Research & Reviews: Journal of Statistics and Mathematical Sciences

Bayesian Estimate of Discrete Burr Distribution with Two Parameters

Halaleh Kamari[1], Hossein Bevrani[1] and Kaniav Kamary[2]
[1]Department of Statistics, Faculty of Mathematical sciences, University of Tabriz, Iran
[2]Department of Statistics, Universit Paris-Dauphine, 75775 Paris Cedex 16, France

RESEARCH ARTICLE

ABSTRACT

In order to generate discrete distributions based on continuous distributions, various techniques have been implemented and their parameters estimate and characteristics have also been discussed through classical methods. But a few number of essays have dealt with such distribution through Bayesian estimate parameters, mainly due to their complexity and increasing number of parameters. This paper dealt with Bayesian estimate discrete Burr distribution with two parameters. In order to simulate the results, Monte-Carlo and Metropolis-Hastings algorithms were used.

INTRODUCTION

Discrete distributions are very important in modeling the real life plan. Many essays have been published on the study and applications of discrete distributions. Using these techniques, one can transform the continuous distributions into the discrete ones. Katz [1], Roknabadi, et al. [2], Krishna and Pundir [3] have introduced different methods for generating discrete families of distributions. The discrete Burr distribution with two parameters is one of them. If the parameters of this distribution are constant, then to estimate them, Maximum Likelihood and Moment Methods [4] are used, but supposing that the variables were random, then the Bayesian method should be used. In fact, one cannot be sure that the parameters of a distribution are constant values to specify them to prior distribution parameters and calculate their Bayesian estimate. In case they are not calculable through the usual methods of Bayesian estimator, the approximate methods such as numerical integration, Monte-Carlo and Laplace's approximation methods [5] will be used. This paper has used these methods to estimate the Bayesian discrete Burr distribution parameters. A lifetime random variable X follows the Burr-XII, or simple Burr, Br (α, β) [1] if its probability density function is as follows:

$$f(x) = \frac{\alpha\beta x^{\alpha-1}}{\left(1+x^{\alpha}\right)^{\beta+1}}, \quad x > 0, \, \alpha, \beta > 0$$

In this case, for each x > 0, the survival function is

$$S(x) = (1 + x^{\alpha})^{-\beta}$$

The risk rate is

$$r(x) = \frac{f(x)}{S(x)} = \frac{\alpha\beta x^{\alpha-1}}{1+x^{\alpha}}$$

The second risk rate is

$$r^*(x) = \log[S(x)/S(x+1)] = -\beta\log[\frac{1+x^\alpha}{1+(1+x)^\alpha}]$$

The r's moment is

$$E(X^r) = \beta B(r\alpha^{-1}+1, \beta-r\alpha^{-1}), \ \alpha, \beta > 0, \ \alpha\beta > r > 0$$

which $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx.$

In special case, if α =1, Pareto distribution Par(β)[2] with its reliability characteristics will be obtained. If times are grouped into unit intervals, the discrete observed variable dX = [X], the largest integer less than or equal to X, will have the probability function as follows:

(1) $P(dX = x) = p(x) = P(x \leq X < x+1) = S(x) - S(x+1)$

which x = 0, 1,2,.....

Krishna and Pundir [1] applied this method to study the discrete Burr and Pareto distributions. Using the equation (1), the probability mass function of discrete Burr distribution DBD (α,θ)[3] can be dened as follows:

$$p(x) = \theta^{\log(1+x^\alpha)} - \theta^{\log\{1+(1+x)^\alpha\}}, \quad x = 0, 1, 2, \ldots$$

It is important to note that S(x) for Br (α, β) and DBD(α,θ) is equal at the correct points of x. Bayesian estimate of parameters in the case n sample will be calculated in next section.

BAYESIAN ESTIMATE OF DISCRETE BURR DISTRIBUTION PARAMETERS IN N SAMPLE CASE

If $X_1,\ldots, X_n$ has discrete Burr distribution with α and θ parameters, then

$$\theta^{\log(1+x_i^{\acute{a}})} - \theta^{\log[1+(1+x_i^{\acute{a}})]}), \ x_i = 0, 1, 2, \ldots 0 < \acute{e} < 1, \acute{a} > 0$$

To calculate Bayesian estimators, three following cases should be considered:

The firrst case (α is known and θ is unknown)

Suppose that θ has beta distribution with a and 1 parameters, so the previous density of θ is:

(3) $\pi(\theta) = a\theta^{a-1}, 0 < \theta < 1$

Then

(4) $\pi(\theta) \propto \theta^{a-1}$

And

(5) $p\theta(\underline{x}) \propto \prod_{i=1}^n (\theta^{w_{i1}} - \theta^{w_{i2}})$

Which $w_{i1} = \log(1+x_i^\alpha)$ and $w_{i2} = \log[1+(1+x_i)^\alpha]$.

Using (4) and (5), posterior density function of θ is obtained as follows:

(6)
$$\pi(\theta|\underline{x}) \propto \left[\prod_{i=1}^n (\theta^{w_{i1}} - \theta^{w_{i2}})\right]\theta^{a-1}$$
$$\propto \theta^{a+\sum_{i=1}^n w_{i1}-1}\left[\prod_{i=1}^n (1-\theta^{w_i})\right]$$

Which $w_i = w_{i2} - w_{i1}$; so

(7) $\pi(\theta|\underline{x}) = Q(\alpha)\theta^{a+\sum_i^n w_{i1}-1}\left[\prod_{i=1}^n (1-\theta^{w_i})\right]$

Now, the marginal distribution of $X_1,\ldots, X_n$ is obtained by integrating (6) over θ :

---

[1]Continuous Burr distribution with parameters α and β
[2]Continuous Pareto distribution with parameter β
[3]Discrete Burr distribution with parameters α and θ

$$Q^{-1}(\alpha) = \int_0^1 [\Pi_{i=1}^n \theta^{w_{i1}+\frac{a-1}{n}} - \theta^{w_{i2}+\frac{a-1}{n}})]d\theta$$

$$= \Pi_{i=1}^n [\frac{w_i}{(w_{i1}+\frac{a-1}{n}+1)(w_{i2}+\frac{a-1}{n}+1)}] = \Pi_{i=1}^n (\frac{w_i}{\delta_i \tau_i})$$

Which $\delta_i = w_{i1}+\frac{a-1}{n}+1$ and $\tau_i = w_{i2}+\frac{a-1}{n}+1$; then

(8) $\quad Q(\alpha) = \dfrac{1}{\Pi_{i=1}^n (\frac{w_i}{\delta_i \tau_i})}$

(9) $\quad \pi(\theta|\underline{x}) = \dfrac{1}{\Pi_{i=1}^n (\frac{w_i}{\delta_i \tau_i})} \theta^{a+\sum_{i=1}^n w_{i1}-1}[\Pi_{i=1}^n (1-\theta^{w_{i1}})], 0 < \theta < 1$

Supposing a squared-error loss function and known prior, the Bayesian estimate θ parameter

will be as follows:

(10) $\quad \theta^* = E(\theta|\underline{x}) = \dfrac{1}{\Pi_{i=1}^n (\frac{w_i}{\delta_i \tau_i})} \Pi_{i=1}^n [\int_0^1 \theta^{w1+\frac{a}{n}} - \theta^{w_{i2}+\frac{a}{n}})d\theta]$

Which

$$A = \int_0^1 [\theta^{w_{i1}+\frac{a}{n}} - \theta^{w_{i2}+\frac{a}{n}})]d\theta$$

$$= \frac{w_i}{(w_{i1}+\frac{a}{n}+1)(w_{i2}+\frac{a}{n}+1)} = \frac{w_i}{\lambda_i \rho_i}$$

Which $\lambda_i = w_{i1}+\frac{a}{n}+1$ and $\rho_i = w_{i2}+\frac{a}{n}+1$.

By putting A in (10) the following will be obtained:

(11) $\quad \theta^* = \dfrac{1}{\Pi_{i=1}^n (\frac{w_i}{\delta_i \tau_i})} \Pi_{i=1}^n (\frac{w_i}{\lambda_i \rho_i})$

So the Bayesian estimator θ parameter in this case is $\theta^* = \dfrac{1}{\Pi_{i=1}^n (\frac{w_i}{\delta_i \tau_i})} \Pi_{i=1}^n (\frac{w_i}{\lambda_i \rho_i})$

The second case (α is unknown and θ is known)

In this case α is assumed to have unknown prior distribution (ie, there is no prior knowledge about the distribution of α), thus, the prior density of α can be written as follow:

(12) $\quad \pi(\alpha) \propto \frac{1}{\alpha}, \alpha > 0$

And

(13) $\quad p_\alpha(\underline{x}) \propto \Pi_{i=1}^n (\theta^{w_{i1}} - \theta^{w_{i2}})$

which $w_{i1} = \log(1 + x_i^\alpha)$ and $w_{i2} = \log[1+(1+x_i^\alpha)$.

using (12) and (13) we get:

(14) $\quad \pi(\alpha|\underline{x}) \propto \frac{1}{\alpha} \theta^{\sum_{i=1}^n w_{i1}}[\Pi_{i=1}^n (1-\theta^{w_{i1}})]$

Thus,

(15) $\quad \pi(\alpha|\underline{x})(\theta) = Z\frac{1}{\alpha} \theta^{\sum_{i=1}^n w_{i1}}[\Pi_{i=1}^n (1-\theta^{w_{i1}})]$

Which $w_i = w_{i2} - w_{i1}$.

$Z(\theta)$ is calculated as follows:

$$Z^{-1}(\theta) = \int_0^\infty \frac{1}{\alpha} \theta^{\sum_{i=1}^{n} w_{i1}} [\Pi_{i=1}^{n}(1 - \theta^{w_{i1}})] d\alpha$$

Since the posterior density obtained in (15) doesn't belong to the known distributions, therefore, the numerical methods will be used through R software. Let's assume $X_1, ..., X_{25}$ is a random sample of discrete Burr distribution and θ = 0.1, 0.2, 0.3, using Metropolis-Hastings algorithm [6] a proper normal distribution will be considered as, q presupposition for parameter α, and applying the aforementioned algorithm through R software, a Markov chain of 10,000 posterior distribution members will be generated. If the loss function is squared error, the mean of this Markov chain, is the Bayesian estimate of the parameters, and if the loss function is the absolute error, then, the median of this Markov chain is the Bayesian estimate of the parame-ters. The results of these simulations are shown in the **Tables 1 and 2** respectively.

Table 1. Simulation results when the loss function is squared error.

$\theta \downarrow$		$\alpha=1$	$\alpha=2$	$\alpha=3$	$\alpha=4$
0.1	$\widehat{\alpha}$	1.2120	1.9623	2.8210	2.9420
	$\mathrm{var}(\widehat{\alpha})$	0.0829	0.2222	0.5887	0.6661
0.2	$\widehat{\alpha}$	1.6285	2.6540	3.8560	3.9910
	$\mathrm{var}(\widehat{\alpha})$	0.1609	0.4266	1.1023	1.2491
0.3	$\widehat{\alpha}$	2.0747	3.3550	4.902	5.0100
	$\mathrm{var}(\widehat{\alpha})$	0.2546	0.6886	1.7974	1.9379

Table 2. Simulation results when the loss function is the absolute error.

$\theta \downarrow$		$\alpha=1$	$\alpha=2$	$\alpha=3$	$\alpha=4$
0.1	$\widehat{\alpha}$	1.1900	1.9096	2.7200	2.8180
	$\mathrm{var}(\widehat{\alpha})$	0.0829	0.2222	0.5887	0.6661
0.2	$\widehat{\alpha}$	1.5926	2.5727	3.7270	3.8140
	$\mathrm{var}(\widehat{\alpha})$	0.1609	0.4266	1.1023	1.2491
0.3	$\widehat{\alpha}$	2.0294	3.2620	4.7120	4.7900
	$\mathrm{var}(\widehat{\alpha})$	0.2546	0.6886	1.7974	1.9379

The third case (α and θ are unknown)

In this case suppose that α has unknown prior, and θ has Beta distribution as in the first case, both α and θ are independent.

α and θ joint prior density will be as follow:

$$\pi(\alpha, \theta) \propto \frac{1}{\alpha} \theta^{a-1}, 0 < \theta < 1, \ \alpha > 0$$

on the other hand, we know that:

(18) $p_{\alpha,\theta}(\underline{x}) \propto \Pi_{i=1}^{n}(\theta^{w_{i1}} - \theta^{w_{i2}})$

Which $w_{i1} = log(1 + x_i^\alpha)$ and $w_{i2} = log[1 + (1 + x_i)^\alpha]$.

using (17) and (18) we have:

(19) $\quad \pi(\alpha,\theta|\underline{x}) \propto \dfrac{1}{\alpha}\theta^{a+\sum_{i=1}^{n}w_{i1}-1}[\Pi_{i=1}^{n}(1-\theta^{w_{i1}})]$

Then

(20) $\quad \pi(\alpha,\theta|\underline{x}) = K\dfrac{1}{\alpha}\theta^{a+\sum_{i=1}^{n}w_{i1}-1}[\Pi_{i=1}^{n}(1-\theta^{w_{i1}})]$

where K is a constant value, it will be calculated as follows:

(21) $\quad K^{-1} = \int_0^{\infty}\dfrac{1}{\alpha}Q^{-1}(\alpha)d\alpha$

which $Q^{-1}(\alpha) = \Pi_{i=1}^{n}(\dfrac{w_i}{\delta_i\tau_i}), w_i = w_{i1} = w_{i2}, \delta_i = w_{i1}+\dfrac{a}{n}+1\ and\ \tau_i = w_{i2}+\dfrac{a}{n}+1.$

where

(22) $\quad K^{-1} = \int_0^{\infty}\dfrac{1}{\alpha}\Pi_{i=1}^{n}(\dfrac{w_i}{\delta_i\tau_i})d\alpha$

Considering the posterior density obtained in (20) does not belong to known distributions, therefore, the numerical methods will be used through R software. Let's assume $X_1,...,X_{25}$ is a random sample of discrete Burr distribution and a = 3, using Metropolis-Hastings algorithm [2], the proper uniform and normal distributions will be considered as q presupposition for θ and α parameters respectively, and applying the aforementioned algorithm through R software, a Markov chain of 10,000 posterior distribution members will be generated. If the loss function is squared error, the mean of this Markov chain is the Bayesian estimate of the parameters, and if the loss function is the absolute error, then, the median of this Markov chain is the Bayesian estimate of the parameters. The results of these simulations are shown in the **Tables 3 and 4** respectively.

Table 3. Simulation results when the loss function is squared error.

$\theta \downarrow$		$\alpha=1$	$\alpha=2$	$\alpha=3$	$\alpha=4$
0.1	$\widehat{\alpha}$	2.1002	10.0120	16.290	26.644
	var$(\widehat{\alpha})$	0.3812	0.6965	0.7592	0.8379
0.2	$\widehat{\alpha}$				
	var$(\widehat{\alpha})$	0.5784	11.7378	36.1284	55.2465
0.3	$\widehat{\alpha}$	0.5784	11.7378	36.1284	55.2465
	var$(\widehat{\alpha})$	0.0069	0.0114	0.0059	0.0029
0.1	$\widehat{\alpha}$				
	var$(\widehat{\alpha})$	0.2046	0.3094	0.3242	0.2523
0.2	$\widehat{\alpha}$				
	var$(\widehat{\alpha})$	2.1918	9.4980	12.7800	25.9300
0.3	$\widehat{\alpha}$	0.4915	0.7429	0.7880	0.8728
	var$(\widehat{\alpha})$				
0.1	$\widehat{\alpha}$	0.5880	14.1243	13.2204	32.2304

θ					
0.2	$\text{var}(\hat{\alpha})$	0.0102	0.0071	0.0039	0.0019
	$\hat{\alpha}$				
0.3	$\text{var}(\hat{\alpha})$	0.0557	0.2473	0.1675	0.1735
	$\hat{\alpha}$				
	$\text{var}(\hat{\alpha})$	3.6080	14.1800	17.5700	17.7600
0.1	$\hat{\alpha}$	0.7012	0.8564	0.8621	0.8422
	$\text{var}(\hat{\alpha})$				
0.2	$\hat{\alpha}$	2.2266	18.6511	32.4364	33.6526
	$\text{var}(\hat{\alpha})$	0.0083	0.0028	0.0027	0.0034
0.3	$\hat{\alpha}$				
	$\text{var}(\hat{\alpha})$	0.1066	0.1883	0.2394	0.2615
0.1	$\hat{\alpha}$				

Table 4. Simulation results when the loss function is the absolute error.

θ ↓		α=1	α=2	α=3	α=4
0.1	$\hat{\alpha}$	1.9970	9.6170	15.1910	25.6790
	$\text{var}(\hat{\alpha})$				
0.2	$\hat{\alpha}$	0.3817	0.7022	0.7703	0.8451
	$\text{var}(\hat{\alpha})$	0.5784	11.7378	36.1284	55.2465
0.3	$\hat{\alpha}$				
	$\text{var}(\hat{\alpha})$	0.0069	0.0114	0.0059	0.0029
0.1	$\hat{\alpha}$				
	$\text{var}(\hat{\alpha})$	0.2046	0.3094	0.3242	0.2523
0.2	$\hat{\alpha}$	2.0560	8.8940	12.6600	27.0000

	var($\hat{\alpha}$)	0.4918	0.7563	0.7966	0.8810
0.3	$\hat{\alpha}$				
	var($\hat{\alpha}$)	0.5880	14.1243	13.2204	32.2304
0.1	$\hat{\alpha}$	0.0102	0.0071	0.0039	0.0019
	var($\hat{\alpha}$)				
0.2	$\hat{\alpha}$	0.0557	0.2473	0.1675	0.1735
	var($\hat{\alpha}$)				
0.3	$\hat{\alpha}$	3.2630	14.8600	17.3100	17.6600
	var($\hat{\alpha}$)	0.7068	0.8681	0.8725	0.8507
0.1	$\hat{\alpha}$				
	var($\hat{\alpha}$)	2.2266	18.6511	32.4364	33.6526
0.2	$\hat{\alpha}$	0.0083	0.0028	0.0027	0.0034
	var($\hat{\alpha}$)				
0.3	$\hat{\alpha}$	0.1066	0.1883	0.2394	0.2615
	var($\hat{\alpha}$)				

CONCLUSION

This study dealt with the Bayesian analysis of discrete Burr distribution parameters with two parameters in three *n* case, and using the proper prior distributions for the parameters, the Metropolis-Hastings method for numerical parameters

REFERENCES

1. Katz L. Characteristics of frequency functions de ned by rst order di erence equations. 1945.

2. Roknabadi AH, Borzadaran GR and Mohammad K. Some aspects of discrete hazard rate function in telescopic families. Economic Quality Control.2009;24:35-42.

3. Krishna H and Pramendra SP. Discrete Burr and discrete Pareto distributions. Statistical Methodology. 2009;6:177-188.

4. Parsian A. Basic concepts of mathematical statistics. Second edition, Isfahan University of Technology. 2004.

5. Robert C and George C. Introducing Monte Carlo Methods with R. Springer Science and Business Media. 2009.

6. Robert C. The Bayesian choice: from decision-theoretic foundations to computational implementation. Springer Science and Business Media. 2007