European Journal of Statistics and Probability

Vol.4, No.5, pp.1-10, October 2016

Published by European Centre for Research Training and Development UK (www.eajournals.org)

On A Family of Discrete Probability Distributions (FDPD)

Md Shakhawat Hossain

Department of Statistics, University of Chittagong, Chittagong-4331, Bangladesh.

ABSTRACT: *In this study, a generalized family of discrete probability distributions is discussed. I have derived a general form of probability mass function (pmf) for the family of discrete probability distributions (FDPD) which belongs all standard discrete probability distributions. From the derived single general pmf any one can easily find the pmf for all the members of FDPD by only suitable choices of some factors given in Table 5.1. Therefore it might be regarded as a convenient, time saving, simplex and generalized device in the theory of probability. Particular cases for standard discrete probability distributions are also illustrated. Further moment generating function and hence some moments, mean and variance of the FDPD discussed as its property.*

KEYWORDS: Family of Discrete Probability Distributions (FDPD), Generalized Device, Inversion Theorem, CF, MGF.

INTRODUCTION

Numerous classical distributions have been extensively used over a long period of time for modeling data in several areas such as engineering, actuarial, environmental and medical sciences, biological studies, demography, economics, finance and insurance etc. In such many applied areas, there is a clear need for extended forms of these distributions. For that reason, several methods for generating new families of distributions have been studied.

Barndorff-Nielsen [1] have developed the one parameter exponential family with probability function is given by

$$f(x;\theta) = \exp[a(x)b(\theta) + c(\theta) + h(x)] \qquad (1.1)$$

The binomial, Poisson, normal, exponential, gamma, geometric, Rayleigh etc. are distributions belonging to this family and can be obtained by suitable choices of a(x), b(θ), c(θ) and h(x). Here b(θ) is a non-trival continuous function of $\theta \epsilon \Omega(d_{1,} d_2)$ (Patel et al., [2]) where d_1 and d_2 are real numbers.

Rahman and Gupta [3], first introduced the transformed chi-square family of distributions which is a subfamily of the exponential family of distributions. According to them irrespective of the form of the distribution of a continuous type random variable X having

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

probability density function of the form (1.1) belonging to the transformed chi-square family. Some members of this family are exponential, gamma, Erlang, normal, lognormal, Maxwell, chi-square, Weibull and Pareto.

Rahman and Nahar [4] obtained a special type of sub family of the exponential family of distributions named "A generalized family of discrete distributions" by considering a(x)=X be a discrete type of random variable having pmf of the form (1.1) and $c(\theta) = kln[1 - \exp\{b(\theta)\}]$, where k is a known constant. Binomial, Poisson, negative binomial and geometric are some of the distributions belonging to generalized family of discrete distributions. They have shown that the characteristic function of the distribution of a(x) is

$$\phi_{a(x)}(t) = \frac{[1-\exp\{b(\theta)\}]^k}{[1-\exp\{b(\theta)+it\}]^k} \qquad (1.2)$$

Some attempts have been made to define new families of probability distributions that extend well-known families of distributions and at the same time provide great flexibility in modeling data in practice. One such example is a broad family of univariate distributions generated from the Weibull distribution introduced by Gurvich et al. [5], by extending the classical Weibull model.

Zografos and Balakrishnan [6] proposed and studied a broad family of univariate distributions through a particular case of Stacy's generalized gamma distribution.

For the ease and convenience to find the probability function of different discrete distributions an attempt is made to have a general form for the probability functions of discrete distributions.

This paper is unfolded as follows. In section 2, we derive a generalized probability mass function for the FDPD. Some particular cases of the family are shown in section 3. In section 4, a few properties of the FDPD are discussed. Finally the study is concluded in section 5.

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

General Form of Probability Function for the FDPD

We know that the characteristic function (CF) of a probability distribution uniquely determines the distribution. Thus we can easily find the probability function for the generalized family of discrete distributions using its characteristic function in (1.2).

Let $\exp\{b(\theta)\} = \frac{1}{ku(\theta)}$, where u(θ) is a function of θ. Then the characteristic function of the distribution of a(x) in (1.2) become as

$$\phi_{a(x)}(t) = \left(1 - \frac{1}{ku(\theta)}\right)^k \left(1 - \frac{\exp(it)}{ku(\theta)}\right)^{-k} \qquad (2.1)$$

Again consider $P = 1 - \frac{1}{ku(\theta)}$ and $Q = 1 - P = \frac{1}{ku(\theta)}$

To have the probability function now use "Inversion Theorem" with the characteristic function in (2.1) and have,

$$\mathfrak{I}_c = \int_{-c}^{c} \exp(-itx) \cdot \phi_{a(x)}(t)dt$$

$$= \int_{-c}^{c} \exp(-itx) \cdot P^k[1 - Qexp(it)]^{-k}dt$$

$$= P^k \int_{-c}^{c} \exp(-itx)\{1 - Qexp(it)\}^{-k}dt$$

$$= P^k \int_{-c}^{c} \exp(-itx) \sum_{j=0}^{\infty} \binom{-k}{j} \{-Qexp(it)\}^j \, dt$$

$$= P^k \sum_{j=0}^{\infty} \binom{-k}{j} \int_{-c}^{c} \exp(-itx)\{-Qexp(itj)\} \quad dt$$

$$= P^k \sum_{j=0}^{\infty} \binom{-k}{j} (-Q)^j \int_{-c}^{c} \exp\{-it(x-j)\} \, dt \qquad (2.2)$$

(i) If $x \neq j$, then from Eq. (2.2) we have

$$\mathfrak{I}_c = P^k \sum_{j=0}^{\infty} \binom{-k}{j} (-Q)^j \left| \frac{\exp\{-it(x-j)\}}{-i(x-j)} \right|_{-c}^{c}$$

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

$$= P^k \sum_{j=0}^{\infty} \binom{-k}{j} (-Q)^j \frac{\exp\{ic(x-j)\} - \exp\{-ic(x-j)\}}{i(x-j)}$$

$$= P^k \sum_{j=0}^{\infty} \binom{-k}{j} (-Q)^j \frac{2isin\{c(x-j)\}}{i(x-j)}$$

$$\therefore \lim_{c\to\infty} \frac{\mathfrak{I}_c}{2c} \to 0 \ \forall x$$

Hence there is no discontinuity in the distribution function when $x \neq j$.

(ii) If *x*=*j*, then from Eq. (2.2) we have

$$\mathfrak{I}_c = P^k \textstyle\sum_{j=0}^{\infty} \binom{-k}{j} (-Q)^j \int_{-c}^{c} dt$$

$$= P^k \cdot (1-Q)^{-k} \cdot 2c = 2c$$

Since $\frac{\mathfrak{I}_c}{2c} \to 1$ at x=j, the distribution function is discontinuous i.e. distribution is discrete and its probability function is

$$P(j) = \binom{-k}{j} (-Q)^j P^k \ ; \text{ j} = \ 0, 1, 2, 3, \cdots \qquad (2.3)$$

Replacing the values of P & Q and *j* by *a(x)* we have the following probability function of the generalized family of discrete distributions:

$$P\big(a(x)\big) = \binom{-k}{a(x)} \left(-\tfrac{1}{ku(\theta)}\right)^{a(x)} \left(1 - \tfrac{1}{ku(\theta)}\right)^k \ ; \ a(x) = 0, 1, 2, 3, \cdots \qquad (2.4)$$

This is the general form of probability mass function (pmf) of the family of discrete probability distributions (FDPD).

Now taking summation on both sides in Eq. (2.4) we have,

$$\sum_{a(x)=0}^{\infty} P\big(a(x)\big) = \sum_{a(x)=0}^{\infty} \binom{-k}{a(x)} \left(-\frac{1}{ku(\theta)}\right)^{a(x)} \left(1 - \frac{1}{ku(\theta)}\right)^k$$

$$= \left(1 - \frac{1}{ku(\theta)}\right)^k \sum_{a(x)=0}^{\infty} \binom{-k}{a(x)} \left(-\frac{1}{ku(\theta)}\right)^{a(x)}$$

$$= \left(1 - \frac{1}{ku(\theta)}\right)^k \left(1 - \frac{1}{ku(\theta)}\right)^{-k}$$

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

$$= \left(1 - \frac{1}{ku(\theta)}\right)^0 = 1$$

which proves that it is a probability function

Particular Cases of the FDPD

The probability functions of discrete probability distributions can be found from the general form probability function of the family of discrete distributions in Eq. (2.4).

3.1 *Bernoulli Probability Distribution*: Let us substitute $a(x) = x, u(\theta) = \frac{\theta-1}{k\theta}, and\ k = -1$ in Eq. (2.4).

$$\therefore P(x) = \left(-\frac{\theta}{\theta-1}\right)^x \left(1 - \frac{\theta}{\theta-1}\right)^{-1}$$

$$= \left(\frac{\theta}{1-\theta}\right)^x \left(\frac{1}{1-\theta}\right)^{-1}$$

$$= \theta^x (1-\theta)^{1-x}\ ; x = 0, 1 \qquad (3.1)$$

which is the probability mass function of Bernoulli distribution with parameter θ.

3.2 *Binomial Probability Distribution*: Let us substitute $a(x) = x, u(\theta) = \frac{\theta-1}{k\theta}, and\ k = -n,$ where n is positive integer, in Eq. (2.4).

$$\therefore P(x) = \binom{n}{x} \left(-\frac{\theta}{\theta-1}\right)^x \left(1 - \frac{\theta}{\theta-1}\right)^{-n}$$

$$= \binom{n}{x} \left(\frac{\theta}{1-\theta}\right)^x \left(\frac{1}{1-\theta}\right)^{-n}$$

$$= \binom{n}{x} \theta^x (1-\theta)^{n-x}\ ; x = 0, 1, 2, \cdots, n \qquad (3.2)$$

which is the probability function of binomial distribution with parameter n and θ.

3.3 *Poisson Probability Distribution*: Let us substitute $a(x) = x, u(\theta) = \frac{1}{\theta}, and\ k = \infty$ in Eq. (2.4).

$$\therefore P(x) = \lim_{k\to\infty} \binom{-k}{x} \left(-\frac{\theta}{k}\right)^x \left(1 - \frac{\theta}{k}\right)^k$$

$$= \lim_{k\to\infty} \frac{(-1)^x(-k)(-k-1)(-k-2)\cdots(-k-x+2)(-k-x+1)}{x!} \left(-\frac{\theta}{k}\right)^x \left(1 - \frac{\theta}{k}\right)^k$$

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

$$= \lim_{k\to\infty} \frac{k^x\left\{1.\left(1+\frac{1}{k}\right)\left(1+\frac{2}{k}\right)\cdots\left(1+\frac{x-2}{k}\right)\left(1+\frac{x-1}{k}\right)\right\}}{x!}\left(-\frac{\theta}{k}\right)^x\left(1-\frac{\theta}{k}\right)^k$$

$$= \frac{\theta^x}{x!}\lim_{k\to\infty}\left\{\left(1+\frac{1}{k}\right)\left(1+\frac{2}{k}\right)\cdots\left(1+\frac{x-2}{k}\right)\left(1+\frac{x-1}{k}\right)\right\}\left(1-\frac{\theta}{k}\right)^k$$

$$= \frac{\theta^x e^{-\theta}}{x!} \;;x = 0, 1, 2, \cdots \quad (3.3)$$

which is the probability function of Poisson distribution with parameter θ.

3.4 *Negative Binomial Probability Distribution*: Let us substitute

$a(x) = x, u(\theta) = \frac{1}{k(1-\theta)}, and\ k = m$, where m is positive integer, in Eq. (2.4).

$$\therefore \;\; P(x) = \binom{-m}{x}\{-(1-\theta)\}^x\{1-(1-\theta)\}^m$$

$$= \binom{-m}{x}(-1)^x\theta^m(1-\theta)^x$$

$$= \binom{x+m-1}{m-1}\theta^m(1-\theta)^x \;; x = 0, 1, 2, \cdots \quad (3.4)$$

which is the probability function of negative binomial distribution with parameter m and θ.

3.5 *Geometric Probability Distribution*: Let us substitute $a(x) = x, u(\theta) = \frac{1}{(1-\theta)}, and\ k = 1$ in Eq. (2.4).

$$\therefore \;\; P(x) = \theta(1-\theta)^x \;; x = 0, 1, 2, \cdots \quad (3.5)$$

which is the probability function geometric distribution with parameter θ.

Some Properties of FDPD

Moment Generating Function (PGF) of family of discrete distributions is

$$M_{a(x)}(t) = E(e^{ta(x)})$$

$$= \sum_{a(x)=0}^{\infty} e^{ta(x)}P(a(x))$$

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

$$= \sum_{a(x)=0}^{\infty} e^{ta(x)} \binom{-k}{a(x)} \left(-\frac{1}{ku(\theta)}\right)^{a(x)} \left(1-\frac{1}{ku(\theta)}\right)^{k}$$

$$= \sum_{a(x)=0}^{\infty} \binom{-k}{a(x)} \left(-\frac{e^t}{ku(\theta)}\right)^{a(x)} \left(1-\frac{1}{ku(\theta)}\right)^{k}$$

$$= \left(1-\frac{1}{ku(\theta)}\right)^{k} \sum_{a(x)=0}^{\infty} \binom{-k}{a(x)} \left(-\frac{e^t}{ku(\theta)}\right)^{a(x)}$$

$$= \left(1-\frac{1}{ku(\theta)}\right)^{k} \left(1-\frac{e^t}{ku(\theta)}\right)^{-k} \qquad (4.1)$$

$First\ raw\ moment,$ $\mu_1^{'} = \left\{\frac{d}{dt} M_{a(x)}(t)\right\}_{t=0}$

$$= \left[\frac{d}{dt}\left\{\left(1-\frac{1}{ku(\theta)}\right)^{k}\left(1-\frac{e^t}{ku(\theta)}\right)^{-k}\right\}\right]_{t=0}$$

$$= \left\{\left(1-\frac{1}{ku(\theta)}\right)^{k}(-k)\left(-\frac{e^t}{ku(\theta)}\right)\left(1-\frac{e^t}{ku(\theta)}\right)^{-k-1}\right\}_{t=0}$$

$$= \left\{\left(1-\frac{1}{ku(\theta)}\right)^{k}\left(\frac{e^t}{u(\theta)}\right)\left(1-\frac{e^t}{ku(\theta)}\right)^{-k-1}\right\}_{t=0}$$

$$= \left(1-\frac{1}{ku(\theta)}\right)^{k}\left(\frac{1}{u(\theta)}\right)\left(1-\frac{1}{ku(\theta)}\right)^{-k-1}$$

$$= \frac{\frac{1}{u(\theta)}}{\left(1-\frac{1}{ku(\theta)}\right)} \qquad (4.2)$$

Now

$Second\ raw\ moment,$

$$\mu_2^{'} = \left\{\frac{s}{dt}\left(1-\frac{1}{ku(\theta)}\right)^{k}\left(\frac{e^t}{u(\theta)}\right)\left(1-\frac{e^t}{ku(\theta)}\right)^{-k-1}\right\}_{t=0}$$

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

$$= \left(1 - \frac{1}{ku(\theta)}\right)^k \left\{\left(\frac{e^t}{u(\theta)}\right)\left(1 - \frac{e^t}{ku(\theta)}\right)^{-k-1} + \left(\frac{e^t}{u(\theta)}\right)(-k-1)\left(-\frac{e^t}{ku(\theta)}\right)\left(1 - \frac{e^t}{ku(\theta)}\right)^{-k-2}\right\}_{t=0}$$

$$= \left(1 - \frac{1}{ku(\theta)}\right)^k \left\{\left(\frac{e^t}{u(\theta)}\right)\left(1 - \frac{e^t}{ku(\theta)}\right)^{-k-1} + \left(\frac{e^t}{u(\theta)}\right)\left(\frac{(k+1)e^t}{ku(\theta)}\right)\left(1 - \frac{e^t}{ku(\theta)}\right)^{-k-2}\right\}_{t=0}$$

$$= \left(1 - \frac{1}{ku(\theta)}\right)^k \left\{\left(\frac{1}{u(\theta)}\right)\left(1 - \frac{1}{ku(\theta)}\right)^{-k-1} + \left(\frac{1}{u(\theta)}\right)\left(\frac{(k+1)}{ku(\theta)}\right)\left(1 - \frac{1}{ku(\theta)}\right)^{-k-2}\right\}$$

$$= \left(\frac{1}{u(\theta)}\right)\left\{\left(1 - \frac{1}{ku(\theta)}\right)^{-1} + \left(\frac{(k+1)}{ku(\theta)}\right)\left(1 - \frac{1}{ku(\theta)}\right)^{-2}\right\}$$

$$= \frac{\frac{1}{u(\theta)}}{\left(1 - \frac{1}{ku(\theta)}\right)} + \frac{\left(\frac{(k+1)}{k(u(\theta))^2}\right)}{\left(1 - \frac{1}{ku(\theta)}\right)^2} \qquad (4.3)$$

$Second\ central\ moment, \mu_2 = \mu_2' - (\mu_1')^2$

$$= \frac{\frac{1}{u(\theta)}}{\left(1 - \frac{1}{ku(\theta)}\right)} + \frac{\left(\frac{(k+1)}{k(u(\theta))^2}\right)}{\left(1 - \frac{1}{ku(\theta)}\right)^2} - \frac{\left(\frac{1}{u(\theta)}\right)^2}{\left(1 - \frac{1}{ku(\theta)}\right)^2}$$

$$= \frac{\frac{1}{u(\theta)}}{\left(1 - \frac{1}{ku(\theta)}\right)} \left\{1 + \frac{\left(\frac{(k+1)}{ku(\theta)}\right)}{\left(1 - \frac{1}{ku(\theta)}\right)} - \frac{\frac{1}{u(\theta)}}{\left(1 - \frac{1}{ku(\theta)}\right)}\right\}$$

$$= \frac{\frac{1}{u(\theta)}}{\left(1 - \frac{1}{ku(\theta)}\right)^2} \qquad (4.4)$$

Therefore,

$$Mean = \mu_1' = \frac{\frac{1}{u(\theta)}}{\left(1 - \frac{1}{ku(\theta)}\right)}$$

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

$$Variance = \mu_2 = \frac{\frac{1}{u(\theta)}}{\left(1 - \frac{1}{ku(\theta)}\right)^2}$$

By putting suitable values of *u(θ)* and *k* given in Table 5.1 in above formulas, one can easily get mean and variance of particular probability distribution.

For example, In case of binomial distribution putting $u(\theta) = \frac{\theta-1}{k\theta}$ and $k = -n$, we have

$$Mean = \mu_1^{'} = \frac{\frac{-n\theta}{\theta - 1}}{\left(1 - \frac{\theta}{\theta - 1}\right)} = \frac{\frac{n\theta}{1-\theta}}{\left(1 + \frac{\theta}{1-\theta}\right)} = n\theta$$

$$Variance = \mu_2 = \frac{\frac{-n\theta}{\theta - 1}}{\left(1 - \frac{\theta}{\theta - 1}\right)^2} = \frac{\frac{n\theta}{1-\theta}}{\left(1 + \frac{\theta}{1-\theta}\right)^2} = n\theta(1-\theta)$$

Similarly it is possible to have mean and variances of other members of the family of discrete distributions.

SUMMARY AND CONCLUSION

Suitable choices for *a(x), u(θ)* and *k* of finding the probability functions of different discrete distributions are summarized in Table 5.1.

Table 5.1: Suitable choices of *a(x), u(θ)* and *k* in Eq. (2.4) for different discrete probability distributions with the possible values of *x*.

Name of distributions	***a(x)***	***u(θ)***	***k***	**Possible values of *x***
Bernoulli	*x*	$\frac{\theta - 1}{k\theta}$	-1	0,1
Binomial	*x*	$\frac{\theta - 1}{k\theta}$	-*n*	0,1,2, …, n
Poisson	*x*	$\frac{1}{\theta}$	∞	0,1,2, …, ∞
Negative binomial	*x*	$\frac{1}{k(1-\theta)}$	*m*	0,1,2, …, ∞
Geometric	*x*	$\frac{1}{(1-\theta)}$	1	0,1,2, …, ∞

ISSN 2055-0154(Print), ISSN 2055-0162(Online)

In this work, I have discussed a special family of discrete probability distributions (FDPD) which belongs all the standard discrete probability distributions. The probability mass function (pmf) is essential in case of any probability distributions. For that reason a general form of pmf of the FDPD is derived in Eq. (2.4). It is a special type of pmf belonging all widely used discrete probability distributions and from which any one can easily find the pmf for different discrete probability distributions by suitable choices of $a(x)$, $u(\theta)$ and k given in Table 5.1. Particular cases for some standard discrete distribution are also illustrated in section 3. Therefore it might be considered as a convenient, time saving, simplex and generalized device in the theory of probability of getting pmf for all the members of FDPD. Properties like moment generating function (MGF) and hence some moments, mean and variance of FDPD are also discussed in this study. We hope this generalization may attract wider applications in the theory of probability and further research work, in theoretical as well as applied, of this field will validate from this study.

Conflict of Interest

The author declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgement
I would like to remember all of my colleagues and students for their suggestions and helps which encourages me very much to do this study.

REFERENCES

[1] Barndorff-Nielsen, O. (1978), Information and Exponential Families in Statistical Theory, John Wiley & Sons, New York.
[2] Patel, J.K., Kapania. C.H. and Owen, D.B. (1976), Hand Book of Statistical Distributions, Marcel Decker.
[3] Rahman, M. S. and Gupta (1993), "Family of Transformed Chi-square Distributions", *Commu. Stat. Theo. Math*., 22, 135-146.
[4] Rahman, M. S. and Nahar, S. (1995), "A Generalized Family of Discrete Distributions", *Journal of Bangladesh Academy of Sciences*, Vol. 19, No.2, Pg:237-240.
[5] Gurvich, M. R., DiBenedetto, A. T. and Ranade, S. V. (1997). A new statistical distribution for characterizing the random strength of brittle materials, *Journal of Materials Science*, 32, 2559-2564.
[6] Zografos, K. and Balakrishnan, N. (2009). On families of beta- and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, 6, 344-362.

ISSN 2055-0154(Print), ISSN 2055-0162(Online)