non-classical asymmetric distributions such as skew-normal and skew Student-t distributions, with support the whole real line, have been proposed in the literature. When it comes to parameter estimation (or in general statistical inference and asymptotic properties) for a given distribution, a convenient class of distributions is the exponential family.

In a conditional setting, when covariates come into play, the interest goes to conditional distributions for given value of the covariate (vector). Also in such a setting the exponential family framework (now in the context of conditional densities) is convenient. A primary interest is often in the conditional mean function, which in the simplest case is assumed to be a linear function of the covariate, possibly only after a transformation (through a link function) leading to the well-known framework of generalized linear models. See for example McCullagh and Nelder (1998).

A (conditional) mean of a (conditional) distribution function is only one of its characteristics. In contrast, a full characterization of the (conditional) distribution is provided by the (conditional) quantile function. Of particular interests herein are the 0.50th-quantile (the median) and extreme quantiles (in case of heavy-tailed distributions). The focus in this paper is in (unconditional as well as conditional) quantile estimation, allowing for possible asymmetry of (unconditional or conditional) distributions.

In the *unconditional* setting there are several approaches for constructing univariate asymmetric distributions, among which these introduced by Azzalini (1985, 1986), Fernández and Steel (1998) and Arellano-Valle and Genton (2005). See also Jones (2015). A starting point for our study is the quantile-based family of asymmetric densities (shortened hereafter as QBA) studied by Gijbels et al. (2018). For a real-valued random variable Y, and index-parameter $\alpha \in (0,1)$, the density of Y, denoted by $f_\alpha(\cdot;\mu,\phi)$, is given by

$$f_\alpha(y;\mu,\phi) = \frac{2\alpha(1-\alpha)}{\phi}\begin{cases} f\left((1-\alpha)\left(\frac{\mu-y}{\phi}\right)\right) & \text{if } y \le \mu \\ f\left(\alpha\left(\frac{y-\mu}{\phi}\right)\right) & \text{if } y > \mu, \end{cases} \tag{1}$$

where f is a given symmetric around 0 density, with support the whole real line, and unimodal, called the reference symmetric density, $\mu \in \mathbb{R}$ is a location parameter and $\phi \in \mathbb{R}_0^+$ is a scale parameter. When the reference symmetric density f is a member of a location-scale family of densities, then the same holds for the family of densities $f_\alpha(\cdot;\mu,\phi)$ in (1). The density $f_\alpha(\cdot;\mu,\phi)$ is a symmetric density if and only if $\alpha = 0.5$, and is a left-skewed (respectively right-skewed) density if α is larger (respectively smaller) than 0.5. This family of asymmetric densities provides a very convenient framework. Firstly, explicit expressions for distribution and quantile functions, characteristic function, skewness, kurtosis, moments (including mean and variance) have been provided in terms of the associated quantities for the reference symmetric density f. In this family the parameter μ equals the αth-quantile of the distribution (i.e. $F_\alpha^{-1}(\alpha;\mu,\phi) = \mu$) which explains the name ‘quantile-based’ family of asymmetric distributions. Secondly, maximum likelihood as well as moment estimators (including for the index-parameter α) have been studied for the general class (1) with explicit expressions for asymptotic variance-covariance matrices in the asymptotic normality result for the estimators. The Fisher information matrix of the maximum likelihood estimators also revealed that the maximum likelihood estimators of μ and ϕ are asymptotically independent; i.e. the parameters μ and ϕ are orthogonal (see Cox and Reid (1987)). Thirdly, the family of densities (1) includes some well-known members studied in the literature. The asymmetric power family of densities proposed by Komunjer (2007) is a subclass of (1) obtained by taking $f(s) = 0.5\left[\Gamma\left(1+\frac{1}{p}\right)\right]^{-1}\exp(-|s|^p)$, for $s \in \mathbb{R}$, with $p > 0$. An element of this subclass (taking $p = 1$) is the asymmetric Laplace density (or double exponential), often appearing in the context of quantile estimation (Kotz et al.; 2001; Koenker; 2005). Taking f in (1) respectively a symmetric normal, Student-t or logistic density leads to the quantile-based asymmetric normal, asymmetric Student-t and asymmetric logistic densities proposed and studied in Gijbels et al. (2018).

In a *conditional* setting, a covariate – say X – comes into play, and the interest is in the conditional distribution of Y given X. A semiparametric context is obtained by allowing the location and scale parameters μ and ϕ to vary with the given covariate value $X = x$, i.e. to consider functions $\mu(x)$ and $\phi(x)$. Keeping the index-parameter α constant, this leads to the *conditional* density

$$f_{Y|X,\alpha}(y;\mu(x),\phi(x)\mid X=x) = \frac{2\alpha(1-\alpha)}{\phi(x)}\begin{cases} f\left((1-\alpha)\left(\frac{\mu(x)-y}{\phi(x)}\right)\right) & \text{if } y \le \mu(x) \\ f\left(\alpha\left(\frac{y-\mu(x)}{\phi(x)}\right)\right) & \text{if } y > \mu(x). \end{cases} \tag{2}$$

In a flexible setting the functions $\mu(\cdot)$ and $\phi(\cdot)$ are unknown (nonparametric elements), and parametric elements are the parameters of the reference density f (e.g. the degrees of freedom in case f is a Student-t density) and possibly the index-parameter α. Estimation of the conditional quantile function, i.e. $q_\beta(x) = F^{-1}_{Y|X,\alpha}(\beta;\mu(x),\phi(x)|x)$, with $\beta \in (0,1)$, in this semiparametric setting has been studied in Gijbels et al. (2019). By definition conditional quantile curves do not cross, i.e. for given $0 < \beta_1 \le \beta_2 < 1$ it holds that $q_{\beta_1}(x) \le q_{\beta_2}(x)$ for all x. A particular advantage of the framework (2) is that estimated conditional quantile curves are quaranteed not to cross each other.

One of the requirements underlying the family of densities in (1) is that the random variable Y is continuous and has support the whole real line. Obviously this is not always the case. For example, if Y is a lifetime variable it takes only nonnegative values. Also variables taking values in a finite interval are of interest, think of data that are proportions (or percentages) within 0 and 1, school grades between 0 and 100 points, visual analogue scales between 0 and 10 cm, etc. Bounded outcomes often have a non-standard distribution which may expose a variety of shapes including unimodal, U-shape, and J-shape. And in particular it is important to allow for asymmetry.

The overall aim of this paper is to extend the family of densities in (1), in the unconditional setting, and the family of conditional densities in (2), in the conditional setting, to allow for any type of continuous variable Y, with support possibly different from the whole real line. In Section 2 we provide a generalization of (1) to obtain a generalized quantile-based asymmetric (GQBA) family of densities. For this a link function is introduced, similar in spirit as in the case of generalized linear models. Recall that in the latter models however the focus is on estimating the mean. In contrast, in this paper the focus is always on estimating quantiles. The generalized quantile-based asymmetric family of densities constitutes a broad class containing as special members other families of densities studied in the literature, such as the tick-exponential family of densities and the asymmetric power family of densities. An aspect of Section 2 is thus also literature review. In a similar fashion, in a conditional setting, we extend the family of conditional densities (2). In the unconditional setting we study maximum likelihood estimation of the parameters in the generalized quantile-based asymmetric family of densities. See Section 3. In the conditional setting, when we focus on the semiparametric situation, we use local maximum likelihood techniques to estimate the unknown functions, and subsequently the conditional quantile function. See Section 4. In Section 5 we illustrate the practical use of the developed methods, both in conditional and unconditional settings. Some further discussions are provided in Section 6. Proofs of the theoretical results are deferred to the Appendix.

2 Generalized quantile-based asymmetric family

Consider Y a real-valued random variable, with support S possibly different from the whole real line, i.e. $S \subseteq \mathbb{R}$. We consider then a function $g : S \mapsto \mathbb{R}$, which is differentiable such that $g'(\cdot) > 0$; and hence g is invertible. In other words the function g is assumed to be a strictly increasing function.

Consider $Z = g(Y)$, which is supported on the whole real line, and assume that Z has a density of the form (1), with location and scale parameters (μ, ϕ). Denoting $\eta = g^{-1}(\mu)$, the density of $Y = g^{-1}(Z)$ is then given by

$$f_\alpha^g(y;\eta,\phi) = \frac{2\alpha(1-\alpha)g'(y)}{\phi} \begin{cases} f\left((1-\alpha)\left(\frac{g(\eta)-g(y)}{\phi}\right)\right) & \text{if } y \le \eta \\ f\left(\alpha\left(\frac{g(y)-g(\eta)}{\phi}\right)\right) & \text{if } y > \eta. \end{cases} \tag{3}$$

The family of densities (3) includes the family (1), where the latter is obtained by taking the identity function $g(y) = y$. We refer to (3) as the *generalized quantile-based asymmetric (GQBA) family of densities*. The function g is called a link function, and is considered known in this paper. The density in (3) depends on the index-parameter α, and on two crucial elements:

- the reference symmetric density f, and
- the monotone strictly increasing 'link' function g.

Expressions for the cumulative distribution and the quantile function of Y are presented in Theorem 1, the proof of which is provided in the Appendix. We write the following assumption.

Assumption (**A**): $g : S \mapsto \mathbb{R}$ is a differentiable function with $g'(\cdot) > 0$.

The cumulative distribution function and the quantile function associated with the reference symmetric density f are denoted by respectively F and F^{-1}.

Theorem 1 *Assume that Y has density (3), where g satisfies Assumption (**A**). The cumulative distribution function of Y equals*

$$F_{\alpha}^{g}(y;\eta,\phi) = \begin{cases} 2\alpha F\left((1-\alpha)\frac{g(y)-g(\eta)}{\phi}\right) & \text{if } y \le \eta \\ 2\alpha - 1 + 2(1-\alpha)F\left(\alpha\frac{g(y)-g(\eta)}{\phi}\right) & \text{if } y > \eta, \end{cases} \tag{4}$$

and for any $\beta \in (0,1)$, the βth-quantile of Y equals

$$\left\{F_{\alpha}^{g}\right\}^{-1}(\beta;\eta,\phi) = \begin{cases} g^{-1}\left(g(\eta) + \frac{\phi}{1-\alpha}F^{-1}\left(\frac{\beta}{2\alpha}\right)\right) & \text{if } \beta \le \alpha \\ g^{-1}\left(g(\eta) + \frac{\phi}{\alpha}F^{-1}\left(\frac{1+\beta-2\alpha}{2(1-\alpha)}\right)\right) & \text{if } \beta > \alpha, \end{cases} \tag{5}$$

with in particular $\left\{F_{\alpha}^{g}\right\}^{-1}(\alpha;\eta,\phi) = \eta$.

Remark 1

1. It suffices to assume that the function g is differentiable and strictly monotone (increasing or decreasing). For simplicity of presentation we focus on the case that g is a strictly increasing function.
2. Using (4) we find that

$$\alpha\,\Pr(Y > \eta) = (1-\alpha)\Pr(Y \le \eta),$$

and hence α controls the amount of mass allocated in the density to the left and right of the mode η.

In the next subsections we discuss various special subclasses and examples of the general family (3).

2.1 Generalized tick-exponential family

We take as the reference symmetric density f in (3), the exponential power type density (or simply power density)

$$f(s) = f_p(s) = \frac{1}{2\Gamma(1+\frac{1}{p})}\exp\left(-|s|^p\right) \qquad -\infty < s < +\infty, \tag{6}$$

where $p \in (0, \infty)$ is an index number. The distribution in (6) is also known as the generalized normal distribution or the generalized error distribution. The density in (6) has a heavier (respectively lighter) tail than a normal density when $p < 2$ (respectively $p > 2$). A lot of research has been done based on this family (see for example, Mineo and Ruggieri (2005)). The package `normalp` in the R software contains software routines for this density. The density in (6) is a Laplace, normal and uniform density respectively for $p = 1$, $p = 2$ and $p \to \infty$. With f as in (6) we obtain from (3) the density

$$
\begin{aligned}
f^g_{\alpha,p}(y;\eta,\phi) &= \frac{\alpha(1-\alpha)g'(y)}{\phi\Gamma(1+\frac{1}{p})} \begin{cases} \exp\left(-(1-\alpha)^p\left[\frac{g(\eta)-g(y)}{\phi}\right]^p\right) & \text{if} \quad y \leq \eta \\ \exp\left(-\alpha^p\left[\frac{g(y)-g(\eta)}{\phi}\right]^p\right) & \text{if} \quad y > \eta \end{cases} \\
&= \frac{\alpha(1-\alpha)g'(y)}{\phi\Gamma(1+\frac{1}{p})} \exp\left(-\frac{\rho_{\alpha,p}(g(y)-g(\eta))}{\phi^p}\right), \qquad (7)
\end{aligned}
$$

where

$$
\rho_{\alpha,p}(u) = |u|^p\left[(1-\alpha)^p\mathbb{I}(u \leq 0) + \alpha^p\mathbb{I}(u > 0)\right], \qquad (8)
$$

with $\mathbb{I}(A)$ the indicator function on $A \subseteq \mathbb{R}$, i.e. $\mathbb{I}(A) = 1$ (respectively 0) if A is true (respectively false).

The family of densities in (7) might be called the generalized quantile-based exponential power family, and is not available in the literature. It could also be called generalized tick-exponential family since it is a generalized version of the existing tick-exponential family which we discuss and review in Section 2.2. The loss function $\rho_{\alpha,p}(u)$ may be termed *power-tick loss function*. A well-known loss function used for nonparametric quantile estimation is the tick loss function, defined as $\rho_\alpha(u) = u(\alpha - \mathbb{I}(u < 0))$ which is a special case of (8) for $p = 1$. See for example Koenker and Bassett Jr (1978) and Koenker (2005) for background information on the tick loss function. A graph of the power-tick loss function $\rho_{\alpha,p}(u)$ for various values of p and for $\alpha \in \{0.25, 0.50\}$ is displayed in Figure 1.

Using the log-link function (i.e. $g(y) = \ln(y)$) and $p = 2$ we obtain from (7) the density

$$
f^g_{\alpha,2}(y;\eta,\phi) = \frac{2\alpha(1-\alpha)}{y\phi\sqrt{\pi}} \begin{cases} \exp\left(-(1-\alpha)^2\left[\frac{\ln(\eta)-\ln(y)}{\phi}\right]^2\right) & \text{if} \quad y \leq \eta \\ \exp\left(-\alpha^2\left[\frac{\ln(y)-\ln(\eta)}{\phi}\right]^2\right) & \text{if} \quad y > \eta, \end{cases} \qquad (9)
$$

which is the density of a log-asymmetric normal distribution, which we denote by Log-AND(η, ϕ, α). As a special case of this, for $\alpha = 0.5$, one obtains the density of a log-normal distribution, denoted by Log-ND(μ, τ^2), with $\mu = \ln(\eta)$ and $\tau^2 = 2\phi^2$. A log-normal density is widely used in applications in financial economics (see, for example, Crow and Shimizu; 1988).

With the identity link function (i.e. $g(y) = y$) the density in (7) results into a density that can be termed an asymmetric exponential power density (AEPD), denoted by AEPD(η, ϕ, α), and given by

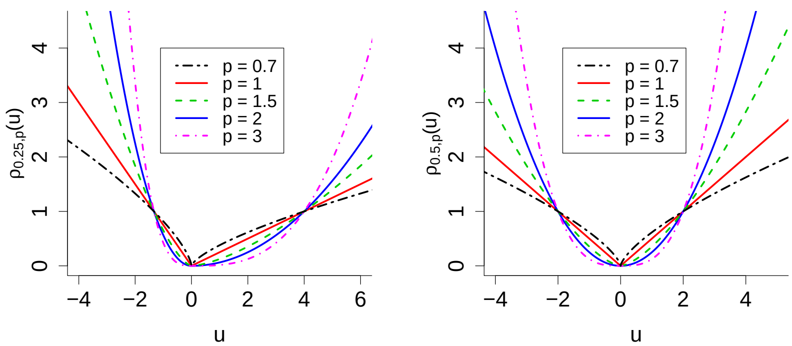

Fig. 1: Power-tick loss function $\rho_{\alpha,p}(u)$ with $\alpha = 0.25$ (left panel) and $\alpha = 0.50$ (right panel) for various values of p.

$$f_{\alpha,p}(y;\eta,\phi) = \frac{\alpha(1-\alpha)}{\phi\Gamma(1+\frac{1}{p})} \begin{cases} \exp\left(-(1-\alpha)^p\left[\frac{\eta-y}{\phi}\right]^p\right) & \text{if } y \le \eta \\ \exp\left(-\alpha^p\left[\frac{y-\eta}{\phi}\right]^p\right) & \text{if } y > \eta. \end{cases} \tag{10}$$

Many examples of AEPD(η,ϕ,α) given in (10) are available in the econometrics literature. For example, for $\eta = 0$ and $\phi = \sqrt[p]{\frac{\alpha^p+(1-\alpha)^p}{2}}$, the density in (10) can be written as

$$f_{\alpha,p}(y) = \begin{cases} \frac{\delta_{\alpha,p}^{\frac{1}{p}}}{\Gamma(1+\frac{1}{p})}\exp\left(-\frac{\delta_{\alpha,p}}{\alpha^p}|y|^p\right) & \text{if } y \le 0 \\ \frac{\delta_{\alpha,p}^{\frac{1}{p}}}{\Gamma(1+\frac{1}{p})}\exp\left(-\frac{\delta_{\alpha,p}}{(1-\alpha)^p}|y|^p\right) & \text{if } y > 0, \end{cases}$$

where $\delta_{\alpha,p} = \frac{2\alpha^p(1-\alpha)^p}{\alpha^p+(1-\alpha)^p}$. This density, some of its probabilistic properties and maximum likelihood estimation of the parameters was studied in Komunjer (2007), where also an application to risk management theory was presented.

Two other examples of AEPD(η,ϕ,α) in (10) are an asymmetric normal (for $p = 2$) and an asymmetric Laplace distribution (for $p = 1$) which were introduced and/or studied in Gijbels et al. (2018). A graph of the density (10) for different values of the index-parameter α and of p are presented in Figure 2. Note from Figure 2 that the density is symmetric for $\alpha = 0.5$ for any value of p, whereas for α larger (respectively smaller) than 0.5 it is a left-skewed (respectively right-skewed) density.

The cumulative distribution function and the quantile function of an exponential power density (6) are given by, respectively,

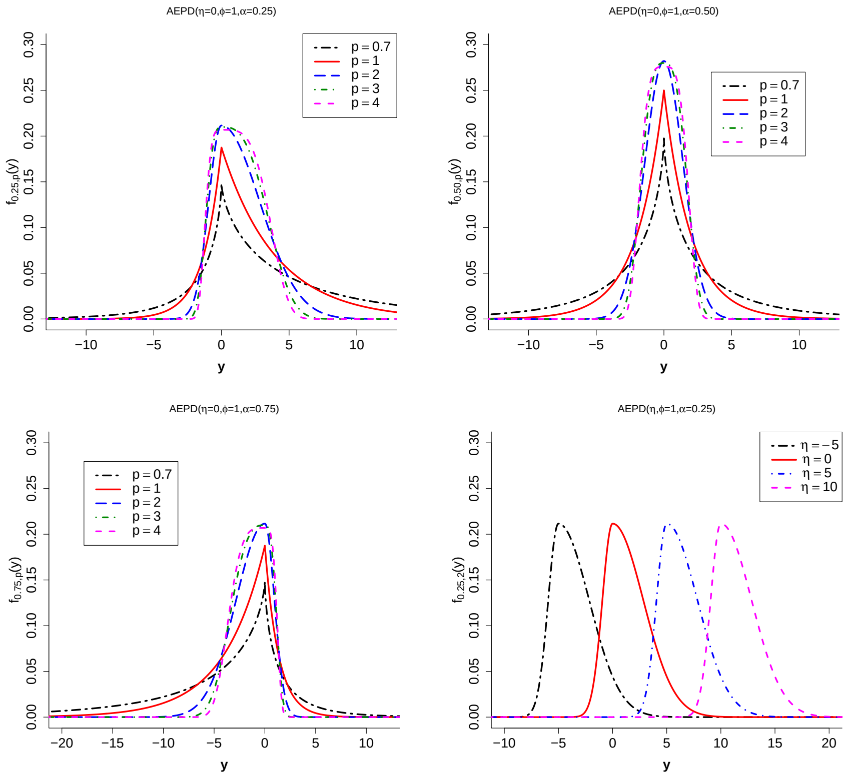

Fig. 2: Asymmetric exponential power densities with $\alpha \in \{0.25, 0.50, 0.75\}$, $\eta = 0$ and $\phi = 1$ for various values of p; and (bottom right) with $\alpha = 0.25$, $\phi = 1$, $p = 2$ and for various values of η.

$$F_p(s) = \frac{1}{2} + \operatorname{sgn}(s)\frac{\gamma\left(\frac{1}{p}, |s|^p\right)}{2\Gamma(\frac{1}{p})} \qquad \text{for } s \in \mathbb{R}$$

$$F_p^{-1}(\beta) = \operatorname{sgn}\left(\beta - \frac{1}{2}\right)\left[\gamma^{-1}\left(\frac{1}{p}, \Gamma\left(\frac{1}{p}\right)\operatorname{sgn}\left(\beta - \frac{1}{2}\right)(2\beta - 1)\right)\right]^{\frac{1}{p}} \qquad \text{for } 0 < \beta < 1,$$

where $\gamma(s, x)$ is the lower incomplete gamma function, i.e. $\gamma(s, x) = \int_0^x t^{s-1}e^{-t}\,dt$, for $x \geq 0$ such that $\gamma(s, 0) = 0$; and where $\gamma^{-1}(s, y)$ is the inverse of the incomplete gamma function, i,e., $x = \gamma^{-1}(s, y)$ is equivalent to $y = \gamma(s, x)$. Substituting these expressions in Equations (4) and (5) in Theorem 1 with $F = F_p$ and $F^{-1} = F_p^{-1}$, one easily deduces the following properties for the generalized tick-exponential family of densities (7).

Corollary 1 *Assume that Y has density (7), where g satisfies Assumption (**A**). For any $p \in (0,\infty)$, the cumulative distribution function $F^g_{\alpha,p}(\cdot;\eta,\phi)$ of Y then equals*

$$F^g_{\alpha,p}(y;\eta,\phi) = \begin{cases} \alpha - \frac{\alpha}{\Gamma(\frac{1}{p})}\gamma\left(\frac{1}{p},(1-\alpha)^p\left(\frac{g(\eta)-g(y)}{\phi}\right)^p\right) & \text{if } y \le \eta \\ \alpha + \frac{1-\alpha}{\Gamma(\frac{1}{p})}\gamma\left(\frac{1}{p},\alpha^p\left(\frac{g(y)-g(\eta)}{\phi}\right)^p\right) & \text{if } y > \eta; \end{cases} \tag{11}$$

and for any $\beta \in (0,1)$, the βth-quantile of Y, denoted by $\left\{F^g_{\alpha,p}\right\}^{-1}(\beta;\eta,\phi)$, is

$$\left\{F^g_{\alpha,p}\right\}^{-1}(\beta;\eta,\phi) = \begin{cases} g^{-1}\left(g(\eta) - \frac{\phi}{1-\alpha}\left[\gamma^{-1}\left(\frac{1}{p},\Gamma(\frac{1}{p})\frac{(\alpha-\beta)}{\alpha}\right)\right]^{\frac{1}{p}}\right) & \text{if } \beta \le \alpha \\ g^{-1}\left(g(\eta) + \frac{\phi}{\alpha}\left[\gamma^{-1}\left(\frac{1}{p},\Gamma(\frac{1}{p})\frac{(\beta-\alpha)}{1-\alpha}\right)\right]^{\frac{1}{p}}\right) & \text{if } \beta > \alpha. \end{cases} \tag{12}$$

For the identity link, i.e. the exponential power densities, the cumulative distribution function (11) and the quantile function (12) are provided, respectively in Figure 3 and Figure 4 for different values of p and α. The right-skewness of the distributions is clearly visible for the case $\alpha = 0.25$.

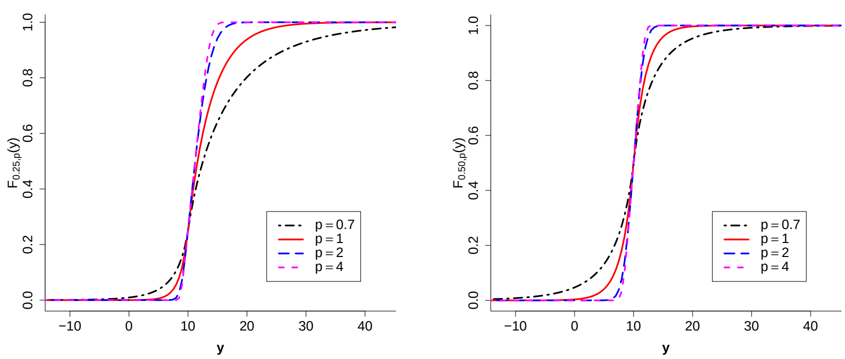

Fig. 3: Cumulative distribution function (11) with identity link, $\eta = 0$ and $\phi = 1$ for different values of p, $\alpha = 0.25$ (left panel) and $\alpha = 0.5$ (right panel).

2.2 Tick-exponential family with scale parameter

A special case of the generalized tick-exponential family of densities (7) is obtained by taking $p = 1$. We then get the tick-exponential family for which the density is of the form

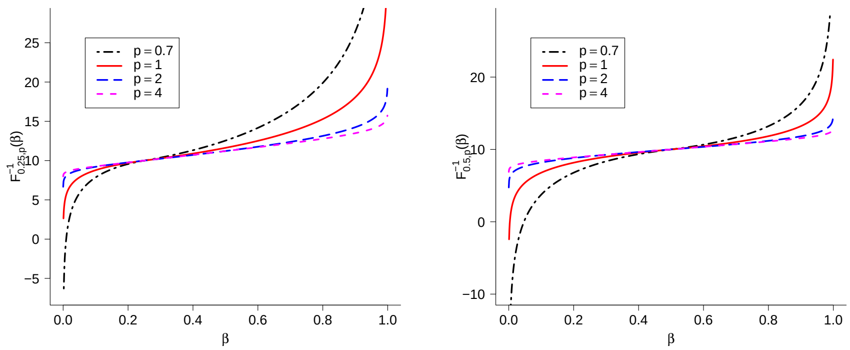

Fig. 4: Quantile function (12) with identity link, $\eta = 0$ and $\phi = 1$ for different values of p, $\alpha = 0.25$ (left panel) and $\alpha = 0.5$ (right panel).

$$f^g_{\alpha,1}(y;\eta,\phi) = \frac{\alpha(1-\alpha)g'(y)}{\phi}\begin{cases}\exp\left(-(1-\alpha)\left[\frac{g(\eta)-g(y)}{\phi}\right]\right) & \text{if} \quad y \le \eta \\ \exp\left(-\alpha\left[\frac{g(y)-g(\eta)}{\phi}\right]\right) & \text{if} \quad y > \eta.\end{cases} \tag{13}$$

A particular form of $f^g_{\alpha,1}(y;\eta,\phi)$ in (13) with $\alpha = 0.5$ has been studied by Jung (1996) in a context of the quasi-likelihood median regression. Komunjer (2005) provided a tick-exponential family in which a density takes on the form:

$$\tilde{f}_\alpha(y;\eta) = \begin{cases}\exp\left[-(1-\alpha)\left(g(\eta)-b(y)\right)\right] & \text{if} \quad y \le \eta \\ \exp\left[\alpha\left(g(\eta)-c(y)\right)\right] & \text{if} \quad y > \eta,\end{cases} \tag{14}$$

where the functions g, b and c are continuous functions and satisfy the following conditions which need to hold for all values of $y \in \mathbb{R}$:

i) $g'(y) > 0$,
ii) $\exp\{-(1-\alpha)[g(y)-b(y)]\} = \alpha(1-\alpha)g'(y)$,
iii) $\exp\{\alpha[g(y)-c(y)]\} = \alpha(1-\alpha)g'(y)$,
iv) $(1-\alpha)b(y) + \alpha c(y) = g(y)$.

Using these conditions, the density in (14) can be rewritten as

$$\tilde{f}_\alpha(y;\eta) = \alpha(1-\alpha)g'(y)\begin{cases}\exp\left[-(1-\alpha)\left(g(\eta)-g(y)\right)\right] & \text{if} \quad y \le \eta \\ \exp\left[\alpha\left(g(\eta)-g(y)\right)\right] & \text{if} \quad y > \eta,\end{cases} \tag{15}$$

which is a special case of (13) with scale parameter $\phi = 1$. The family of densities (15) is studied in for example Komunjer (2005), and a special case in Gourieroux et al. (1984).

From Corollary 1, taking $p = 1$, and using that $\gamma(1,x) = \int_0^x e^{-t}dt = 1 - e^{-x}$ and $\gamma^{-1}(1,y) = -\ln(1-y)$, the cumulative distribution function of Y with density (13) and the βth-quantile of Y ($\beta \in (0,1)$) are respectively

$$F^g_{\alpha,1}(y;\eta,\phi) = \begin{cases} \alpha \exp\left((1-\alpha)\left[\frac{g(y)-g(\eta)}{\phi}\right]\right) & \text{if} \quad y \le \eta \\ 1-(1-\alpha)\exp\left(-\alpha\left[\frac{g(y)-g(\eta)}{\phi}\right]\right) & \text{if} \quad y > \eta, \end{cases}$$

$$\left\{F^g_{\alpha,1}\right\}^{-1}(\beta;\eta,\phi) = \begin{cases} g^{-1}\left(g(\eta) + \frac{\phi}{1-\alpha}\ln\left(\frac{\beta}{\alpha}\right)\right) & \text{if } \beta \le \alpha \\ g^{-1}\left(g(\eta) - \frac{\phi}{\alpha}\ln\left(\frac{1-\beta}{1-\alpha}\right)\right) & \text{if } \beta > \alpha. \end{cases}$$

We now look into some specific link functions, different from the identity link. Komunjer (2005) reported that the link function $g(y) = \frac{1}{\alpha(1-\alpha)}\operatorname{sgn}(y)\ln[1+|y|^q]; q \in \mathbb{N}^+$ can be used for standard nonlinear quantile estimation. With this link function the density in (13) takes the form

$$f_{\alpha,1}(y;\eta,\phi) = \begin{cases} \frac{q(-y)^{q-1}}{\phi(1+(-y)^q)}\left(\frac{1+(-y)^q}{1+(-\eta)^q}\right)^{-\frac{1}{\phi\alpha}} & \text{if} \quad y \le \eta \\ \frac{qy^{q-1}}{\phi(1+y^q)}\left(\frac{1+\eta^q}{1+y^q}\right)^{\frac{1}{\phi(1-\alpha)}} & \text{if} \quad y > \eta. \end{cases} \tag{16}$$

A graph of this density, with $\eta = 0$, $\phi = 1$, for various values of α and q is depicted in Figure 5. For more details on density (16) and its use, see Komunjer (2005).

Gneiting (2011) considered generalized piecewise linear loss functions. Related to this is choosing a log-link function $g(y) = \ln(y)$ in (13), giving the density

$$f_{\alpha,1}(y;\eta,\phi) = \frac{\alpha(1-\alpha)}{\phi y}\begin{cases} \left(\frac{y}{\eta}\right)^{\frac{(1-\alpha)}{\phi}} & \text{if} \quad y \le \eta \\ \left(\frac{\eta}{y}\right)^{\frac{\alpha}{\phi}} & \text{if} \quad y > \eta. \end{cases} \tag{17}$$

This density is called the log-asymmetric Laplace distribution, denoted by Log-ALaD(η,ϕ,α). Reparametrization of the density in (17) using the parameters $\phi = (\gamma+\zeta)^{-1}$ and $\alpha = \gamma(\gamma+\zeta)^{-1}$, leads to the density

$$f_\gamma(y;\eta,\zeta) = \frac{1}{\eta}\frac{\gamma\zeta}{(\gamma+\zeta)}\begin{cases} \left(\frac{y}{\eta}\right)^{\zeta-1} & \text{if} \quad y \le \eta \\ \left(\frac{\eta}{y}\right)^{\gamma+1} & \text{if} \quad y > \eta, \end{cases} \tag{18}$$

which was proposed in (Kozubowski and Podgórski; 2003, eq. (3)). A plot of (17) is presented in Figure 6. A particular form of Log-ALaD(η,ϕ,α) in (17) with $\alpha = 0.5$ is called a log-Laplace distribution which was considered in, for example, Lindsey (2004).

Another interesting link function is one of a logit type $g(y) = \operatorname{logit}(F_0(y)) = \ln\left[\frac{F_0(y)}{1-F_0(y)}\right]$ where F_0 is any continuous distribution function, strictly increasing on $S \subseteq \mathbb{R}$. Note that such a $g : S \mapsto \mathbb{R}$ satisfies Assumption (**A**). As a first example consider F_0 the cumulative distribution function of a uniform random variable on a finite interval $[a,b] \subset \mathbb{R}$. Then $F_0(y) = \frac{y-a}{b-a}$ for $y \in [a,b]$, and is strictly increasing

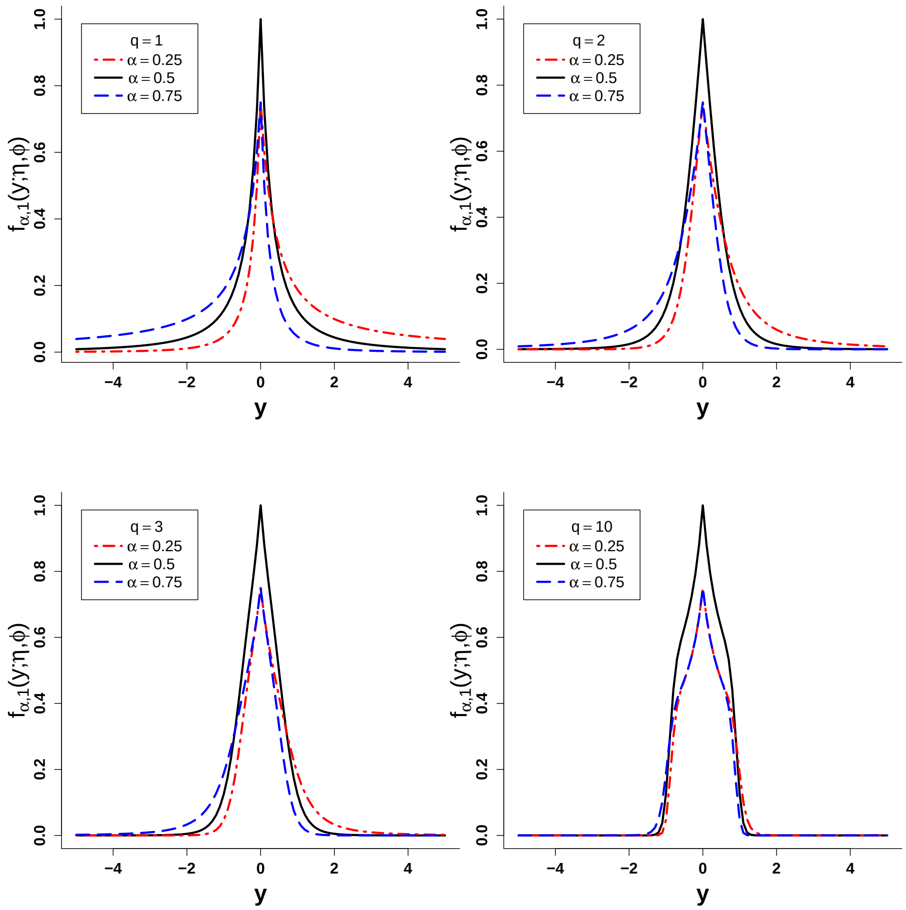

Fig. 5: Density $f_{\alpha,1}(y;\eta,\phi)$ in (16) with $\eta = 0$ and $\phi = 1$, and various values of q and α.

on $S = [a,b]$. The resulting link function is $g(y) = \ln\left(\frac{y-a}{b-y}\right)$, for $y \in [a,b]$; which is very appropriate for modeling a continuous random variable Y that takes on values in the bounded interval $[a,b]$. Several authors considered this link function in quantile estimation, including Bottai et al. (2010) and Columbu and Bottai (2016). Using the link function $g(y) = \ln\left(\frac{y-a}{b-y}\right)$; $y \in [a,b]$, in (13) leads to the density

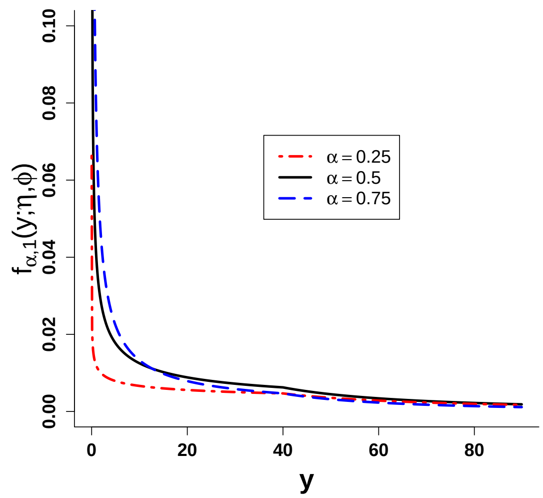

Fig. 6 Density log-ALaD(η, ϕ, α) in (17) for $\eta = 40$ and $\phi = 1$.

$$f_{\alpha,1}(y;\eta,\phi) = \frac{\alpha(1-\alpha)}{\phi}\frac{(b-a)}{(y-a)(b-y)}\begin{cases}\left(\frac{(b-\eta)(y-a)}{(\eta-a)(b-y))}\right)^{\frac{(1-\alpha)}{\phi}} & \text{if} \quad y \leq \eta \\ \left(\frac{(\eta-a)(b-y))}{(b-\eta)(y-a)}\right)^{\frac{\alpha}{\phi}} & \text{if} \quad y > \eta,\end{cases}$$

for which a graph is presented in the left panel of Figure 7.

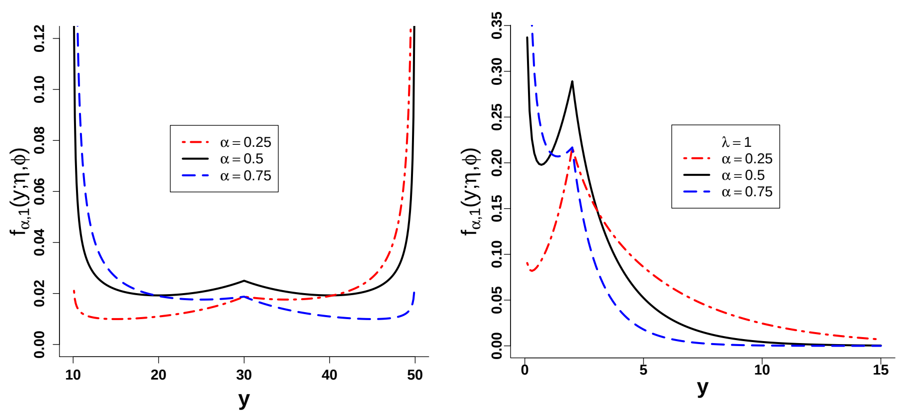

Fig. 7: Densities (13) with link function $g(y) = \ln\left[\frac{F_0(y)}{1-F_0(y)}\right]$, for $\eta = 30$ and $\phi = 1$. Left panel: F_0 from a U$[a, b]$; Right panel: F_0 from a standard exponential.

As a second example take F_0 equal to the cumulative distribution function of an exponential distribution with parameter $\lambda > 0$, i.e. $F_0(y) = 1 - \exp(-\lambda y)$, which is strictly increasing on $S = [0, +\infty)$. This leads to the link function $g(y) = \ln(e^{\lambda y} - 1)$, for $y \in [0, +\infty)$. With this link function we obtain from (13) with $\phi = 1$ the density

$$f_{\alpha,1}(y; \eta, 1) = \alpha(1-\alpha)\lambda e^{\lambda y} \begin{cases} \left(e^{\lambda\eta} - 1\right)^{-(1-\alpha)} \left(e^{\lambda y} - 1\right)^{-\alpha} & \text{if} \quad y \leq \eta \\ \left(e^{\lambda\eta} - 1\right)^{\alpha} \left(e^{\lambda y} - 1\right)^{-(1+\alpha)} & \text{if} \quad y > \eta, \end{cases}$$

for which a plot is depicted in the right panel of Figure 7. For real data applications an important issue is to find an appropriate density model. Goodness-of-fit tests can be used here, or model selection tools, among others. See also Section 5.

2.3 Generalized quantile-based asymmetric family: conditional setting

In the conditional setting we follow the same reasoning as we did when passing from (1) to (2). Keeping in mind the general family in (3) this leads to the general family of conditional densities

$$\begin{aligned} & f^g_{Y|X,\alpha}(y; \eta(x), \phi(x) \mid X = x) \\ & \quad = \frac{2\alpha(1-\alpha)g'(y)}{\phi(x)} \begin{cases} f\left((1-\alpha)\left(\frac{g(\eta(x)) - g(y)}{\phi(x)}\right)\right) & \text{if } y \leq \eta(x) \\ f\left(\alpha\left(\frac{g(y) - g(\eta(x))}{\phi(x)}\right)\right) & \text{if } y > \eta(x), \end{cases} \end{aligned} \tag{19}$$

with index-parameter $\alpha \in (0, 1)$, given link function g, and unknown location and scale functions $\eta(\cdot)$ and $\phi(\cdot)$.

All subclasses and special examples discussed in Sections 2.1 and 2.2 can also be considered in this conditional setting. Results similar to these in Theorem 1 straightforwardly hold. As an example the generalized tick-exponential family of *conditional* densities of Y given $X = x$ is

$$\begin{aligned} & f^g_{Y|X,\alpha,p}(y; \eta(x), \phi(x) \mid X = x) \\ & \quad = \frac{\alpha(1-\alpha)g'(y)}{\phi(x)\Gamma(1+\frac{1}{p})} \exp\left(-\frac{\rho_{\alpha,p}(g(y) - g(\eta(x)))}{(\phi(x))^p}\right). \end{aligned} \tag{20}$$

In the next sections we turn to statistical estimation in both settings: the unconditional one and the conditional one.

3 Unconditional setting: maximum likelihood estimation

Let $Y_1, \ldots, Y_n$ be an i.i.d. sample from Y with density from the GQBA family (3). The main objective is to estimate, based on this sample, the parameter vector $\boldsymbol{\theta}^g = (\eta, \phi, \alpha)^T$, where the dependence on the link function g comes in via $\eta = g^{-1}(\mu)$. The estimated parameters can then be substituted into the expression provided in Theorem 1 to get the estimated quantile function. Since the link function g is known, we can obtain the i.i.d. sample $Z_1, \ldots, Z_n$ (with $Z_i = g(Y_i)$, for $i = 1, \ldots, n$) from Z with density (1), and location and scale parameters (μ, ϕ). Maximum likelihood estimation of the parameter vector $\boldsymbol{\theta} = (\mu, \phi, \alpha)^T$ can then be carried out, as discussed in Gijbels et al. (2018). The maximum likelihood estimator (MLE) of $\boldsymbol{\theta}^g = (\eta, \phi, \alpha)^T$ is then easily obtained via the relationship $\eta = g^{-1}(\mu)$. The likelihood function for the parameter vector $\boldsymbol{\theta} = (\mu, \phi, \alpha)^T$ of the density in (1), based on the calculated sample $Z_1, \ldots, Z_n$, is

$$\begin{aligned} & L_n(\mu, \phi, \alpha) \\ &= \left[\frac{2\alpha(1-\alpha)}{\phi}\right]^n \prod_{i=1}^n \left[f\left((1-\alpha)\left(\frac{\mu - Z_i}{\phi}\right)\right)\right]^{\mathbb{I}(Z_i \leq \mu)} \times \left[f\left(\alpha\left(\frac{Z_i - \mu}{\phi}\right)\right)\right]^{\mathbb{I}(Z_i > \mu)}, \end{aligned}$$

leading to the log-likelihood function

$$\begin{aligned} \ell_n(\mu, \phi, \alpha) = n \ln[2\alpha(1-\alpha)] - n \ln(\phi) &+ \sum_{i=1}^n \mathbb{I}(Z_i \leq \mu) \ln\left[f\left((1-\alpha)\left(\frac{\mu - Z_i}{\phi}\right)\right)\right] \\ &+ \sum_{i=1}^n \mathbb{I}(Z_i > \mu) \ln\left[f\left(\alpha\left(\frac{Z_i - \mu}{\phi}\right)\right)\right]. \end{aligned}$$

The MLE of $\boldsymbol{\theta}$ is obtained as a solution to $\max_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \ell_n(\mu, \phi, \alpha)$; where $\boldsymbol{\Theta} = \mathbb{R} \times \mathbb{R}^+ \times (0, 1)$ is the parameter space of $\boldsymbol{\theta}$. A detailed study on how to solve this optimization problem, and statistical properties of the estimator can be found in Section 3 of Gijbels et al. (2018). Denote by $\widehat{\boldsymbol{\theta}}_n^{\text{MLE}} = (\widehat{\mu}_n^{\text{MLE}}, \widehat{\phi}_n^{\text{MLE}}, \widehat{\alpha}_n^{\text{MLE}})^T$ the maximum likelihood estimator of $\boldsymbol{\theta} = (\mu, \phi, \alpha)^T$. The plug-in estimator of $\boldsymbol{\theta}^g = (\eta, \phi, \alpha)^T$ is then $\widehat{\boldsymbol{\theta}_n^g}^{\text{MLE}} = (\widehat{\eta}_n^{\text{MLE}}, \widehat{\phi}_n^{\text{MLE}}, \widehat{\alpha}_n^{\text{MLE}})^T$, where $\widehat{\eta}_n^{\text{MLE}} = g^{-1}(\widehat{\mu}_n^{\text{MLE}})$.

The asymptotic properties of the maximum likelihood estimator $\widehat{\boldsymbol{\theta}_n^g}^{\text{MLE}}$ follow easily (by applying the Delta method) from the asymptotic normality result for $\widehat{\boldsymbol{\theta}}_n^{\text{MLE}}$ established in Theorem 3.4 of Gijbels et al. (2018). The following assumptions are needed.

Assumptions:

(B1) Let $\boldsymbol{\Theta}_R = [-\mu_u, \mu_u] \times [\phi_l, \phi_u] \times [\alpha_l, \alpha_u]$, where $0 < \mu_u < \infty$, $0 < \phi_l \leq \phi \leq \phi_u < \infty$, and $0 < \alpha_l \leq \alpha \leq \alpha_u < 1$, be a compact subset of $\boldsymbol{\Theta}$, and assume that $\boldsymbol{\theta} \in \mathring{\boldsymbol{\Theta}}_R$, with $\mathring{\boldsymbol{\Theta}}_R$ the interior of $\boldsymbol{\Theta}_R$.

(B2) $\int_0^\infty \left|\ln f(s)\right| f(s) ds < \infty$; where $f(s)$ is the reference symmetric density.

(B3) $\gamma_r = \int_0^\infty s^{r-1} \cdot \frac{(f'(s))^2}{f(s)} ds < \infty \quad \text{for } r = 1, 2, 3.$

(B4) $\lim_{s\to\infty} s f(s) = 0$ or $\int_0^\infty s f'(s) ds = -\frac{1}{2}$.

Theorem 2 *Suppose Assumptions (**B1**)—(**B4**)and Assumption (**A**) hold. Then the MLE* $\widehat{\boldsymbol{\theta}_n^g}^{MLE}$ *of* $\boldsymbol{\theta}^g = (\eta, \phi, \alpha)^T$ *centered according to* $\boldsymbol{\theta}^g$ *is asymptotically trivariate normally distributed with mean vector* $\mathbf{0}$ *and asymptotic variance-covariance matrix* $\mathcal{I}(\boldsymbol{\theta}^g)^{-1}$*:*

$$\sqrt{n}[\widehat{\boldsymbol{\theta}_n^g}^{MLE} - \boldsymbol{\theta}^g] \xrightarrow{d} \mathcal{N}_3\left(\mathbf{0}, \mathcal{I}(\boldsymbol{\theta}^g)^{-1}\right),$$

where

$$\mathcal{I}(\boldsymbol{\theta}^g)^{-1} = \begin{bmatrix} \frac{\gamma_3\phi^2}{2\alpha(1-\alpha)(\gamma_1\gamma_3-\gamma_2^2)}\left(\left(g^{-1}\right)'(\mu)\right)^2 & \frac{(1-2\alpha)\gamma_2\phi^2}{2\alpha(1-\alpha)(\gamma_1\gamma_3-\gamma_2^2)}\left(g^{-1}\right)'(\mu) & \frac{\gamma_2\phi}{2(\gamma_1\gamma_3-\gamma_2^2)}\left(g^{-1}\right)'(\mu) \\ \frac{(1-2\alpha)\gamma_2\phi^2}{2\alpha(1-\alpha)(\gamma_1\gamma_3-\gamma_2^2)}\left(g^{-1}\right)'(\mu) & [\mathcal{I}(\boldsymbol{\theta}^g)^{-1}]_{22} & \frac{(1-2\alpha)\gamma_1\phi}{2(\gamma_1\gamma_3-\gamma_2^2)} \\ \frac{\gamma_2\phi}{2(\gamma_1\gamma_3-\gamma_2^2)}\left(g^{-1}\right)'(\mu) & \frac{(1-2\alpha)\gamma_1\phi}{2(\gamma_1\gamma_3-\gamma_2^2)} & \frac{\alpha(1-\alpha)\gamma_1}{2(\gamma_1\gamma_3-\gamma_2^2)} \end{bmatrix},$$

with the $(2,2)$*th element of the matrix being*

$$[\mathcal{I}(\boldsymbol{\theta}^g)^{-1}]_{22} = \frac{\left(6\,\alpha^2\gamma_1\gamma_3 + 2\,\gamma_2{}^2\alpha^2 - 4\,\alpha^2\gamma_1 - 6\,\alpha\,\gamma_1\gamma_3 - 2\,\gamma_2{}^2\alpha + 4\,\alpha\,\gamma_1 + 2\,\gamma_1\gamma_3 - \gamma_1\right)\phi^2}{2\alpha(1-\alpha)(2\gamma_3 - 1)(\gamma_1\gamma_3 - \gamma_2^2)}.$$

If the index-parameter α *is known, then the asymptotic variance-covariance matrix of the MLE* $\widehat{\boldsymbol{\theta}_n^g}^{MLE} = (\widehat{\eta}_n^{MLE}, \widehat{\phi}_n^{MLE})^{sT}$ *of* $(\eta, \phi)^T$ *is*

$$\mathcal{I}(\boldsymbol{\theta}^g)^{-1} = \begin{bmatrix} \frac{\phi^2}{2\alpha(1-\alpha)\gamma_1}\left(\left(g^{-1}\right)'(\mu)\right)^2 & 0 \\ 0 & \frac{\phi^2}{2\gamma_3-1} \end{bmatrix}.$$

Remark 2

1. Note that, for an identity link function $g(y) = y$, we have $\left(g^{-1}\right)'(\mu) = 1$. In this special case, the asymptotic normality result in Theorem 2 reduces to that provided in Theorem 3.4 of Gijbels et al. (2018).
2. For the log-asymmetric Laplace density with parametrization as in (18), Kozubowski and Podgórski (2003) provide the Fisher information matrix

$$\mathcal{I}(\eta, \gamma, \zeta) = \begin{bmatrix} \frac{\gamma\zeta}{\eta^2} & -\frac{1}{\eta}\frac{\zeta}{\gamma+\zeta} & \frac{1}{\eta}\frac{\gamma}{\gamma+\zeta} \\ -\frac{1}{\eta}\frac{\zeta}{\gamma+\zeta} & \frac{1}{\gamma^2} - \frac{1}{(\gamma+\zeta)^2} & -\frac{1}{(\gamma+\zeta)^2} \\ \frac{1}{\eta}\frac{\gamma}{\gamma+\zeta} & -\frac{1}{(\gamma+\zeta)^2} & \frac{1}{\zeta^2} - \frac{1}{(\gamma+\zeta)^2} \end{bmatrix}. \tag{21}$$

Exploiting the connections between the reparametrizations in (18) and (17), we can show that the expression of the inverse of the Fisher information matrix in (21) coincides with $\mathcal{I}(\boldsymbol{\theta}^g)^{-1}$ if f is a Laplace ($i.e.$, $f(s) = \frac{1}{2}e^{-|s|}$) and g is a log-link function (i.e. $g(y) = \ln(y)$).

We next discuss estimation in case of a conditional density function taking the form (3) where the parameters are unknown functions of a covariate. We then study semiparametric conditional quantile curve estimation.

4 Conditional setting: local-likelihood estimation

We turn to the conditional setting, assuming that the conditional density of Y given the covariate value $X = x$ is as in (19). In this section we assume for simplicity that the index-parameter α is known. The main objective is then the estimation of the unknown location function $\eta(x)$ and scale function $\phi(x)$. From the conditional version of Theorem 1 we get that the βth-conditional quantile function of Y given $X = x$ (with $(0 < \beta < 1)$) is

$$\left\{F^g_{Y|X,\alpha}\right\}^{-1}(\beta;\eta(x),\phi(x)|x) = g^{-1}\left(g(\eta(x)) + \phi(x)\cdot C_\alpha(\beta)\right), \tag{22}$$

where

$$C_\alpha(\beta) = \frac{1}{1-\alpha}F^{-1}\left(\frac{\beta}{2\alpha}\right)\mathbb{I}(\beta < \alpha) + \frac{1}{\alpha}F^{-1}\left(\frac{1+\beta-2\alpha}{2(1-\alpha)}\right)\mathbb{I}(\beta \geq \alpha).$$

The quantity $C_\alpha(\beta)$ is a known constant and is a monotonic function of β. For estimating the βth-conditional quantile function of Y given $X = x$, we first obtain estimators for the functions $\eta(x)$ and $\phi(x)$, and then substitute these into expression (22). Given that structure and that g^{-1} is a monotonic function, as well as $C_\alpha(\beta)$ (looked upon as a function of β) it is granted that the estimator for $\left\{F^g_{Y|X,\alpha}\right\}^{-1}(\beta;\eta(x),\phi(x)|x)$ obtained as such satisfies the property of non-crossing estimated quantile functions.

Let $(Y_1, X_1), \ldots, (Y_n, X_n)$ be an i.i.d. sample from (Y, X) where the conditional density of Y given $X = x$ is as in (19). From this sample we form the sample $(Z_1, X_1), \ldots, (Z_n, X_n)$ from (Z, X) where the conditional density of $Z = g(Y)$ given $X = x$ is of the form (2). This conditional density depends on the unknown functions $\mu(x)$ and $\phi(x)$ which will be estimated using local maximum likelihood techniques, as proposed in Gijbels et al. (2019). We briefly discuss this method.

The semiparametric estimation procedure consists of approximating $\mu(x)$ and $\ln(\phi(x))$ locally by a polynomial function, i.e. performing a local polynomial fit. See for example Fan and Gijbels (1996) for background information on this smoothing technique. Hereby focus on $\ln(\phi(x))$ instead of on $\phi(x)$ is done to ensure that the estimator of $\phi(x) = \exp\{\ln(\phi(x))\}$ is always positive. For simplicity of presentation, denote $\theta_1(x) = \mu(x)$ and $\theta_2(x) = \ln(\phi(x))$. Suppose we want to estimate $\theta_r(x_0)$, for a given $x_0 \in \mathbb{R}$. For each observation X_i in a neighbourhood of x_0 we can approximate $\theta_r(X_i)$ by using a Taylor expansion with a polynomial of order $p_r \in \mathbb{N}$; $(r = \{1, 2\})$:

$$\begin{aligned}\theta_r(X_i) &\approx \theta_r(x_0) + \theta_r'(x_0)(X_i - x_0) + \cdots + \frac{\theta_r^{(p_r)}(x_0)}{p_r!}(X_i - x_0)^{p_r} \\ &\equiv \sum_{j=0}^{p_r} \theta_{rj}(X_i - x_0)^j = \mathbf{X}_{i,p_r}^T \boldsymbol{\theta}_r, \end{aligned} \tag{23}$$

where $\mathbf{X}_{i,p_r} = (1, (X_i - x_0), \cdots, (X_i - x_0)^{p_r})^T$, $\boldsymbol{\theta}_r = (\theta_{r0}, \cdots, \theta_{rp_r})^T$ with $\theta_{rv} = \frac{\theta_r^{(v)}(x_0)}{v!}$; $v = 0, 1, \ldots, p_r$.

Since the Taylor expansion in (23) is only valid for X_i close to x_0, this needs to be taken into account for the contribution of each datum (Z_i, X_i) to the log-likelihood

$$\ell(\theta_1(X_i), \theta_2(X_i); Z_i) = \ln f_{Z|X,\alpha}(Z_i; \theta_1(X_i), \theta_2(X_i)|X_i).$$

In local likelihood estimation techniques this is done by introducing a weight function that only gives a non-zero weight to the contribution if X_i is indeed close to x_0. More precisely, let K be a symmetric probability density with compact support, and denote by $K_h(\cdot) = K(\cdot/h)/h$ the rescaled version of $K(\cdot)$, where $h > 0$ is a bandwidth parameter (determining the size of the neighbourhood of x_0). Each entry in the log-likelihood function is given the weight $K_h(X_i - x_0)$ and the resulting *conditional local kernel-weighted log-likelihood* is

$$\mathcal{L}_n(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2; h, x_0) = \sum_{i=1}^{n} \ell(\mathbf{X}_{i,p_1}^T \boldsymbol{\theta}_1, \mathbf{X}_{i,p_2}^T \boldsymbol{\theta}_2; Z_i) K_h(X_i - x_0). \tag{24}$$

The unknown vector of function values $(\boldsymbol{\theta}_1(x_0), \boldsymbol{\theta}_2(x_0))$ is then estimated by

$$(\widehat{\boldsymbol{\theta}}_1(x_0), \widehat{\boldsymbol{\theta}}_2(x_0)) = \arg\max_{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2} \sum_{i=1}^{n} \ell(\mathbf{X}_{i,p_1}^T \boldsymbol{\theta}_1, \mathbf{X}_{i,p_2}^T \boldsymbol{\theta}_2; Z_i) K_h(X_i - x_0). \tag{25}$$

The estimator $\widehat{\theta}_r^{(v)}(x_0)$ for $\theta_r^{(v)}(x_0)$, $v = 0, 1, \ldots, p_r$ is then given by $\widehat{\theta}_r^{(v)}(x_0) = v!\widehat{\theta}_{rv}(x_0)$.

The entire function $\theta_r(\cdot)$ (for $r \in \{1, 2\}$) is estimated by considering a grid of x_0-values and solving maximization problem (25) for each point in the grid. Finally, the estimated βth-conditional quantile function of Z (respectively Y) at the point x_0 is

$$\widehat{F}_{Z|X,\alpha}^{-1}(\beta; \mu(x_0), \phi(x_0)|x_0) = \widehat{\theta}_1(x_0) + \exp(\widehat{\theta}_2(x_0)) \cdot C_\alpha(\beta) \tag{26}$$

$$\left\{\widehat{F^g{}_{Y|X,\alpha}}\right\}^{-1}(\beta; \eta(x_0), \phi(x_0)|x_0) = g^{-1}\left(\widehat{\theta}_1(x_0) + \exp(\widehat{\theta}_2(x_0)) \cdot C_\alpha(\beta)\right). \tag{27}$$

We next discuss an example of the above semiparametric estimation procedure. Consider the conditional generalized tick-exponential family of conditional densities

(20). Based on the (calculated) data $(Z_1, X_1), \ldots, (Z_n, X_n)$ we write the local kernel-weighted conditional log-likelihood function (24) and obtain

$$\mathcal{L}_n(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2; h, x_0) = \ln\left[\frac{\alpha(1-\alpha)}{\Gamma(1+\frac{1}{p})}\right]\sum_{i=1}^{n} K_h(X_i - x_0) - \sum_{i=1}^{n} X_{i,p_2}^T \boldsymbol{\theta}_2 K_h(X_i - x_0) \\ - \sum_{i=1}^{n}\left(\frac{\rho_{\alpha,p}(Z_i - \mathbf{X}_{i,p_1}^T \boldsymbol{\theta}_1) K_h(X_i - x_0)}{\{\exp[\mathbf{X}_{i,p_2}^T \boldsymbol{\theta}_2]\}^p}\right). \tag{28}$$

A special situation occurs when we take $p_2 = 0$, and hence approximate $\boldsymbol{\theta}_2(x_0)$ locally by a constant, and $\boldsymbol{\theta}_2(x_0) = \theta_{20}(x_0)$. In this case the solution to maximization problem (28) has the explicit expression

$$\begin{cases} \widehat{\boldsymbol{\theta}}_1(x_0) = \arg \min\limits_{\boldsymbol{\theta}_1 \in \mathbb{R}^{(d+1)\times 1}} \sum\limits_{i=1}^{n} \rho_{\alpha,p}(Z_i - \mathbf{X}_{i,p_1}^T \boldsymbol{\theta}_1) K_h(X_i - x_0) \\ \widehat{\theta}_{20}(x_0) = \left(\frac{1}{p}\right) \ln\left[\frac{p \sum_{i=1}^{n} \rho_{\alpha,p}(Z_i - \mathbf{X}_{i,p_2}^T \boldsymbol{\theta}_1) K_h(X_i - x_0)}{\sum_{i=1}^{n} K_h(X_i - x_0)}\right]. \end{cases}$$

We refer to Gijbels et al. (2019) for a detailed study and for approaches to choose the bandwidth parameter h.

For the identity link function, an asymptotic normality result for the estimators of $(\theta_1(x_0), \theta_2(x_0))$ is established in Gijbels et al. (2019). From this an asymptotic distributional result for the local log-likelihood estimators of $(\eta(x_0), \theta_2(x_0))$ and for the conditional quantile estimator $\left\{F_{Y|X,\alpha}^{g}\right\}^{-1}(\cdot; \eta(x_0), \phi(x_0))|X = x_0)$ can be derived. Due to the technicality, we do not elaborate on this here.

5 Real data applications

In this section we illustrate the use of the proposed methodology in data analysis. The data that we consider are data on locomotor performance in small and large terrestrial mammals. A detailed description of these data is available in Iriarte-Díaz (2002). For $n = 142$ species of mammals measurements on their body length, body mass (in kg) and maximum relative running speed were recorded. The maximum relative running speed measurement takes into account the body length of the mammals, and was obtained by dividing the maximum speed of the mammal species by its body length. In Section 5.1 we are interested in finding an appropriate density to describe the distribution of the maximum relative running speed of terrestrial mammals (Y). Of interest is also to find out the relationship between the maximum relative running speed and the body mass of the mammals. One of the findings in Iriarte-Díaz (2002) was that the relationship differs, in mean behaviour, when focusing on small or on large mammals. In Section 5.2 we go beyond investigation of a conditional mean, and study instead conditional quantiles, which allows to have a more complete understanding of the conditional distribution of Y for given log(body mass). Note that here the response variable Y takes on only positive values. An appropriate link function in this real data analysis turns out to be log-link function i.e. $g(y) = \ln(y)$.

5.1 Parametric quantile estimation

We first would like to find an appropriate distribution for the maximum relative running speed. To get an idea about the possible asymmetry of the distribution we plot in Figure 8 the histogram of the log-transformed data (the $Z_i = \ln(Y_i)$ observations), together with a kernel density estimate. In this and later kernel density estimates we use a Gaussian kernel with Silverman's rule-of-thumb bandwidth (the default in the R command `density`). As can be seen from Figure 8 the log-transformed data are left-skewed.

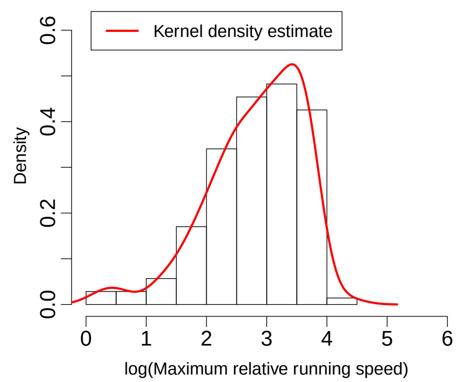

Fig. 8 Histogram and kernel density estimate (solid line) of log(Maximum relative running speed) data.

For the reference symmetric density f in (3) we consider the following densities: a standard normal density, a Student-t density with ν degrees of freedom, a standard logistic density and a standard Laplace density. We refer to the resulting asymmetric densities as the Log-asymmetric normal, Student-t, logistic and Laplace densities, abbreviated as Log-AND(η, ϕ, α), Log-ATD$(\eta, \phi, \alpha, \nu)$, Log-ALD$(\eta, \phi, \alpha)$ and Log-ALaD(η, ϕ, α), respectively. For selecting an appropriate density for Y, in the set of considered densities, we look into the equivalent problem of selecting the appropriate density for $Z = \ln(Y)$, in the family (1) which involves the parameter $\boldsymbol{\theta} = (\mu, \phi, \alpha)^T$ as well as possible other parameters. The corresponding set of candidate densities consists of the asymmetric normal, Student-t, logistic and Laplace densities, abbreviated as AND(μ, ϕ, α), ATD(μ, ϕ, α, ν), ALD(μ, ϕ, α) and ALaD(μ, ϕ, α), respectively. The full parameter vector in each model is estimated by using maximum likelihood estimation as described in Section 3. We perform a Kolmogorov-Smirnov (KS) goodness-of-fit test for each model in testing the hypotheses

$$H_0 : \text{Sample data come from the } F_\alpha(\cdot;\mu,\phi) \text{ distribution} \quad (\mu \text{ and } \phi \text{ unspecified})$$
$$H_1 : \text{Sample data do not come from the } F_\alpha(\cdot;\mu,\phi) \text{ distribution.}$$

We calculate the realized value of the KS-test statistic, denoted by $D_n = \sup_z |F_0(z) - F_n(z)|$, where $F_0(\cdot)$ is the cumulative distribution function under H_0 (with estimated parameters) and $F_n(\cdot)$ is the empirical distribution function. We retain a candidate distribution if and only if the P-value of D_n is larger than the fixed significance level 0.05 which indicates that there is no strong evidence against H_0. In order to select the most parsimonious density model among all retained candidates, we compute the value of Akaike's information criteria (AIC):

$$\text{AIC} = -2\ln\left(L_n\left(\widehat{\boldsymbol{\theta}}_n^{\text{MLE}}\right)\right) + 2k,$$

where k is the number of estimated parameters in the model, and $L_n\big(\widehat{\boldsymbol{\theta}}_n^{\text{MLE}}\big)$ is the realized maximal likelihood value. The best model among all retained candidates is chosen based on the smallest AIC value.

For each candidate quantile-based asymmetric model we list in Table 1 the maximum likelihood estimates of the parameters, the realized maximal value of the log-likelihood $\ln\big(L_n\big(\widehat{\boldsymbol{\theta}}_n^{\text{MLE}}\big)\big)$, the AIC-value, and the value of the test statistic D_n with the associated P-value for the test. If parameters are not involved in a model, we mention this by NAP = Not Applicable in Table 1.

Table 1: Maximum likelihood estimates, maximal log-likelihood and corresponding AIC-value, Kolmogorov-Smirnov test statistic value and corresponding P-value.

Density	AND(μ,ϕ,α)	ATD(μ,ϕ,α,ν)	ALD(μ,ϕ,α)	ALaD(μ,ϕ,α)
$\widehat{\mu}$	3.5596	3.5945	3.5854	3.6702
$\widehat{\phi}$	0.1914	0.1793	0.1143	0.1056
$\widehat{\alpha}$	0.8372	0.8435	0.8521	0.8892
$\widehat{\nu}$	NAP	8.9212	NAP	NAP
LogLik	−150.2551	−150.7701	−150.8381	−150.8381
AIC	306.5102	307.3200	307.6762	307.6763
D_n	0.0446	0.0503	0.0632	0.0635
P-Value	0.9294	0.8492	0.5969	0.6558

NAP = Not Applicable

From Table 1 it is seen that, among the considered models, the asymmetric normal density AND(μ,ϕ,α) has the smallest value for the KS-statistics D_n with the largest associated P-value, as well as the lowest AIC value (among all retained models, which are all models here). Hence an asymmetric normal

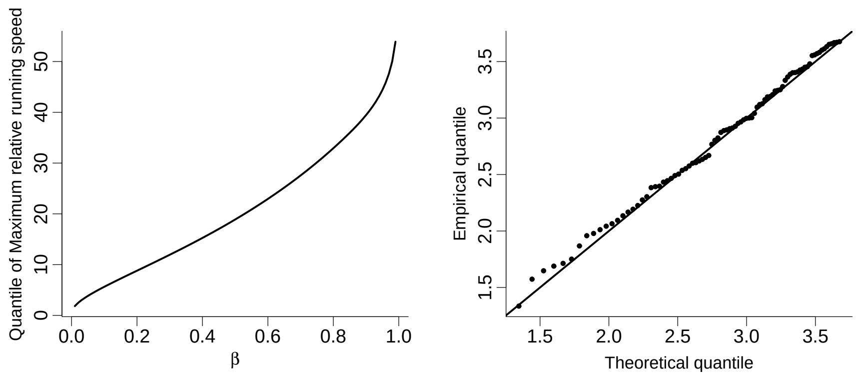

Fig. 9: Left: Estimated quantile function of Log-AND$(\widehat{\eta}, \widehat{\sigma}, \widehat{\alpha})$; Right: Q-Q plot for the log-asymmetric normal distribution Log-AND$(\widehat{\eta}, \widehat{\phi}, \widehat{\alpha})$.

model is the selected model for Z=log(maximum relative running speed). For this selected model, the estimated value for the index-parameter is $\widehat{\alpha} = 0.8372$, which coincides with our earlier finding that the distribution is left-skewed. For Y, the maximum relative running speed, the selected model is thus the density in (9) with index-parameter 0.8374. The maximum likelihood estimate for η is $\widehat{\eta} = g^{-1}(\widehat{\mu}) = 35.1491$. The estimated quantile function is easily obtained from (5) with $F^{-1}(\beta) = \text{sgn}(\beta - 0.5)\sqrt{2\gamma^{-1}(0.5, \text{sgn}(\beta - 0.5)\sqrt{\pi}(2\beta - 1))}$ for the quantile function of a standard normal density f. The estimated βth-quantile function of Log-AND(η, ϕ, α) is

$$F_{\widehat{\alpha}}^{-1}(\beta) = \exp\left(\log(\widehat{\eta}) - \frac{\widehat{\phi}}{1-\widehat{\alpha}}\sqrt{2\gamma^{-1}\big(\frac{1}{2}, \frac{\sqrt{\pi}(\widehat{\alpha}-\beta)}{\widehat{\alpha}}\big)}\mathbb{I}(\beta \le \widehat{\alpha}) \right.$$
$$\left. + \frac{\widehat{\phi}}{\widehat{\alpha}}\sqrt{2\gamma^{-1}\big(\frac{1}{2}, \frac{\sqrt{\pi}(\beta-\widehat{\alpha})}{1-\widehat{\alpha}}\big)}\mathbb{I}(\beta > \widehat{\alpha})\right),$$

which is depicted in the left panel of Figure 9. The right panel of Figure 9 presents the Quantile-Quantile (Q-Q) plot comparing the fitted quantiles (using Log-AND$(\widehat{\eta}, \widehat{\phi}, \widehat{\alpha})$) and the empirical quantiles of the Y_i data, together with a 45-degree reference line (the solid line). It is observed that most of the Q-Q values are close to the 45-degree reference line which indicates that the Log-AND$(\widehat{\eta}, \widehat{\phi}, \widehat{\alpha})$ distribution fits quite well the maximum relative running speed data.

5.2 Semiparametric quantile estimation

In this section we consider the maximum relative running speed as the response variable Y and log(Body Mass) as a covariate X. Similar to the conditional mean estimation in Iriarte-Díaz (2002), we again consider a logarithmic transformation of Y, denoted by Z (i.e., $Z = \ln(Y)$). A scatter plot of the observations of log(Maximum relative running speed) against the observations of log(Body mass) is provided in Figure 10 (left panel). Presented is also a nonparametric estimate $\widehat{m}$ of the conditional mean of $\mathrm{E}(Z|X = x)$ obtained by local linear fitting using a Gaussian kernel. The scatterplot with the conditional mean estimate reveals that the maximum relative running speed decreases with increasing body mass and that the average maximum relative running speed behaves differently over different ranges of body mass. Overall, it decreases with increasing body mass.

We next use the semiparametric method exposed in Section 4 to investigate in more detail the conditional distribution of the maximum relative running speed for given log(body mass). In the proposed methodology, we assumed that the index-parameter α is known which may not be the case in a real data application, as here. We proceed as follows to select a reference symmetric density f and to approximate the index-parameter α. Using the local linear regression estimate $\widehat{m}$ we consider the conditional mean residuals $Z_i - \widehat{m}(X_i)$. A histogram for these residuals is provided in the right panel of Figure 10, together with a kernel density estimate. From this figure it is seen that the residuals are slightly left-skewed.

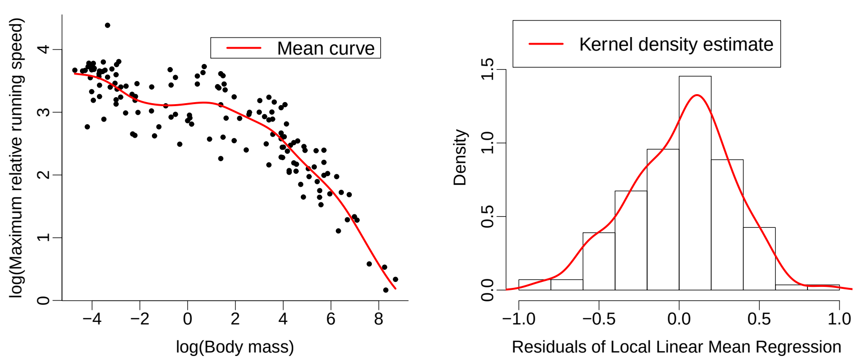

Fig. 10: Left: Scatterplot and conditional mean estimate of log(Maximum relative running speed); Right: Histogram of the residual obtained by local linear mean regression fitting.

We then look for an appropriate asymmetric density which describes well the distribution of the residuals. This distribution will then be used in our semiparametric analysis. As candidate densities for the distribution of the residuals we consider asymmetric Laplace, normal, Student-t and logistic densities. Similarly as in the analysis of Section 5.1 we evaluate the appropriateness of a model via the Kolmogorov Smirnov goodness-of-fit test, and only consider models for which the P-value of the test is larger than the significance level 0.05. The Kolmogorov Smirnov test gives the largest P-value (0.7597) for the asymmetric normal density (denoted by AND). In addition, this density also appears to be the most parsimonious one among all candidates since it has the smallest AIC-value (79.1487). For this AND model the maximum likelihood estimator for α is $\widehat{\alpha} = 0.5937$, which confirms the slight left-skewness of the distribution of residuals observed in Figure 10.

Based on the above preliminary analysis we then consider a conditional asymmetric normal density (20) with $g(y) = \ln(y)$ (and power $p = 2$) and index-parameter $\alpha = 0.5937$. We apply the semiparametric method of Section 4 to estimate $\boldsymbol{\theta}_1(x_0)$ and $\boldsymbol{\theta}_2(x_0)$. We use local linear fitting for both unknown functions, i.e. $p_1 = p_2 = 1$ and a bandwidth value $h = 0.9030$. This bandwidth was determined in a data-driven manner, using a rule of thumb bandwidth selector, discussed in detail in Gijbels et al. (2019). From the estimates of $\boldsymbol{\theta}_1(x_0)$ and $\boldsymbol{\theta}_2(x_0)$ we then obtain the estimated βth-conditional quantile of Z and Y from (26) and (27), respectively. Figure 11 displays the estimated βth-conditional quantile functions of Z (in left panel) and Y (in right panel), for values $\beta \in \{0.1, 0.5, \alpha, 0.9\}$.

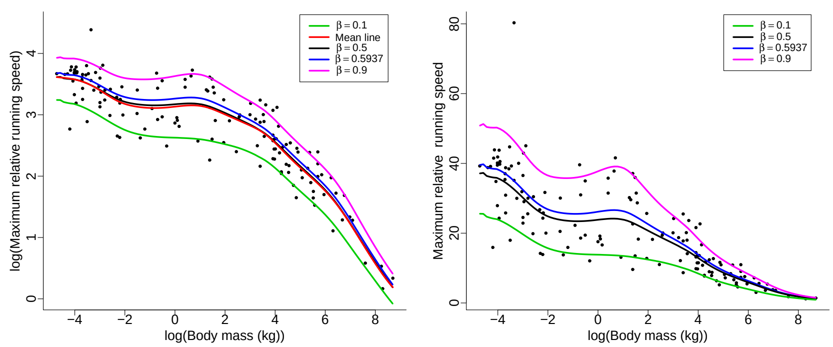

Fig. 11: Estimated semiparametric quantile curves of log(Maximum relative running speed) (left panel); and of Maximum relative running speed (right panel).

From the right panel of Figure 11 it is clearly seen that the upper conditional quantile curve of the maximum relative running speed has a different behaviour (shape) than the lower conditional quantile curve. All presented estimated quantile curves show a decreasing trend. The estimated 0.90th-conditional quantile curve is rapidly decreasing for all values of log(Body mass) except in the interval (−1.5, 2). In this interval, the upper extreme quantile curve is slightly increasing for increasing log(Body mass). It would be interesting to further investigate these findings, and relate this back to the species of terrestrial mammals.

Figure 12 depicts the estimated function $\widehat{\theta}_2(\cdot)$. It is clearly seen that the estimated log-scale function $\widehat{\theta}_2(\cdot)$ is increasing with increasing log(Body mass) up to 1 and then decreasing.

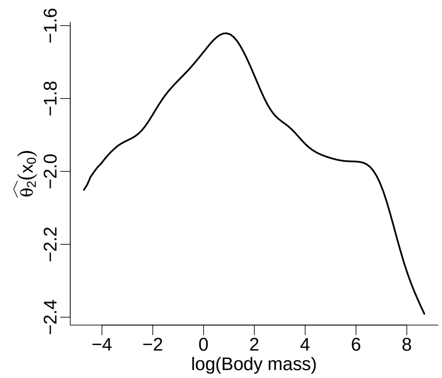

Fig. 12 Local maximum log-likelihood estimate of $\theta_2(x_0)$ with $p_2 = 1$, using the conditional asymmetric normal likelihood.

6 Conclusion and further discussion

In this paper a new broad class of asymmetric densities is proposed, where these densities are appropriate for modelling a continuous random variable Y with arbitrary (bounded or unbounded) support in $\mathbb{R}$. The class depends on an index-parameter α (with $0 < \alpha < 1$), and the location parameter η of the class coincides with the αth-quantile of the distribution. The class of densities depends on two crucial elements: (i) the reference symmetric density f and (ii) the monotone link function g. The class of densities includes several examples studied separately in the literature. In this paper we provide results for statistical inference for all the members of the whole broad class in one single track.

We also consider a regression setting, when the interest is in the impact of a covariate X on the variable of interest Y. Assuming that the index-parameter is not changing with the realized value of X, we consider a similar broad class of conditional densities, where the location and scale parameters are allowed to vary with the realized value of X, and are left unspecified, leading to a semiparametric framework. In both settings, the unconditional and conditional ones, we discuss estimation methods and establish asymptotic properties. A specific merit of this study is that the results can be employed to the many examples that are available in the literature.

An R package `QBAsyDist` has been written by the authors, providing codes for simulating data from the general (conditional) density families, for plotting densities and presenting model characteristics, for parametric and semiparametric estimation, including goodness-of-fit testing and some model selection tools for choosing an appropriate model. The package will be available soon, and requests for it can be addressed to the authors.

In this paper we assume g to be a known link function. In real data applications however this link function might not be known, and estimation of the link function might be needed. One approach is then to consider a broad parametric class of appropriate transformations, and to estimate the link function by estimating the parameter(s) describing the broad class. An example of a possible parametric class of link functions would be a Box-Cox power transformation. See for example Mu and He (2007).

We only discuss the univariate covariate case. In case of a d-dimensional covariate vector $\mathbf{X} = (X_1, \ldots, X_d)$ the methodology presented in Section 4 can still be employed, but due to the possible curse of dimensionality it might be necessary to put some more structure on the d-variate location and scale functions. A possible approach is to consider additive modelling structures for these d-variate functions. This topic is studied in current research.

Acknowledgements This research was supported by the FWO research project G.0826.15N (Flemish Science Foundation). The first author gratefully acknowledges support from the GOA/12/014 project of the Research Fund KU Leuven. The third author acknowledges support from the Flemish Science Foundation (FWO research grant 1518917N), and from the Special Research Fund (Bijzonder Onderzoeksfonds) of Hasselt University.

Appendix:

A.1 Proof of Theorem 1

If Z is a random variable with asymmetric density $f_\alpha(\cdot; \mu, \phi)$ in (1), then the cumulative distribution function of Z is given by

$$F_\alpha(z; \mu, \phi) = \begin{cases} 2\alpha F\big((1-\alpha)(\frac{z-\mu}{\phi})\big) & \text{if } z \le \mu \\ 2\alpha - 1 + 2(1-\alpha) F\big(\alpha(\frac{z-\mu}{\phi})\big) & \text{if } z > \mu, \end{cases} \tag{A.1}$$

and for any $\beta \in (0, 1)$, the βth-quantile of Z is

$$F_\alpha^{-1}(\beta) = \begin{cases} \mu + \frac{\phi}{1-\alpha} F^{-1}\big(\frac{\beta}{2\alpha}\big) & \text{if } \beta \le \alpha \\ \mu + \frac{\phi}{\alpha} F^{-1}\Big(\frac{1+\beta-2\alpha}{2(1-\alpha)}\Big) & \text{if } \beta > \alpha, \end{cases}$$

with $F_\alpha^{-1}(\alpha) = \mu$. These results are given in Corollary 2.1 of Gijbels et al. (2018). Using Expression (A.1), we find

$$
\begin{aligned}
F_{\alpha}^{g}(y;\eta,\phi) &= \Pr(Y \le y) = \Pr\left(g^{-1}(Z) \le y\right) = \Pr\left(Z \le g(y)\right) \\
&= \begin{cases} 2\alpha F\left((1-\alpha)(\frac{g(y)-g(\eta)}{\phi})\right) & \text{if } y \le \eta \\ 2\alpha - 1 + 2(1-\alpha)F\left(\alpha(\frac{g(y)-g(\eta)}{\phi})\right) & \text{if } y > \eta. \end{cases}
\end{aligned}
$$

From this we then easily obtain (5).

A.2 Proof of Theorem 2

Theorem 3.4 of Gijbels et al. (2018) states that under the assumptions (**B1**)—(**B4**), the MLE $\widehat{\boldsymbol{\theta}}_n^{(\text{MLE})}$ centered with $\boldsymbol{\theta}$ is asymptotically normally distributed with mean vector $\mathbf{0}$ and variance-covariance matrix $[\mathcal{I}(\boldsymbol{\theta})]^{-1}$:

$$
\sqrt{n}(\widehat{\boldsymbol{\theta}}_n^{\text{MLE}} - \boldsymbol{\theta}) \xrightarrow{d} \mathcal{N}_3(\mathbf{0}, \mathcal{I}(\boldsymbol{\theta})^{-1}) \qquad \text{as} \quad n \to \infty,
$$

where $\mathcal{I}(\boldsymbol{\theta})$ is the Fisher information matrix given in Proposition 3.2 of Gijbels et al. (2018), with inverse

$$
\mathcal{I}(\boldsymbol{\theta})^{-1} = \begin{bmatrix} \frac{\gamma_3\phi^2}{2\alpha(1-\alpha)(\gamma_1\gamma_3-\gamma_2^2)} & \frac{(1-2\alpha)\gamma_2\phi^2}{2\alpha(1-\alpha)(\gamma_1\gamma_3-\gamma_2^2)} & \frac{\gamma_2\phi}{2(\gamma_1\gamma_3-\gamma_2^2)} \\ \frac{(1-2\alpha)\gamma_2\phi^2}{2\alpha(1-\alpha)(\gamma_1\gamma_3-\gamma_2^2)} & [\mathcal{I}(\boldsymbol{\theta})^{-1}]_{22} & \frac{(1-2\alpha)\gamma_1\phi}{2(\gamma_1\gamma_3-\gamma_2^2)} \\ \frac{\gamma_2\phi}{2(\gamma_1\gamma_3-\gamma_2^2)} & \frac{(1-2\alpha)\gamma_1\phi}{2(\gamma_1\gamma_3-\gamma_2^2)} & \frac{\alpha(1-\alpha)\gamma_1}{2(\gamma_1\gamma_3-\gamma_2^2)} \end{bmatrix},
$$

with $[\mathcal{I}(\boldsymbol{\theta})^{-1}]_{22} = [\mathcal{I}(\boldsymbol{\theta}^g)^{-1}]_{22}$ where the latter quantity is stated in Theorem 2.

We want to find an asymptotic distribution for $\widehat{\boldsymbol{\theta}_n^{g}}^{\text{MLE}} = (\widehat{\eta}_n^{\text{MLE}}, \widehat{\phi}_n^{\text{MLE}}, \widehat{\alpha}_n^{\text{MLE}})^T$, where $\widehat{\eta}_n^{\text{MLE}} = g^{-1}(\widehat{\mu}_n^{\text{MLE}})$, which is a function of $\widehat{\boldsymbol{\theta}}_n^{\text{MLE}}$. Using the multivariate delta method, we obtain

$$
\sqrt{n}[\widehat{\boldsymbol{\theta}_n^{g}}^{\text{MLE}} - \boldsymbol{\theta}^g] \xrightarrow{d} \mathcal{N}_3\left(\mathbf{0}, \mathcal{I}(\boldsymbol{\theta}^g)^{-1}\right),
$$

with $\mathcal{I}(\boldsymbol{\theta}^g)^{-1}$ as given in the statement of Theorem 2. Similarly, the results in (2) can be obtained if α is known.

References

Arellano-Valle, R. B. and Genton, M. G. (2005). On fundamental skew distributions, *Journal of Multivariate Analysis* **96**(1): 93–116.

Azzalini, A. (1985). A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics* **12**(2): 171–178.

Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones, *Statistica* **46**(2): 199–208.

Bottai, M., Cai, B. and McKeown, R. E. (2010). Logistic quantile regression for bounded outcomes, *Statistics in Medicine* **29**(2): 309–317.

Columbu, S. and Bottai, M. (2016). Logistic quantile regression to model cognitive impairment in Sardinian cancer patients, *in* T. Di Battista, M. E. and R. W. (eds), *Topics on Methodological and Applied Statistical Inference*, Springer, pp. 65–73.

Cox, D. R. and Reid, N. (1987). Parameter orthogonality and approximate conditional inference, *Journal of the Royal Statistical Society, Series B (Methodological)* **49**(1): 1–39.

Crow, E. L. and Shimizu, K. (eds) (1988). *Lognormal Distributions: Theory and Applications*, CRC Press.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its applications*, Vol. 66, CRC Press.

Fernández, C. and Steel, M. F. (1998). On Bayesian modeling of fat tails and skewness, *Journal of the American Statistical Association* **93**(441): 359–371.

Gijbels, I., Karim, R. and Verhasselt, A. (2018). On quantile-based asymmetric family of distributions: properties and inference, *Under review* .

Gijbels, I., Karim, R. and Verhasselt, A. (2019). Semiparametric quantile regression using quantile-based asymmetric family of densities, *Submitted* .

Gneiting, T. (2011). Quantiles as optimal point forecasts, *International Journal of Forecasting* **27**(2): 197–207.

Gourieroux, C., Monfort, A. and Trognon, A. (1984). Pseudo maximum likelihood methods: theory, *Econometrica* **52**(3): 681–700.

Iriarte-Díaz, J. (2002). Differential scaling of locomotor performance in small and large terrestrial mammals, *Journal of Experimental Biology* **205**(18): 2897–2908.

Jones, M. (2015). On families of distributions with shape parameters, *International Statistical Review* **83**(2): 175–192.

Jung, S.-H. (1996). Quasi-likelihood for median regression models, *Journal of the American Statistical Association* **91**(433): 251–257.

Koenker, R. (2005). *Quantile Regression*, Cambridge University Press.

Koenker, R. and Bassett Jr, G. (1978). Regression quantiles, *Econometrica* **46**(1): 33–50.

Komunjer, I. (2005). Quasi-maximum likelihood estimation for conditional quantiles, *Journal of Econometrics* **128**(1): 137–164.

Komunjer, I. (2007). Asymmetric power distribution: theory and applications to risk measurement, *Journal of Applied Econometrics* **22**(5): 891–921.

Kotz, S., Kozubowski, T. J. and Podgórski, K. (2001). Asymmetric Laplace distributions, *The Laplace Distribution and Generalizations*, Springer, chapter 3, pp. 133–178.

Kozubowski, T. J. and Podgórski, K. (2003). Log-Laplace distributions, *International Mathematical Journal* **3**(4): 467–495.

Lindsey, J. K. (2004). *Statistical Analysis of Stochastic Processes in Time*, Cambridge University Press.

McCullagh, P. and Nelder, John, A. (1998). *Generalized Linear Models*, Chapman and Hall.

Mineo, A. and Ruggieri, M. (2005). A software tool for the exponential power distribution: The normalp package, *Journal of Statistical Software* **12**(4): 1–24.

Mu, Y. and He, X. (2007). Power transformation toward a linear regression quantile, *Journal of the American Statistical Association* **102**(477): 269–279.