International Journal of Mathematics, Statistics and Operations Research
Volume 1; Number 2; 2021 : pp. 175-185

ARF INDIA
Academic Open Access Publishing
www.arfjournals.com

TWO NEW PROBABILITY DISTRIBUTIONS

T.A.Rather and N.A.Rather

Department of Mathematics, University of Kashmir, Srinagar, 190006, India.
Email: tariqrather@redimail.com

ARTICLEINFO

Received: 12 November 2021
Revised: 30 November 2021
Accepted: 7 December 2021
Online: 30 December 2021

To cite this paper:
T.A. Rather & N.A. Rather (2021). Two New Probability Distributions. *International Journal of Mathematics, Statistics and Operations Research.* 1(2): pp. 175-185.

ABSTRACT

In this paper, we present two new statistical probability models connecting $n + 1$, non-negative real parameters, $p_0, p_1, p_2, \ldots, p_n$ such that $0 < p_n \geq p_n \geq \ldots p_2 \geq p_1 \geq 0$. The rst model concerns with random variable of discrete type and the second model is its analogue for the random variable of continuous type.General formulas the mean, variance, m.g.f., the rth moment, the skewness and the kurtosis of the discrete model are obtained. In case of continuous model for $n = 1$, we also obtain the mean, variance and the m.g.f.

Key words: Probability density function, statistical model, moments, kurtosis, skewness and moment generating functions.

1. INTRODUCTION

The probability distributions plays a dominant role in modelling processes and it is often used to derive more realistic models that need characteristic observed data. "The theory of probability which originated in consideration of games of chance should have become the most important object of human knowledge". Blaise Pascal and Pierre Fermat are credited with founding mathematical probability because they solved the problem of points, the problem of equitably dividing the stakes when a fair game is halted before either player has enough points to win. This problem had been discussed for several centuries before 1654 but Pascal and Fermat were the first to give the solution we now consider correct. During the century that followed this work, other authors including James and Nicholas Bernoulli, Pierre Remond de Montmort and Abraham De Moivre developed more powerful techniques in order to calculate odds in more complicated games. De Moivre, Thomas Simpson and others also used the theory to calculate fair prices for annuities and insurance policies (for reference see S. Ross (2006), G. Shefer (1976, 1982, 1987, 1990, 1992) and T.M. Apostol (1969)).

Statistical models describe a phenomenon in the form of mathematical equations. Thus a large number of observations say 100 or 1000 can be summarized in an equation with say two unknown quantities (called parameters of the model). Such reduction is certainly necessary for human mind. Out of a large number of methods and tools developed so far for analysing data on the life of science etc.), the statistical models are the latest innovations. In the literature Hogg and Crag(1970), Johnson and Kotz (1970), Lawless (1982), S.C. Malik (1984), DeGroot and Schervish (1995), P. Fan (2006), McCulah and Nelder(1989), J.V. Uspensky (1937), we come across dierent types of models, e.g., Linear models, Non-linear models, Generalized linear models, etc. Statistical models form a basic and promising field of study in the domain of statistics and have many important applications in a wide variety of disciplines such as social sciences, biological and medical sciences, physical sciences, operation research, quality-control, engineering, agriculture and so on (Chakraborti. S (2015), D. Gupta (1993), G. Deniz (2010), I.S. Kozubowski (2006), C.D. Lai (2012), A. Jamjoom (2013), G. Deniz (2014)).

Let $f(x) = p_x x_n + p_{n-1}x^{n-1} + ... + p_1 x + p_0$ be a polynomial of degree $n \geq 1$ with real coefficients. It is interesting to ask under what conditions of the coefficient $p_0, p_1, ...+ p_n$, the polynomial $f(x)$ represents a statistical model. In case some of the individual coefficients p_i, $0 \leq i \leq n$, of the polynomial $f(x)$ are negative, it is more dicult to guaranty the necessity of positive values for the probability distyribution $f(x)$. Here, we consider the case when the coefficients $p_0, p_1, p_2 ..., p_{n-1}$ of the polynomial $f(x)$ are non-negative with $p_n > 0$.

In fact, the main aim of this paper is to present two new statistical probability models connecting $n + 1$, non-negative real parameters, $p_0, p_1, p_2, ..., p_n$ such that $0 < p_n \geq p_{n-1} \geq ... \geq p_2 \geq p_1 \geq 0$. The first model concerns with random variable of discrete type and the second model is its analogue for the random variable of continuous type.

Consider the probability distribution of the tossing of a coin, if we denote tails by $x = 0$ and heads by $x = 1$, then the distribution of the probability of this experiment can be written as

X	0	1
$P(X = x)$	1/2	1/2

which further can be briey written as

$$f(x) = 1/2, x = 0, 1 \quad (1.1)$$
$$= 0, \text{elsewhere}, x = 0, 1$$

Here, we prove the following result which involves $(n + 1)$ parameters and includes (1:1) as a special case.

Theorem 1: If $p_0, p_1, p_2, \ldots, p_n$ be non-negative real numbers such that $0 < p_n \geq p_{n-1} \geq \ldots \geq p_2 \geq p_1 \geq p_0 \geq 0$ and if

$$f(x) = f(x, p_0, p_1, p_2, \ldots, p_n)$$

$$= \frac{1}{p_0 + p_n}[(p_n - p_{n-1})x^n + (p_{n-1} - p_{n-2})x^{n-1} + \ldots + (p_1 - p_0)x + p_0], \qquad (1.2)$$

$$= 0, \text{ elsewhere, } x = 0, 1,$$

then the function $f(x)$ is probability mass function (p.m.f) of the random variable X of discrete type.

Remark: Taking $p_1 = p_2 =, \ldots, = p_n$ in Theorem 1, we get (1:1).

Proof of Theorem 1: We have

$$f(0) = \frac{p_0}{p_0 + p_n} \geq 0,$$

$$f(1) = \frac{p_n}{p_0 + p_n} \geq 0,$$

so that

$$\sum_{x=0}^{1} f(x) = f(0) + f(1) = 1.$$

Since $f(x) \geq 0$ for all real x. Therefore, $f(x)$ is a probability density function of random variable X of the Discrete type.

Mean, Variance and m.g.f of Theorem 1

The mean is given by

$$\mu = E(X) = \sum_{x=0}^{1} x f(x)$$

$$= 0f(0) + 1f(1)$$

$$= \frac{p_n}{p_0 + p_n}.$$

Now

$$E(X^2) = \sum_{x=0}^{1} x^2 f(x)$$

$$= f(1) = \frac{p_n}{p_0 + p_n},$$

so that

$$V\,ar(X) = \alpha^2 = E(X^2) - E(X)^2$$

$$= \left(\frac{p_n}{p_0 + p_n}\right) - \left(\frac{p_n}{p_0 + p_n}\right)^2$$

$$= \left(\frac{p_n}{p_0 + p_n}\right)\left(1 - \frac{p_n}{p_0 + p_n}\right)$$

$$= \left(\frac{p_n}{p_0 + p_n}\right)\left(\frac{p_0}{p_n + p_n}\right)$$

$$= \frac{p_0 p_n}{(p_0 + p_n)^2}.$$

The r^{th} moment of the distribution 1.1 is given by

$$E(X^r) = \sum_{x=0}^{1} x^r f(x)$$

$$= f(1) = \frac{p_n}{p_0 + p_n},\ r = 1, 2, \ldots$$

The m.g.f of the Theorem 1 is given by

$$M(t) = E(t^X)$$

$$= \sum_{x=0}^{1} x^{tx} f(x)$$

$$= f(0) + e^t f(1)$$

$$= \left(\frac{p_n}{p_0 + p_n}\right) + \left(\frac{p_n e^t}{p_0 + p_n}\right)$$

$$= \frac{p_0 + p_n e^t}{p_0 + p_n}, \text{ for all real } t.$$

Skewness and Kurtosis of Theorem 1

Skewness is given by

$$E\left(\frac{X - \mu}{\sigma}\right)^3 = E\left(\frac{(X - \mu)^3}{\sigma^3}\right)$$

$$= \frac{1}{\sigma^3}(E(X^3 - 3X^2\mu + 3X\mu^2 - \mu^3))$$

$$= \frac{1}{\sigma^3}(E(X^3) - 3\mu E(X^2) + 3\mu^2 E(X) - \mu^3)$$

$$= \frac{1}{\sigma^3}(\mu - 3\mu \cdot \mu + 3\mu^2 \cdot \mu - \mu^3)$$

$$= \frac{1}{\sigma^3}(1 - 3\mu^2 + 3\mu^2 - \mu^3)$$

$$= \frac{\mu}{\sigma^3}(1 - 3\mu^2 + 2\mu^2)$$

$$= \frac{\mu}{\sigma^3}\left(1 - 3\frac{p_n}{p_0 + p_n} + 2\frac{p_n^2}{p_0 + p_n}\right)$$

$$= \frac{\mu}{\sigma^3}\left(\frac{(p_0 + p_n)^2 - 3p_n(p_0 - p_n) + 2p_n^2}{(p_0 + p_n)^2}\right)$$

$$= \frac{\mu}{\sigma^3}\left(\frac{p_0^2 + 2p_0p_n + p_n^2 - 3p_0p_n - 3p_n^2 + 2p_n^2}{(p_0 + p_n)^2}\right)$$

$$= \frac{\mu}{\sigma^3}\left(\frac{(p_0^2 - p_np_0)}{(p_0 + p_n)^2}\right)$$

$$= \frac{\mu}{\sigma^3}\left(\frac{p_0(p_0 - p_n)}{(p_0 + p_n)^2}\right)$$

$$= \left(\frac{p_n}{(p_0 + p_n)}\right)\left(\frac{(p_0 + p_n)^3}{(p_0p_n)^{\frac{3}{2}}}\right)\left(\frac{p_0(p_0 - p_n)}{(p_0 + p_n)^2}\right)$$

$$= \frac{(p_0 - p_n)}{(p_0p_n)^{\frac{1}{2}}}.$$

Similarly, kurtosis is given by

$$E\left(\frac{X - \mu}{\sigma}\right)^4 = E\left(\frac{(X - \mu)^4}{\sigma^4}\right)$$

$$= \frac{1}{\sigma^4}(E(X^4 - 4X^3\mu + 6X^2\mu^2 - 4X\mu^3 + \mu^4))$$

$$= \frac{1}{\sigma^4}(E(X^4) - 4\mu E(X)^3 + 6\mu^2 E(X)^2 - 4\mu^3 E(X) + \mu^4)$$

$$= \frac{1}{\sigma^4}(\mu - 4\mu\cdot\mu + 6\mu^2\cdot\mu - 4\mu^3\cdot\mu + \mu^4)$$

$$= \frac{1}{\sigma^4}(1 - 4\mu^2 + 6\mu^2 - 4\mu^4 + \mu^4)$$

$$= \frac{\mu}{\sigma^4}(1 - 4\mu + 6\mu^2 - 3\mu^3)$$

$$= \frac{\mu}{\sigma^4}\left(1 - \frac{4p_n}{(p_0 + p_n)} + \frac{6p_n^2}{(p_0 + p_n)^2} - \frac{3p_n^2}{(p_0 + p_n)^3}\right)$$

$$= \frac{\mu}{\sigma^4}\left(\frac{(p_0 + p_n)^2 - 4p_n(p_0 \_ p_n)^2 + 6p_n^2(p_0 + p_n) - 3p_n^3}{(p_0 + p_n)^3}\right)$$

$$= \frac{\mu}{\sigma^4(p_0 + p_n)^3}\left(p_0^3 - p_0^2 p_n + p_0 p_n^2\right)$$

$$= \frac{\mu p_0}{\sigma^4(p_0 + p_n)^3}\left(p_0^2 - p_0 p_n + p_n^2\right)$$

$$= \left(\frac{p_0 p_n}{(p_0 + p_n)^4}\right)\left(\frac{(p_0 + p_n)^4}{p_0^2 p_n^2}\right)\left(p_0^2 - p_0 p_n + p_n^2\right)$$

$$= \frac{p_0^2 - p_0 p_n + p_n^2}{p_0 p_n}.$$

A random variable X of continuous type is said to be uniform distributed over the interval (0, 1), if its p.d.f. is given by

$$f(x) = 1,\ 0 < x < 1 \qquad (1.3)$$
$$= 0,\ \text{elsewhere.}$$

Finally, we present the following result which is a generalization of (1:3) for the random variable X of the continuous type involving $(n + 1)$ parameters.

Theorem 2: Let $p_0, p_1, p_2, \ldots, p_n$ be non-negative real numbers such that $0 < p_n \geq p_{n-1} \geq \ldots \geq p_2 \geq p_1 \geq p_0 \geq 0$. Then the function

$$f(x) = f(x, p_0, p_1, p_2, \ldots, p_n)$$
$$= \frac{1}{p_n}\,[(n+1)(p_n - p_{n-1})x^n + n(p_{n-1} - p_{n-2})x^{n-1} + \ldots + 2(p_1 - p_0)x + p_0], \qquad (1.4)$$
$$= \text{elsewhere},\ 0 < x < 1,$$

is probability density function (p.d.f) of the random variable X of the continuous type.

Remark: Taking $p_0 = p_1 =, \ldots, = p_n$ in Theorem 2, we get (1.3).

Proof of Theorem 2: Clearly $f(x) \geq 0$ for all x, we show

$$\int_0^1 f(x)dx = 1.$$

We have

$$\int_0^1 f(x)dx = \int_0^1 \frac{1}{p_n}((n+1)(p_n - p_{n-1})x^n + \ldots + 2(p_1 - p_0)x + p_0)dx$$

$$= \frac{1}{p_n}\left((p_n - p_{n-1})\int_0^1 (n+1)x^n dx + \ldots + (p_1 - p_0)\int_0^1 2x dx + p_0 \int_0^1 dx\right)$$

$$= \frac{1}{p_n}((p_n - p_{n-1})[x^{n+1}]_0^1 + \ldots + (p_1 - p_0)[x^2]_0^1 + p_0[x]_0^1)$$

$$= \frac{1}{p_n}(p_n - p_{n-1} + \ldots + p_1 - p_0 + p_0)$$

$$= \frac{1}{p_n}(p_n) = \frac{p_n}{p_n} = 1.$$

Hence $f(x)$ is a p.d.f. of the random variable X of continuous type. This completes the proof of Theorem 2.

A Special case

If we take $n = 1$, in Theorem 2, we get

$$f(x) = \frac{1}{p_1}(2(p_1 - p_0)x + p_0),\ 0 < x < 1,\ 0 < p_2 > p_1 > p_0 \geq 0 \qquad (1.5)$$

$= 0$, elsewhere.

Here mean of the distribution (1.3) is given by

$$\mu = E(X) = \frac{1}{p_1}\int_0^1 x(2(p_1 - p_0)x + p_0)dx$$

$$= \frac{1}{p_1}\int_0^1 (2(p_1 - p_0)x^2 + p_0 x)dx$$

$$= \frac{1}{p_1}\int_0^1 \left[\frac{2}{3}(p_1 - p_0)x^3 + p_0\frac{x^2}{2}\right]_0^1$$

$$= \frac{1}{p_1}\left(\frac{2}{3}(p_1 - p_0) + \frac{p_0}{2}\right)$$

$$= \frac{1}{p_1}\left(\frac{4p_1 - p_0}{6}\right)$$

$$= \frac{4p_1 - p_0}{6p_1}.$$

Now

$$E(X^2) = \frac{1}{p_1}\int_0^1 x^2(2(p_1 - p_0)x + p_0)dx$$

$$= \frac{1}{p_1}\int_0^1 (2(p_1 - p_0)x^3 + p_0 x^2)dx$$

$$= \frac{1}{p_1}\left[\frac{2}{3}(p_1 - p_0)x^4 + p_0\frac{x^3}{3}\right]_0^1$$

$$= \frac{1}{p_1}\left(\frac{p_1 - p_0}{2} + \frac{p_0}{3}\right)$$

$$= \frac{3p_1 - p_0}{6p_1}.$$

So that, variance of (1.3) is given by

$$\sigma^2 = E(X - \mu)^2$$

$$= E(X^2) - \sigma^2$$

$$= \left(\frac{3p_1 - p_0}{6p_1}\right) - = \left(\frac{4p_1 - p_0}{6p_1}\right)^2$$

$$= \left(\frac{6p_1(3p_1 - p_0) - (4p_1 - p_0)^2}{36p_1^2}\right)$$

$$= \frac{18p_1^2 - 6p_1p_0 - (16p_1^2 - 8p_1p_0 + p_0^2)}{36p_1^2}$$

$$= \frac{18p_1^2 - 6p_1p_0 - 16p_1^2 - 8p_1p_0 + p_0^2}{36p_1^2}$$

$$= \frac{2p_1^2 - 2p_1p_0 - p_0^2}{36p_1^2}$$

Again, m.g.f. of the distribution 1.3 is

$$M(t) = E(e^{tX})$$

$$= \int_0^1 e^{tx} \frac{1}{p_1}(2(p_1 - p_0)x + p_0)dx$$

$$= \frac{2(p_1 - p_0)}{p_1}\int_0^1 xe^{tx}dx + \frac{p_0}{p_1}\int_0^1 e^{tx}dx$$

$$= \frac{2(p_1 - p_0)}{p_1}\left(\left[x\frac{e^{tx}}{t}\right]_0^1 - \int_0^1 \frac{e^{tx}}{t}dx\right) - \frac{p_0}{p_1}\left[x\frac{e^{tx}}{t}\right]_0^1$$

$$= \frac{2(p_1 - p_0)}{p_1}\left(\left[\frac{e^t}{t}\right] - \frac{1}{t}\left[\frac{e^{tx}}{t}\right]_0^1\right) - \frac{p_0}{p_1}\left(\frac{e^t}{t} - \frac{1}{t}\right)$$

$$= \frac{2(p_1 - p_0)}{p_1}\left(\frac{e^t(t-1)+1}{t^2}\right) - \frac{p_0}{p_1}\left(\frac{e^t}{t} - \frac{1}{t}\right),$$

for all real values of t.

Remark 1: As in the case $n = 1$, a number of dierent interesting models can be obtained from Theorem 2, for other values of $n = 2, 3, 4...$

CONCLUSIONS

In this paper the author concludes the following:

1. For $p_0 = p_1 =, ..., = p_n$ the statistical probability model connecting $n + 1$ non-negative real numbers $p_0, p_1, ..., p_n$ reduces to probability model of a tossing of a coin.
2. For $p_0 = p_1 =, ..., = p_n$ the statistical probability model connecting $n + 1$ non-negative real numbers $p_0, p_1, ..., p_n$ reduces to Uniform distribution of the continuous type.
3. For various numerical values of the parameters $p_0, p_1, ..., p_n$, the two models presented in the paper yield a number of statistical models which forms the bases of further investigations.

Acknowledgements

The author is highly grateful to the referee whose commendable suggestions and comments added greatly to the quality of this paper.

COMPETING INTERESTS

Author has declared that no competing interests exist.

References

1. Malik, S.C, Mathematical Analysis, Wiley Eastern Limited, New Delhi, 1984.
2. DeGroot M.H. and Schervish M.J., Probability and Statistics, 3rd Edition, Addison-Wiley, 1995.
3. Shafer G., A Mathematical Theory of Evidence, Princeton Univ. Press, Princeton N.J., 1976.
4. Shafer G., Belief functions and parametric models (with discussion), J. Roy. Stat. Soc. Ser. B44, 322-352, 1982.
5. Shafer G., Probability judgement in articial intelligence and expert systems (with discussion), Stat. Sci. 2, 3-44, 1987.
6. Shafer G., The unity and diversity of probability (with discussion), Stat. Sci. 5, 435-462,1990.
7. Shafer G., Rejoinders to Comments on Perspective on the Theory and Practice of Belief Functions, International Journal of Approximate Reasoning, 6, 445-480,1992.
8. Sheldon Ross., A First Course in Probability, Pearson Prentice Hall, USA(2006).
9. Apostol. T.M., A short history of probability, 2nd ed. John Willey and Sons, 1969.
10. Yeates, M.P., Tolkamp, B.J., Allcroft, D.J and Kyriazakis, I, The use of mixed distribution models to determine bout criteria for analysis of animal behaviour, J. Theor. Biol. 213, 413-425, 2001.
11. Lawless J.F., Statistical Models and Methods for Life-Time Data, Wiley,New York, 1982.
12. Johnson N.L. and Kotz S., Continuous Univariate Distributions, Houghton Min Company, Boston, 1 and 2, 1970.
13. Uspensky J.V., Introduction to Mathematical Probability, McGraw-Hill, New York,1937.
14. Fan P., Stochastic Process: Theory and Applications,Press of I Singhua University, pp 52, 2006.
15. McCullah P. and Nelder J.A., Generalized Linear Models, 2nd Edition, Chapman and Hall, 1989.
16. Chakraborty S. and Bhati D., Transmuted Geometric Distribution with Applications in Modeling and Regression Analysis of Count Data, under Review, 2015.
17. Chakraborty S., Generating discrete analogues of continuous probability distributions-A survey of methods and constructions, Journal of Statistical Distributions and Applications (2015).
18. Gupta D.R., Cauchy Equation on discrete domain and some characterizations, Theory. Probability Appl. 38, 318-328, 1993.
19. Gomez-Deniz E., Another Generalization of Geometric Distribution, Test 19, 399-415, 2010.

20. Inusah S. and Kozubowski T.J., A Discrete Analogue of the Laplace Distribution, J. Stat. Planning. Inference. 136, 1090-1102, 2006.
21. Lai C.D., Constructions and Applications of Life Time Distributions, Appl. Stochastic. Models. Bus. Ind., 29, 127-140, 2012.
22. Jamjoom A.A., Order Statistics from Discrete Gamma Distribution, J.A. Sci. 9(7), 487-498, 2013.
23. Gomez-Deniz E.,Vazquez P. and Garcia G.V., A Discrete Version of the Half Normal Distribution and its Generalization with Applications, Stat Papers 55(2), 497-511, 2014.