Journal of Statistics Applications & Probability Letters
An International Journal

http://dx.doi.org/10.18576/jsapl/040204

Discrete Three Parameter Burr Type III Distribution in the Epidemiology of Dental Caries

B. A. Para * *and T. R. Jan*

Department of statistics, University of Kashmir, Srinagar, J&K, India.

Received: 30 Nov. 2016, Revised: 7 Jan. 2017, Accepted: 12 Jan. 2017.
Published online: 1 May 2017.

Abstract: In this paper we propose a discrete analogue of three parameter Burr type III distribution as a model in the epidemiology of dental caries. Numerical Illustrations of data recorded by Grainger and Reid [2] regarding carious teeth data has been discussed. It has been observed that three parameter discrete Burr type III distribution is the best fitted model for caries process in contrast with the classical models discussed. It may be worth exploring the possibility of developing a discrete version of three parameter Burr type III distribution, so that same can be used for modeling a discrete medical count data. Discrete Burr type III distribution is also suggested as a suitable reliability model to fit a range of discrete life time data, as it is shown that hazard rate function can attain monotonic increasing (decreasing) shape for certain values of parameters. The equivalence of discrete Burr type III (DBD-III) and continuous Burr type III (BD-III) distributions has been established. Various theorems relating discrete Burr type III distribution with other statistical distributions have also been proved.

Method: Discretised model of three parameter Burr type III distribution is used to fit dental data recordings of Grainger and Reid [2] ,distribution of the number of smooth surfaces affected by caries for the children of 11 years age in a sample of 146. The mean count is 1.97 carious surfaces per child with the standard deviation of 2.55 carious surfaces.

Results: Discretised model of three parameter Burr type III provides a better fit as compared to other models discussed. The log likelihood value of model is –278.528 and Akaike's Information Criterion (AIC) value is 563.057 and BIC value is 572.008, which are minimum as compared to other models discussed.

Conclusions: A three parameter discrete Burr type III model was used to test data collected from 11 year old children. The probability was estimated by applying fitdistr techniques by using R studio statistical software. The model is a reasonably good approximation of the caries process. A goodness of fit test was performed using chi-square distribution. The null hypothesis that the data come from a three parameter discrete Burr type III distribution is accepted at the 5% level of significance.

Keywords: Dental caries, Burr type III distribution, epidemiology, discrete models, reliability, failure rate, Akaike Information Criterion.

1 Introduction

Statistical models describe a phenomenon in the form of mathematical equations. Out of large number of methods and tools developed so far, for analyzing data in the medical sciences etc., the statistical models are the latest innovations. Current clinical problems are subjected to effective mathematical inquiring with special emphasis on how to set up the basic mathematical equations that govern the models that describe the phenomena. Dental science is one of the fields where the statisticians have contributed in buildings the models to analyse caries process. Tooth decay is common among humans and is one of the most prevalent oral diseases.

Grainger and Reid [2] made an attempt obtain the statistical distributions of dental caries. Their method consists of fitting the well known distributions such as the negative binomial to sample data and selecting the ones which give close fits. Knutson [10] reported on the relationship between prevalence of caries and caries severity, DMFT (Mean number of decayed, missing or filled teeth) in populations. Knutson used data from the first series of fluoridation trials in the United States, which covered 6-15 year olds. In the epidemiology of caries, a chain binomial model of the caries process was defined and tested with data that consisted of the total number of various deciduous molars in 10 and 11 year old children

*Corresponding author e-mail:parabilal@gmail.com

byPhyo [1]. A truncated Poisson distribution by Singh and William [5] was shown to provide a closer fit to observed caries in a sample of 10 to 11 year old children than a previously reported chain binomial distribution by Phyo [1]. Operations research techniques were used in estimating a parameter. Spencer and Lewis [11] discussed two types of curvilinear equations which were fitted to the prevalence/severity data in order to define the relationships between prevalence of caries, percentage with active caries and caries severity. Krishna and Pundir [3] proposed a discrete version of Pareto and Burr type XII distributions as the closer fit models for the recordings of Phyo [1] of the total number of carious teeth among the four deciduous molars in a sample of 100 children 10 and 11 years old. Symmetry between right and left molars is presumed and only the right molars are considered with a time unit of two years. Lingström and Borrman [6] discussed the Distribution of Dental Caries in an Early 17th Century Swedish Population with Special Reference to Diet. Lewsey and Thomson [7] obtained data on dental caries occurrence at ages 5, 18 and 26 years from the Dunedin Multidisciplinary Health and Development Study (DMHDS). Zero-inflated Poisson (ZIP) and zero-inflated negative binomial (ZINB) models were fitted to the cross-sectional (n=745) and longitudinal (n=809) data sets using Stata (Intercooled Stata 7.0). The dependent variables for the three cross-sectional analyses were the DMFS (Mean number of decayed, missing or filled teeth surfaces) indices at age 5, 18, and 26 years, and net decayed and filled tooth surfaces (DFS) increment was the dependent variable for the longitudinal analysis.

In the present paper we propose a three parameter discrete Burr type III (DBD-III) model which provides a better fit to carious teeth data recorded by Grainger and Reid [2] as compared to some well known classical models and discrete class of continuous models.

Burr [1] introduced a family of distributions includes twelve types of cumulative distribution functions, which yield a variety of density shapes. The two important members of the family are Burr type III and Burr type XII distributions. Types III and XII are the simplest functionally and therefore, the two distributions are the most desirable for statistical modeling.

A continuous random variable X is said to follow a Burr type III distribution if its pdf is given by

$$f(x) = \begin{cases} \dfrac{ck\theta}{x^{c+1}(\theta x^{-c}+1)^{k+1}} & , x > 0, c > 0\,, k > 0, \theta > 0 \\ 0 & \text{elsewhere} \end{cases} \tag{1.1}$$

and its cumulative distribution function is given by

$$F(x) \;\; = (1+\theta x^{-c})^{-k}$$

$$x > 0, k > 0, c > 0, \theta > 0 \tag{1.2}$$

Fig.1 to Fig.4 gives the pdf plot for (1.1.0) for different values of parameters. It is evident that the distribution of the rv X is right skewed.

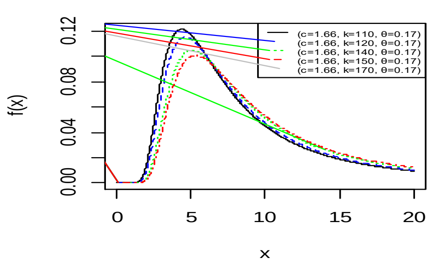

Fig.1. pdf plot for BD-III(c,k,θ)

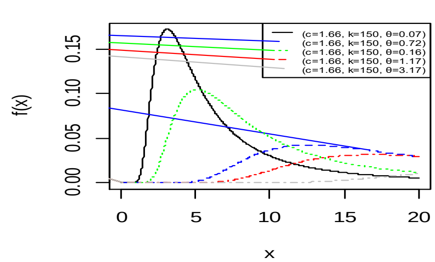

Fig.2. pdf plot for BD-III(c,k,θ)

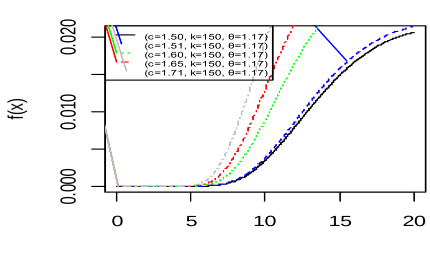

Fig.3. pdf plot for BD-III(c,k,θ)

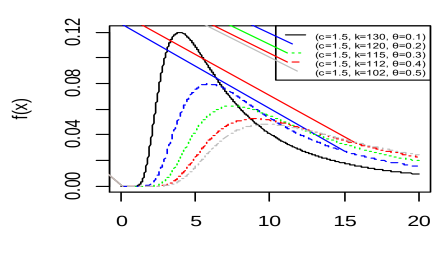

Fig.4. pdf plot for BD-III(c,k,θ)

2 The Various Reliability Measures ofa Random Variable X are Given by

(a) Survival function

$$s(x) = 1 - \int_0^x f(x)dx$$

$$= 1 - \int_0^x \frac{ck\theta}{x^{c+1}(1+\theta x^{-c})^{k+1}} dx$$

$$= 1 - (1+\theta x^{-c})^{-k} \qquad x > 0;\ c > 0;\ k > 0;\ \theta > 0$$

(b) The failure rate is given by

$$r(x) = \frac{ck\theta}{[x^{c+1}(1+\theta x^{-c})^{k+1}][1-(1+\theta x^{-x})^{-k}]}$$

$$x > 0;\ c > 0;\ k > 0;\ \theta > 0$$

(c) The second rate of failure is given by

$$SRF(x) = \log\left(\frac{s(x)}{s(x+1)}\right) = \log\left(\frac{1-(1+\theta x^{-c})^{-k}}{1-(1+\theta(x+1)^{-c})^{-k}}\right)$$

$$x > 0;\ c > 0;\ k > 0;\ \theta > 0$$

Note that for second rate of failure

$$SRF(0) = SRF(1) \Rightarrow c = -\log\left[\frac{\left((1+\theta)^{-k}(2-(1+\theta)^{-k})\right)^{-1/k} - 1}{\theta\log 2}\right] = \alpha\ (\text{say})$$

It could be seen that SRF(x) is decreasing in x if c <α and for c >α, SRF(0) < SRF(1) and for all other values x > 1, SRF(x) decreases for all.

(d) The rth moment is

$$E(x^r) = \int_0^\infty x^r f(x)dx$$

$$= k\theta^{\frac{r}{c}}\beta(1-\frac{r}{c}, k+\frac{r}{c}) \text{ Where } \beta(a,b) = \int_0^\infty \frac{x^{a-1}}{(1+x)^{a+b}}dx\ , \quad x > 0;\ c > 0;\ k > 0;\ \theta > 0$$

3 Three Parameter Discrete Burr Type III Model

Roy [4] pointed out that the univariate geometric distribution can be viewed as a discrete concentration of a corresponding exponential distribution in the following manner:

$$p\,[X = x] = s(x) - s\,(x+1) \quad \text{When } x = 0, 1, 2, \ldots..$$

Where X is discrete random variable, following geometric distribution with probability mass functions as

$$p\,(x) = \theta^x(1-\theta) \quad x = 0,1,2,\ldots\ldots.$$

Where s(x) represents the survival function of an exponential distribution of the form s(x) = exp(-λx) clearly

θ = exp(-λ), 0 <θ< 1.

Thus, one to one correspondence between the geometric distribution and the exponential distribution can be established, the survival functions being of the same form.

The general approach of dicretising a continuous variable is to introduce a greatest integer function of X i.e., [X] (the greatest integer less than or equal to X till it reaches the integer), in order to introduce grouping on a time axis. On the same pattern discrete versions of the normal and rayleigh distributions were also proposed by Roy [8] and Roy[9] respectively.

A discrete Burr type III variable, dX can be viewed as the discrete concentration of the continuous Burr type III variable X, where the corresponding probability mass function of dX can be written as:

$$P(dX = x) = p(x) = s(x) - s(x+1)$$

The probability mass function takes the form

$$P(x) = \begin{cases} q^{\log(1+\theta)} & x=0 \\ q^{\log(1+\theta(x+1)^{-c})} - q^{\log(1+\theta(x)^{-c})} & x = 1,2,3 \ldots . \end{cases} \quad (3.1)$$

Where $q = e^{-k} 0 < q < 1;\ \theta > 0; c > 0$

The parameters q and θ completely determines the pmf (3.1) at x = 0. It should be also noted that the p(x) is always monotonic decreasing for x = 1,2,3,4,….

When $\log q < \frac{\log 2}{\log(1+\theta 2^{-c}) - \log(1+\theta)}$

P(0) < P(1) and then p(x) decreases ∀ x = 1, 2,3,… i.e., p(x) is a unimodal (with mode at 1). The shape parameter c has more influence on the pmf than q and θ after x = 0, also as the c becomes smaller, the tail of the pmf becomes longer. Parameter θ has also influence on the model value of the distribution. Fig.5 to Fig.10 gives the pmf plot for (3.1) for different values of parameters.

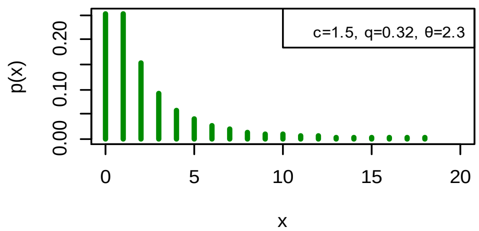

Fig.5 pmf plot for DBD-III(c,q,θ)

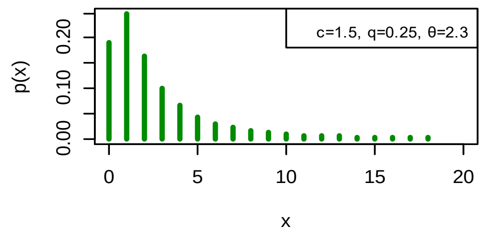

Fig.6 pmf plot for DBD-III(c,q,θ)

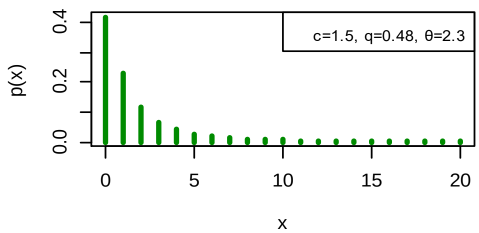

Fig.7 pmf plot for DBD-III(c,q,θ)

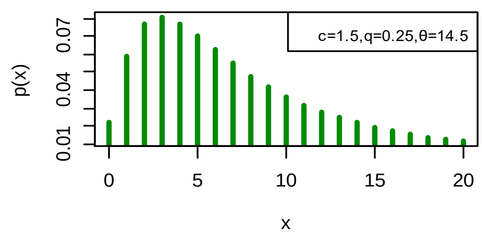

Fig.8 pmf plot for DBD-III(c,q,θ)

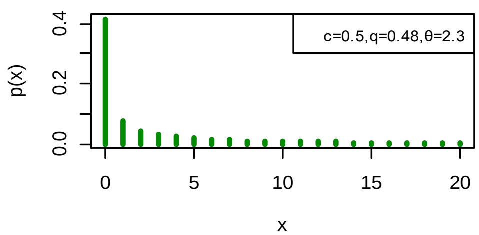

Fig.9 pmf plot for DBD-III(c,q,θ)

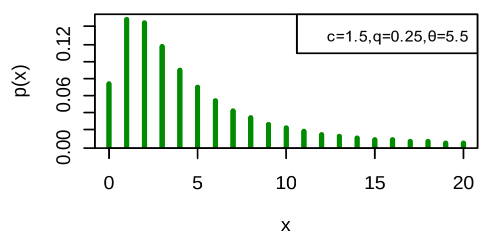

Fig.10 pmf plot for DBD-III(c,q,θ)

4 Reliability Measures of Discrete Burr Type III Random Variable dX are given by

(e) Survival function

$$s(x) = p(dX \geq x) = 1 - q^{\log(1+\theta x^{-c})} x = 0,1,2,3 \ldots.$$

$$c > 0; 0 < q < 1; \theta > 0$$

s(x) is same for continuous Burr type III distribution and discrete Burr type III distribution at the integer points of x.

(f) Rate of failure, r(x) is given by

$$r(x) = \frac{p(x)}{s(x)} = \frac{q^{\log(1+\theta(x+1)^{-c})} - q^{\log(1+\theta x^{-c})}}{1 - q^{\log(1+\theta x^{-c})}} \quad x = 0,1,2,3 \ldots.$$

$$c > 0; 0 < q < 1; \theta > 0$$

(g) Second rate of failure is given by

$$\mathrm{SRF}(x) = \log\left(\frac{1-q^{\log(1+\theta x^{-c})}}{1-q^{\log(1+\theta(x+1)^{-c})}}\right) \quad x = 0,1,2,3 \ldots$$

$$c > 0; 0 < q < 1; \theta > 0$$

It could be seen that r(x) and SRF(x) are completely determined by q *and* θ.

When $c < -log\left[(e^{\emptyset(q,\theta)} - 1)/\theta\right]/\log 2 \ = \alpha(\text{say})$

$$\text{where } \emptyset(q,\theta) = \frac{\log[1 - \left(1 - q^{\log(1+\theta)}\right)^2]}{\log q}$$

r(0) > r(1) and SRF(0) > SRF(1) $\qquad \theta > 0\,;\ 0 < q < 1;\ c>0$

Fig.11 to fig.16 illustrates the second rate of failure plot for DBD-III for different values of parameters. For c > α; r(0) < r(1) and SRF(0) < SRF(1), clearly the hazard rates of continuous model and the discrete modal shows the same monotonocity.

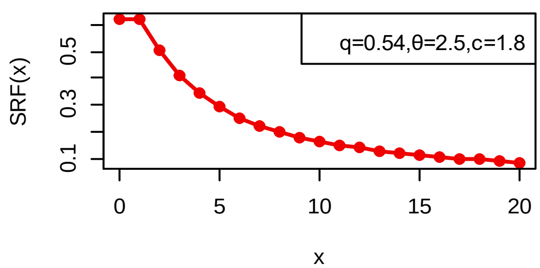

Fig.11 SRF(x) plot for DBD-III(q,θ,c)

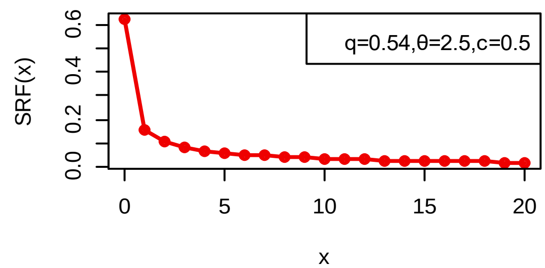

Fig.12 SRF(x) plot for DBD-III(q,θ,c)

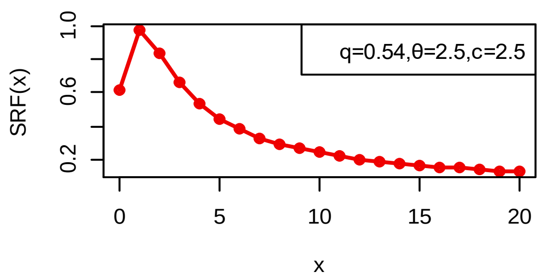

Fig.13 SRF(x) plot for DBD-III(q,θ,c)

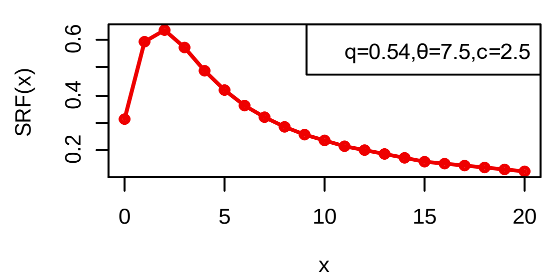

Fig.14 SRF(x) plot for DBD-III(q,θ,c)

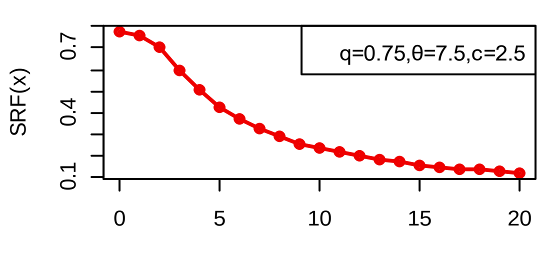

Fig.15 SRF(x) plot for DBD-III(q,θ,c)

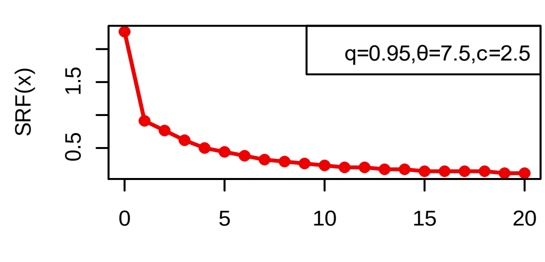

Fig.16 SRF(x) plot for DBD-III(q,θ,c)

4.1 Moments of Discrete Burr Type III Distribution

$$E(x^r) = \sum_{x=0}^{\infty} x^r p(x)$$
$$= \sum_{x=1}^{\infty} [x^r - (x-1)^r]\, s(x)$$
$$\leq \sum_{x=1}^{\infty} r\, x^{r-1}\, s(x)$$
$$= \sum_{x=1}^{\infty} r\, x^{r-1} (1 - \theta^{\log(1+\theta x^{-c})})$$

R.H.S expression is finite if c > r

Now

$$E(x) = \sum_{1}^{\infty} s(x) = \sum_{1}^{\infty} (1 - \theta^{\log(1+\theta x^{-c})})$$

is finite if c > 1

Similarly for the convergence of variance, c must be greater than 2.

There is a one to one correspondence between three parameter continuous Burr type III distribution and three parameter discrete Burr type III distribution, as the expressions for survival function, failure rate function, second rate of failure function for DBD-III(q, θ,c) can be directly obtained from continuous Burr type XII distribution by replacing k=-log(q). In both the cases $E(X^r)$ will exist iff c>r ;c>0,r>0.

5 Estimation of the Parameters of Three Parameter Discrete Burr Type III Distribution

Estimation of the parameters based on the ML method: Let $X_{1,}X_2X_{3,} \ldots \ldots X_n$ be a random sample of size n. If these $X_{i.}$'s are assumed to be iid random variables following discrete Burr type III distribution i.e., $DBD-III(q, \theta, c)$, their likelihood function is given by

$$L(c, \theta; x) = \prod_{i=1}^{n} p(x_i)$$
$$= \prod_{i=1}^{n} (q^{\log(1+\theta(x_i+1)^{-c})} - q^{\log(1+\theta(x_i)^{-c})}) \tag{5.1}$$

And (5.1) can be rewritten as follows

$$L(c, \theta; x) = \prod_{i=1}^{n} q^{\log(1+\theta(x_i)^{-c})} (q^{\emptyset(x_i,c,\theta)} - 1) \tag{5.2}$$

where $\emptyset(x_i, c, \theta) = \log[\frac{(1+\theta(x_i+1)^{-c})}{(1+\theta x_i{}^{-c})}]$

$$\log L = \sum[\log(1+\theta x^{-c}) \log q + \log(q^{\emptyset(x,c,\theta)} - 1)] \tag{5.3}$$

Now to find out the estimates of three parameters by ML technique, we have to solve the below equations.

$$\frac{\partial \log L}{\partial q} = \sum_{i=1}^{n} [\frac{\log(1+\theta x_i{}^{-c})}{\hat{q}} + \frac{\emptyset(x_i,c,\theta)\hat{q}^{\emptyset(x_i,c,\theta)-1)}}{\hat{q}^{\emptyset(x_i,c,\theta)}-1)}] = 0 \tag{5.4}$$

$$\frac{\partial \log L}{\partial c} = \sum_{i=1}^{n} [\frac{(-\theta x_i{}^{-\hat{c}}) \log q \log x_i}{(1+\theta x_i{}^{-\hat{c}})} + \frac{\log q \emptyset\prime(x_i,\hat{c},\theta) q^{\emptyset(x_i,\hat{c},\theta)}}{q^{\emptyset(x_i,\hat{c},\theta)}-1}] = 0 \tag{5.5}$$

$$\text{where } \emptyset'(x_i, \hat{c}, \theta) = \frac{\partial \emptyset(x_i, c, \theta)}{\partial c}$$

$$\frac{\partial \log L}{\partial \theta} = \sum_{i=1}^{n} [\frac{(x_i{}^{-c}) \log q}{(1+\theta x_i{}^{-c})} + \frac{\log q \emptyset\prime(x_i,c,\theta) q^{\emptyset(x_i,c,\theta)}}{q^{\emptyset(x_i,c,\theta)}-1}] = 0 \tag{5.6}$$

$$\text{where } \emptyset'(x_i, c, \theta) = \frac{\partial \emptyset(x_i, c, \theta)}{\partial \theta}$$

Case I: All the three parameters are unknown.

By using numerical computation, the solution of the three log-likelihood equations (5.4), (5.5) and (5.6) will rovide the MLE of q,θ and c.

Case II: When q is known

In this case, the MLEs of c and θ can be obtained by solving the likelihood equations (5.5) and (5.6) using Newton Raphson method.

Case III: When c is known

In this case, the MLEs of q and θ can be obtained in a similar way as in case ii, by solving the likelihood equations (5.4) and (5.6) using Newton Raphson method.

Case IV: When θ is known

In this case, the MLE of the unknown parameter c and q is obtained by solving the likelihood equations (5.4) and (5.5). With an observed sample these equations can be solved using an iterative numerical method.

In the example 1, we will discuss these cases on the basis of random sample generated from three parameter discrete Burr

Type III distribution.

Example 1: A random sample of size 100 is generated from three parameter Burr type III distribution with c=3, q=0.3 and $\theta = 2$. The sample is

X	0	1	2	3	4	5	7	Total
Frequency	**29**	**45**	**18**	**4**	**2**	**1**	**1**	**100**

The summary statistics for the randomly generated discrete Burr type III sample is given by

Min	1st Quartile	Median	Mean	3rd Quartile	Max
0.00	0.00	1.00	1.13	2.00	7.00

In order to estimate the parameters of three parameter discrete Burr type III distribution on the bases of generated random sample we discuss here three cases.

Case I: When c,q and θ are unknown.

By using numerical computation, the solution of three log-likelihood equations (5.4), (5.5) and (5.6) provide the MLE of c,q and θ . For this purpose, using fitdistr procedure in R studio statistical software we get the estimates as $\hat{c}$ =3.40, $\hat{q}$=0.55 and $\hat{\theta}$=6.71. The corresponding standard errors for the parameters are (0.82, 0.21 and 9.7) respectively.

Case II: When q is known (q=0.3)

In this case numerical computation, the solution of two log-likelihood equations (5.5) and (5.6) provide the MLE of c and θ. In this case estimates of the c and θ based on the observed sample are given by $\hat{c}$ =1.86 and $\hat{\theta}$=2.83 with the corresponding standard errors (0.37, 0.29)

Case III: When θ is known (θ=0.3)

Here, again proceeding in a similar way by solving two log-likelihood equations (5.4) and (5.5) provide the MLE of c and θ. ML estimates for c and q are given by $\hat{c}$ =2.85, $\hat{q}$=0.32 with the standard errors for c and q as (0.26,0.05) respectively.

6 Burr Type III as A Model in the Epidemiology of Dental Caries

Here we consider the recordings of Grainger and Reid [2], the distribution of the number of smooth surfaces affected by caries for the children of 11 years of age. His results are reproduced in the below table 1.

Table 1: Distribution of the number of smooth surfaces affected by caries for the children of 11 years age in a sample of 146.

Total number of carious teeth	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Frequency	52	34	23	9	5	7	6	2	2	2	3	0	1

For the purpose of parameter estimation, we make use of the fitdistr procedure in R studio statistical software to find out the estimates of the parameters. The ML estimates and their standard errors provided by the fitdistr procedure are given in the table 2.

We compute the expected frequencies for fitting Poisson, Geometric, DBD-XII, DPareto, DIWD, DBD-III and DRaylei distributions with the help of R studio statistical software and Pearson's chi-square test is applied to check the goodness fit of the models discussed. The calculated figures are given in the table 3.

Fig.17 gives an overview of Chi-grams of fitted distributions namely DBD-III, DBD-XII, discrete Rayleigh, Poisson, DIWD and Pareto distributions. It is clear from the figure that overrepresentation and underrepresentation of carious teeth frequencies given by fitted distributions is minimum in case of DBD-III model.

Table 2: Estimated parameters by ML method for fitted distributions.

Distribution	Parameter Estimates	Standard Error of the estimates	Model function
Poisson	λ =1.97	λ =0.116	$\frac{e^{-\lambda}\lambda^x}{x!}$ $\lambda > 0; x = 0,1,2,...$
DRayleigh	q=0.92	q=0.006	$q^{x^2} - q^{(x+1)^2} 0 < q < 1;\ x = 0,1,2,....$
DPareto	q=0.36	q=0.031	$q^{\log(1+x)} - q^{\log(2+x)} 0 < q < 1; x = 0,1,2,...$
DIWD	c=1.27,q=0.33	c=0.11,q=0.03	$q^{(1+x)^{-c}} - q^{(x)^{-c}}$ $x = 0,1,2,...$ $0 < q < 1; c > 0$
DBD-XII	c=1.95,q=0.52	c=0.275,q=0.046	$q^{\log(1+x^c)} - q^{\log(1+(x+1)^c)}$ $x = 0,1,2,...$ $0 < q < 1; c > 0$
DBD-III	c=2.4,q=0.74, θ=28.79	c=0.55,q=0.09, θ=41.79	$q^{\log(1+\theta(x+1)^{-c})} - q^{\log(1+\theta(x)^{-c})}$ $x = 0,1,2,3$ $0 < q < 1;\ \theta > 0; c > 0$

Table 3: Expected frequencies for carious teeth data with Chi-square p-values.

X	Observed	Poisson	DRayleigh	DPareto	DIWD	DBD-XII	DBD-III
0	52	20.31	11.25	73.43	49.64	52.95	52.66
1	34	40.06	28.80	24.36	43.77	40.98	31.36
2	23	39.51	34.99	12.14	18.42	18.49	21.32
3	9	25.98	30.47	7.27	9.52	9.59	13.71
4	5	12.81	20.81	4.84	5.67	5.67	8.60
5	7	5.05	11.53	3.45	3.71	3.68	5.43
6	6	1.66	5.27	2.58	2.60	2.55	3.52
7	2	0.47	2.01	2.01	1.90	1.86	2.35
8	2	0.12	0.64	1.60	1.45	1.40	1.62
9	2	0.03	0.17	1.31	1.13	1.10	1.15
10	3	0.00	0.04	1.09	0.91	0.88	0.84
11	0	0.00	0.01	0.92	0.74	0.71	0.63
12	1	0.00	0.00	10.99	6.52	6.14	2.79
Total	146	146	146	146	146	146	146
P-value	-	0.00	0.00	0.00	0.0037	0.0258	0.1964

X: Number of smooth surfaces affected by caries
Observed: Number of children

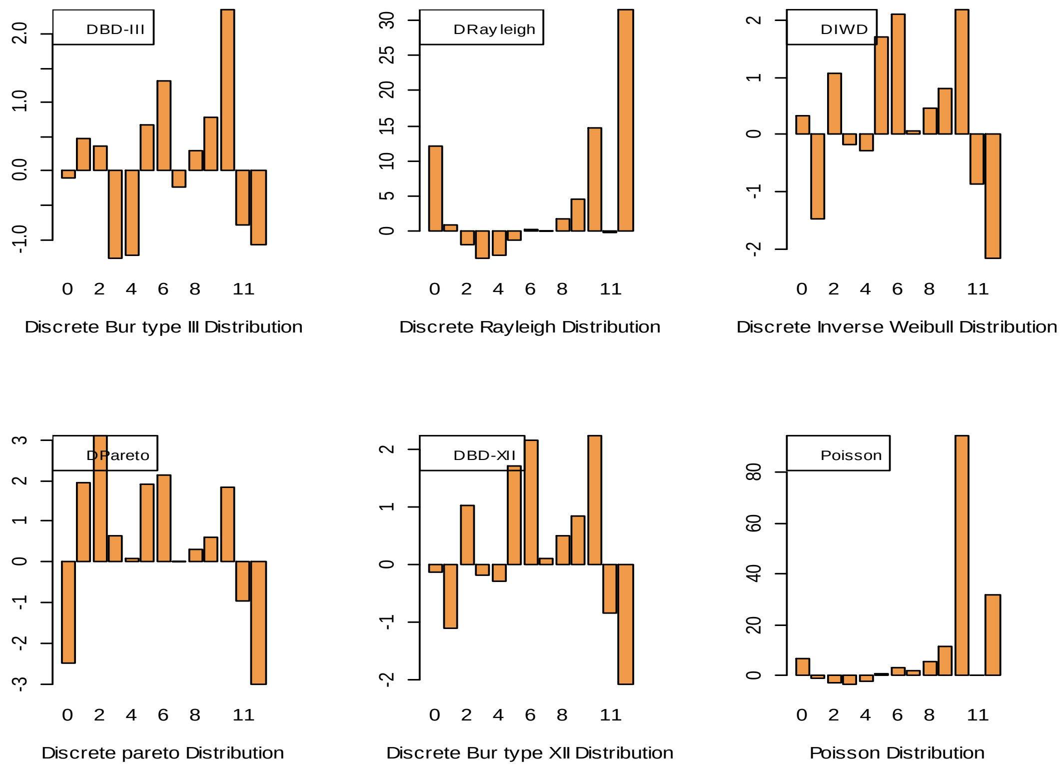

Fig. 17: Chi-grams for Carious Teeth Data

Fig. 18 gives the graphical overview of the observed and model frequencies of carious teeth data.

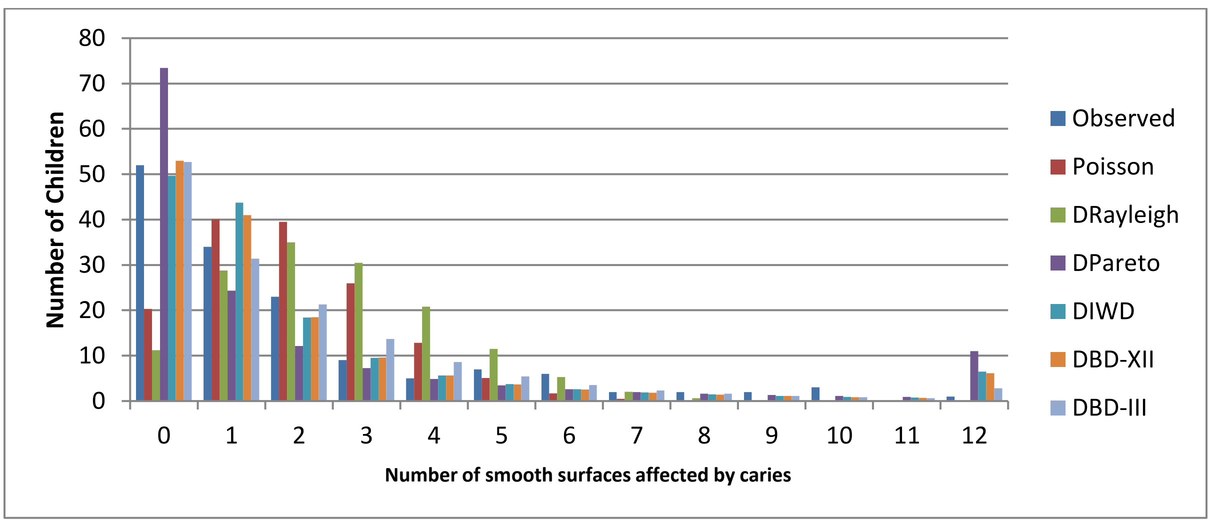

Fig. 18: Distribution of carious teeth data

The p-values of Pearson's Chi-square statistic are 0.0001, 0.000, 0.000, 0.0037, 0.0258 and 0.1964 for Poisson, DRayleigh, DPareto, DIWD, DBD-XII and DBD-III distributions, respectively (see Table 3). This reveals that Poisson and DRayleigh ,DBD-XII, DIWD and DPareto distributions are not good fit at all, where as DBD-III distribution fits the distribution of caries well.

From AIC,BIC , negative loglikelihood and AICC measures it is observed that the three parameter discrete Burr type III provides us a better fit to carious teeth data. The AIC,BIC , negative loglikelihood and AICC figures are given in the table 5.

Table 5: AIC, BIC, Negative log-likelihood and AICC values of fitted distributions of carious teeth data.

Criteria	Poisson	DRayleigh	DPareto	DIWD	DBD-XII	DBD-III
Neg-loglik	342.456	352.137	295.201	283.674	282.930	278.528
AIC	686.912	706.274	592.403	571.348	569.860	563.057
BIC	689.896	709.257	595.386	577.315	575.828	572.008
AICC	686.953	706.315	592.443	571.472	569.984	563.307

Fig.19 gives the graphical description of table 5. It is clear that the three parameter Burr type III attains minimum for AIC,BIC, AICC and Negative log likelihood values , hence it proves that DBD-III is reasonably good model in the caries process.

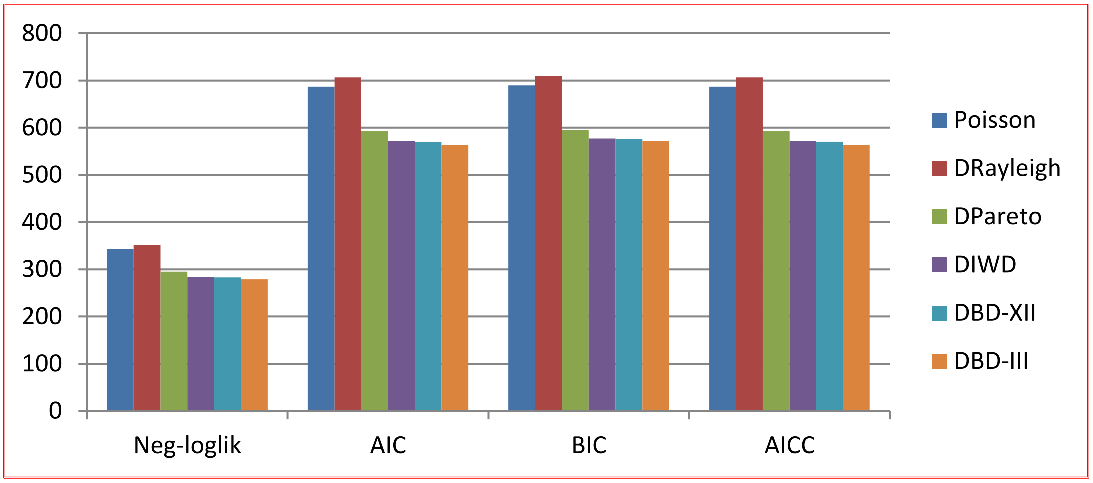

Fig. 19: Goodness of fit criterion for carious teeth data

7 Some Theorems Related to Three Parameter Discrete Burr Type III Distribution

Theorem 1: If X follows a three parameter Burr type III distribution with parameters (c, k, θ) then Y= [X] follows a three parameter discrete Burr type III distribution with parameters (c,q, θ)

Where $q = e^{-k}; 0 < q < 1; c > 0; k > 0; \theta > 0$

Proof:- Consider

$$P(Y \geq y) = P[[X] \geq y]$$

$$= P[X \geq y] because\ [X] \geq Y \overset{s}{\Leftrightarrow} X \geq Y$$

$$= 1 - (1 + \theta y^{-c})^{-k}$$

$$= 1 - q^{\log(1+\theta y^{-c})}$$

Which is the survival function of a three parameter discrete Burr type III distribution i.e., *DBD-III* (c, q, θ)

Theorem 2: If X ~ BD-III (c,k, θ) then $Y = \left[[\log(1 + \theta X^{-c})]^{-1/c}\right]$ follows discrete inverse Weibull distribution i.e., DIW (c, q)

$$q = e^{-k};\ \ 0 < q < 1\,; c > 0\,;\ \theta > 0; k > 0$$

Proof:-

$$P[Y \geq y] = P\left[\left[[\log(1 + \theta X^{-c})]^{-1/c}\right] \geq y\right]$$

$$= P\left[[\log(1 + \theta X^{-c})]^{-1/c} \geq y\right]$$

$$= P\left[X \geq [(e^{y^{-c}} - 1)/\theta]^{-1/c}\right]$$

$$= 1 - q^{log\left[1+\theta\left[[(e^{y^{-c}}-1)/\theta]^{-1/c}\right]^{-c}\right]}$$

$$= 1 - q^{\log e^{y^{-c}}} = 1 - q^{y^{-c}}$$

Which is the survival function of a discrete inverse Weibull distribution.

Theorem 3: If $X \sim$ BD-III (c, k, θ), then $Y = \left[\{\log(1+\theta X^{-c})\}^{1/c}\right]$ follows discrete Weibull distribution i.e., DWD (c, q) $q = e^{-k}$ $0 < q < 1$, k>0 , c>0; $\theta > 0$

Proof: Consider

$$P[Y \geq y] = P\left[\left[\{\log(1+X^{-c})\}^{1/c}\right] \geq y\right]$$

$$= 1 - P\left[X \geq [(e^{y^c}-1)/\theta]^{-1/c}\right]$$

$$= 1 - \left[1 - \left[(1+\left[\theta[(e^{y^c}-1)/\theta]^{-1/c}\right]^{-c}\right]^{-k}\right]$$

$$= q^{y^c} \text{ Where } q = e^{-k}$$

Which is the survival function of a discrete weibull distribution .Hence $Y \sim DW(c, q)$

Theorem 4: Let X be random variable following three parameter continuous Burr type III distribution having parameters (c,k, θ) with $E(X^r) < \infty$ $\forall\, r = 1,2,3 \ldots.$. Then $E(Y^r) < \infty$ where $Y = [X]$ following three parameter discrete Burr type III distribution with parameters (c,q, θ)

Proof: Proof is straight forward, since $0 \leq [X] \leq X$, so clearly if $E(X^r) < \infty$ $\forall$ r = 1,2,3……

Then $E([X]^r) < \infty$

Theorem 5: If X is a non-negative rv and t is the positive number, then $X_t = [X^t] \sim \mathrm{DBD-III}\left(\frac{c}{t}, q, \theta\right)$ iff $X \sim \mathrm{BD-III}\ (c, k, \theta)$

$q = e^{-k}; 0 < q < 1;\ \theta > 0; c > 0$

Proof: Let $X \sim \mathrm{BD-III}(c, k)$ then $\forall\, x = 0,1,2, \ldots$

$$P[X_t \geq x] = P\left[[X^t] \geq x\right] = P[X^t \geq x]$$

$$= P\left[X \geq x^{1/t}\right]$$

$$= 1 - \theta^{\log(1+\theta x^{-c/t})}$$

$\Rightarrow X_t \sim \mathrm{DBD-III}(\frac{c}{t}, q, \theta)$

References

[1] Burr, I. W. (1942). Cumulative Frequency Functions. *Ann. Math. Statist*, *13*(2), 215-232. doi:10.1214/aoms/1177731607

[2] Grainger, R. W., & Reid, D. (1954). Distribution of Dental Caries in Children. *Journal of Dental Research*, *33*(5), 613-623. doi:10.1177/00220345540330050501

[3] Krishna, H., & Singh Pundir, P. (2009). Discrete Burr and discrete Pareto distributions. *Statistical Methodology*, *6*(2), 177-188. doi:10.1016/j.stamet.2008.07.001

[4] Roy, D. (1993). Reliability Measures in the Discrete Bivariate Set-Up and Related Characterization Results for a Bivariate Geometric Distribution. *Journal of Multivariate Analysis*, *46*(2), 362-373. doi:10.1006/jmva.1993.1065

[5] Singh, C., & Snyder, C. W. (1975). Truncated Poisson Distribution with Optimization Techniques in the Epidemiology of Caries. *Journal of Dental Research*, *54*(5), 1046-1048. doi:10.1177/00220345750540050701

[6] Lingström, P., &Borrman, H. (1999). Distribution of dental caries in an early 17th century swedish population with special reference to diet. *International Journal of Osteoarchaeology, 9(6), 395-403.*

[7] Lewsey, J. D., & Thomson, W. M. (2004). The utility of the zero-inflated Poisson and zero-inflated negative binomial models: a case study of cross-sectional and longitudinal DMF data examining the effect of socio-economic status. *Commun Dent Oral Epidemiol*,32(3), 183-189. doi:10.1111/j.1600-0528.2004.00155.x

[8] Roy, D. (2003). The Discrete Normal Distribution. *Communications in Statistics - Theory and Methods*, *32*(10), 1871-1883. doi:10.1081/sta-120023256

[9] Roy, D. (2004). Discrete Rayleigh Distribution. *IEEE Transactions on Reliability*, *53*(2), 255-260. doi:10.1109/tr.2004.829161

[10] Knutson, J. W. (1958). Epidemiological trend patterns of dental caries prevalence data. *The Journal of the American Dental Association*, *57*(6), 821-829. doi:10.14219/jada.archive.1958.0214

[11] Spencer, A. J., & Lewis, J. (1987). The relationships between prevalence of caries, percentage with active caries, and caries severity. *Australian Dental Journal*, *32*(6), 441-445. doi:10.1111/j.1834-7819.1987.tb01296.x

B. A. Para is a research scholar in the Department of Statistics, University of Kashmir, Srinagar, India. His area of research is Statistical models in medical sciences. He has remained as a business analyst for two years in TCS Bangalore under CSSI Nielsen. Europe Project. He completed his M.phil on topic entitled "Some Burr type distributions" from University of Kashmir, Srinagar, India in 2014. He has published several research articles in reputed international journals of statistical and mathematical sciences.

T. R. Jan has done his post-graduation and M.Phil from Aligarh Muslim University and obtained Doctorial from University of Kashmir in 2004. He is currently working as Senior Assistant Professor in the Department of Statistics, University of Kashmir, Srinagar, India. He has visited some foreign universities in USA and Malaysia as well. His research interests are in the areas of Bio-statistics, Reliability theory and Generalised models in biological sciences. He has published several research articles in reputed international journals of mathematical and biosciences.