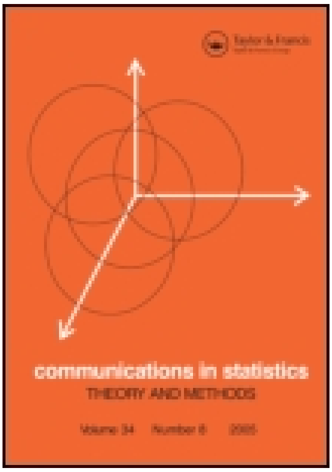

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/lsta20

A Three-Parameter Asymmetric Laplace Distribution and Its Extension

Keming Yu [a] & Jin Zhang [b]

[a] School of Mathematics and Statistics , University of Plymouth , UK

[b] University of Manitoba , Manitoba, Canada

Published online: 02 Sep 2006.

Communications in Statistics—Theory and Methods, 34: 1867–1879, 2005

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920500199018

Taylor & Francis
Taylor & Francis Group

Distributions and Applications

A Three-Parameter Asymmetric Laplace Distribution and Its Extension

KEMING YU[1] AND JIN ZHANG[2]

[1]School of Mathematics and Statistics, University of Plymouth, UK
[2]University of Manitoba, Manitoba, Canada

In this article, a new three-parameter asymmetric Laplace distribution and its extension are introduced. This includes as special case the symmetric Laplace double-exponential distribution. The distribution has established a direct link to estimation of quantile and quantile regression. Properties of the new distribution are presented. Application is made to a flood data modeling example.

Keywords Asymmetric Laplace distribution; Parameter estimation; Quantile.

1. Introduction and Notation

Recently, two papers by Koenker and Machado (1999) and Yu and Moyeed (2001) have appeared in statistical journals applying a new skew distribution for quantile regression. This distribution has applications beyond quantile estimation and can be generalized and named as asymmetric Laplace distribution (ALD), which has the following probability density function (pdf):

$$f(x; \mu, \sigma, p) = \frac{p(1-p)}{\sigma} \exp\left(- \frac{(x-\mu)}{\sigma}[p - I(x \leq \mu)]\right), \tag{1}$$

where $0 < p < 1$ is the skew parameter, $\sigma > 0$ is the scale parameter, $-\infty < \mu < \infty$ is the location parameter, and $I(\cdot)$ is the indication function. The range of x is $(-\infty, \infty)$. We write through it as ALD(μ, σ, p) or ALD in the article.

ALD(μ, σ, p) is skewed to left when $p > \frac{1}{2}$, and skewed to right when $p < \frac{1}{2}$. Figure 1 shows a variety of ALD densities. The Laplace double exponential distribution we usually call is a special case of ALD(μ, σ, p) with $p = \frac{1}{2}$.

Received September 10, 2004; Accepted April 15, 2005
Address correspondence to Keming Yu, School of Mathematics and Statistics, University of Plymouth, PL4 8AA, UK; E-mail: kyu@plymouth.ac.uk

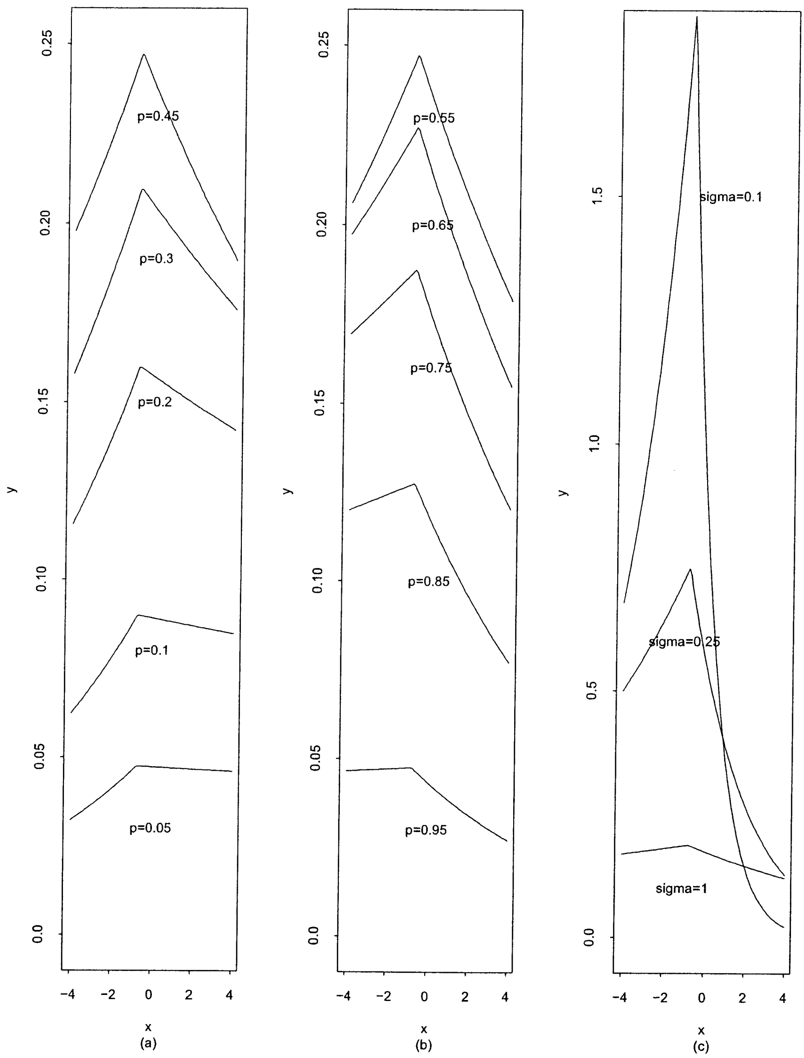

Figure 1. ALD densities (a) for $0 < p < 0.5$ with $\sigma = 1$; (b) for $0.5 < p < 1$ with $\sigma = 1$; (c) for $0.1 \leq \sigma \leq 1$ with $p = 0.75$.

Actually, when $p = \frac{1}{2}$, we have the density function of double-exponential distribution

$$g\left(x; \mu, \sigma, \frac{1}{2}\right) = \frac{1}{4\sigma} \exp\left(-\frac{|x - \mu|}{2\sigma}\right). \tag{2}$$

The Laplace double exponential distribution is a symmetric distribution. There were several asymmetric extensions in generalizing the double exponential distribution in the literature; see the details from Hsu (1978), Ch. 3 of Kotz et al. (2001), Ch. 24 of Johnson et al. (1995), and Kotz et al. (2002). Basically, all these existing asymmetric Laplace distributions are either a percent mixture of the double exponential distribution or a split of it. For example, the distribution in

Johnson et al. (1995) with density

$$p_X(x; \mu, p, \phi) = \begin{cases} p\phi^{-1} \exp\left(-\frac{1}{\phi}(x-\mu) \right), & \text{if } x \geq \mu \\ (1-p)\phi^{-1} \exp\left(\frac{1}{\phi}(x-\mu) \right), & \text{if } x < \mu \end{cases}$$

can be generated from $g(x; \mu; 2\phi)$ by a variate X taking negative values with probability p and positive ones with probability $(1-p)$, and the distribution in Kotz et al. (2002) with density

$$p_X(x; \mu, k, \sigma) = \frac{\sqrt{2}}{\sigma} \frac{k}{1+k^2} \begin{cases} \exp\left(-\frac{\sqrt{2}k}{\sigma}(x-\mu) \right), & \text{if } x \geq \mu \\ \exp\left(\frac{\sqrt{2}}{\sigma k}(x-\mu) \right), & \text{if } x < \mu \end{cases}$$

can be regarded as a split of $g(x)$ via $\frac{2}{k+\frac{1}{k}}(g(k\sqrt{2}x)I(0,\infty)(x-\mu) + g(\frac{\sqrt{2}x}{k})I(\infty, 0)$ $(x-\mu))$ (Fernández and Steel, 1998).

The ALD considered in this article, however, is different from these existing asymmetric distributions except $p = \frac{1}{2}$, which reduces to the double-exponential distribution. Specifically, if a random variable $X \sim \text{ALD}(\mu, \sigma, p)$, then $Pr(X < \mu) = p$ and $Pr(X > \mu) = 1 - p$, which shows that the parameters μ and p in ALD satisfy: μ is the pth quantile of the distribution. This important feature of ALD has been generally adopted for quantile inference (Yu et al., 2003) and made it more popular than other asymmetric Laplace distributions. However, a deep investigation of ALD has not been addressed anywhere else.

If a random variable $X \sim \text{ALD}(\mu, \sigma, p)$, then the distribution function, quantile function and moment generating function $\phi(t) = E\exp(tX)$ of X are given, respectively, by

$$F(x; \mu, \sigma, p) = \begin{cases} p\exp\left(\frac{1-p}{\sigma}(x-\mu) \right), & \text{if } x \leq \mu \\ 1-(1-p)\exp\left(-\frac{p}{\sigma}(x-\mu) \right), & \text{if } x > \mu \end{cases} \tag{3}$$

$$F^{-1}(x; \mu, \sigma, p) = \begin{cases} \mu + \frac{\sigma}{1-p}\log\left(\frac{x}{p} \right), & \text{if } 0 \leq x \leq p \\ \mu - \frac{\sigma}{p}\log\left(\frac{1-x}{1-p} \right), & \text{if } p < x \leq 1 \end{cases} \tag{4}$$

and

$$\phi(t) = p(1-p)\frac{\exp(\mu t)}{(p-\sigma t)(\sigma t + 1 - p)}, \tag{5}$$

where $\frac{p-1}{\sigma} < t < \frac{p}{\sigma}$.

Furthermore, the ALD has the flexibility of different proportions and exponential coefficients in each part, and there is, presumably, the further (four parameter) generalization in which all parameters are free. In fact, this four-parameter generalized ALD has pdf

$$f_G(x;\mu,p_1,p_2,\sigma)=\frac{1}{\sigma(\frac{1}{p_1}+\frac{1}{p_2})}\begin{cases}\exp\left\{-\frac{p_1}{\sigma}(x-\mu)\right\}, & \text{if } x\geq\mu\\ \exp\left\{\frac{p_2}{\sigma}(x-\mu)\right\}, & \text{if } x<\mu\end{cases}$$

Figure 2 displays a variety of $f_G(x;\mu,p_1,p_2,\sigma)$. The distribution is skewed to the left when $p_1>p_2$, and skewed to the right when $p_1<p_2$. Due to $Pr(X<\mu)=\frac{p_1}{p_1+p_2}$ and $Pr(X>\mu)=\frac{p_2}{p_1+p_2}$, the location parameter μ in four-parameter generalized ALD is now the $\frac{p_1}{p_1+p_2}$th quantile of the distribution.

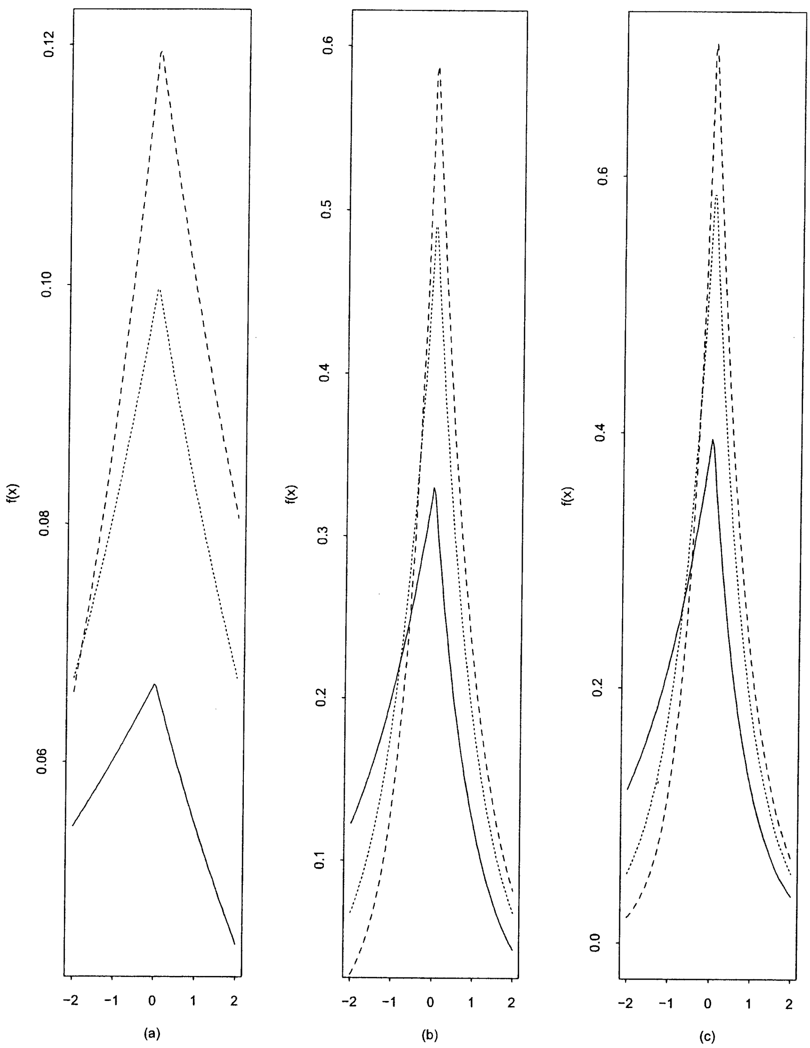

Figure 2. The generalized ALD densities for $\sigma=0.5$ and $p_2=cp_1$ with $c=0.5$ (solid line), 1 (dotted line), and 1.5 (dashed line), having increasing amount of right skewness from left, in the cases (a) $p_1=0.1$ (b) $p_1=0.5$, and (c) $p_1=0.6$.

Corresponding to Eqs. (3), (4), and (5) of the three-parameter ALD, the distribution function, quantile function, and the moment generating function $\phi(t) = E\exp(tX)$ of the four-parameter generalized ALD are given, respectively, by

$$F_G(x; \mu, p_1, p_2, \sigma) = \begin{cases} \dfrac{p_1}{p_1 + p_2} \exp\left(\dfrac{p_2}{\sigma}(x - \mu)\right), & \text{if } x \leq \mu \\ 1 - \dfrac{p_2}{p_1 + p_2} \exp\left(-\dfrac{p_1}{\sigma}(x - \mu)\right), & \text{if } x > \mu \end{cases} \tag{6}$$

$$F_G^{-1}(x; \mu, p_1, p_2, \sigma) = \begin{cases} \mu + \dfrac{\sigma}{p_2} \log\left(\dfrac{(p_1 + p_2)x}{p_1}\right), & \text{if } 0 \leq x \leq \frac{p_1}{p_1+p_2} \\ \mu - \dfrac{\sigma}{p_1} \log\left(\dfrac{(p_1 + p_2)(1 - x)}{p_2}\right), & \text{if } \frac{p_1}{p_1+p_2} < x \leq 1 \end{cases} \tag{7}$$

and

$$\phi(t) = p_1 p_2 \frac{\exp(\mu t)}{(p_1 - \sigma t)(p_2 + \sigma t)}, \tag{8}$$

where $-\frac{p_2}{\sigma} < t < \frac{p_1}{\sigma}$.

Hence, for the four-parameter generalized ALD, we have

$$E(X - \mu)^k = k!\, \sigma^k \frac{p_1 p_2}{p_1 + p_2}\left(\frac{1}{p_1^{k+1}} + \frac{(-1)^k}{p_2^{k+1}}\right), \tag{9}$$

and the skewness and kurtosis coefficients are

$$\text{skewness}_G = \frac{2(p_1^3 - p_2^3)}{(p_1^2 + p_2^2)^{3/2}},$$

and

$$\text{kurtosis}_G = \frac{9p_1^4 + 6p_1^2 p_2^2 + 9p_2^4}{(p_1^2 + p_2^2)^2},$$

respectively.

We mainly discuss the estimation and application of ALD in this article. Some properties of ALD are present in Sec. 2. Proofs of the results are given in Appendix A. Although ALD is mainly in quantile or quantile regression studies, we present a data analysis example by fitting ALD in Sec. 3. Finally, in Sec. 4, some closing comments are made.

2. Basic Properties and Parameter Estimation of ALD

2.1. *Derivation and Simulation of ALD*

We can simulate an ALD distribution via the qunatile function (4). Alternatively, like double-exponential distribution (2) which is the differences between two independent exponential variates, the ALD occurs as a simple linear combination of two independent exponential variates. Actually, we can prove that, if ξ and η are independent and identical standard exponential distributions, then $\frac{\xi}{p} - \frac{\eta}{1-p} \sim$ ALD(0, 1, p).

Similarly, the four-parameter generalized ALD with density $f_G(x; 0, p_1, p_2, 1)$ can be generated via an alternative linear combination of two independent and identical exponential variates: $\frac{\xi}{p_1} - \frac{\eta}{p_2}$ $(0 < p_1, p_2 < 1)$.

Like normal distribution, any ALD can be derived from the standard ALD(0, 1, p). That is, if $X \sim \text{ALD}(0, 1, p)$, then $Y = \mu + \sigma X \sim \text{ALD}(\mu, \sigma, p)$. Further, if $X \sim \text{ALD}(\mu, \sigma, p)$, $Y = \alpha + \beta X$, then $Y \sim \text{ALD}(\alpha + \beta\mu, \beta\sigma, p)$.

2.2. *Maximum Likelihood Estimation of Parameters*

First, we present the moments of ALD.

The kth central moment of a random variable $X \sim \text{ALD}(\mu, \sigma, p)$ is given by

$$E(X - \mu)^k = k!\,\sigma^k\,p(1-p)\left(\frac{1}{p^{k+1}} + \frac{(-1)^k}{(1-p)^{k+1}}\right). \tag{10}$$

In particular, the mean and variance of X are given by $E(X) = \mu + \frac{\sigma(1-2p)}{p(1-p)}$ and $\text{var}(X) = \frac{\sigma^2(1-2p+2p^2)}{(1-p)^2p^2}$ respectively.

The skewness and kurtosis of X are given by

$$\text{skewness} = \frac{2(p^3 - (1-p)^3)}{((1-p)^2 + p^2)^{3/2}},$$

and

$$\text{kurtosis} = \frac{9p^4 + 6p^2(1-p)^2 + 9(1-p)^4}{(1-2p+2p^2)^2},$$

respectively. These quantities do not depend on the value of the scale parameter σ, although σ controls the range of the density of ALD. Furthermore, Fig. 3 illustrates the dependence of both skewness and kurtosis on p.

As skewness characterizes the degree of asymmetry of a distribution around its mean. Figure 3(a) shows that positive skewness and negative skewness occur for $p > 0.5$ and $p < 0.5$, respectively, which indicates a distribution with an asymmetric tail extending towards more positive values and more negative values, respectively.

Similarly, kurtosis characterizes the relative peakedness or flatness of a distribution compared to the normal distribution. Positive kurtosis indicates a relatively peaked distribution. Clearly, Fig. 3(b) gives ALD a relatively peaked distribution, even for $p = 0.5$.

In addition,

$$E|X - \mu| = \sigma\frac{1 - 2p + 2p^2}{p(1-p)}.$$

These expressions provide a simple way to the moment estimation of three parameters μ, σ, and p, although these estimator are usually not efficient. However, they provide some initial values for the calculation of maximum likelihood

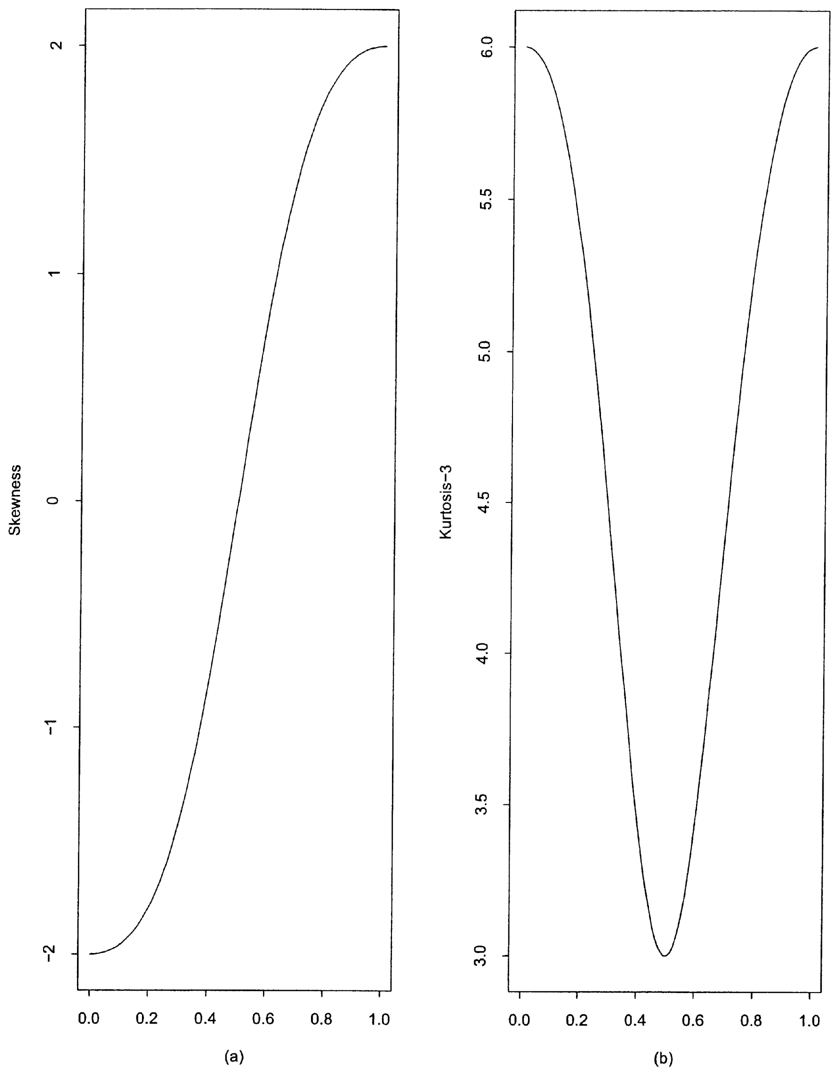

Figure 3. The skewness and kurtosis-3 of ALD against p $(0 < p < 1)$.

estimators (MLEs) below. Many computing packages can solve the moment estimation equations conveniently. For example, the LINGO (Winston) can do this computing very simple and efficiently. However, if one wants to make his own code for the moment estimation via numerical approach, he could choose the initial value $p = 1/2$, or from

$$\text{skewness} = 2(2p - 1)/g(x), \tag{11}$$

where $g(x) = (1 - 2x)^{3/2}/(1 - x)$, $x = p(1 - p)$ and $0 < x < 1/4$. Substituting $g(x)$ by $g(1/8) + g(1/8)(x - 1/8)$ into (11), a quadratic equation for p would be obtained, from which he can solve p and then solve μ and σ from the equations decided by mean and variance.

To begin the MLE, let $\rho_p(t) = t(p - I(t < 0))$, and let $\hat{\mu}, \hat{\sigma}, \hat{p}$ be the MLEs of parameters (μ, σ, p), respectively, then we have the following likelihood equations:

$$\begin{aligned} \hat{\mu} &= \text{Argmin}_{\mu} \sum_{i=1}^{n} \rho_{\hat{p}}(x_i - \mu) \\ \hat{\sigma} &= \frac{1}{n} \sum_{i=1}^{n} \rho_{\hat{p}}(x_i - \hat{\mu}) \\ \hat{p} &= \frac{a + \sqrt{a^2 - (\bar{x} - \hat{\mu})a}}{\bar{x} - \hat{\mu}}, \end{aligned} \tag{12}$$

where $\hat{a} = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu}) I(x_i \leq \hat{\mu})$.

Equation (12) provides alternative solutions for parameter μ, p, and σ: $\hat{\mu}$ is the $\hat{p}$th order statistic. That is, the likelihood function about μ attains its largest value when $\hat{\mu}$ is the mth order statistic $x_{(m)}$ with $m = [n\hat{p}] + 1$, where $[c]$ is the integer part of c. And

$$\begin{aligned} \hat{p} &= \frac{\hat{\sigma} + \hat{a}}{\bar{x} - \hat{\mu}} \\ \hat{\sigma} &= \frac{\hat{p}(1 - \hat{p})(\bar{x} - \hat{\mu})}{1 - 2\hat{p}}. \end{aligned} \tag{13}$$

The iteration algorithm is needed to solve the $\hat{\mu}, \hat{\sigma}, \hat{p}$ in the equations above. And our experience found that the algorithm converges fast given the initial values provided by the moment estimation. The iteration process may alternatively accompany a simple way to find $\hat{\mu}$; we have n ordered observations which divide the scale into $n + 1$ parts: below the lowest observation, above the highest and between each adjacent pair. The proportion of the distribution which lies below the ith observation is estimated by $i/(n + 1)$. We set this equal to p and get $i = p(n + 1)$. If i is an integer, the ith observation is the required order statistic. If not, let j be the integer part of i, the part before the decimal point. The order statistic will lie between the jth and $j + 1$th observations. We estimate it by $x_j + (x_{j+1} - x_j) \times (i - j)$.

2.3. *Confidence Interval Estimation of Parameters*

First, we discuss the confidence interval of p.

Appendix A has proved that $\sqrt{n}(\hat{p} - p) \sim N\big(0, \frac{p^2(1-p)^2}{(1-2p)^2}\big)$ asymptotically. Thus, after using the consistent estimators to substitute the parameters μ and σ, we may get the asymptotic $(1 - \alpha)\%$ confidence interval of p as

$$\left(\hat{p} - \Phi^{-1}(\alpha/2)\frac{\hat{p}(1 - \hat{p})}{|1 - 2\hat{p}|}/\sqrt{n}, \quad \hat{p} - \Phi^{-1}(1 - \alpha/2)\frac{\hat{p}(1 - \hat{p})}{|1 - 2\hat{p}|}/\sqrt{n}\right).$$

Alternatively, we could get an asymptotical χ^2 distribution based confidence interval for p as follows:

$$Pr\left[\frac{\hat{\sigma}\,\chi^2_{\frac{\alpha}{2}}(2n)/n + \hat{a}}{\bar{x} - \hat{\mu}} < p < \frac{\hat{\sigma}\,\chi^2_{1-\frac{\alpha}{2}}(2n)/n + \hat{a}}{\bar{x} - \hat{\mu}}\right] = 1 - \alpha.$$

In fact, if $X \sim \text{ALD}(x; \mu, \sigma, p)$, let $Y = \rho_p(X - \mu)$, then it is easy to prove that Y has the exponential distribution. That is, if $X \sim \text{ALD}(x; \mu, \sigma, p)$, then $Y = \rho_p(X - \mu) \sim Exp(\sigma)$.

Now given a set of sample from X, let $a = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)I(X_i \leq \mu)$ and $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n}\rho_p(X_i\mu)$, then $n\overline{Y}/\sigma \sim \chi^2(2n)$, thus

$$Pr\left[\chi^2_{\frac{\alpha}{2}}(2n) < \frac{n\overline{Y}}{\sigma} < \chi^2_{1-\frac{\alpha}{2}}(2n)\right] = 1 - \alpha$$

for given $0 < \alpha < 1$.

Note that $\overline{Y} = p(\overline{X} - \mu) - a$, then

$$Pr\left[\frac{\sigma\chi^2_{\frac{\alpha}{2}}(2n)/n + a}{\bar{x} - \mu} < p < \frac{\sigma\chi^2_{1-\frac{\alpha}{2}}(2n)/n + a}{\bar{x} - \mu}\right] = 1 - \alpha.$$

Second, the asymptotic confident interval of μ could be built through the asymptotical normality of $\sqrt{n}(\hat{\mu} - \mu)$, which is given by

$$\left(\hat{\mu} - \hat{\sigma}\frac{\Phi^{-1}(\alpha/2)}{\sqrt{n}}, \quad \hat{\mu} - \hat{\sigma}\frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{n}}\right).$$

In addition, note that μ is the pth quantile of $\text{ALD}(x; \mu, \sigma, p)$, we can estimate confidence intervals for it using the binomial distribution (Bland, 2000). This is also a large sample method. The 95% confidence interval for the pth quantile can be found by an application of the binomial distribution. The number of observations less than the pth quantile will be an observation from a binomial distribution with parameters n and p, and hence has mean np and standard deviation root $(np(1-p))$. We calculate j and k such that: $j = np - 1.96\sqrt{(np(1-p))}$, $k = np + 1.96\sqrt{(np(1-p))}$. We round j and k up to the next integer. Then the 95% confidence interval is between the jth and the kth observations in the ordered data.

Finally, the asymptotic $(1-\alpha)\%$ confidence interval of $\hat{\sigma}$ is given by

$$Pr\left(\frac{\hat{\sigma}}{1 + \frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{n}}} < \sigma < \frac{\hat{\sigma}}{1 + \frac{\Phi^{-1}(\alpha/2)}{\sqrt{n}}}\right) = 1 - \alpha.$$

3. An Example of the Application of ALD

Section 1 has mentioned that ALD has been used in modeling the model errors of quantile regression models. In this section we demonstrate that ALD can also fit real data directly. In fact, double-exponential distribution has often been used to fit data in hydrology and finance for symmetric distribution (Fernández and Steel, 1998; Hsu, 1978; Puig and Stephens, 2000). We illustrate the application of ALD by fitting an asymmetric distribution of a flood data taken from the Table 1.1 of Gilchrist (2000). The data are based on the maximum flow of flood water on a river for 20 periods of 4 years, in million cu. ft. per sec. Gilchrist (2000) made a details analysis of this data by introducing a number of summary statistics and graphs. Here we use the ALD to fit it straightway. The relative peakedness and skewness of the histogram in Fig. 4(a) compared to the normal distribution shows ALD's

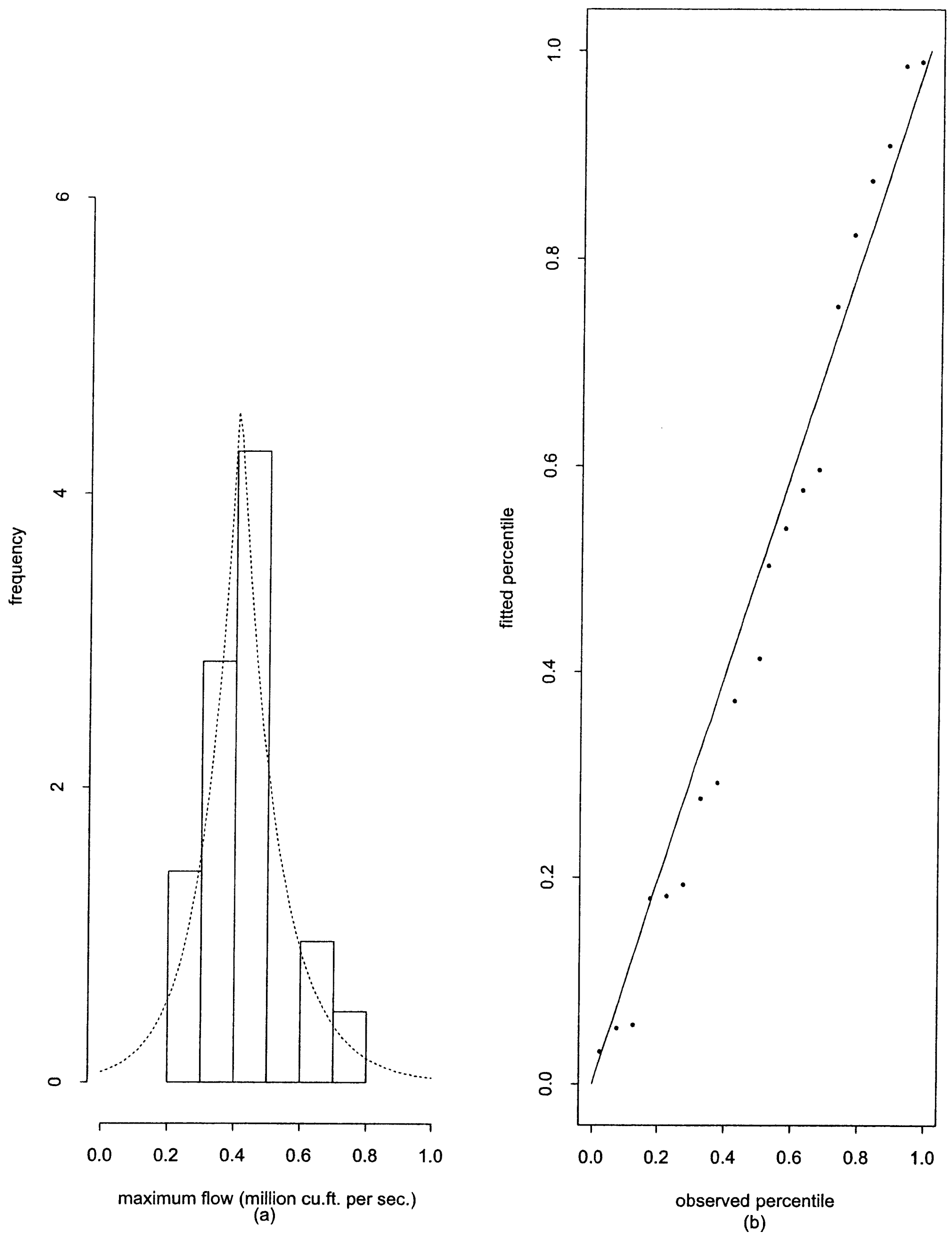

Figure 4. (a) Histogram of the flood data (solid line), together with the fitted ALD (dash line) and (b) PP-plot for the same data and fit.

suitability for the distribution of the data. With this data the maximum likelihood estimates of the three parameters of ALD and their estimated standard errors (in brackets) are $\mu = 0.407\,(0.065)$, $p = 0.453\,(0.082)$, and $\sigma = 0.053\,(0.021)$. The right skewness of this data set is from $p < 0.5$ with evidence from the straggling tail to the right. A likelihood ratio test based on twice the log-likelihood ratio, which gives value of 7.39 thus has approximately p-value 0.007 from 1 degree of freedom of chi-squared distribution, is in favor of ALD. A *PP*-plot of the fitted distribution is given in Fig. 4(b); the reasonable fit of the model to the data is apparent.

It should point out that the asymmetric Laplace distributions $p_X(x; \mu, p, \phi)$ of Johnson et al. (1995) and $p_X(x; \mu, k, \sigma)$ of Kotz et al. (2002) can also be used to fit the data. A comparison of the three distributions for fitting the data is illustrated in Fig. 5. Clearly, the difference of fitting by ALD and $p_X(x; \mu, k, \sigma)$ lies in the peak and ALD does better. Except peak, ALD fits the distribution of flood data better than $p_X(x; \mu, p, \phi)$.

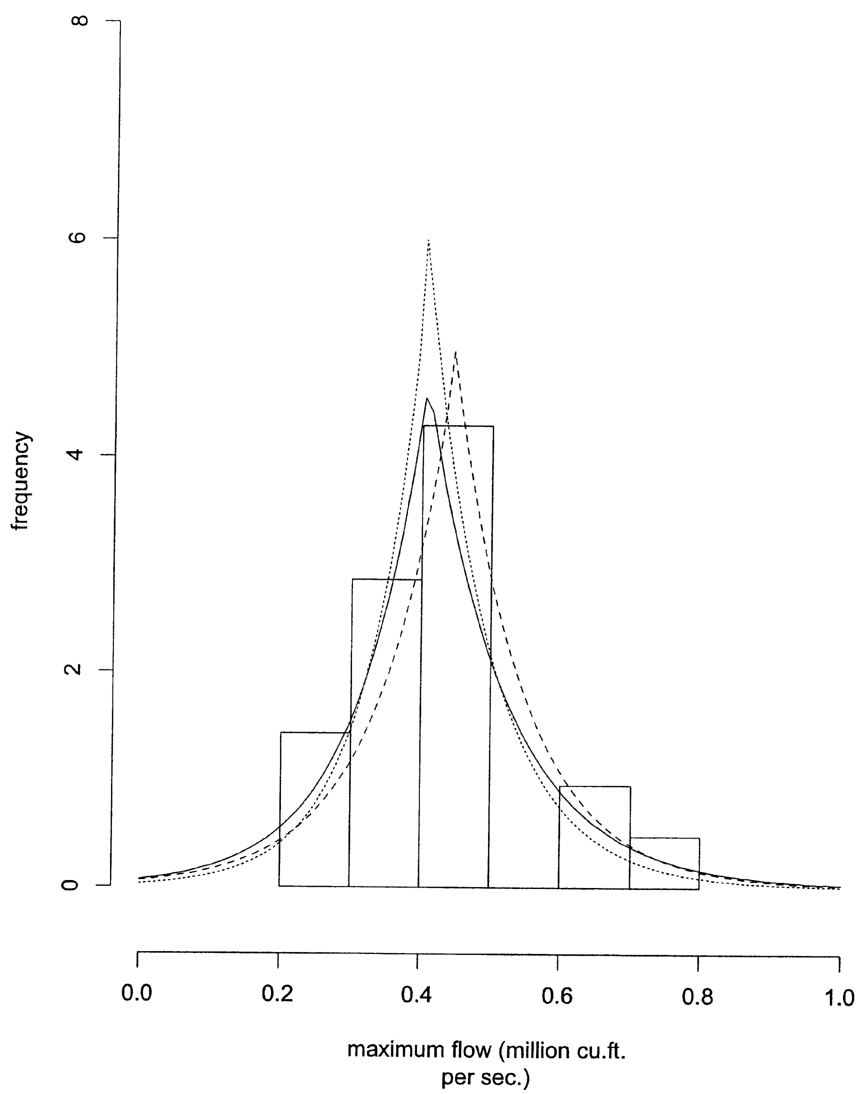

Figure 5. Fitting the flood data by ALD (solid line), $p_X(x; \mu, p, \phi)$ (dash line), and $p_X(x; \mu, k, \sigma)$ (dotted line).

4. Closing Remarks

We have made a full investigation of the asymmetric Laplace distribution. The distribution is new to the class of asymmetric Laplace distributions, and is useful for fitting quantile and quantile regression as well as for data analysis in general. The distribution is sufficiently tractable that simple simulation and likelihood theory for it can be easily made.

Appendix A: Asymptotic Properties of Parameter Estimates

We investigate the consistency and asymptotic normality of MLE.

$$\text{If } X \sim f(x;\mu,\sigma,p) = \frac{p(1-p)}{\sigma}\begin{cases}\exp\left(\frac{1-p}{\sigma}(x-\mu)\right), & \text{if } x \le \mu\\ \exp\left(-\frac{p}{\sigma}x\right), & \text{if } x > \mu\end{cases}$$

then both $XI(X<\mu)$ and $XI(X\ge\mu)$ have the exponential distribution densities $\frac{1-p}{\sigma}\exp(\frac{1-p}{\sigma}(x-\mu))$, $x<\mu$, and $\frac{p}{\sigma}\exp(-\frac{p}{\sigma}(x-\mu))$, $x>\mu$, respectively, and independently. Also

$$EXI(X<0) = -\sigma\frac{p}{1-p},$$
$$EXI(X>0) = \sigma\frac{1-p}{p},$$
$$EX^2I(X<0) = -\sigma^2\frac{2p}{(1-p)^2},$$
$$EXI(X>0) = \sigma^2\frac{2(1-p)}{p^2}.$$

Hence if $\rho_p(x) = pxI(x>0)-(1-p)x(x<0)$, then

$$E\rho_p(X) = pEX - EXI(X<0) = \sigma,$$
$$E\rho_p(X)^2 = p^2EX^2 + (1-2p)EXI(X<0) = 2\sigma^2,$$
$$\operatorname{var}\rho_p(X) = \sigma^2,$$

asymptotically.

Clearly, $\hat{p}\to\frac{1}{2}$ if and only if $\hat{\mu}=\bar{x}$, $\hat{p}<\frac{1}{2}$ if and only if $\hat{\mu}<\bar{x}$ and $\hat{p}>\frac{1}{2}$ if and only if $\hat{\mu}>\bar{x}$.

The efficiency of MLE can be investigated by checking the consistency and asymptotic normality. But, the usual regularity conditions stated in theorems concerning the asymptotic normality of MLE does not hold for the ALD here, as all derivative functions of the likelihood function are not differentiable with respect to μ, so using the information matrix to give asymptotic variances of MLE may not be reasonable.

However, as a sample order statistic, $\hat{\mu}$ is a consistent estimate of μ. That is, $\lim_{n\to\infty} Pr(|\hat{\mu}-\mu|\ge\epsilon)=0$ for any $\epsilon>0$, and we denote it as $\hat{p}\to^P p$.

$\hat{a} < 0$ and $\hat{a} \to -\frac{p}{1-p}\sigma$ following in the large number theorem when $n \to \infty$. Similarly, $\bar{x} \to \mu + \frac{(1-2p)\sigma}{p(1-p)}$, thus $\hat{\sigma}$ and $\hat{p}$ are the consistent estimators of σ and p, respectively. We will discuss the asymptoic normality of parameter estimates below.

As $\hat{\mu}$ is the $\hat{p}$th order statistic, the asymptotic normality holds for $\hat{\mu}$. Moreover, if $a'(n)$ and $b'(n)$ are chosen so that $a(n)^{-1}[F(b'(n) + ta'(n)) - b(n)] \to t$, $n \to \infty$, $Pr[a'(n)^{-1}(x_{(m)} - b'(n)) \leq t] = \Phi(t) + o(1)$, where $\Phi(t)$ is the standard normal distribution function. See Sec. 4.1 of Reiss (1989) for details. In particular, by taking $a(n) = p(1-p)/\sqrt{n}$, $b(n) = p$, $a'(n) = a(n)/F'(F^{-1}(p))$ and $b'(n) = F^{-1}$, and notice that $F^{-1}(p) = \mu$ and $F'(F^{-1}(p)) = \frac{p(1-p)}{\sigma}$, we have $\sqrt{n}\frac{\hat{\mu}-\mu}{\sigma} \sim N(0, 1)$ asymptotically. Next, using the consistency of $\hat{p}$ and $\hat{\sigma}$, we have $\sqrt{n}\frac{\hat{\mu}-\mu}{\hat{\sigma}} \sim N(0, 1)$ asymptotically.

From the likelihood equation for $\hat{\sigma}$, using the central limit theorem and note that $E\rho_p(X - \mu) = \sigma$, $var\,\rho_p(X - \mu) = \sigma^2$, we know that $\hat{\sigma}$ is asymptotic normality for known p and μ. Then taking advantage of $\hat{p} \to^P p$, $\hat{\mu} \to^P \mu$, we have that $\sqrt{n}\frac{\hat{\sigma}-\sigma}{\hat{\sigma}} \sim N(0, 1)$ asymptotically.

When $p \neq \frac{1}{2}$, using $\hat{p} = \frac{\hat{\sigma}}{\bar{x}-\hat{\mu}} + \frac{\hat{a}}{\bar{x}-\hat{\mu}}$, and note that $\bar{x} - \hat{\mu} \to^P \frac{\sigma(1-2p)}{p(1-p)}$, $\hat{a} \to^P -\frac{p}{1-p}\sigma$, so that both $\hat{p} - p$ and $\frac{p(1-p)}{1-2p}(\frac{\hat{\sigma}}{\sigma} - 1)$ have asymptotically identical distributions, thus $\sqrt{n}(\hat{p} - p) \sim N(0, \frac{p^2(1-p)^2}{(1-2p)^2})$ asymptotically.

Acknowledgments

We thank Dr. Alan Watkins and a referee for thoughtful comments which led to a substantial improvement of the article.

References

Bland, M. (2000). *An Introduction to Medical Statistics.* 3rd ed. Oxford: Oxford University Press.

Fernández , G., Steel, M. F. J. (1998). On Bayesian modelling of fat tails and skewness. *J. Amer. Statist. Assoc.* 93(1):359–371.

Gilchrist, W. G. (2000). *Statistical Modelling with Quantile Function*. London: Chapman & Hall.

Hsu, D. A. (1978). Long-tailed distributions for position errors in navigation. *Appl. Statist.* 28(1):62–72.

Johnson, N. L., Kotz, S., Balakrishnan, N. (1995). *Distributions in Statistics, Volume 2: Continuous Univariate Distributions*. 2nd ed. New York: John Wiley.

Koenker, R., Machado, J. (1999). Goodness of fit and related inference processes for quantile regression. *J. Amer. Statist. Assoc.* 94(3):1296–1309.

Kotz, S., Kozabowski, T. J., Podgórski, K. (2001). *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Boston: Birkhäuser.

Kotz, S., Kozabowski, T. J., Podgórski, K. (2002). Maximum likelihood estimation of asymmetric Laplace parameters. *Ann. Inst. Statist. Math.* 54(2):816–826.

Puig, P., Stephens, M. A. (2000). Tests of fit for the Laplace distribution, with applications. *Technometrics* 42:417–424.

Yu, K., Moyeed, R. A. (2001). Bayesian quantile regression. *Statist. Probab. Lett.* 54:437–447.

Yu, K., Lu, Z., Stander, J. (2003). Quantile regression: applications and current research area. *The Statistician* 52(2):331–350.