Two-Piece Normal-Laplace Distribution: Properties, Estimation and Applications

A. Ardalan

Department of Statistics
Shiraz University
arash@stat.auckland.ac.nz
Joint with:
Sadooghi-Alvandi S. M. , Nematollahi A. R. and Mardnifard, H.A.

Introduction

- In the last several decades, various form of skew and asymmetric distribution have appeared in the literature, see McGill (1962), Azzalini(1985,1986), John(1982), Fernandez & Steel (1998), Zhu and Walsh (2009).
- In fact, this topic has many various applications in the theory and applied statistics such as biostatistics, economics, physics, geology, and even psychology and etc.
- Two important classes of these split families are,
  - the two-piece normal family.
  - the two-piece Laplace family.

  Here we introduce,
- two-piece normal-Laplace (TPNL) distribution as a two-piece skew distribution, and we investigate some properties, applications and estimation of it .

Two-Piece Normal Family

- John (1982) introduced the two-piece split normal distribution, also Mudholkar and Hutson (2000) considered another representation of this family, with density

Epsilon-Skew Normal (ESN)

$$f(x, \mu, \sigma, \epsilon) = \frac{1}{\sigma\sqrt{2\pi}} \begin{cases} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2(1+\epsilon)^2}\right), & x \leq \mu \\ \exp\left(-\frac{(x-\mu)^2}{2\sigma^2(1-\epsilon)^2}\right), & x > \mu \end{cases} \tag{1}$$

where $\sigma > 0$, $-1 \leq \epsilon \leq 1$, $\mu \in \mathbb{R}$.

- The ESN can also be reparametrized, $p = (1+\epsilon)/2$, in the form μ is the $p\mathrm{th}$ quantile of X; i.e. $P(X \leq \mu) = p$. Here we have called it two-piece normal (TPN) distribution.
- This distribution can be interpreted as a skew distribution with short tails. The most important application of this class is in "*Expectile Regression*", see Newey and Powell (1987).

Two-Piece Laplace Family

Statistical literature seems to reveal many asymmetric forms of the Laplace distribution.

- One of the oldest forms considered by McGill (1962). In addition, Holla and Bhattacharya (1968) and also Hinkley and Revankar (1977) have considered another form of the two-piece Laplace distribution.
- One of the important representation, considered by Koeneker and Machado (1999) in the context of "*Quantile Regression*",

Two-Piece Laplace (TPL)

$$f(x,\mu,\sigma,p) = \frac{p(1-p)}{\sigma}\begin{cases} \exp\left(-\frac{|x-\mu|}{\sigma}(1-p)\right), & x \le \mu \\ \exp\left(-\frac{|x-\mu|}{\sigma}p\right), & x > \mu \end{cases}, \tag{2}$$

where $0 < p < 1$, $\sigma > 0$, and $\mu \in \mathbb{R}$, is the $p\mathrm{th}$ quantile of X.

- For other representations of the two-piece Laplace distribution you can see Kotz et al. (2001) and also Kozubowski and Nadarajah (2010).
- The most important application of this class is in *Quantile Regression*.

General Form of Two-Piece Generalized Exponential

- The more general class of two-piece exponential distribution has been defined:

Two-Piece Generalized Exponential

$$f(x;k_1,k_2,p,\mu,\sigma)=\frac{C}{\sigma}\begin{cases}\exp\{-\left|\frac{x-\mu}{\sigma}\right|^{k_1}g_1(p)\} & x\le\mu,\\ \exp\{-\left|\frac{x-\mu}{\sigma}\right|^{k_2}g_2(p)\} & x>\mu,\end{cases}\tag{3}$$

where $-\infty<\mu<\infty$, $\sigma>0$, $g_i(p)>0$, $i=1,2$, $k_i>0$, $i=1,2$

- The functions $g_i(p)>0, i=1,2$ are the measures that by use of them one can determine the location parameter μ as a statistical tendency such as the *quantiles* (median), the *expectiles* (mean) and etc.
- C, is the normalized coefficient that must be satisfy the following condition:

$$\frac{1}{C}=\frac{\Gamma(1/k_1)}{k_1g_1(p)^{1/k_1}}+\frac{\Gamma(1/k_2)}{k_2g_2(p)^{1/k_2}}$$

- Zhu and Walsh (2009) have discussed some properties of this distribution and also investigated some application of it in economic.

Two-Piece Normal-Laplace (TPNL)

- Here, we introduce the two-piece normal-Laplace (TPNL) distributions, with density

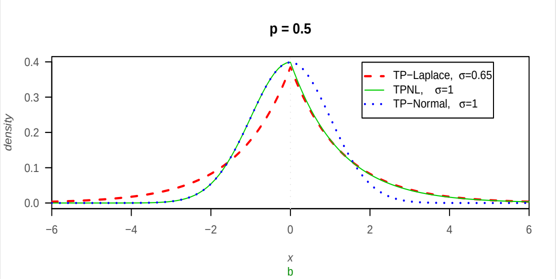
Two-Piece Normal-Laplace (TPNL)

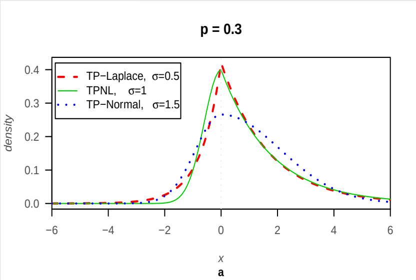
$$f(x;\mu,\ \sigma,p)=\frac{1}{\sigma\sqrt{2\pi}}\begin{cases}\exp\left(-\frac{(x-\mu)^2}{8p^2\sigma^2}\right), & x\le\mu\\ \exp\left(-\frac{(x-\mu)}{(1-p)\sigma\sqrt{2\pi}}\right), & x>\mu,\end{cases}\tag{4}$$

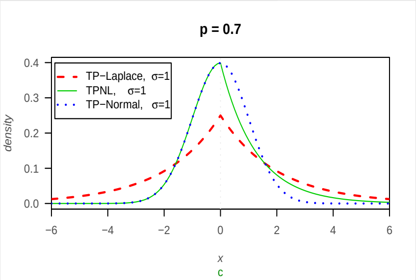
where $0<p<1$, $\sigma>0$, and $\mu\in\mathbb{R}$.

- This distribution consist of two pieces, one piece is half normal and other piece is exponential distribution.
- It provides a better fit in many applications.
- Note that in this parametrization, μ is the $p\text{th}$ quantile of X; i.e.
- $P(X\le\mu)=p$.
- It is equivalent to $k_1=2$, $k_2=1$, $g_1(p)=1/(8p^2)$ and $g_2(p)=1/(\sqrt{\pi}(1-p))$ in two-piece generalized exponential.

TPNL Graph and Comparison with Two-Piece Normal and Laplace







Basic Statistical Properties

- The standard two pieces normal Laplace distribution, $TPNL(0,1,p)$, is a mixture of a half-normal distribution and an exponential distribution.
- $T = (1-U)(1-p)E - U2p|N|$ where U, N, E are independent, and $P(U=1) = p = 1 - P(U=0)$, and N is a standard normal and E is a standard exponential distribution with parameter $\sqrt{2\pi}$.
- If $X \sim TPNL(\mu,\sigma,p)$ then $W = -X$ is distributed as $TPNL(-\mu,\sigma,1-p)$, and we namely it two-pieces Laplace-normal (TPLN) distribution.
- Furthermore, if $Y \sim TPNL(0,1,p)$ we can consider the model $X = \sigma Y + \mu$ and X has distribution $TPNL(\mu,\sigma,p)$.
- In addition if $X \sim TPNL(\mu,\sigma,p)$ and, $W = \alpha + \beta X$ then if $\beta > 0$, $W \sim TPNL(\alpha+\beta\mu,\beta\sigma,p)$, if $\beta < 0$, $W \sim TPLN(\alpha+\beta\mu,|\beta|\sigma,1-p)$.

Basic Statistical Properties

$$E\left[Y\right] = -\frac{4p^2}{\sqrt{2\pi}} + (1-p)^2\sqrt{2\pi}, \tag{5}$$

$$Var\left(Y\right) = 4\left(p^3 + \pi(1-p)^3\right) - \left[-\frac{4p^2}{\sqrt{2\pi}} + (1-p)^2\sqrt{2\pi}\right]^2, \tag{6}$$

$$E\left[Y^m\right] = \frac{1}{\sqrt{2\pi}}\left\{\frac{(-1)^m}{2}\Gamma\left(\frac{m+1}{2}\right)\left(8p^2\right)^{\frac{m+1}{2}} + \Gamma\left(m+1\right)\ \left[(1-p)\sqrt{2\pi}\ \right]^{m+1}\right\}$$

for $m = 1,\ 2, \ldots$.
In addition

$$-1 < \text{skewness} < 2 \qquad 0.59 < \text{kurtosis} < 6$$

also

$$\text{skewness}\left(0.72\right) = 0$$

Illustrate the usefulness of the TPNL

- In this section, we illustrate the usefulness of the TPNL family by showing that it provides a suitable fit to some real data.
- Comparison TPNL with the two-piece normal, given by (1), and the two-piece Laplace distribution, given by (2).

- For each family, the three parameters were estimated by the method of maximum likelihood and the fit was assessed by the Kolmogorov-Smirnov (KS) statistic.

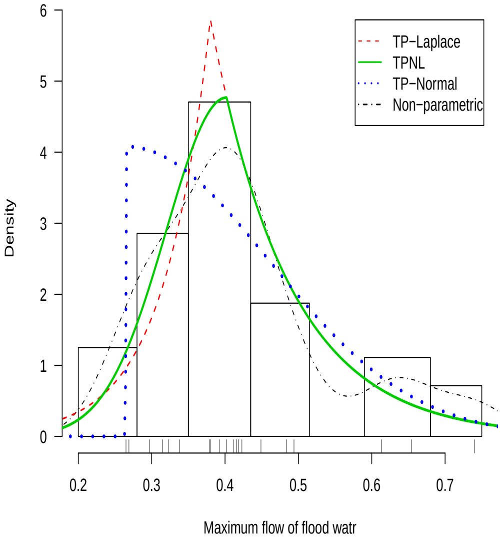
- For each example we also considered the non-parametric estimator of the density function and then compared the fit of the estimated density with the nonparametric density estimator by computing absolute error distance (AED).
- $$\mathrm{AED} = \sum_{i=1}^{n} |\hat{f}(x_i) - \tilde{f}(x_i)| \qquad (7)$$

  where $\widehat{f}$ and $\tilde{f}$ are the parametric and non-parametric estimates of the density.

Flow Data

- The data concern the maximum flow of flood water on a river for a period of 20 years., in million cu. ft. per second. The flood data taken from Gilchrist (2000, Table 1.1).
- Yu and Zhang (2006) applied the two-piece Laplace distribution to this data.
- Although all the three families seem to fit the data, it appears that the TPNL gives the best fit.
-

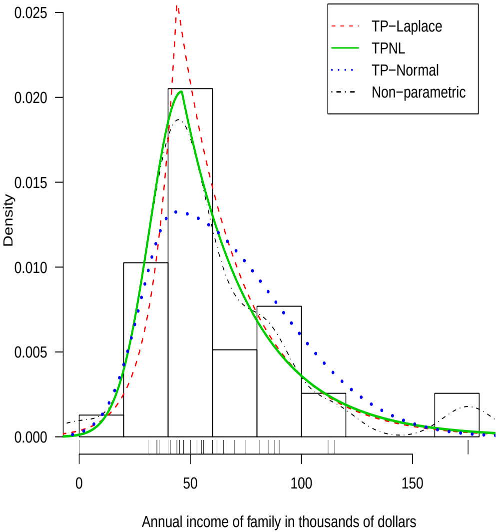
Dist	K-S	AED
A-Laplace	0.81	199.43
TPNL	0.92	145.26
A-normal	0.84	247.72



Happy Data

- This set of data were collected to test of conditions that "love" and "work" are the important factors for an individual's happiness (Happy data) that was used by George and McCulloch (1993).
- We use the distribution of "income" (the annual income of 40 families in thousands of dollars).
- The TPNL distribution is more suitable for this data.
-

Dist	K-S	AED
A-Laplace	0.20	0.53
TPNL	0.66	0.28
A-normal	0.10	0.70

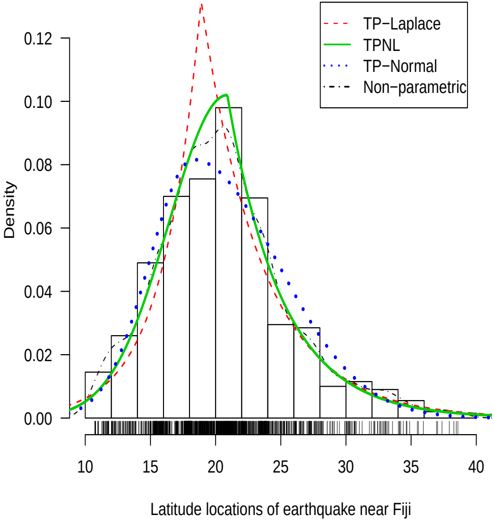


Earthquack Data

- The data is the distribution on latitude of 1000 seismic events of MB > 4.0. The events occurred in a cube near Fiji since 1964, see Wasserman (2006).

- It seems TPNL is more suitable for this data.

Dist	K-S	AED
A-Laplace	0.02	2.62
TPNL	0.21	1.16
A-normal	0.01	2.20



Maximum Likelihood Estimation

In this part:

- First, we discuss log-likelihood function of TPNL distribution and we investigate the properties of it.
- After that, we present an algorithm for finding Maximum Likelihood Estimation (MLE) for the parameters of the TPNL distribution.
- Then, we will prove the MLEs are consistent.
- Finally, we will prove the MLEs have the asymptotically normal distribution.

Likelihood Function and Properties

- $$\ell(\mathbf{X},\theta) = -n\ln(\sigma) \; - \sum_{i=1}^{n}\left[\frac{(X_i-\mu)^2}{8\sigma^2 p^2} I_{[X_i\leq\mu]} + \frac{(X_i-\mu)}{\sigma(1-p)\sqrt{2\pi}} I_{[X>\mu]}\right]$$
- The log-likelihood function is not well behaved (in particular it is not differentiable at $\mu = X_i$), the same problem is occurred in two-piece Laplace distribution, see Hinkley and Revankar(1977), and Kotz et all(2002).

Likelihood Function and Properties

- $$\ell\left(\mathbf{X},\theta\right) = -n\ln(\sigma) \ - \sum_{i=1}^{n}\left[\frac{\left(X_i-\mu\right)^2}{8\sigma^2 p^2} I_{[X_i\le\mu]} + \frac{\left(X_i-\mu\right)}{\sigma\left(1-p\right)\sqrt{2\pi}} I_{[X>\mu]}\right]$$
- The log-likelihood function is not well behaved (in particular it is not differentiable at $\mu = X_i$), the same problem is occurred in two-piece Laplace distribution, see Hinkley and Revankar(1977), and Kotz et all(2002).
- Another problem here is that the $\ell\left(\mathbf{X},\theta\right)$ is not a convex function in $\theta = (\mu,\ \sigma, p)$. However, it will be convex in each parameter for given other two parameters.

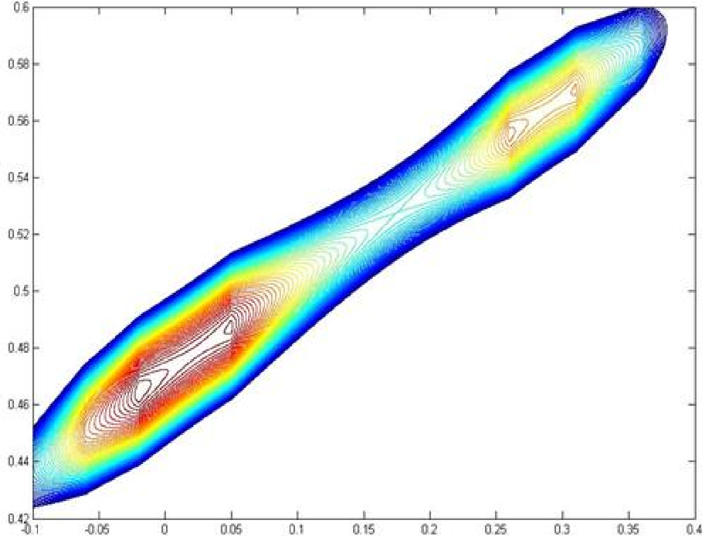
Likelihood Function and Properties

- $$\ell\left(\mathbf{X},\theta\right)=-n\ln(\sigma)\ -\sum_{i=1}^{n}\left[\frac{\left(X_i-\mu\right)^2}{8\sigma^2p^2}I_{[X_i\leq\mu]}+\frac{\left(X_i-\mu\right)}{\sigma\left(1-p\right)\sqrt{2\pi}}I_{[X>\mu]}\right]$$
- The log-likelihood function is not well behaved (in particular it is not differentiable at $\mu=X_i$), the same problem is occurred in two-piece Laplace distribution, see Hinkley and Revankar(1977), and Kotz et all(2002).
- Another problem here is that the $\ell\left(\mathbf{X},\theta\right)$ is not a convex function in $\theta=(\mu,\ \sigma,p)$. However, it will be convex in each parameter for given other two parameters.
- Therefore we have difficulty for finding the ML estimators. Zhu and Walsh (2009) have proposed that "*fmincon*" command in Matlab package, can find the MLE for parameters of the AEP distribution (which contains the TPNL family, too), there are many examples which show that "*fmincon*" command fails to find the global maximum even with excellent initial value.

Likelihood Function and Properties

- $$\ell(\mathbf{X},\theta) = -n\ln(\sigma) \ - \sum_{i=1}^{n}\left[\frac{(X_i-\mu)^2}{8\sigma^2 p^2} I_{[X_i\leq\mu]} + \frac{(X_i-\mu)}{\sigma(1-p)\sqrt{2\pi}} I_{[X>\mu]}\right]$$
- The log-likelihood function is not well behaved (in particular it is not differentiable at $\mu = X_i$), the same problem is occurred in two-piece Laplace distribution, see Hinkley and Revankar(1977), and Kotz et all(2002).
- Another problem here is that the $\ell(\mathbf{X},\theta)$ is not a convex function in $\theta = (\mu,\ \sigma, p)$. However, it will be convex in each parameter for given other two parameters.
- Therefore we have difficulty for finding the ML estimators. Zhu and Walsh (2009) have proposed that "*fmincon*" command in Matlab package, can find the MLE for parameters of the AEP distribution (which contains the TPNL family, too), there are many examples which show that "*fmincon*" command fails to find the global maximum even with excellent initial value.
- In the following, we present an efficient algorithm for finding MLE.

A Contour Plot of the Log-Likelihood Function for Given σ



An Algorithm for Finding MLE

We propose an algorithm, for finding MLE, but before we need a lemma, some notations and a Theorem.

- Lemma 4.1. Let $X_1, \ldots, X_n$ be a sample from $TPNL(\mu,\ \sigma,\ p)$; and for given $\sigma,\ p$ $\widehat{\mu}$ should lie in the interval $\left[\mathrm{x}_{(1)},\ \ \mathrm{x}_{(\mathrm{n})}\right]$.

An Algorithm for Finding MLE

We propose an algorithm, for finding MLE, but before we need a lemma, some notations and a Theorem.

- Lemma 4.1. Let $X_1, \ldots, X_n$ be a sample from $TPNL(\mu,\ \sigma,\ p)$; and for given $\sigma,\ p$ $\widehat{\mu}$ should lie in the interval $\left[\mathrm{x}_{(1)},\ \mathrm{x}_{(\mathrm{n})}\right]$.

- In addition consider:

An Algorithm for Finding MLE

We propose an algorithm, for finding MLE, but before we need a lemma, some notations and a Theorem.

- Lemma 4.1. Let $X_1, \ldots, X_n$ be a sample from $TPNL(\mu,\ \sigma,\ p)$; and for given $\sigma,\ p$ $\widehat{\mu}$ should lie in the interval $\left[\mathrm{x}_{(1)},\ \ \mathrm{x}_{(\mathrm{n})}\right]$.

- In addition consider:
- $\overline{x}_{(j)} = j^{-1} \sum_{i=1}^{j} x_{(i)}$, $x_{(i)}$ is the ith order statistics

An Algorithm for Finding MLE

We propose an algorithm, for finding MLE, but before we need a lemma, some notations and a Theorem.

- Lemma 4.1. Let $X_1, \ldots, X_n$ be a sample from $TPNL(\mu,\ \sigma,\ p)$; and for given $\sigma,\ p\ \widehat{\mu}$ should lie in the interval $\left[\mathrm{x}_{(1)},\ \ \mathrm{x}_{(\mathrm{n})}\right]$.

- In addition consider:
- $\overline{x}_{(j)} = j^{-1}\sum_{i=1}^{j} x_{(i)}$, $x_{(i)}$ is the ith order statistics

- ℓ_{μ^-} and ℓ_{μ^+} are the left and right derivatives of the log-likelihood function with respect to μ, respectively.

An Algorithm for Finding MLE

We propose an algorithm, for finding MLE, but before we need a lemma, some notations and a Theorem.

- Lemma 4.1. Let $X_1, \ldots, X_n$ be a sample from $TPNL\,(\mu,\ \sigma,\ p)$; and for given $\sigma,\ p$ $\widehat{\mu}$ should lie in the interval $\left[\mathrm{x}_{(1)},\ \mathrm{x}_{(\mathrm{n})}\right]$.

- In addition consider:
- $\overline{x}_{(j)} = j^{-1} \sum_{i=1}^{j} x_{(i)}$, $x_{(i)}$ is the ith order statistics

- ℓ_{μ^-} and ℓ_{μ^+} are the left and right derivatives of the log-likelihood function with respect to μ, respectively.

- $A_n = A_n\,(\mu, p) = \frac{1}{8p^2} \sum_{i=1}^{n} (x_i - \mu)^2 I_{[x_i \le \mu]}$

An Algorithm for Finding MLE

We propose an algorithm, for finding MLE, but before we need a lemma, some notations and a Theorem.

- Lemma 4.1. Let $X_1, \ldots, X_n$ be a sample from $TPNL\left(\mu,\ \sigma,\ p\right)$; and for given $\sigma,\ p\ \widehat{\mu}$ should lie in the interval $\left[\mathrm{x}_{(1)},\ \ \mathrm{x}_{(\mathrm{n})}\right]$.
- In addition consider:
- $\overline{x}_{(j)} = j^{-1}\sum_{i=1}^{j} x_{(i)}$, $x_{(i)}$ is the ith order statistics
- ℓ_{μ^-} and ℓ_{μ^+} are the left and right derivatives of the log-likelihood function with respect to μ, respectively.
- $A_n = A_n\left(\mu, p\right) = \frac{1}{8p^2}\sum_{i=1}^{n}\left(x_i - \mu\right)^2 I_{[x_i \leq \mu]}$
- $B_n = B_n\left(\mu, p\right) = \frac{1}{(1-p)\sqrt{2\pi}}\sum_{i=1}^{n}\left(x_i - \mu\right) I_{[x_i > \mu]}$.

An Algorithm for Finding MLE

We propose an algorithm, for finding MLE, but before we need a lemma, some notations and a Theorem.

- Lemma 4.1. Let $X_1, \ldots, X_n$ be a sample from $TPNL(\mu,\ \sigma,\ p)$; and for given $\sigma,\ p$ $\widehat{\mu}$ should lie in the interval $\left[\mathrm{x}_{(1)},\ \ \mathrm{x}_{(\mathrm{n})}\right]$.

- In addition consider:
- $\overline{x}_{(j)} = j^{-1}\sum_{i=1}^{j} x_{(i)}$, $x_{(i)}$ is the ith order statistics

- ℓ_{μ^-} and ℓ_{μ^+} are the left and right derivatives of the log-likelihood function with respect to μ, respectively.

- $A_n = A_n(\mu, p) = \frac{1}{8p^2}\sum_{i=1}^{n}(x_i - \mu)^2 I_{[x_i \leq \mu]}$
- $B_n = B_n(\mu, p) = \frac{1}{(1-p)\sqrt{2\pi}}\sum_{i=1}^{n}(x_i - \mu)\, I_{[x_i > \mu]}$.
- $W_n = \frac{\sqrt{2\pi}}{4\sigma}\frac{\sum_{i=1}^{n}(x_i-\mu)^2 I_{[x_i \leq \mu]}}{\sum_{i=1}^{n}(x_i-\mu) I_{[x_i > \mu]}} \geq 0.$

Theorem 4.2. Let $X_1, X_2, \ldots, X_n$ be a sample from TPNL distribution with parameters μ, σ, and p. Then

(a). If σ and p are known, then

$$\widehat{\mu} = \widehat{\mu}_0(\sigma, p) = \begin{cases} x_{(j)} & \ell_{\mu^+}\left(x_{(j)}\right) \geq 0 \\ \frac{(n-j+1)}{j-1}\frac{4p^2\sigma}{(1-p)\sqrt{2\pi}} + \overline{x}_{(j-1)} & \ell_{\mu^-}\left(x_{(j)}\right) < 0 \end{cases}$$

j is the first number for which $\ell_{\mu^+}\left(x_{(j)}\right) < 0$.

(b). If μ and p are known, then

$$\widehat{\sigma} = \widehat{\sigma}_0(\mu, p) = \frac{1}{2n}\left(B_n + \sqrt{B_n^2 + 8nA_n}\right)$$

(c). If μ and σ are known, then the MLE of p is given by the (unique) root of the equation $p^3 - W_n(1-p)^2 = 0$. More precisely,

$$\hat{p} = \hat{p}_0(\mu, \sigma) = \begin{cases} 0 & \mu = x_{(1)} \\ \left\{p :\ p^3 - W_n(1-p)^2 = 0, \quad 0 < p < 1\right\} & x_{(1)} < \mu < x_{(n)} \\ 1 & \mu = x_{(n)} \end{cases},$$

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

1. $(\mu_{0,j}, p_{0,j}) \leftarrow \left(x_{(j)}, \frac{j}{n}\right)$

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

1. $(\mu_{0,j}, p_{0,j}) \longleftarrow \left(x_{(j)}, \frac{j}{n}\right)$
2. $\sigma_{0,j} \longleftarrow \widehat{\sigma}_0\left(\mu_{0,j}, p_{0,j}\right)$

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

1. $\left(\mu_{0,j}, p_{0,j}\right) \longleftarrow \left(x_{(j)}, \frac{j}{n}\right)$
2. $\sigma_{0,j} \longleftarrow \widehat{\sigma}_0\left(\mu_{0,j}, p_{0,j}\right)$
3. $i \longleftarrow 1$

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

1. $\left(\mu_{0,j}, p_{0,j}\right) \longleftarrow \left(x_{(j)}, \frac{j}{n}\right)$
2. $\sigma_{0,j} \longleftarrow \widehat{\sigma}_0\left(\mu_{0,j}, p_{0,j}\right)$
3. $i \longleftarrow 1$
4. While [stable solutions] do the following:

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

1. $(\mu_{0,j}, p_{0,j}) \leftarrow \left(x_{(j)}, \frac{j}{n}\right)$
2. $\sigma_{0,j} \leftarrow \widehat{\sigma}_0\left(\mu_{0,j}, p_{0,j}\right)$
3. $i \leftarrow 1$
4. While [stable solutions] do the following:

    a) $\mu_{i,j} \leftarrow \widehat{\mu}_0\left(\sigma_{i-1,j}, p_{i-1,j}\right)$

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

1. $(\mu_{0,j}, p_{0,j}) \longleftarrow \left(x_{(j)}, \frac{j}{n}\right)$
2. $\sigma_{0,j} \longleftarrow \widehat{\sigma}_0\left(\mu_{0,j}, p_{0,j}\right)$
3. $i \longleftarrow 1$
4. While [stable solutions] do the following:
   a) $\mu_{i,j} \longleftarrow \widehat{\mu}_0\left(\sigma_{i-1,j}, p_{i-1,j}\right)$
   b) $\sigma_{i,j} \longleftarrow \widehat{\sigma}_0\left(\mu_{i,j}, p_{i-1,j}\right)$

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

1. $\left(\mu_{0,j}, p_{0,j}\right) \longleftarrow \left(x_{(j)}, \frac{j}{n}\right)$
2. $\sigma_{0,j} \longleftarrow \widehat{\sigma}_0\left(\mu_{0,j}, p_{0,j}\right)$
3. $i \longleftarrow 1$
4. While [stable solutions] do the following:

   a) $\mu_{i,j} \longleftarrow \widehat{\mu}_0\left(\sigma_{i-1,j}, p_{i-1,j}\right)$

   b) $\sigma_{i,j} \longleftarrow \widehat{\sigma}_0\left(\mu_{i,j}, p_{i-1,j}\right)$

   c) $p_{i,j} \longleftarrow \hat{p}_0\left(\mu_{i,j}, \sigma_{i,j}\right)$

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

1. $\left(\mu_{0,j}, p_{0,j}\right) \leftarrow \left(x_{(j)}, \frac{j}{n}\right)$
2. $\sigma_{0,j} \leftarrow \widehat{\sigma}_0\left(\mu_{0,j}, p_{0,j}\right)$
3. $i \leftarrow 1$
4. While [stable solutions] do the following:
    a) $\mu_{i,j} \leftarrow \widehat{\mu}_0\left(\sigma_{i-1,j}, p_{i-1,j}\right)$
    b) $\sigma_{i,j} \leftarrow \widehat{\sigma}_0\left(\mu_{i,j}, p_{i-1,j}\right)$
    c) $p_{i,j} \leftarrow \hat{p}_0\left(\mu_{i,j}, \sigma_{i,j}\right)$
    d) $i \leftarrow i+1$

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

1. $(\mu_{0,j}, p_{0,j}) \leftarrow \left(x_{(j)}, \frac{j}{n}\right)$
2. $\sigma_{0,j} \leftarrow \widehat{\sigma}_0 \left(\mu_{0,j}, p_{0,j}\right)$
3. $i \leftarrow 1$
4. While [stable solutions] do the following:
   a) $\mu_{i,j} \leftarrow \widehat{\mu}_0 \left(\sigma_{i-1,j}, p_{i-1,j}\right)$
   b) $\sigma_{i,j} \leftarrow \widehat{\sigma}_0 \left(\mu_{i,j}, p_{i-1,j}\right)$
   c) $p_{i,j} \leftarrow \hat{p}_0 \left(\mu_{i,j}, \sigma_{i,j}\right)$
   d) $i \leftarrow i + 1$
5. $(\widehat{\mu}_j, \widehat{\sigma}_j, \hat{p}_j) \leftarrow \left(\mu_{i-1,j}, \sigma_{i-1,j}, p_{i-1,j}\right)$

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

1. $(\mu_{0,j}, p_{0,j}) \longleftarrow \left(x_{(j)}, \frac{j}{n}\right)$
2. $\sigma_{0,j} \longleftarrow \widehat{\sigma}_0\left(\mu_{0,j}, p_{0,j}\right)$
3. $i \longleftarrow 1$
4. While [stable solutions] do the following:
   a) $\mu_{i,j} \longleftarrow \widehat{\mu}_0\left(\sigma_{i-1,j}, p_{i-1,j}\right)$
   b) $\sigma_{i,j} \longleftarrow \widehat{\sigma}_0\left(\mu_{i,j}, p_{i-1,j}\right)$
   c) $p_{i,j} \longleftarrow \hat{p}_0\left(\mu_{i,j}, \sigma_{i,j}\right)$
   d) $i \longleftarrow i+1$
5. $(\widehat{\mu}_j, \widehat{\sigma}_j, \hat{p}_j) \longleftarrow (\mu_{i-1,j}, \sigma_{i-1,j}, p_{i-1,j})$
6. $\ell_j \longleftarrow \ell\left(\mathbf{x};; \widehat{\mu}_j, \widehat{\sigma}_j, \hat{p}_j\right)$

Algorithm.

I- for $j = 2, \dots, n-1$ do the following:

1. $(\mu_{0,j}, p_{0,j}) \leftarrow \left(x_{(j)}, \frac{j}{n}\right)$
2. $\sigma_{0,j} \leftarrow \widehat{\sigma}_0\left(\mu_{0,j}, p_{0,j}\right)$
3. $i \leftarrow 1$
4. While [stable solutions] do the following:
   a) $\mu_{i,j} \leftarrow \widehat{\mu}_0\left(\sigma_{i-1,j}, p_{i-1,j}\right)$
   b) $\sigma_{i,j} \leftarrow \widehat{\sigma}_0\left(\mu_{i,j}, p_{i-1,j}\right)$
   c) $p_{i,j} \leftarrow \hat{p}_0\left(\mu_{i,j}, \sigma_{i,j}\right)$
   d) $i \leftarrow i + 1$
5. $(\widehat{\mu}_j, \widehat{\sigma}_j, \hat{p}_j) \leftarrow \left(\mu_{i-1,j}, \sigma_{i-1,j}, p_{i-1,j}\right)$
6. $\ell_j \leftarrow \ell\left(\mathbf{x};; \widehat{\mu}_j, \widehat{\sigma}_j, \hat{p}_j\right)$

II- -$\ell_1 \leftarrow \ell\left(\mathbf{x}; x_{(1)}, \widehat{\sigma}_1, 0\right)$ & $\ell_n \leftarrow \ell\left(\mathbf{x}; x_{(n)}, \widehat{\sigma}_n, 1\right)$

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

1. $(\mu_{0,j}, p_{0,j}) \leftarrow \left(x_{(j)}, \frac{j}{n}\right)$
2. $\sigma_{0,j} \leftarrow \widehat{\sigma}_0\left(\mu_{0,j}, p_{0,j}\right)$
3. $i \leftarrow 1$
4. While [stable solutions] do the following:
   - a) $\mu_{i,j} \leftarrow \widehat{\mu}_0\left(\sigma_{i-1,j}, p_{i-1,j}\right)$
   - b) $\sigma_{i,j} \leftarrow \widehat{\sigma}_0\left(\mu_{i,j}, p_{i-1,j}\right)$
   - c) $p_{i,j} \leftarrow \hat{p}_0\left(\mu_{i,j}, \sigma_{i,j}\right)$
   - d) $i \leftarrow i+1$
5. $(\widehat{\mu}_j, \widehat{\sigma}_j, \hat{p}_j) \leftarrow \left(\mu_{i-1,j}, \sigma_{i-1,j}, p_{i-1,j}\right)$
6. $\ell_j \leftarrow \ell\left(\mathbf{x};; \widehat{\mu}_j, \widehat{\sigma}_j, \hat{p}_j\right)$

II- -$\ell_1 \leftarrow \ell\left(\mathbf{x}; x_{(1)}, \widehat{\sigma}_1, 0\right)$ & $\ell_n \leftarrow \ell\left(\mathbf{x}; x_{(n)}, \widehat{\sigma}_n, 1\right)$

III- - $j \leftarrow \operatorname{argmin}\left(\{\ell_1, \ell_2, \ldots, \ell_n\}\right)$

Algorithm.

I- for $j = 2, \ldots, n-1$ do the following:

1. $(\mu_{0,j}, p_{0,j}) \leftarrow \left(x_{(j)}, \frac{j}{n}\right)$
2. $\sigma_{0,j} \leftarrow \widehat{\sigma}_0\left(\mu_{0,j}, p_{0,j}\right)$
3. $i \leftarrow 1$
4. While [stable solutions] do the following:
   - a) $\mu_{i,j} \leftarrow \widehat{\mu}_0\left(\sigma_{i-1,j}, p_{i-1,j}\right)$
   - b) $\sigma_{i,j} \leftarrow \widehat{\sigma}_0\left(\mu_{i,j}, p_{i-1,j}\right)$
   - c) $p_{i,j} \leftarrow \hat{p}_0\left(\mu_{i,j}, \sigma_{i,j}\right)$
   - d) $i \leftarrow i+1$
5. $(\widehat{\mu}_j, \widehat{\sigma}_j, \hat{p}_j) \leftarrow \left(\mu_{i-1,j}, \sigma_{i-1,j}, p_{i-1,j}\right)$
6. $\ell_j \leftarrow \ell\left(\mathbf{x};; \widehat{\mu}_j, \widehat{\sigma}_j, \hat{p}_j\right)$

II- $-\ell_1 \leftarrow \ell\left(\mathbf{x}; x_{(1)}, \widehat{\sigma}_1, 0\right)$ & $\ell_n \leftarrow \ell\left(\mathbf{x}; x_{(n)}, \widehat{\sigma}_n, 1\right)$

III- - $j \leftarrow \operatorname{argmin}\left(\{\ell_1, \ell_2, \ldots, \ell_n\}\right)$

IV- - $(\widehat{\mu}, \widehat{\sigma}, \hat{p}) \leftarrow (\widehat{\mu}_j, \widehat{\sigma}_j, \hat{p}_j)$

Asymptotic Behavior of MLE

- In the literature, there are many theorems that establish consistency and normality of the MLE in the classical way. They usually depend on restrictions:

Asymptotic Behavior of MLE

- In the literature, there are many theorems that establish consistency and normality of the MLE in the classical way. They usually depend on restrictions:
  - Differentiability of log-likelihood,

Asymptotic Behavior of MLE

- In the literature, there are many theorems that establish consistency and normality of the MLE in the classical way. They usually depend on restrictions:
  - Differentiability of log-likelihood,
  - Convexity of log-likelihood,

Asymptotic Behavior of MLE

- In the literature, there are many theorems that establish consistency and normality of the MLE in the classical way. They usually depend on restrictions:
  - Differentiability of log-likelihood,
  - Convexity of log-likelihood,
  - Compactness of parameter space.

Asymptotic Behavior of MLE

- In the literature, there are many theorems that establish consistency and normality of the MLE in the classical way. They usually depend on restrictions:
  - Differentiability of log-likelihood,
  - Convexity of log-likelihood,
  - Compactness of parameter space.
- The $\ell(\mathbf{X}, \theta)$ has a sharp point at $x = \mu$, (it has a cusp at $x = \mu$) and so some of the differentiability conditions are violated and also the classical regularity conditions break down.

Asymptotic Behavior of MLE

- In the literature, there are many theorems that establish consistency and normality of the MLE in the classical way. They usually depend on restrictions:
  - Differentiability of log-likelihood,
  - Convexity of log-likelihood,
  - Compactness of parameter space.
- The $\ell(\mathbf{X}, \theta)$ has a sharp point at $x = \mu$, (it has a cusp at $x = \mu$) and so some of the differentiability conditions are violated and also the classical regularity conditions break down.
- Another problem here is that the $\ell(\mathbf{X}, \theta)$ is not a convex function in $\theta = (\mu, \sigma, p)$.

Asymptotic Behavior of MLE

- In the literature, there are many theorems that establish consistency and normality of the MLE in the classical way. They usually depend on restrictions:
  - Differentiability of log-likelihood,
  - Convexity of log-likelihood,
  - Compactness of parameter space.
- The $\ell(\mathbf{X}, \theta)$ has a sharp point at $x = \mu$, (it has a cusp at $x = \mu$) and so some of the differentiability conditions are violated and also the classical regularity conditions break down.
- Another problem here is that the $\ell(\mathbf{X}, \theta)$ is not a convex function in $\theta = (\mu, \sigma, p)$.
- We prove consistency of MLE, by modified Wald's approach for nonregular case.

Asymptotic Behavior of MLE

- In the literature, there are many theorems that establish consistency and normality of the MLE in the classical way. They usually depend on restrictions:
  - Differentiability of log-likelihood,
  - Convexity of log-likelihood,
  - Compactness of parameter space.
- The $\ell(\mathbf{X}, \theta)$ has a sharp point at $x = \mu$, (it has a cusp at $x = \mu$) and so some of the differentiability conditions are violated and also the classical regularity conditions break down.
- Another problem here is that the $\ell(\mathbf{X}, \theta)$ is not a convex function in $\theta = (\mu, \sigma, p)$.
- We prove consistency of MLE, by modified Wald's approach for nonregular case.
- The asymptotic normality of MLE, will be established according to Huber (1967).

Modified Wald Consistency of the MLE

- Wald (1949) handles the non-compact set of parameters by some assumptions which ensure that the MLE $\hat{\theta}_n$ is inside a compact set eventually, almost surely, a modified of Walds consistency is given by Ghosh and Ramamoorthi (2003)

Modified Wald Consistency of the MLE

- Wald (1949) handles the non-compact set of parameters by some assumptions which ensure that the MLE $\hat{\theta}_n$ is inside a compact set eventually, almost surely, a modified of Walds consistency is given by Ghosh and Ramamoorthi (2003)
- Let $X_1, \ X_2, \ \ldots, \ X_n$ be a random sample from $f(x, \ \theta)$ and $\theta_{\mathbf{0}}$ be the true value from a parameter space $\mathbf{\Theta}$ and let $T(\theta, x) = \log\left(\frac{f(x, \ \theta)}{f(x, \ \theta_0)}\right)$ be a real value function on $\mathbf{\Theta} \times \mathbb{R}$. Suppose the following conditions satisfy:
  1. Let $\mathbf{\Theta} = \cup U_i$ where the $U_i s$ are compact and $U_1 \subset U_2 \subset \ldots$.For any sequence

  $$\theta_i \in U^c_{(i-1)} \cap U_i, \quad \lim_i f(x, \ \theta_i) = 0,$$

  2. For each x, $T(\theta, x)$ is continuous in θ, and also for each θ, $T(\theta, x)$ is measurable,
  3. Let $\varphi_i(x) = \sup_{\theta \in U^c_{(i-1)}} (T(\theta, x))$, then $E_{\theta_0}[\varphi_i^+(X)] < \infty$, for some i,

  then any MLE $\widehat{\theta}_{\mathbf{n}}$ is consistent at $\theta_{\mathbf{0}}$.

Modified Wald Consistency of MLE

- We can not use directly this theorem for the distribution which have more than one parameter, with one observation.
- So, we should reorganize the sample $x_1, \ldots, x_n$ into $\mathbf{y}_1, \ldots, \mathbf{y}_{N-1}, \mathbf{y}_N$ where $N = \left[\frac{n}{2}\right]$ and $\mathbf{y}_i = (x_{2i-1}, x_{2i})'$, $i = 1, 2, \ldots, N-1$ and

$$\mathbf{y}_N = \begin{cases} (x_{n-1}, x_n)' & n \quad is \quad even, \\ (x_{n-2}, x_{n-1}, x_n)' & n \quad is \quad odd. \end{cases}$$

  Note that the likelihood functions of the both samples are the same.
- Theorem 4.3. The $\widehat{\theta}_n$ is a consistent ML estimator for θ_0, where θ_0 is the true value parameter $\theta_0 = (\mu_0, \ \sigma_0, p_0)$, i.e., $\widehat{\theta}_n \xrightarrow{a.s} \theta_0$.

Proof:

We should check the conditions of **(Ghosh and Ramamoorthi, 2003)** for $\mathbf{y}_1, \ldots, \mathbf{y}_{N-1}, \mathbf{y}_N$.

Normality of MLE by Huber Conditions

In this part we want to show $\widehat{\theta}_n$ has asymptotic normal distribution, by the Huber approach, i.e.

$$\sqrt{n}\left(\widehat{\theta}-\theta_0\right) \xrightarrow{L} Z \in N\left(0, \mathbf{I}^{-1}(\theta_0)\right),$$

where $\mathbf{I}\left(\theta_{\mathbf{0}}\right)$ is the Fisher information matrix of TPNL. Huber (1967) assumes that, $\boldsymbol{\Theta}$ is an open subset of m-dimensional Euclidean space $\mathbb{R}^m$, $(\Omega, \{, P)$ is a probability space, and $\boldsymbol{\Psi} : \Omega \times \boldsymbol{\Theta} \to \mathbb{R}^m$ is some function. Assuming that $x_1, x_2, \ldots$ are independent identically distributed random variables with values in Ω and common distribution P, he gives sufficient conditions ensuring that every sequence $T_n = T_n\left(x_1, \ldots, x_n\right)$ satisfying,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \boldsymbol{\Psi}\left(x_i, T_n\right) \xrightarrow{P} \mathbf{0}, \tag{8}$$

is asymptotically normal.
Huber's Theorem requires that (8) and the assumptions (N-1)-(N-4) be satisfied (Refer to Huber, 1967, Section4).

Normality of MLE by Huber Conditions

We define

$$\mathbf{\Psi}(x,\theta)=\begin{bmatrix}\psi_1(x,\theta)\\ \psi_2(x,\theta)\\ \psi_3(x,\theta)\end{bmatrix}=\begin{bmatrix}\frac{1}{2}\left(g_{\mu^-}(x,\theta)+g_{\mu^+}(x,\theta)\right)\\ g_\sigma(x,\theta)\\ g_p(x,\theta)\end{bmatrix},$$

Where $g(x,\theta)=\ln f(x,\theta)$; and g_{μ^-} and g_{μ^+} be the left and right partial derivative of g with respect to μ; and g_σ and g_p be the partial derivatives of g with respect to σ and p, respectively. Note that

$$g_{\mu^-}(x,\theta)=\frac{x-\mu}{4p^2\sigma^2}I_{[x\leq\mu]}+\frac{1}{(1-p)\,\sigma\sqrt{2\pi}}I_{[x>\mu]}\ ,$$

$$g_{\mu^+}(x,\theta)=\frac{(x-\mu)}{4p^2\sigma^2}I_{[x<\mu]}+\frac{1}{(1-p)\,\sigma\sqrt{2\pi}}I_{[x\geq\mu]}.$$

Lemma 4.4. Let $\mathbf{\Psi}(x,\theta)$ is defined above, then,

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{\Psi}\left(X_i,\widehat{\theta}_n\right)\xrightarrow{p}\mathbf{0}.$$

The Sketch of Proof for Normality by Huber Approach

Let

$$\overline{\Psi^*}_{\mathrm{n}}(\theta) = \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\left[\Psi(X_i, \theta)\right]$$

has been named it smoothed objective function.
Using the mean value theorem for smoothed objective function and expanding it in the neighborhood of $\widehat{\theta}$, we have

$$0 = \overline{\Psi^*}_{\mathrm{n}}(\theta_0) = \overline{\Psi^*}_{\mathrm{n}}\left(\widehat{\theta}\right) + \frac{\partial \overline{\Psi^*}_{\mathrm{n}}\left(\overline{\theta}\right)}{\partial \theta^{\mathrm{T}}}(\theta_0 - \widehat{\theta}),$$

where $\overline{\theta}$ lies on the line segment joining θ_0 and $\widehat{\theta}$ and hence

$$\left(\theta_0 - \widehat{\theta}\right) = -\left[\frac{\partial \overline{\Psi^*}_{\mathrm{n}}\left(\overline{\theta}\right)}{\partial \theta^{\mathrm{T}}}\right]^{-1} \overline{\Psi^*}_{\mathrm{n}}\left(\widehat{\theta}\right).$$

Under suitable assumption on $\Psi(X, \theta)$, one obtains $\frac{\partial \overline{\Psi^*}_{\mathrm{n}}\left(\overline{\theta}\right)}{\partial \theta^{\mathrm{T}}} \xrightarrow{p} M$.
For the iid case, $M = -\mathbf{I}(\theta_0)$, see Huber(1969, p-231).

The Sketch of Proof for Normality by Huber Approach

Therefore, we have

$$\sqrt{n}\left(\widehat{\theta}-\theta_0\right)=\left(\mathbf{I}^{-1}+\mathrm{o_p}(1)\right)\sqrt{n}\ \overline{\Psi^*}_{\mathrm{n}}\left(\widehat{\theta}\right),$$

where $\mathrm{o_p}\,(1)$ converges in probability to zero as $n \to \infty$.
And if $\sqrt{n}\ \overline{\Psi^*}_{\mathrm{n}}\left(\widehat{\theta}\right) \xrightarrow{L} Z \in N(\mathbf{0}, \mathbf{I}(\theta_0))$, then by Slutsky Theorem

$$\sqrt{n}\left(\widehat{\theta}-\theta_0\right) \xrightarrow{L} Z \in N\left(0, \mathbf{I}^{-1}(\theta_0)\right),$$

to show it, we can write

$$-\sqrt{n}\ \overline{\Psi^*}_{\mathrm{n}}\left(\widehat{\theta}\right)=-\left(\sqrt{n}\ \overline{\Psi}_{\mathrm{n}}\left(\theta_{\mathbf{0}}\right)+\sqrt{n}\ \overline{\Psi^*}_{\mathrm{n}}\left(\widehat{\theta}\right)\right)+\sqrt{n}\ \overline{\Psi}_{\mathrm{n}}\left(\theta_{\mathbf{0}}\right)$$

By ordinary CLT $\sqrt{n}\ \overline{\Psi}_{\mathrm{n}}\left(\theta_{\mathbf{0}}\right) \xrightarrow{L} N(\mathbf{0}, \mathbf{I}(\theta_{\mathbf{0}}))$ and so if we show

$$\left(\sqrt{n}\ \overline{\Psi}_{\mathrm{n}}\left(\theta_{\mathbf{0}}\right)+\sqrt{n}\ \overline{\Psi^*}_{\mathrm{n}}\left(\widehat{\theta}\right)\right) \xrightarrow{p} 0,$$

then by Slutsky Theorem we can conclude that $-\sqrt{n}\ \overline{\Psi^*}_{\mathrm{n}}\left(\widehat{\theta}\right) \xrightarrow{L} N(\mathbf{0}, \mathbf{I}(\theta_{\mathbf{0}}))$.

Summary

- In this article we have introduced a family of split skew distribution that can be used for analyzing skew data.
- Note that this distribution can be applied for the population which from one side is shrunk (short tail) and from other side is stretched (heavy tail) and also two-piece normal and two-piece laplace unable cover it.
- Obviously much additional work is needed.

References:

- Azzalini, A. (1985) A class of distributions which includes the normal ones. Scand. J. Statist. 12, 171-178.
- Azzalini, A. (1986) Further results on a class of distributions which includes the normal ones. Statistica 46,199-208.
- Fernandez, C. and Steel, M. F. J. (1998) On Bayesian modeling of fat tails and skewness. J. Amer. Statist. Assoc., 93, 359–371.
- Gilchrist, W. G. (2000) Statistical modeling with quantile functions. Texts in Statistical Science Series. Chapman & Hall/CRC, Boca Raton, FL Distribution (Probability theory), Sampling (Statistics).
- George, E. I. and McCulloch, R. E. (1993) Variable selection Via Gibbs sampling. J. Amer. Statist. Assoc., 84 :881–889.
- Ghosh, J.K. and Ramamoorthi, R.V. (2003), *Bayesian nonparametric*, Springer-Verlag, New York, Inc
- Hinkley, D.V. and Revankar, M.S. (1977), Estimation on the Pareto law from underrereported data. Journal of Econometrics, 5, l-l1.
- Holla, M.S. and S.K. Bhattacharya, (1968), On a compound Gaussian distribution, Annals of the Institute of Statistical Mathematics Tokyo 20, 331-336.

References:

- Huber, P. (1967) The Behavior of Maximum Likelihood Estimates Under Nonstandard Conditions, in: L.M. LeCam and J. Neyman, eds., Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Berkeley: University of California Press.
- John, S. (1998) The three parameter two pieces normal family of distributions and its fitting. Comm. Statist. Theory Methods., 11, pp. 879–885.
- Koenker, R. and Machado, J. A. F. (1999) Goodness of fit and related inference processes for quantile regression. J. Amer. Statist. Assoc., 94:1296–1310.
- Kotz, S. and Kozubowski, T. J. and Podgorski K. (2001) The Laplace distribution and generalizations. Birkhauser Boston Inc., Boston, MA. Revisit with applications to communications, economics, engineering, and finance.
- Kotz, S. and Kozubowski, T. J. and Podgorski K. (2002) Maximum likelihood estimation of asymmetric Laplace parameters. Ann. Inst. Static. Math., Vol. 54, No. 4, 816-826.
- McGill, W.J., 1962, Random fluctuations of response rate, Psychometrika 27, 3-17.

References:

- Mudholkar, G. S. and Hutson, A. D. (2000) The epsilon-skew-normal distribution for analyzing near-normal data. J. Statist. Plann. Inference, 83:291–309.
- Newey, W. K. and Powell, J. L. Asymmetric least squares estimation and testing. Econ,ometrica, 55(4):819847, 1987.
- Newey, W. K. and McFadden, D. L. (1994) Large sample estimation and hypothesis testing. In Handbook of Econometrics,Vol. 4, Engle RF, McFadden DL (eds). Elsevier Science: Amsterdam; 2113–2247.
- Poiraud-Casanova, S. and Thomas-Agnan. C. (2000) About monotone regression quantiles. Statist. Probab. Lett., 48:101–104.
- Kozubowski, T. and Nadarajah, S. , 2010. "Multitude of Laplace distributions," Statistical Papers, Springer, vol. 51(1), pages 127-148.
- Wasserman, W. (2006) All of nonparametric statistics. Springer Texts in Statistics. Springer, New York.

References:

- Wald, A. (1949) Note on the consistency of maximum likelihood estimate, Annals of Mathematical Statistics 20, 595-601.
- Yu, K. and Zhang, J. Erratum (2005) A three-parameter asymmetric Laplace distribution and its extension" [Comm. Statist. Theory Methods 34 (2005), no. 9-10, 1867–1879; mr2205240]. Comm. Statist. Theory Methods, 35 (7-9):1359, 2006.
- Zhu, D. and Walsh, V. Z. (2009), Properties and Estimation of Asymmetric Exponential Power Distribution, Journal of Econometrics, 148, pp. 86-99.

Thank You