International Journal of Statistics and Systems
ISSN 0973-2675 Volume 19, Number 1 (2024), pp.13-21

https://dx.doi.org/10.37622/IJSS/19.1.2024.13-21

A Study on Discrete Pearsonian Family of Distributions

Anusha.R[1], Bhatracharyulu N. Ch[2] and *Jayasree. G[2]

[1]TGSWRDC (W), Warangal West, Kakatiya University, Warangal
[2]Department of Statistics, University College of Science, Osmania University, Hyderabad-7

Abstract

The Continuous Probability distributions belonging to this family are derived from a Pearson differential equation and Discrete Probability distributions are derived from a Pearson difference equation. Several authors made attempts in defining, deriving and studying this family of distributions. In this paper, we made an attempt to Study the discrete Pearsonian family of distributions and identifying the new members of discrete Pearsonian family of distributions.

Keywords: Hyper geometric distribution, Beta Pascal distribution, discrete student's t distribution, Mean and Variance.

1. INTRODUCTION

The Pearson system of continuous distributions is defined by the differential equation

$$\frac{df(x)}{dx}=\frac{(a-x)f(x)}{b_0+b_1 x+b_2 x^2},$$

where, f(x) is the probability density function of the random variable X and a, b0, b1, b2 are the parameters.

This Pearson system of continuous distributions offers a structure for describing about many significant continuous distributions. Similarly, the use of difference equation helps us to create a class of discrete distributions based on a unit-width lattice. The difference equation is,

$$\Delta f_{x-1} = \frac{a-x}{b_0+b_1 x+b_2 x(x-1)} f_{x-1}, \quad x \in T$$

where Δf_{x-1} is the finite difference and a, b0, b1, b2 are parameters.

First, Carver (1919) and Katz (1948) applied this technique to an actuarial data. Later, Ord (1967) developed a system of discrete distributions using the above difference equation based on the criterion $k=\frac{(b_1-b_2-1)^2}{4b_2(b_0+a)}$ and $I=\frac{\mu_2}{\mu_1'}$.

2. Study on Discrete Pearsonian Family of Distributions

Type – I: A discrete probability distribution satisfying the Criteria $I<1$ and for $k>1$; $k<0$; $k>1$; then the corresponding distribution belongs to the Type–I Discrete Pearson Family of Distributions.

Example:

Hyper Geometric Distribution: Let X be a discrete random variable following Hyper geometric distribution with parameters (N, M, n) then the Probability mass function is,

$$f(x)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}; \quad x=0,1,2\ldots\min(n,M);N,M,n>0$$

which can be written in the form of Discrete Pearson difference equation

$$\Delta f_{x-1}=\frac{a-x}{b_0+b_1 x+b_2 x(x-1)}\, f_{x-1} \text{ as } \Delta f_{x-1}=\left[\frac{\frac{(M+1)(n+1)}{N+2}-x}{\frac{(N-M-n+1)}{N+2} x+\frac{1}{N+2} x(x-1)}\right] f_{x-1};$$

where $a=\frac{(M+1)(n+1)}{N+2}$; $b_0=0$; $b_1=\frac{(N-M-n+1)}{N+2}$; $b_2=\frac{1}{N+2}$.

Consider the Ord Criteria $k=\frac{(b_1-b_2-1)^2}{4b_2 (b_0+a)}$ and $I=\frac{\mu_2}{\mu_1'}$

$$k=\frac{(b_1-b_2-1)^2}{4b_2 (b_0+a)}=\frac{\left(\frac{(N-M-n+1)}{N+2}-\frac{1}{N+2}-1\right)^2}{4\frac{1}{N+2}\left(0+\frac{(M+1)(n+1)}{N+2}\right)}=\frac{(M+n+2)^2}{4(M+1)(n+1)}$$

as N, M, n > 0 then k > 1

and for the criterion $I=\frac{\mu_2}{\mu_1'}$; consider the Mean and Variance of Hyper geometric distribution

Mean = (Mn) / N and Variance = [Mn(N-n) (N-M)] / N2(N-1)

then, $I=\frac{\frac{Mn(N-n)(N-M)}{N^2 (N-1)}}{Mn/N}=\frac{(N-n)(N-M)}{N(N-1)}$

as N, M, n > 0 then I < 1

According to Ord (1967), if I < 1 and k > 1 then the corresponding distribution belongs to Type I Discrete Pearsonian family.

Hence, Hyper Geometric distribution is a member of Discrete Pearsonian family of distributions.

Type – II: If the Criteria $I<1$ and for $k=1$; $k=0$; $k=1$; is satisfied then the corresponding distribution belongs to the Type–II Discrete Pearsonian Family of Distributions

Example: The Symmetric forms of Type – I

Type – III: If the Criteria $I<1$ and $k\to\infty$; $I>1$ and $k\to\infty$; $I=1$ and $k\to\infty$; is satisfied, then the corresponding distribution belongs to Type–III Discrete Pearson Family of Distributions.

Example: Poisson Distribution: Let X be a Discrete random variable following Poisson distribution with parameter λ then the probability mass function is, $f(x)=(e^{-\lambda}$

$\lambda^x)/x!; x=0,1,2\ldots..; \lambda>0;$

which can be written in the form of Pearson Difference equation $\Delta f_{(x-1)}=\frac{a-x}{b_0+b_1 x+b_2 x(x-1)} f_{(x-1)}$ as $\Delta f_{(x-1)}=\left[\frac{\lambda-x}{x}\right] f_{(x-1)}$; where, $a=\lambda$; $b0=0$; $b1=1$; $b2=0$.

Consider the Ord Criteria $k=\frac{(b_1-b_2-1)^2}{4b_2 (b_0+a)}$ and $I=\frac{\mu_2}{\mu_1'}$

$k=\frac{(b_1-b_2-1)^2}{4b_2 (b_0+a)}=\frac{(1-0-1)^2}{4(0)(0+\lambda)} \to \infty$

and for the criterion $I=\frac{\mu_2}{\mu_1'}$;

the Mean and Variance of Poisson distribution are Mean= λ and Variance = λ;

then, $I=\lambda/\lambda=1$

According to Ord (1967), if the Criteria $I=1$ and $k\to\infty$ then the corresponding distribution belongs to Type III Discrete Pearson family of distributions.

Hence, Poisson distribution is a member of Discrete Pearsonian family of distribution.

Type – IV: If the Criteria $0 < k < 1$ then the corresponding distribution belongs to Type–IV Discrete Pearson Family of Distributions.

Example: Let X be a discrete random variable with probability mass function

$f(x) = \frac{\alpha Q(x,a,d)}{Q(x,r+a,b)}$, $x>0$ and $f(x) = \frac{\alpha Q(x,a,d)}{Q(x,r+a,b)}$, $x<0$

where $Q(x,a,d)=(a^2+d^2)\{(a+1)^2+d^2\}\ldots\{(a+x)^2+d^2\}$

Type – V: If the Criteria $k = 0$ is satisfied then the corresponding distribution belongs to Type–V Discrete Pearson Family of Distributions.

Example: $b = 0$ in Type – IV.

Type – VI: If the Criteria $I > 1$ and $k > 1$ satisfied then the corresponding distribution comes under the Type–VI Discrete Pearson Family of Distributions.

Example: Beta Pascal Distribution: Let X be a discrete random variable following Beta Pascal distribution with parameters (n, t, s), then the probability mass function is

$f(x)=\binom{n-1+x}{x} \frac{B(n+t,s+x)}{B(t,s)}; x=0,1,2,\ldots; n,t,s>0;$

which can be written in the form of Pearson Difference equation $\Delta f_{(x-1)}=\frac{a-x}{b_0+b_1 x+b_2 x(x-1)} f_{(x-1)}$ as $\Delta f_{(x-1)}=\left[\frac{\frac{(s-1)(n-1)}{1+t}-x}{\frac{n+t+s}{1+t}x+\frac{1}{1+t}x(x-1)}\right] f_{(x-1)}$

where, $a = \frac{(s-1)(n-1)}{1+t}$; $b0 = 0$; $b1 = \frac{n+t+s}{1+t}$; $b2 = \frac{1}{1+t}$

Consider, the Ord Criteria $k=\frac{(b_1-b_2-1)^2}{4b_2 (b_0+a)}$ and $I=\frac{\mu_2}{\mu_1'}$; then

$k=\frac{(b_1-b_2-1)^2}{4b_2 (b_0+a)}=\frac{\left(\frac{n+t+s}{1+t}-\frac{1}{1+t}-1\right)^2}{4\frac{1}{1+t}\left(0+\frac{(s-1)(n-1)}{1+t}\right)}=\frac{(n+s-2)^2}{4(s-1)(n-1)}$

as n, s > 0 then k > 1

and for the criterion I=μ_2/(μ_1^'); the Mean and Variance of Beta Pascal distribution is

Mean = ns/(t-1) and Variance = (ns(n+t-1)(s+t-1))/((t-2)⟦(t-1)⟧^2)

then, I=((ns(n+t-1)(s+t-1))/((t-2)⟦(t-1)⟧^2))/(ns/(t-1))=((n+t-1)(s+t-1))/((t-2)(t-1))

as n, t, s > 0 then I > 1.

According to Ord (1967), if the Criteria I > 1 and k > 1 then the corresponding distribution belongs to Type VI Discrete Pearsonian family of distributions.

Hence, Beta Pascal distribution is a member of Discrete Pearsonian family of distribution.

Type – VII: If the Criteria 0 < k < 1 is satisfied, then the corresponding distribution belongs to Type–VII Discrete Pearsonian Family of Distributions.

Example: Symmetric form of Type – IV.

Table 1: Discrete Pearsonian Family of Distributions Types, Probability MassFunction, Criteria, and Range

Type		Probability Mass function	Distribution	Criteria $I = \frac{\mu_2}{\mu_1'}$; $k = \frac{(b_1 - b_2 - 1)^2}{\{4b_2(b_0 + a)\}}$	Range
I	A	$\binom{M}{x}\binom{N-M}{n-x}/\binom{N}{n}$	Hyper geometric	$I < 1$, $k > 1$	$[0, \min(n, M)]$
	B	$\binom{x+r-1}{x}\binom{N-x-r}{M-x}/\binom{N}{M}$	Negative hypergeometric or beta-binomial	$k < 0$	$[0, M]$
	C	$\binom{A}{x}\binom{C}{B-x}/\binom{A+C}{B}$	--	$k > 1$	$[0, \infty)$
	D	$\alpha\left\{\binom{A}{C+x}\binom{B}{D-x}\right\}^{-1}$	--	$k > 1$	$[0, n]$
II	A	**As for type I (.)**	--	$I < 1$, $k = 1$	**As type I (.)**
	B			$k = 0$	

	D			$k = 1$	
III	A	$\binom{n}{x} p^x (1-p)^{n-x}$	Binomial	$I < 1,$ $k \to \infty$	$[0,\ n]$
	B	$\binom{x+r-1}{x} p^r (1-p)^x$	Negative binomial or Pascal	$I > 1,$ $k \to \infty$	$[0, \infty)$
	C	$e^{-\lambda}\lambda^x / x!$	Poisson	$I = 1,$ $k \to \infty$	$[0, \infty)$
IV		$\frac{\alpha Q(x,a,d)}{Q(x,r+a,b)},\ x > 0$ $\frac{\alpha Q(x,a,d)}{Q(x,r+a,b)},\ x < 0$ $Q(x,a,d) = (a^2+d^2)\{(a+1)^2 + d^2\} \ldots \{(a+x)^2 + d^2\}$	--	$0 < k < 1$	$(-\infty,\ \infty)$
V		As type IV, but b=0	--	$k = 0$	$[0,\ \infty)$ or $(-\infty,\ \infty)$
VI		$\binom{n-1+x}{x} \frac{B(n+a, b+x)}{B(a,b)}$	Beta Pascal	$I > 1,$ $k > 1$	$[0,\ \infty)$
VII		$\alpha \left[\prod_{i=1}^{r} \{(x+i+a)^2 + b^2\} \right]^{-1}$	Discrete Student's t	$0 < k < 1$	$(-\infty,\ \infty)$

3. New Members of Discrete Pearsonian Family of Distributions

1. **Geometric Distribution:** Let X be a discrete random variable following Geometric distribution with parameters (1, q) then the probability mass function is,

$$f(x) = pq^{x-1};\ x = 1,2,3 \ldots;\ 0 < p,\ q<1,\ p+q=1.$$

which can be written in the form of Pearson Difference equation

$$\Delta f_{x-1} = \frac{a-x}{b_0 + b_1 x + b_2 x(x-1)}\ f_{x-1} \text{ as } \Delta f_{x-1} = \left[\frac{0-x}{\frac{1}{p}x}\right] f_{x-1};\text{ where, a=0; } b_0 = 0;\ b_1 = \frac{1}{p};\ b_2 = 0.$$

Consider the Ord Criteria $k = \frac{(b_1 - b_2 - 1)^2}{\{4b_2(b_0 + a)\}}$ and $I = \frac{\mu_2}{\mu_1'}$

$$k = \frac{(b_1 - b_2 - 1)^2}{\{4b_2(b_0 + a)\}} = \frac{\left(\frac{1}{p} - 0 - 1\right)^2}{\{4(0)(0+0)\}} \to \infty$$

And for $I = \frac{\mu_2}{\mu_1'}$, the Mean and Variance of Beta Pascal distribution are Mean = $\frac{q}{p}$ and Variance = $\frac{q}{p^2}$,

then, $I = \frac{\frac{q}{p^2}}{\frac{q}{p}} = \frac{1}{p}$; for 0 < p < 1 then I>1.

According to Ord (1967), if the Criteria I>1 and k→∞ is satisfied, then the corresponding distribution comes under Type III Discrete Pearsonian family of distributions.

Hence, Geometric distribution is a member of Discrete Pearsonian family of distribution.

2. **Beta Geometric Distribution:** Let X be a Discrete random Variable which is said to follow beta Geometric distribution with Parameters (1, t, s) then the probability Mass function is, $f(x) = \frac{B(1+t, s+x)}{B(t,s)}; x = 0,1,2, \ldots;\ t, s > 0$

which can be written in the form of Pearson Differential equation

$$\Delta f_{x-1} = \frac{a - x}{b_0 + b_1 x + b_2 x(x-1)} f_{x-1} \text{ as } \Delta f_{x-1} = \left[\frac{0 - x}{\frac{1+t+s}{1+t}x + \frac{1}{1+t}x(x-1)}\right] f_{x-1}$$

where, a = 0; b_0 = 0; $b_1 = \frac{1+t+s}{1+t}$; $b_2 = \frac{1}{1+t}$

Consider the Ord Criteria $k = \frac{(b_1 - b_2 - 1)^2}{\{4b_2(b_0 + a)\}}$ and $I = \frac{\mu_2}{\mu_1'}$

$$k = \frac{(b_1 - b_2 - 1)^2}{\{4b_2(b_0 + a)\}} = \frac{\left(\frac{1+t+s}{1+t} - \frac{1}{1+t} - 1\right)^2}{\left\{4\frac{1}{1+t}(0+0)\right\}} \to \infty$$

and for $I = \frac{\mu_2}{\mu_1'}$, the Mean and Variance of Beta Pascal distribution are Mean = $\frac{s}{t-1}$ and

Variance = $\frac{ts(s+t-1)}{(t-2)(t-1)^2}$

then, $I = \frac{\frac{ts(s+t-1)}{(t-2)(t-1)^2}}{\frac{s}{t-1}} = \frac{t(s+t-1)}{(t-2)(t-1)}$ as s, t > 0 then I>1.

According to Ord (1967), if the Criteria I>1 and k→∞ is satisfied then the corresponding distribution belongs to Type III Discrete Pearsonian family of distributions.

Hence, Beta Geometric distribution is a member of Discrete Pearsonian family of distribution.

Table 2: New Members of Discrete Pearsonian Family of Distributions Types, Probability MassFunction, Criteria, and Range

Type	Probability Mass function	Distribution	Criteria $I = \frac{\mu_2}{\mu_1'}; k = \frac{(b_1 - b_2 - 1)^2}{\{4b_2(b_0 + a)\}}$	Range
III	pq^{x-1}	Geometric	$I > 1$, $k \to \infty$	$[1, \infty]$
	$\frac{B(1+t, s+x)}{B(t,s)}$	Beta Geometric	$I > 1$, $k \to \infty$	$[0, \infty]$

References

[1] Carver, H. C. (1919). On the graduation of frequency distributions. Proc. Cas8ualty Act. Soc. Am. 6, pp. 52-72.

[2] J.K. Ord (1967): On a system of discrete distributions, Biometrika 54: pp 649-656

[3] J.K. Ord (1968): The discrete student's t distribution, Annals of Math. Stat., 39: 1513-1516.

[4] Jorge Ollero and Hector M. Ramos (1995): Description of a sub family of the discrete Pearson system as Generalized binomial distribution, Ital. Statist. Soc.: pp. 235-249.

[5] Katz, L. (1948): Frequency functions defined by the Pearson difference equation. Ann. Math. Statist. Vol. 19, 120.

[6] Noack, A. (1950): A Class of random variables with discrete distributions. Ann. Math. Statist. Vol. 21, pp. 127-132.

[7] Norman L. Johnson, Samuel Kotz and N. Balakrishnan (1994): "Continuous univariate distributions", Vol-1 second Edition, A Wiley Interscience Publication.

[8] Norman L. Johnson, A.W. Kemp and Samuel Kotz (2005): "Univariate Discrete Distributions" Third Edition, A Wiley Interscience Publication.

[9] Pearson K (1895): "Contributions to the mathematical theory of evolution, II: Skew variation in homogeneous material." Philos. Trans. Royal Soc. London. A, Vol. 186, pp. 343-414.

[10] Pearson K (1901): "Mathematical contributions to the theory of evolution, X: supplement to a memoir on skew variation", Philosophical Transactions of the Royal Society of London. A, Vol. 68, pp. 442-450.

[11] Pearson K (1916): "Mathematical contributions to the theory of evolution, XIX: Second supplement to a memoir on skew variation". Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, Vol. 216, pp. 429-457.