Discrete Laplace distribution and Process

Kuttykrishnan A.P “Laplace autoregressive time series models”, Department of Statistics, University of Calicut, 2006
6.1. Introduction

Kemp (1997) studied a class of discrete distribution, namely discrete normal distribution, supported on the set of integers and analogue of the normal distribution. The probability mass function of the discrete normal random variable is obtained using the method

$$P(Y = k) = \frac{f(k)}{\sum_{j=-\infty}^{\infty} f(j)}, \quad k = 0, \pm 1, \pm 2, \ldots \quad (6.1.1)$$

where $f(x)$ is the probability density function of a normal random variable.

Corresponding to any continuous random variable $X$, we can construct a discrete random variable $Y$ on integers with probability mass function given by the equation (6.1.1). Hence, a discrete random variable corresponding to a classical Laplace random variable with probability density function (1.3.1) can be defined and the resulting discrete random variable $Y$ has the probability mass function

$$P(Y = k) = \frac{\frac{1}{2\sigma} e^{-|k|/\sigma}}{\sum_{j=-\infty}^{\infty} \frac{1}{2\sigma} e^{-|j|/\sigma}},$$

On simplification, we get
Definition 6.1.1.

A random variable $Y$ is said to follow discrete Laplace distribution with parameter $p \in (0,1)$ if its probability mass function is

\[ P(Y = k) = \frac{1-p}{1+p} p^{|k|}, \text{ where } p = e^{-1/\sigma}. \]

Inusah and Kozubowski (2006) studied various properties of the discrete Laplace distribution. Some properties of discrete Laplace distribution are

(1) Let $X \sim DL(p)$, then mean and variance are given by

\[ E(X) = 0 \text{ and } V(X) = \frac{2p}{(1-p)^2} \text{ respectively.} \]

The absolute moment is equal to $E|X| = \frac{2p}{(1+p)}$.

The characteristic function is equal to $\phi(t) = \frac{(1-p)^2}{(1-e^{itp})(1-e^{-itp})}$.

(2) Among the discrete distributions on integers with non-vanishing probability density function and $E|X| = c > 0$, the discrete Laplace distribution has the maximum entropy.

(3) The discrete Laplace distribution is infinitely divisible, geometrically infinitely divisible and stable with respect to geometric compounding.
We know that a Laplace random variable has the same distribution as the difference of two independent and identically distributed exponential random variables and an analogue property holds in the case of discrete Laplace distribution. Hence a discrete Laplace random variable has the same distribution of difference of two independent and identically distributed geometric random variables. This property leads to applications of the discrete Laplace distribution in analysis of uncertainty in hydro climatic systems. The hydro climatic episodes such as droughts, floods, warm spells and cold spells are commonly quantified in terms of their duration and magnitude. The durations of episodes above and below the reference level known as positive and negative episodes respectively are frequently modeled by geometric distribution. Hence, the difference of positive and negative episodes may be modeled using the discrete Laplace model (see Biondi et al. (2002) Inusah and Kozubowski (2006)).

6.2. First order discrete Laplace autoregressive process

Statistical data expressed in terms of counts taken sequentially in time and which are correlated arise in many contexts. Examples of this process are the number of patients in a hospital at a specific point of time or the number of persons waiting in a queue for certain moment. In each of these examples, an element of the process at time $t$ can be either the survival of an element of the process at previous time or an arrival (innovation) sequence, which has a certain discrete distribution. The statistical data corresponding to such processes constitute a discrete time series. Recently much effort has been put in for the study of discrete time series. Al-Osh
and Alzaid (1987), Alzaid and Al-Osh (1990), Jin Guan and Yuan (1991) and Pillai and Jayakumar (1995) have developed discrete time series models with binomial, Poisson, geometric and discrete Mittag-Leffler marginal distributions. Now we develop an integer valued autoregressive time series model with discrete Laplace distribution as marginal distribution.

Let \( \{X_n, n \geq 1\} \) be a sequence of random variables defined by the autoregressive equation

\[
X_n = \begin{cases} 
\varepsilon_n & \text{w.p. } \theta \\
X_{n-1} + \varepsilon_n & \text{w.p. } 1-\theta
\end{cases}
\]

(6.2.1)

where \( 0 < \theta < 1 \) and \( \{\varepsilon_n\} \) be a sequence of independent and identically distributed random variables such that \( X_m \) and \( \varepsilon_n \) are independent when \( m < n \).

**Theorem 6.2.1.**

Let \( \{\varepsilon_n\} \) be a sequence of independent and identically distributed discrete Laplace random variables such that \( \varepsilon_n \sim DL(\delta) \), then the first order autoregressive process \( \{X_n, n \geq 1\} \) given by (6.2.1) defines a stationary time series of DL(p) random variables, where

\[
\delta = 1 - \frac{1-p}{2\theta p} \left( \sqrt{(1-p)^2 + 4\theta p - 1 + p} \right).
\]
Proof:

Let the characteristic functions of \( \{X_n\} \) and \( \{\varepsilon_n\} \) are denoted by \( \phi_{X_n}(t) \) and \( \phi_{\varepsilon_n}(t) \) respectively, then from (6.2.1) we get

\[
\phi_{X_n}(t) = \theta \phi_{\varepsilon_n}(t) + (1-\theta) \phi_{X_{n-1}}(t) \phi_{\varepsilon_n}(t) .
\]

(6.2.2)

Assume \( \{X_n\} \) is stationary with discrete Laplace marginal distribution then from (6.2.2) we have

\[
\phi_{\varepsilon_n}(t) = \frac{\phi_{X_n}(t)}{\theta + (1-\theta) \phi_{X_n}(t)}
\]

\[
= \frac{(1-p)^2}{(1-e^{it}p)(1-e^{-it}p)} = \frac{\theta + (1-\theta)}{(1-p)^2} \frac{(1-p)^2}{(1-e^{it}p)(1-e^{-it}p)}
\]

Hence

\[
\phi_{\varepsilon_n}(t) = \frac{(1-p)^2}{\theta(1-e^{it}p)(1-e^{-it}p) + (1-\theta)(1-p)^2}.
\]

(6.2.3)

Now we show that the characteristic function (6.2.3) is same as the characteristic function of DL(\( \delta \)) random variable where \( \delta \) is given by

\[
\delta = 1 - \frac{1-p}{2 \theta p} \left( \sqrt{(1-p)^2 + 4 \theta p - 1} + p \right).
\]

181
Set (6.2.3) equal to \( \frac{(1-\delta)^2}{(1-e^{it}\delta)(1-e^{-it}\delta)} \), which is the characteristic function of the DL(\( \delta \)) random variable, we have

\[
\frac{(1-p)^2}{\theta(1-e^{it}p)(1-e^{-it}p)+(1-\theta)(1-p)^2} = \frac{(1-\delta)^2}{(1-e^{it}\delta)(1-e^{-it}\delta)}
\]

This produces the equation

\[(1-p)^2(1-e^{it}\delta)(1-e^{-it}\delta) = (1-\delta)^2 \theta(1-e^{it}p)(1-e^{-it}p)+(1-\theta)(1-p)^2(1-\delta)^2\]

and hold for each \( t \in (-\infty, \infty) \).

The above equation is satisfied provided

\[p \theta (1-\delta)^2 = \delta (1-p)^2\]

and

\[(1+p)^2 (1-\delta)^2 \theta = (1+\delta^2)(1-p)^2 - (1-p)^2 (1-\theta)(1-\delta)^2.\]

On simplification, we get that these two equations are equivalent to the quadratic equation in \( \delta \) where \( \delta \in (0,1) \) and given by

\[h(\delta) = \theta p \delta^2 - \delta (2 \theta p + (1-p)^2) + \theta p = 0\]

where \( p, \theta \in (0,1) \).

Since \( h(0) = p \theta > 0 \) and \( h(1) = -(1-p)^2 < 0 \), there exists a unique solution of \( \delta \) in the interval \((0,1)\).
Hence

\[ \delta = \frac{(2 \theta p + (1-p)^2) - \sqrt{(1-p)^2 + 4 \theta p}}{2 \theta p} \]

\[ = 1 - \frac{1-p}{2 \theta p} (\sqrt{(1-p)^2 + 4 \theta p} - 1+p). \]

Therefore, from the above discussion equation (6.2.3) can be written as

\[ \phi_{\varepsilon_n}(t) = \frac{(1-\delta)^2}{(1-e^{-it\delta})(1-e^{it\delta})}. \]

So

\[ \varepsilon_n \sim DL(\delta) \quad \text{where} \quad \delta = 1 - \frac{1-p}{2 \theta p} (\sqrt{(1-p)^2 + 4 \theta p} - 1+p). \]

Hence the theorem.

The autocovariance function of the process (6.2.1) is given by

\[ \text{Cov}(X_n, X_{n-h}) = \theta \text{Cov}(\varepsilon_n, X_{n-h}) + (1-\theta) \text{Cov}(X_{n-1} + \varepsilon_n, X_{n-h}) \]

\[ = (1-\theta) \text{Cov}(X_{n-1}, X_{n-h}) \]

\[ = (1-\theta)^h \text{Var}(X_{n-h}), \text{by repeating the procedure.} \]

Hence

\[ \rho(h) = (1-\theta)^h. \]

(6.2.4)
So the autocorrelation function is positive and decays exponentially and has the same form as the Yule-Walker equation of the first order autoregressive process.

The conditional expectation of the process \( \{ X_n \} \) is linear in \( x \) and given by the expression \( E(X_n / X_{n-1} = x) = (1 - \theta) x \).

Now \( P(X_n > X_{n-1}) \) of the process is obtained as follows:

By definition of the process

\[
P(X_n > X_{n-1}) = \theta P(\varepsilon_n > X_{n-1}) + (1 - \theta) P(\varepsilon_n > 0). \tag{6.2.5}
\]

Consider

\[
P(\varepsilon_n > X_{n-1}) = \sum_{x = -\infty}^{\infty} P(\varepsilon_n > X_{n-1} = x) P(\varepsilon_{n-1} = x)
\]

\[
= \sum_{x = -\infty}^{\infty} P(\varepsilon_n > x) P(X_{n-1} = x)
\]

\[
= \sum_{x = -\infty}^{\infty} \frac{1 - \delta}{1 + \delta} \delta^{|x|} \frac{1 - p}{1 + p} \delta^{|x|}
\]

\[
= \frac{1 - \delta}{1 + \delta} \frac{1 - p}{1 + p} \sum_{|x| = -\infty}^{\infty} (\delta p)^{|x|}
\]

\[
= \frac{1 - \delta}{1 + \delta} \frac{1 - p}{1 + p} \frac{1 + \delta p}{1 - \delta p}
\]

Also note that
Hence, using (6.2.5) we have

\[
P(\varepsilon_n > 0) = \sum_{x=1}^{\infty} P(\varepsilon_n = x) = \sum_{x=1}^{\infty} \frac{1 - \delta}{1 + \delta} \delta^x = \frac{\delta}{1 + \delta}.
\]

The probability that \( X_n \) is greater than \( X_{n-1} \) for the discrete Laplace autoregressive model is a function of \( \theta \) and \( p \) (note that \( \delta \) is a function of \( p \)).

Now we can estimate the parameters \( \theta \) and \( p \) from the equations (6.2.4) and (6.2.6) using the first order sample autocorrelation function \( \hat{\rho}(1) \) and the relative frequency of the number of up-runs in the sample \( x_0, x_1, ..., x_n \) of \( n+1 \) observations.

Suppose \( \hat{\rho}(1) \) be the first order sample autocorrelation function. Then

\[
\hat{\rho}(1) = \frac{\sum_{j=0}^{n-1} (x_{j+1} - \overline{x})(x_j - \overline{x})}{\sum_{j=0}^{n} (x_j - \overline{x})^2},
\]

where \( \overline{x} = \frac{1}{n+1} \sum_{j=0}^{n} x_j \).
Let us define

\[ I(X_i > X_{i-1}) = \begin{cases} 1 & \text{if } X_i > X_{i-1} \\ 0 & \text{otherwise.} \end{cases} \]

Then the estimate of \( P(X_n > X_{n-1}) \), denoted by \( \hat{P} \) is given by

\[ \hat{P} = \frac{1}{n} \sum_{i=1}^{n} I(X_i > X_{i-1}). \] (6.2.8)

Hence, the estimate of \( \theta \) is \( \hat{\theta} = 1 - \hat{\rho}(1) \) and the estimate \( \hat{p} \) of \( p \) is obtained by substituting \( \hat{P} \) for \( P(X_n > X_{n-1}) \) and \( \hat{\theta} \) for \( \theta \) in (6.2.6) where the values of \( \hat{\rho}(1) \)

and \( \hat{P} \) are available from (6.2.7) and (6.2.8) respectively.

### 6.3. First order skewed discrete Laplace process

A discrete random variable \( Y \) on set of integers corresponding to the asymmetric Laplace distribution with probability density function (1.4.13) can be defined using the equation (6.1.1). The distribution of such random variable is known as skewed discrete Laplace distribution. Kozubowski and Inusah (2006) studied skewed discrete Laplace distributions.

**Definition 6.3.1.**

A random variable \( Y \) is said to follow skewed discrete Laplace distribution

with parameter \( p_1 \in (0,1) \) and \( p_2 \in (0,1) \) if its probability mass function is
if \( k = 0, 1, 2, \ldots \),

\[
f_{p_1, p_2}(k) = \begin{cases} 
   p_1^k & \text{if } k = 0, 1, 2, \ldots \\
   p_2^{-k} & \text{if } k = -1, -2, \ldots 
\end{cases}
\]  

(6.3.1)

and it is represented by \( Y \sim \text{SDL}(p_1, p_2) \).

It may be noted that when \( p_1 = p_2 = p \), we obtain symmetric discrete Laplace distribution with probability mass function (6.1.2).

Kozubowski and Inusah (2006) studied various properties including unimodality, infinite divisibility, geometric infinitely divisibility and maximum entropy property and obtained expressions for mean, variance, characteristic function etc. of the skewed discrete Laplace distribution.

The characteristic function of \( X \sim \text{SDL}(p_1, p_2) \) is given by

\[
\phi(t) = \frac{(1-p_1)(1-p_2)}{(1-e^{it}p_1)(1-e^{-it}p_2)}. 
\]  

(6.3.2)

The mean and variance of \( Y \sim \text{SDL}(p_1, p_2) \), may be determined from (6.3.2) or directly using (6.3.1). The mean and variance are obtained as

\[
E(X) = \frac{p_1}{1-p_1} - \frac{p_2}{1-p_2}
\]

and

\[
V(X) = \frac{1}{(1-p_1)^2 (1-p_2)^2} \left[ p_2 (1-p_1)^3 (1+p_2) + p_1 (1-p_2)^3 (1+p_1) - (p_1 - p_2)^2 \right],
\]

respectively.
Similar to the representation of a skewed Laplace random variable as the difference of two exponential random variables with different parameters an analogue result for the skewed discrete Laplace random variable is possible. Hence a skewed discrete Laplace random variable can be considered as the difference of two independent geometric random variables. Similar to discrete Laplace distribution skewed discrete Laplace distribution can be used to model in the analysis of uncertainty in the hydro climatic episodes.

Now we develop a stationary time series model using skewed discrete Laplace random variable.

**Theorem 6.3.1.**

Let \( \{\epsilon_n\} \) be a sequence of independent and identically distributed skewed discrete Laplace random variables such that \( \epsilon_n \sim \text{SDL}(\delta_1, \delta_2) \), then the first order autoregressive process \( \{X_n, n \geq 1\} \) given by (6.2.1) defines a stationary time series of SDL\((p_1, p_2)\) random variables, where

\[
\delta_1 = \frac{2 p_1 \theta}{(p_1 + p_2) \theta + (1 - p_1)(1 - p_2) + \sqrt{((p_1 + p_2) \theta + (1 - p_1)(1 - p_2))^2 - 4 p_1 p_2 \theta^2}} \tag{6.3.3}
\]

and

\[
\delta_2 = \frac{p_2 \delta_1}{p_1} \tag{6.3.4}
\]
Proof:

Assume \( \{X_n\} \) is stationary with skewed discrete Laplace marginal distribution then from (6.2.2) we have

\[
\phi_{e_n}(t) = \frac{(1-p_1)(1-p_2)}{(1-e^{i\theta}p_1)(1-e^{-i\theta}p_2)} \frac{(1-p_1)(1-p_2)}{\theta + (1-\theta)\frac{(1-p_1)(1-p_2)}{(1-e^{i\theta}p_1)(1-e^{-i\theta}p_2)}}
\]

\[
= \frac{(1-p_1)(1-p_2)}{\theta(1-e^{i\theta}p_1)(1-e^{-i\theta}p_2) + (1-\theta)(1-p_1)(1-p_2)}.
\]

Hence

\[
\phi_{e_n}(t) = \frac{(1-p_1)(1-p_2)}{\theta(1-e^{i\theta}p_1)(1-e^{-i\theta}p_2) + (1-\theta)(1-p_1)(1-p_2)}.
\] (6.3.5)

Now we show that the characteristic function \( \phi_{e_n}(t) \) is same as the characteristic function of SDL(\( \delta_1, \delta_2 \)) random variable where

\[
\delta_1 = \frac{2p_1\theta}{(p_1+p_2)\theta + (1-p_1)(1-p_2) + \sqrt{(p_1+p_2)\theta + (1-p_1)(1-p_2)}^2 - 4p_1p_2\theta^2}
\]

\[
\delta_2 = \frac{p_2\delta_1}{p_1}.
\]
Set \( \phi_n (t) \) equal to \( \frac{(1-\delta_1)(1-\delta_2)}{(1-e^{it}\delta_1)(1-e^{-it}\delta_2)} \), which is the characteristic function of the SDL(\( \delta_1, \delta_2 \)) random variable, we have

\[
\frac{(1-p_1)(1-p_2)}{\theta(1-e^{it}p_1)(1-e^{-it}p_2)+(1-\theta)(1-p_1)(1-p_2)} = \frac{(1-\delta_1)(1-\delta_2)}{(1-e^{it}\delta_1)(1-e^{-it}\delta_2)}
\]

This produces the equation

\[
(1-p_1)(1-p_2)(1-e^{it}\delta_1)(1-e^{-it}\delta_2) = (1-\delta_1)(1-\delta_2)\theta(1-e^{it}p_1)(1-e^{-it}p_2) + (1-\theta)(1-p_1)(1-p_2)(1-\delta_1)(1-\delta_2)
\]

and hold for each \( t \in (-\infty, \infty) \).

The above equation is satisfied provided

\[
p_2 (1-\delta_1)(1-\delta_2) \theta = \delta_2 (1-p_1)(1-p_2)
\]

and

\[
p_1 (1-\delta_1)(1-\delta_2) \theta = \delta_1 (1-p_1)(1-p_2).
\]

Dividing the corresponding sides of the above equation leads to \( \frac{p_2}{p_1} = \frac{\delta_2}{\delta_1} \).

Substituting \( \delta_2 = \frac{p_2 \delta_1}{p_1} \) in the first equation resulted into a quadratic equation of the form
\[ h(\delta_1) = p_1 \delta_1^2 \theta - \delta_1 \left[ (p_1 + p_2) \theta + (1-p_1)(1-p_2) \right] + p_1 \theta = 0 \]

Since \( h(0) = p_1 \theta > 0 \) and \( h(1) = -(1-p_1)(1-p_2) < 0 \), there exists a unique solution of \( \delta_1 \) in the interval \((0,1)\).

Hence

\[
\delta_1 = \frac{(p_1 + p_2) \theta + (1-p_1)(1-p_2)}{2} \sqrt{\left( (p_1 + p_2) \theta + (1-p_1)(1-p_2) \right)^2 - 4p_1p_2 \theta^2}
\]

and

\[
\delta_2 = \frac{p_2 \delta_1}{p_1}.
\]

Therefore, from the above discussion equation (6.3.5) can be written as

\[
\phi_{\xi_n}(t) = \frac{(1-\delta_1)(1-\delta_2)}{(1-e^{it\delta_1})(1-e^{-it\delta_2})}.
\]

Hence

\[
\xi_n \sim \text{SDL}(\delta_1, \delta_2),
\]

where \( \delta_1 \) and \( \delta_2 \) are given by (6.3.3) and (6.3.4) respectively.

Hence the theorem.
From the definition (6.2.1) of the stationary first order autoregressive process with skewed discrete Laplace marginal distribution, we can show that the autocorrelation function is

\[ \rho(h) = (1 - \theta)^h \]

and the conditional expectation of the process \( \{X_n\} \) is given by the expression

\[ E(X_n / X_{n-1} = x) = (1 - \theta) x + \frac{\delta_1 - \delta_2}{(1 - \delta_1)(1 - \delta_2)}. \]