World Applied Sciences Journal 11 (11): 1446-1449, 2010
ISSN 1818-4952

Characterization of Subfamily in Gaussian Hypergeometric Distributions

[1]*Masood Anwar and* [2]*Munir Ahmad*

[1]Department of Mathematics,
COMSATS Institute of Information Technology, Park Road, Chak Shahzad Islamabad, Pakistan
[2]National College of Business Administration and Economics, Gulberg III, Lahore, Pakistan

Abstract: A characterization theorem based on first moment is given for the subfamily of distributions generated by the Gaussian hypergeometric function. The theorem is then applied to some discrete probability distributions, providing specific characterization theorems for each of them.

Key words: Characterization . discrete distributions . gaussian hypergeometric function

INTRODUCTION

Characterization theorems in statistics are of great interest and widely appreciated for their role in clarifying the structure of the families of probability distributions. They form an essential tool of statistical inference and their role have as a natural, logical and effective starting point for constructing goodness-of-fit tests. The most of papers since 1960 on characterizations of statistical distributions deal with exponential and geometric distributions because the associated mathematics is often simple. The survey by Kotz [1] covers a substantial number of results in the field. Sampson [2] considers certain results on characterizing exponential family distributions through their mean and moment generating functions. He [2] also gives necessary and sufficient conditions so that a function can be the mean value function of an exponential family distribution. Gokhale [3] gives the mean-variance result by proving that within power series distributions, μ_2 = m (1-mc) if and only if X has a binomial, Poisson, or negative binomial distribution according to whether c is a positive integer, zero, or negative integer, respectively. Consul [4] shows that discrete probability distributions, with mass functions defined as functions of their mean can be characterized by their variance only. Consul [4] also considers characterization theorems when variance is a linear, quadratic and cubic function of mean. Letac [5] also characterizes several distributions by considering variance as a function of mean. Consul [6] proves a general characterization theorem based on conditional expectation for the exponential class of distributions. Discrete distributions whose probability generating function is given by the Gaussian hypergeometric function,

$$ {}_2F_1(\alpha,\beta;\gamma;\lambda z)=\sum_{r=0}^{\infty}\frac{(\alpha)_r(\beta)_r}{(\gamma)_r}\frac{(\lambda z)^r}{r!} \tag{1}$$

where $\alpha, \beta, \gamma, \lambda \in \mathbb{R}$ and

$$(a)_r = a(a+1)(a+2)\cdots(a+r-1), \quad r=1,2,\ldots,$$

$$(a)_0 = 1$$

have been widely studied in the field of Mathematical Statistics [7]. Most well-known discrete distributions, such as the Binomial, Negative binomial, Geometric, Hypergeometric, Poisson, Beta-Binomial, Crow-Bardwell or hyper-Poisson, Waring, Yule, etc (a detailed table appears in Dacey [8]), belong to this family. Rodríguez-Avi *et al*. [9] consider the problem of the estimation of the parameters in the context of the subfamily of distributions generated under the assumption that $\lambda = 1$, when the Gauss summation theorem provides a result for computing the exact value of probabilities and moments.

In this article a characterization theorem based on the first moment is presented for the subfamily of distributions generated by Gaussian hypergeometric function when $\lambda = 1$. The theorem is then applied to some discrete probability distributions, providing specific characterization theorems for each of them.

GAUSSIAN HYPERGEOMETRIC DISTRIBUTIONS

The discrete version of the Pearson system is given by

Corresponding Author: Dr. Masood Anwar, Department of Mathematics, COMSATS Institute of Information Technology, 30, Sector H-8/1, Islamabad, Pakistan, Cell: +92-321-5316007, E-mail: masoodanwar@comsats.edu.pk

$$G(r)f_{r+1} - L(r)f_r = 0, \ r = 0,1,2,...., \quad (2)$$

where f_r = Pr [X = r] is the probability mass function and L and G are given functions.
If we consider quadratic polynomials L and G given by

$$L(r) = (\alpha + r)(\beta + r)\lambda$$

$$G(r) = (\gamma + r)(r + 1)$$

in which α, β, γ and λ are real numbers that provide a probability mass function (*pmf*), then the solution of Eq. (2) is:

$$f_r = f_0 \frac{(\alpha)_r (\beta)_r}{(\gamma)_r} \frac{\lambda^r}{r!}, \ r = 0,1,2,.... \quad (3)$$

where $f_0^{-1} = {}_2F_1(\alpha, \beta; \gamma; \lambda)$ and the probability generating function (*pgf*) is

$$g(z) = \frac{{}_2F_1(\alpha,\beta;\gamma;\lambda z)}{{}_2F_1(\alpha,\beta;\gamma;\lambda)}.$$

Distributions generated in this manner are named *Gaussian hypergeometric distributions* (*GHD*), [7,10,11] and are denoted by Kemp and Kemp [12] as GHD (α, β, γ, λ). In order to obtain the values of Eq. (3), summation results of Eq. (1) are necessary. In this sense the only known general result is the Gauss's summation theorem, when λ = 1.

$${}_2F_1(\alpha,\beta;\gamma;1) = \frac{\Gamma(\gamma-\alpha-\beta)\Gamma(\gamma)}{\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)}, \ \text{when } \gamma>\alpha+\beta$$

In this article we shall characterize the triparametric family of distributions, GHD (α, β, γ, 1) through first moment. This family includes some well-known distributions, such as the Hypergeometric, Negative or inverse hypergeometric, Beta-Pascal, Pólya, Beta-Binomial, Waring, Generalized Waring, Yule, etc. The general case 0<λ≤1 is under consideration.

THE CHARACTERIZATION THEOREM

Theorem 1: A non-negative discrete random variable *X* defined over a given domain, belongs to the GHD (α, β, γ, 1) with probability mass function (*pmf*),

$$f_r = f_0 \frac{(\alpha)_r (\beta)_r}{(\gamma)_r \, r!}, \ r = 0,1,2,.... \quad (4)$$

if and only if

$$\mu_1' = \frac{\alpha\beta}{\gamma - (\alpha + \beta + 1)}, \text{for } \gamma > (\alpha + \beta + 1) \quad (5)$$

where $f_0^{-1} = {}_2F_1(\alpha, \beta; \gamma; 1)$, μ_1' denotes first moment about zero, α, β, γ are real numbers that provide a *pmf*.

Proof: Suppose (5) holds, then we have

$$(\gamma-1)\mu_1' = \alpha\beta + (\alpha+\beta)\mu_1',$$

$$(\gamma-1)\sum_{r=0}^{\infty} rf_r + \sum_{r=0}^{\infty} r^2 f_r = \alpha\beta \sum_{r=0}^{\infty} f_r + (\alpha+\beta)\sum_{r=0}^{\infty} rf_r + \sum_{r=0}^{\infty} r^2 f_r,$$

$$(\gamma-1)\sum_{r=1}^{\infty} rf_r + \sum_{r=1}^{\infty} r^2 f_r = \alpha\beta \sum_{r=0}^{\infty} f_r + (\alpha+\beta)\sum_{r=0}^{\infty} rf_r + \sum_{r=0}^{\infty} r^2 f_r,$$

$$(\gamma-1)\sum_{r=0}^{\infty} (r+1)f_{r+1} + \sum_{r=0}^{\infty} (r+1)^2 f_{r+1} = \sum_{r=0}^{\infty}\left[\alpha\beta + (\alpha+\beta)r + r^2\right] f_r,$$

$$\sum_{r=0}^{\infty} (\gamma+r)(r+1)f_{r+1} = \sum_{r=0}^{\infty} (\alpha+r)(\beta+r)f_r.$$

Hence, we can get the following equivalent set of equations (for similar results see also [13, 14, 15]):

$$(\gamma + r)(r+1)f_{r+1} = (\alpha + r)(\beta + r)f_r, \ r = 0,1,2,...., \quad (6)$$

Eq. (6) is a discrete version of the Pearson system (2) and its solution gives f_r.
The converse is straightforward.

Special Cases 1: Putting α = -n, β = -N_p, γ = N-N_p-n+1 in Theorem 1; we have

Corollary 3.1: A nonnegative discrete random variable *X* has *hypergeometric* distribution with *pmf*

$$f_r = \binom{N_p}{r}\binom{N - N_p}{n - r} \Big/ \binom{N}{n}$$

where

$$n, N \in \mathbb{Z}^+, 0 < p < 1,$$

and $\max(0, n - N + N_p) \le r \le \min(n, N_p)$
if and only if

$$\mu_1' = \frac{nN_p}{N} \qquad (3.1)$$

2. Putting α = -n, β = v+1, γ = -w-n, in Theorem 1; we have

Corollary 3.2: A nonnegative discrete random variable *X* has *negative or inverse hypergeometric* distribution with *pmf*

$$f_r = \binom{-v-1}{r}\binom{-w-1}{n-r}\bigg/\binom{-v-w-2}{n}$$

if and only if

$$\mu_1' = \frac{n(v+1)}{(w+v+2)}, \text{for } w+v+2<0 \qquad (3.2)$$

3. Putting α = -n, β = w/c, γ = -b/c-n+1, in Theorem 1; we have

Corollary 3.3: A nonnegative discrete random variable *X* has *Pólya* distribution with *pmf*

$$f_r = \binom{-w/c}{r}\binom{-b/c}{n-r}\bigg/\binom{-(w+b)/c}{n}, w>0, b>0, n \in \mathbb{Z}^+$$

if and only if

$$\mu_1' = \frac{nw}{(w+b)}, \text{for } w+b<0 \qquad (3.3)$$

4. Putting α = -n, β = a, γ = -b-n+1, in Theorem 1; we have

Corollary 3.4: A nonnegative discrete random variable *X* has *beta-binomial* distribution with *pmf*

$$f_r = \binom{-a}{r}\binom{-b}{n-r}\bigg/\binom{-a-b}{n}, a \ge 0, b \ge 0, n \in \mathbb{Z}^+$$

if and only if

$$\mu_1' = \frac{na}{(a+b)}, \text{ for } a+b<0 \qquad (3.4)$$

when a = b = 1, the out-come is a discrete rectangular distribution.

5. Putting α = a, β = k, γ = k+ρ+a, in Theorem 1; we have

Corollary 3.5: A nonnegative discrete random variable *X* has Generalized Waring (*beta-Pascal distribution*) distribution with *pmf*

$$f_r = \binom{-a}{r}\binom{\rho+a-1}{-k-r}\bigg/\binom{\rho-1}{-k}, a>0, k>0, \rho>0$$

if and only if

$$\mu_1' = \frac{ak}{(\rho-1)}, \text{ for } \rho>1 \qquad (3.5)$$

6. Putting a = 1 in corollary 3.5; corollary 3.6 results.

Corollary 3.6: A nonnegative discrete random variable *X* has *Waring* distribution with *pmf*

$$f_r = \binom{-1}{r}\binom{\rho}{-k-r}\bigg/\binom{\rho-1}{-k}, k>0, \rho>0$$

if and only if

$$\mu_1' = \frac{k}{(\rho-1)}, \text{ for } \rho>1 \qquad (3.6)$$

7. Putting a = k = 1, in corollary 3.5; corollary 3.7 results.

Corollary 3.7: A nonnegative discrete random variable *X* has *Yule* distribution with *pmf*

$$f_r = \binom{-1}{r}\binom{\rho}{-1-r}\bigg/\binom{\rho-1}{-1}, \ \rho>0,$$

if and only if

$$\mu_1' = \frac{1}{(\rho-1)}, \text{ for } \rho>1. \qquad (3.7)$$

8. Putting α = a-1, β = 1, γ = a+b, in Theorem 1; we have

Corollary 3.8: A discrete random variable *X* has *geometric compound* distribution with *pmf*

$$f_r = \frac{\Gamma(a+b)\ \Gamma(a+r-1)\ \Gamma(b+1)}{\Gamma a\ \Gamma b\ \Gamma(a+b+r)}, a>0, b>0, r=1,2,....,$$

if and only if

$$\mu_1' = \frac{(a-1)}{(b-1)}, \text{ for } b>1. \qquad (3.8)$$

9. Putting α = n-λ+1, β = 1, γ = λ+1, in Theorem 1; we have

Corollary 3.9: A nonnegative discrete random variable *X* has *Factorial* distribution with *pmf*

$$f_r = \frac{\lambda\lambda!}{(2\lambda - n - 1)(n - \lambda)!}\frac{(n - \lambda + r)!}{(\lambda + r)!}, r = 0, 1, 2, \ldots,$$

if and only if

$$\mu_1' = \frac{(n - \lambda + 1)}{(2\lambda - n - 2)}, \text{ for } \lambda > (n/2) + 1 \qquad (3.9)$$

REFERENCES

1. Kotz, S., 1974. Characterizations of Statistical Distributions: A Supplement to Recent Surveys. International Statistical Review, 42: 39-65.
2. Sampson, A.R., 1975. Characterization of Exponential Family Distributions by Moment Generating Functions. Annals of Statistics, 3 (3): 747-753.
3. Gokhale, D.V., 1980. On Power Series Distributions with Mean-variance Relationships. Journal of the Indian Statistical Association, 18: 81-84.
4. Consul, P.C., 1990. New Class of Location-Parameter Discrete Probability Distributions and their Characterizations. Communications in Statistics-Theory and Methods, 19: 4653-4666.
5. Letac, G. and M. Mora, 1990. Natural Real Exponential Families with Cubic Variance Functions. Annals of Statistics, 18 (1): 1-37.
6. Consul, P.C., 1995. Some Characterizations for the Exponential Class of Distributions. IEEE Transactions on Reliability, 44 (3): 403-407.
7. Johnson, N.L., S. Kotz and A.W. Kemp, 2005. Univariate Discrete Didtributions 3rd Edn. John Wiley and Sons, New York.
8. Dacey, M.F., 1972. A Family of Discrete Probability Distributions Defined by the Generalized Hypergeometric Series. Sankhya, B (34): 243-250.
9. Rodríguez-Avi, J., A. Conde-Sánchez, A.J. Sáez Castillo and M.J. Olmo-Jiménez, 2003. Estimation of Parameters in Gaussian Hypergeometric Distributions. Communications in Statistics-Theory and Methods, 32 (6): 1101-1119.
10. Kemp, A.W., 1968. A Wide Class of Discrete Distributions and the Associated Differential Equations. Sankhya, A (30): 401-410.
11. Kemp, A.W. and C.D. Kemp, 1975. Models for Gaussian Hypergeometric Distributions. In: Patil, G.P., Ed. A Modern Course on Statistical Distributions in Scientific Work, Dordrecht: Reidel, pp: 31-40.
12. Kemp, A.W. and C.D. Kemp, 1971. On Mixing Processes and Lost-games Distribution. Zastosowania Matematyki, 12: 167-173.
13. Osaki, S. and X. Li, 1988. Characterization of Gamma and Negative Binomial Distributions. IEEE Transactions on Reliability, 37 (4): 379-382.
14. Ahmed, A.N., 1991. Characterizations of the Beta, Binomial and Poisson Distributions. IEEE Transactions on Reliability, 40: 403-407.
15. Kemp, A.W. and C.D. Kemp, 2004. Factorial Moment Characterizations for Certain Binomial-type Distributions. Communications in Statistics-Theory and Methods, 33 (12): 3059-3068.