A Three-Parameter Binomial-Lindley Distribution: Properties

Xiaoqing Zhang

Supervisor: **Dr. Dianliang Deng**

26^{th} November, 2020

Outline

- Introduction.
  1. Binomial Distribution
  2. Two-Parameter Lindley Distribution
- Proposed Model: Three-Parameter Binomial-Lindley Distribution.
- Properties of the New Generalized Distribution.
  1. Shape of the Probability Function
  2. Generating Function: MGF and PGF
- Future Research.

Binomial Distribution

Definition: A random variable X is said to have a binomial distribution based on m trials with success probability p if and only if

$$Pr(X = x) = \binom{m}{x} p^x (1-p)^{m-x};$$

$$x = 0, 1, 2, ..., m \text{ and } 0 \leq p \leq 1.$$

The Lindley Distribution

D. V. Lindley introduced a one-parameter distribution at 1958, known as Lindley distribution(LD).

- Sankaran(1970) proposed the poisson-lindley distribution by compounding the Poisson distribution with LD.
- At 2010, negative binomial-lindley distribution by mixing NB and LD was proposed by Zamani and Ismail(2010).
- Sharma(2014) introduced the new continuous distribution, the beta-lindley distribution that extends LD.
- ...

The p.d.f. of Lindley distribution is

$$f(x;\theta) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}; \quad x > 0,\ \theta > 0$$

It can be seen that this distribution is a mixture of exponential(θ) and gamma($2,\theta$) distributions.

$$f(x;\theta) = pf_1(x) + (1-p)f_2(x)$$

where $p = \frac{\theta}{\theta+1}$, $f_1(x) = \theta e^{-\theta x}$ and $f_2(x) = \theta^2 x e^{-\theta x}$.

$$f(x;\theta) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}; \quad x > 0,\ \theta > 0$$

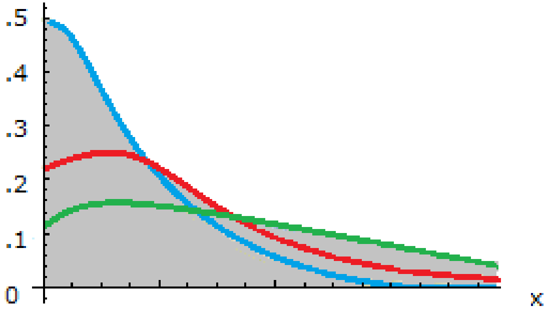


Figure: The shape of p.d.f.

Two-Parameter Lindley Distribution(TPLD)

The probability density function and cumulative distribution function of this two-parameter Lindley distribution(TPLD), introduced by Shanker(2013) are given by

$$f(x;\alpha,\theta) = \frac{\theta^2}{(\theta+\alpha)}(1+\alpha x)e^{-\theta x}; \quad x>0,\ \theta>0,\ \theta+\alpha>0$$

$$F(x;\alpha,\theta) = 1-\left[1+\frac{\alpha x}{\theta+\alpha}\right]e^{-\theta x}; \quad x>0,\ \theta>0,\ \theta+\alpha>0$$

Two-Parameter Lindley Distribution

Its p.d.f can be shown as a mixture of exponential(θ) and gamma($2, \theta$) distributions as follow:

$$f(x; \alpha, \theta) = pf_1(x) + (1 - p)f_2(x)$$

where $p = \frac{\theta}{\theta+\alpha}$, $f_1(x) = \theta e^{-\theta x}$ and $f_2(x) = \theta^2 x e^{-\theta x}$.

- It can easily be seen that at $\alpha = 1$, the two-parameter Lindley Distribution reduces to the one parameter Lindley Distribution.
- At $\alpha = 0$, it reduces to the exponential distribution with parameter θ.

Proposed Model: Three-Parameter Binomial-Lindley Distribution

Definition: A random variable X follows a three-parameter Binomial-Lindley distribution if it follows the stochastic representation

$$X \mid \lambda \sim \text{Binomial}(m, p = 1 - e^{-\lambda})$$

and

$$\lambda \sim \text{TPLD}(\alpha, \theta)$$

where $x = 0, 1, ..., m$, $\lambda > 0$, $\theta > 0$ and $\theta + \alpha > 0$.

Three-Parameter Binomial-Lindley Distribution

Theorem: Let X be a random variable which follows a three-parameter Binomial-Lindley distribution with parameters m, α and θ. Then, the pmf of X is given by

$$Pr(X=x)=\binom{m}{x}\frac{\theta^2}{\theta+\alpha}\sum_{j=0}^{x}\binom{x}{j}(-1)^j\frac{\theta+j+m-x+\alpha}{(\theta+j+m-x)^2};$$

where $x=0,1,...,m$, $\theta>0$ and $\theta+\alpha>0$.

Proof:

$$
\begin{aligned}
Pr(X=x) &= \int_0^\infty Pr(X=x \mid \lambda) f(\lambda;\alpha,\theta) \; d\lambda \\
&= \int_0^\infty \binom{m}{x} (1-e^{-\lambda})^x (e^{-\lambda})^{m-x} f(\lambda;\alpha,\theta) \; d\lambda \\
&= \int_0^\infty \binom{m}{x} \left[\sum_{j=0}^{x} \binom{x}{j} (-1)^j e^{-\lambda j} \right] e^{-\lambda(m-x)} f(\lambda;\alpha,\theta) \; d\lambda \\
&= \binom{m}{x} \sum_{j=0}^{x} \binom{x}{j} (-1)^j \int_0^\infty e^{-\lambda(j+m-x)} f(\lambda;\alpha,\theta) \; d\lambda \\
&= \binom{m}{x} \sum_{j=0}^{x} \binom{x}{j} (-1)^j \cdot M_\lambda \left[-(j+m-x) \right]
\end{aligned}
$$

where $M_\lambda(t)$ is the moment generating function(mgf) of two-parameter Lindley distribution.

The mgf of TPLD

$$
\begin{aligned}
M_\lambda(t) = E(e^{t\lambda}) &= \int_0^\infty e^{t\lambda} \frac{\theta^2}{\theta+\alpha}(1+\alpha\lambda)e^{-\theta\lambda} \; d\lambda \\
&= \frac{\theta^2}{\theta+\alpha}\left(\int_0^\infty e^{-\lambda(\theta-t)} \; d\lambda + \alpha \int_0^\infty \lambda e^{-\lambda(\theta-t)} \; d\lambda\right) \\
&= \frac{\theta^2}{\theta+\alpha}\left(\frac{1}{\theta-t} + \frac{\alpha}{(\theta-t)^2}\right) \\
&= \frac{\theta^2}{\theta+\alpha} \cdot \frac{\theta-t+\alpha}{(\theta-t)^2} \qquad\qquad \textit{where} \;\; t < \theta
\end{aligned}
$$

Thus,

$$
M_\lambda\left[-(j+m-x)\right] = \frac{\theta^2}{\theta+\alpha} \cdot \frac{\theta+j+m-x+\alpha}{(\theta+j+m-x)^2}
$$

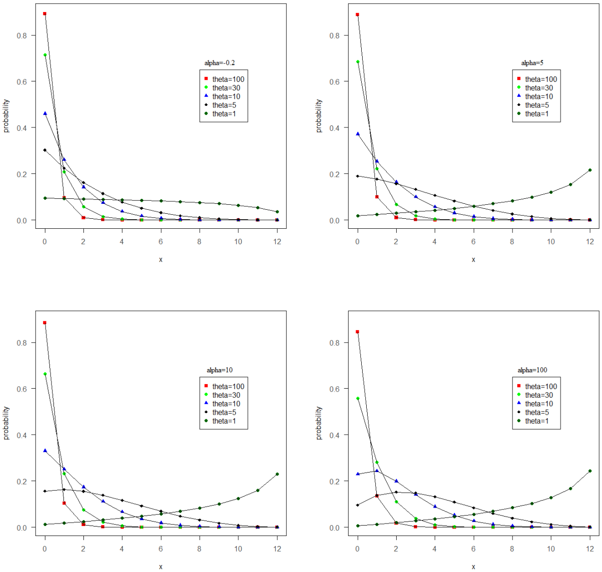
Three-Parameter Binomial-Lindley Distribution

Now we have the pmf of binomial-Lindley distribution as follows

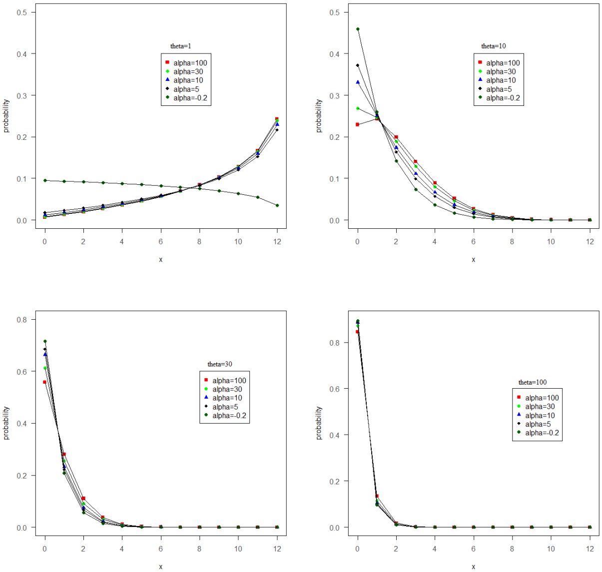
$$
\begin{aligned}
Pr(X=x) &= \binom{m}{x}\sum_{j=0}^{x}\binom{x}{j}(-1)^{j}\cdot M_{\lambda}\left[-(j+m-x)\right] \\
&= \binom{m}{x}\sum_{j=0}^{x}\binom{x}{j}(-1)^{j}\frac{\theta^{2}}{\theta+\alpha}\,\frac{\theta+j+m-x+\alpha}{(\theta+j+m-x)^{2}} \\
&= \binom{m}{x}\frac{\theta^{2}}{\theta+\alpha}\sum_{j=0}^{x}\binom{x}{j}(-1)^{j}\frac{\theta+j+m-x+\alpha}{(\theta+j+m-x)^{2}}
\end{aligned}
$$

with $X = 1, 2, ..., m; \theta > 0$ and $\alpha + \theta > 0$.

Properties of the new generalized distribution



Properties of the new generalized distribution



The Moment Generating Function

Now we derive the moment generating function for binomial-Lindley distribution as follows:

$$
\begin{aligned}
M_X(t) &= E(e^{tx}) \\
&= \sum_{x=0}^{m} e^{tx} Pr(X = x) \\
&= \sum_{x=0}^{m} e^{tx} \binom{m}{x} \frac{\theta^2}{\theta+\alpha} \sum_{j=0}^{x} \binom{x}{j} (-1)^j \frac{\theta + j + m - x + \alpha}{(\theta + j + m - x)^2}
\end{aligned}
$$

The Moment Generating Function

Using the law of total expectation that $E(E(X \mid Y)) = E(X)$, where $E(\mid X \mid) < \infty$.

$$M_X(t) = E(e^{tx}) = E_\lambda \left[E(e^{tx} \mid \lambda) \right]$$

The Moment Generating Function

For binomial distribution, its moment generating function is

$$M_X(t) = (q + pe^t)^m$$

Thus, for $X \mid \lambda \sim Binomial(m, p = 1 - e^{-\lambda})$,

$$E(e^{tx} \mid \lambda) = M_{X|\lambda}(t) = \left[e^{-\lambda} + \left(1 - e^{-\lambda}\right) e^t\right]^m$$

The Moment Generating Function

$$
\begin{aligned}
M_X(t) = E(e^{tx}) = E_\lambda\left[E(e^{tx} \mid \lambda)\right] &= \int_0^\infty \left[e^t + e^{-\lambda}(1-e^t)\right]^m f_\lambda(\lambda) \;\; d\lambda \\
&= \int_0^\infty (1-e^t)^m \left[\frac{e^t}{1-e^t} + e^{-\lambda}\right]^m f_\lambda(\lambda) \;\; d\lambda \\
&= \int_0^\infty (1-e^t)^m \sum_{y=0}^{m} \binom{m}{y} \left(\frac{e^t}{1-e^t}\right)^{m-y} e^{-\lambda y} f_\lambda(\lambda) \;\; d\lambda \\
&= (1-e^t)^m \sum_{y=0}^{m} \binom{m}{y} \left(\frac{e^t}{1-e^t}\right)^{m-y} \int_0^\infty e^{-\lambda y} f_\lambda(\lambda) \;\; d\lambda \\
&= (1-e^t)^m \sum_{y=0}^{m} \binom{m}{y} \left(\frac{e^t}{1-e^t}\right)^{m-y} \frac{\theta^2}{\theta+\alpha} \frac{\theta+y+\alpha}{(\theta+y)^2} \\
&= e^{mt} \frac{\theta^2}{\theta+\alpha} \sum_{y=0}^{m} \binom{m}{y} \left(e^{-t} - 1\right)^y \frac{\theta+y+\alpha}{(\theta+y)^2}
\end{aligned}
$$

The Probability Generating Function

The probability generating function of binomial-Lindley distribution has the following form:

$$
\begin{aligned}
G_X(t) &= E(t^x) = E(e^{x\log t}) = M_X(\log t) \\
&= e^{m\log t}\frac{\theta^2}{\theta+\alpha}\sum_{y=0}^{m}\binom{m}{y}\left(e^{-\log t}-1\right)^y\frac{\theta+y+\alpha}{(\theta+y)^2} \\
&= t^m\frac{\theta^2}{\theta+\alpha}\sum_{y=0}^{m}\binom{m}{y}\left(\frac{1}{t}-1\right)^y\frac{\theta+y+\alpha}{(\theta+y)^2}
\end{aligned}
$$

The Expectation $E(X)$

Next we compute the expectation and variance of binomial-Lindley distribution by using its MGF.

$$M'_X(t) = me^{mt}\frac{\theta^2}{\theta+\alpha}\sum_{y=0}^{m}\binom{m}{y}\left(e^{-t}-1\right)^{y}\frac{\theta+y+\alpha}{(\theta+y)^2}$$
$$+\ e^{(m-1)t}\frac{\theta^2}{\theta+\alpha}\sum_{y=1}^{m}\binom{m}{y}(-y)\left(e^{-t}-1\right)^{y-1}\frac{\theta+y+\alpha}{(\theta+y)^2}$$

$$E(X) = M'_X(0)$$
$$= \frac{m\theta^2}{\theta+\alpha}\left[\left(e^{-0}-1\right)^0\frac{\theta+0+\alpha}{(\theta+0)^2} - \left(e^{-0}-1\right)^0\frac{\theta+1+\alpha}{(\theta+1)^2}\right]$$
$$= m - m\frac{\theta^2(\theta+1+\alpha)}{(\theta+\alpha)(\theta+1)^2} = m\frac{\theta^2+\theta+2\alpha\theta+\alpha}{(\theta+\alpha)(\theta+1)^2}$$

$E(X^2)$

$$
\begin{aligned}
M_X''(t) = {} & m^2 e^{mt} \frac{\theta^2}{\theta+\alpha} \sum_{y=0}^{m} \binom{m}{y} \left(e^{-t}-1\right)^{y} \frac{\theta+y+\alpha}{(\theta+y)^2} \\
& - (2m-1) e^{(m-1)t} \frac{\theta^2}{\theta+\alpha} \sum_{y=1}^{m} \binom{m}{y} y \left(e^{-t}-1\right)^{y-1} \frac{\theta+y+\alpha}{(\theta+y)^2} \\
& + e^{(m-2)t} \frac{\theta^2}{\theta+\alpha} \sum_{y=2}^{m} \binom{m}{y} y(y-1) \left(e^{-t}-1\right)^{y-2} \frac{\theta+y+\alpha}{(\theta+y)^2}
\end{aligned}
$$

$$
\begin{aligned}
E(X^2) &= M_X''(0) \\
&= m^2 - (2m^2-m) \frac{\theta^2(\theta+1+\alpha)}{(\theta+\alpha)(\theta+1)^2} + (m^2-m) \frac{\theta^2(\theta+2+\alpha)}{(\theta+\alpha)(\theta+2)^2}
\end{aligned}
$$

The Variance $V(X)$

Now we can compute the variance using $E(X)$ and $E(X^2)$.

$$V(X) = E(X^2) - [E(X)]^2$$

$$= m^2 - (2m^2 - m)\frac{\theta^2(\theta+1+\alpha)}{(\theta+\alpha)(\theta+1)^2} + (m^2 - m)\frac{\theta^2(\theta+2+\alpha)}{(\theta+\alpha)(\theta+2)^2} - \left[m - m\frac{\theta^2(\theta+1+\alpha)}{(\theta+\alpha)(\theta+1)^2}\right]^2$$

$$= m\frac{\theta^2}{\theta+\alpha}\left[\frac{\theta+1+\alpha}{(\theta+1)^2} - \frac{\theta+2+\alpha}{(\theta+2)^2}\right] + m^2\frac{\theta^2}{\theta+\alpha}\left[\frac{\theta+2+\alpha}{(\theta+2)^2} - \frac{\theta^2(\theta+1+\alpha)^2}{(\theta+\alpha)(\theta+1)^4}\right]$$

E(X) from conditional expectation method

Since $X \mid \lambda \sim Binomial(m, p = 1 - e^{-\lambda})$, $E(X \mid \lambda) = m(1 - e^{-\lambda})$, $V(X \mid \lambda) = me^{-\lambda}(1 - e^{-\lambda})$ and by using conditional expectation method, we have

$$\begin{aligned}
E(X) &= E\left[E(X \mid \lambda)\right] \\
&= E\left[m(1 - e^{-\lambda})\right] \\
&= m - mE(e^{-\lambda}) \\
&= m - mM_{\lambda}(-1) \\
&= m - m\frac{\theta^2(\theta + 1 + \alpha)}{(\theta + \alpha)(\theta + 1)^2} \\
&= m\frac{\theta^2 + \theta + 2\alpha\theta + \alpha}{(\theta + \alpha)(\theta + 1)^2}
\end{aligned}$$

V(X) from conditional expectation method

The theorem using here is that $V(X) = E\left[V(X \mid Y)\right] + V\left[E(X \mid Y)\right]$, where $E(X^2) < \infty$

$$
\begin{aligned}
V(X) &= E\left[V(X \mid \lambda)\right] + V\left[E(X \mid \lambda)\right] \\
&= E\left[me^{-\lambda}(1-e^{-\lambda})\right] + V\left[m(1-e^{-\lambda})\right] \\
&= mE(e^{-\lambda}) - mE(e^{-2\lambda}) + m^2 V(e^{-\lambda}) \\
&= mM_\lambda(-1) - mM_\lambda(-2) + m^2\left[E(e^{-2\lambda}) - \left(E(e^{-\lambda})\right)^2\right] \\
&= m\frac{\theta^2}{\theta+\alpha}\left[\frac{\theta+1+\alpha}{(\theta+1)^2} - \frac{\theta+2+\alpha}{(\theta+2)^2}\right] \\
&\qquad + m^2\frac{\theta^2}{\theta+\alpha}\left[\frac{\theta+2+\alpha}{(\theta+2)^2} - \frac{\theta^2(\theta+1+\alpha)^2}{(\theta+\alpha)(\theta+1)^4}\right]
\end{aligned}
$$

The Index of Dispersion

It is also called Variance-to-Mean Ratio(VMR). It is very similar to the Coefficient of Variation ($CV = \frac{\sigma}{\mu}$), but they are not same.

$$
\begin{aligned}
d &= \frac{\sigma^2}{\mu} \\
&= \frac{E(X^2) - [E(X)]^2}{E(X)} \\
&= \frac{E(X^2)}{E(X)} - E(X)
\end{aligned}
$$

The Index of Dispersion

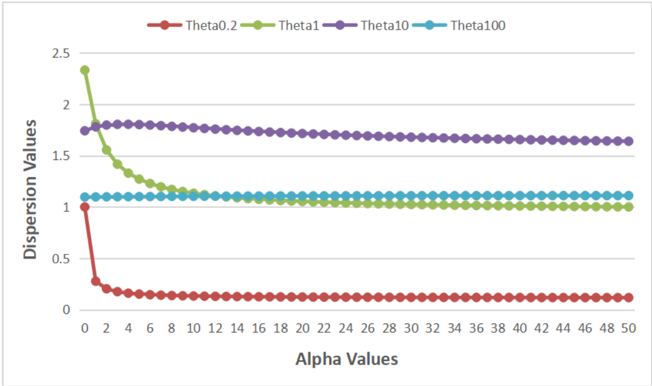
Differences between VMR and CV:

- Similar but different equations.
- VMR has a dimension, CV is always dimensionless.
- VMR is not scale invariant, CV is scale invariant.

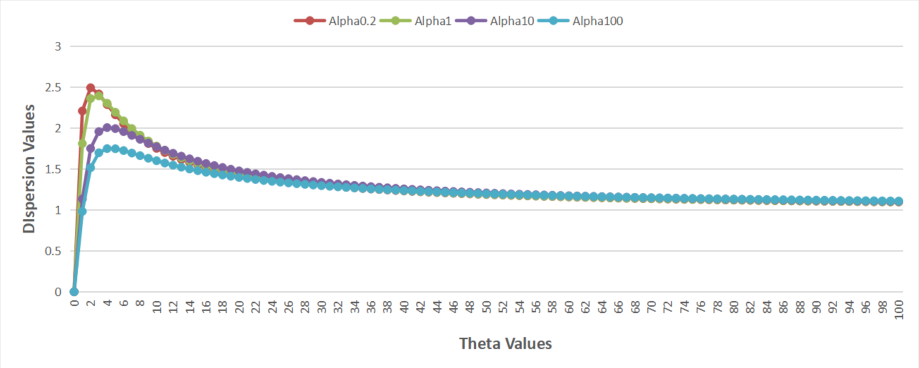
The Index of Dispersion

Distribution	$E(X)$	$V(X)$	$d = \frac{\sigma^2}{\mu}$
Degenerate	k	0	0
Binomial(m,p)	mp	mpq	$0 < q < 1$
Poisson(λ)	λ	λ	1
Negative Binomial(r,p)	$\frac{rp}{q}$	$\frac{rp}{q^2}$	$\frac{1}{q} > 1$

The Index of Dispersion



The Index of Dispersion



Future Research

- Estimating the parameters
  1. Maximum likelihood estimation
  2. EM algorithm
  3. Method of moments
- Application to real data set
- Compounding the binomial distribution with a three-parameter lindley distribution

Thank you for attending