Brazilian Journal of Probability and Statistics
2009, Vol. 23, No. 1, 49–56
DOI: 10.1214/08-BJPS003

A new class of discrete distributions

C. Satheesh Kumar

University of Kerala

Abstract. A new class of discrete distributions is introduced by extending the generalized hypergeometric recast distribution; some of its properties are studied. It is shown that all its moments exist finitely. A genesis, probability mass function, mean and variance are obtained. Certain recurrence relations for probabilities, moments and factorial moments are derived. Certain mixtures and limiting cases are also considered.

1 Introduction

Discrete distributions whose probability generating function (PGF) is given in terms of generalized hypergeometric function have been studied extensively in the field of Mathematical Statistics. For details, see Johnson, Kemp and Kotz (2005). Johnson, Kemp and Kotz (2005), page 90, commented that many common discrete probability distributions have PGF's of the form

$$A(t) = \frac{{}_pF_q(\underline{a}; \underline{b}; \theta t + \xi)}{{}_pF_q(\underline{a}; \underline{b}; \theta + \xi)}, \tag{1.1}$$

where ${}_pF_q(\underline{a}; \underline{b}; \theta)$ is the generalized hypergeometric series, in which a's, b's, θ and ξ are appropriate reals such that the general term in the series is non-negative. For a detailed account on generalized hypergeometric series and conditions on their convergence, see Slater (1966) or Mathai and Saxena (1973). The family of discrete distributions with PGF (1.1) is known as generalized hypergeometric recast distributions (GHRD). Two broad classes of discrete distributions, namely generalized hypergeometric probability distributions (GHPD) of Kemp (1968) and generalized hypergeometric factorial moment distributions (GHFMD) of Kemp and Kemp (1969), are special cases of GHRD. An extended version of GHPD is considered in Kumar (2002) through the following PGF and call it as the extended generalized hypergeometric probability distributions (EGHPD):

$$P(t) = \frac{{}_pF_q(\underline{a}; \underline{b}; \theta_1 t + \theta_2 t^m)}{{}_pF_q(\underline{a}; \underline{b}; \theta_1 + \theta_2)}. \tag{1.2}$$

The EGHPD contains several popular families of discrete distributions such as the Hermite distribution of Kemp and Kemp (1965), the generalized Hermite

Key words and phrases. Generalized Gegenbauer distribution, generalized Hermite distribution, generalized hypergeometric function, generalized hypergeometric recast distribution, mixtures, probability generating function.

Received September 2007; accepted February 2008.

distribution of Gupta and Jain (1974), the Gegenbaur distribution of Plunkett and Jain (1975) and the generalized Gegenbaur distribution of Medhi and Borah (1984). Kemp and Kemp (1974) has shown that GHFMD includes not only some of the most common discrete distributions such as binomial, Poisson, negative binomial, hypergeometric, Poisson-beta etc. but also a number of important matching and occupancy distributions as its special cases. Kumar (2008) introduced and studied a bivariate version of GHFMD.

Through the present paper, we introduce a new class of discrete distributions by extending GHRD which we call the extended generalized hypergeometric recast distributions (EGHRD) and study some of its properties. In Section 2 we present a genesis of the EGHRD and derive its probability mass function, mean and variance. There we also show that all the moments of EGHRD exist finitely. Certain recurrence formula for probabilities, raw moments and factorial moments of the EGHRD are obtained in Section 3. In Section 4 it is shown that certain mixtures of EGHRD are again EGHRD and in Section 5 EGHRD is obtained as a limit of another EGHRD.

The genesis presented in Section 2 shows that EGHRD possess a random sum structure. As such, EGHRD is applicable wherever such a structure arises. Random sum structure arises in several areas of research. Chapter 9 of Johnson, Kemp and Kotz (2005) is fully devoted to such random sum distributions.

2 A genesis and special cases of EGHRD

Consider the sequence $\{Z_n, n \geq 1\}$ of independent and identically distributed random variables, where Z_1 is assumed to have the PGF

$$H(t) = \alpha t + (1-\alpha)t^m, \tag{2.1}$$

where $0 < \alpha \leq 1$ and m is a positive integer. Let N be a non-negative integer valued random variable with PGF (1.1), in which θ takes values in Θ, an open interval of one of the sets $(0,\infty)$, $(-\infty, 0)$. Consider reals θ_1, θ_2 which are such that $\theta_1 > 0$, $\theta_2 \geq 0$ and $\theta = \theta_1 + \theta_2$ is an element of Θ if $\Theta \subset (0,\infty)$ and are such that $\theta_1 < 0$, $\theta_2 \leq 0$ and θ is an element of Θ if $\Theta \subset (-\infty, 0)$. Set $\alpha = \frac{\theta_1}{\theta}$. Assume that $\{Z_n, n \geq 1\}$ and N are statistically independent. Define $V_0 = Z_0 = 0$. Consider V_N, the random sum of random variables defined by $V_N = \sum_{k=0}^{N} Z_k$. Then the PGF of V_N is the following, in which T_0 is as defined in (2.5):

$$\begin{aligned} G(t) &= A(H(t)) \\ &= T_0^{-1}\,{}_pF_q(\underline{a}; \underline{b}; \theta_1 t + \theta_2 t^m + \xi). \end{aligned} \tag{2.2}$$

A distribution with PGF (2.2) we call the extended generalized hypergeometric recast distribution or in short EGHRD. Obviously EGHRD with $\theta_2 = 0$ is the GHRD. Further, EGHRD with $\xi = 0$ is the EGHPD of Kumar (2002) with PGF

(1.2) and EGHRD with $\xi = -(\theta_1 + \theta_2)$ is an extended version of GHFMD of Kemp and Kemp (1969), with PGF

$$G_1(t) = {}_pF_q[\underline{a}; \underline{b}; \theta_1(t-1) + \theta_2(t^m - 1)].$$

Thus EGHRD contains extended versions of several well-known discrete distributions such as binomial, Poisson, displaced Poisson, hyper-Poisson, negative binomial, hypergeometric, inverse hypergeometric, negative hypergeometric, Poisson-beta, Waring, Yule, Naor, Feller, etc.

Lemma 1. *Let N be a non-negative integer valued random variable having GHRD with PGF given in* (1.1), *then $E(N^r) < \infty$ for every positive integer r.*

The proof is simple and hence omitted.

Result 1. *Let Y be a random variable having EGHRD with PGF* (2.2). *Then $\mu_r = E(Y^r) < \infty$ for every positive integer r.*

Proof. Since Y and V_N of the genesis given above are identically distributed,

$$0 \le \mu_r = E(V_N^r) \le m^r E(N^r) < \infty,$$

by Lemma 1. □

For $n = 0, 1, 2, \ldots$ let $Q_n(\underline{a}; \underline{b})$ denote the probability mass function (p.m.f.) of EGHRD. Then we have the following from (2.2):

$$\begin{aligned} G(t) &= \sum_{n=0}^{\infty} Q_n(\underline{a}; \underline{b}) t^n \\ &= T_0^{-1} {}_pF_q(\underline{a}; \underline{b}; \theta_1 t + \theta_2 t^m + \xi). \end{aligned} \tag{2.3}$$

On expanding (2.3) and equating the coefficients of t^n on both sides, we have the following result.

Result 2. *Let Y be a random variable distributed as EGHRD with PGF* (2.2). *Then the following is the PMF of EGHRD for $n = 0, 1, 2, \ldots,$ in which $\delta = n - (m-1)j$:*

$$Q_n(\underline{a}; \underline{b}) = T_0^{-1} \sum_{j=0}^{[n/m]} \frac{\prod_{r=1}^{p} (a_r)_\delta}{\prod_{s=1}^{q} (b_s)_\delta} {}_pF_q(\underline{a} + \underline{\delta}_p; \underline{b} + \underline{\delta}_q; \xi) \frac{\theta_1^{n-mj} \theta_2^j}{(n-mj)!\, j!}, \tag{2.4}$$

where $[v]$ denote the integer part of v,

$$(a)_0 = 1; \qquad (a)_k = a(a+1)(a+k-1), \qquad \textit{for } k \ge 1,$$

$$\underline{a} + \underline{\delta}_p = a_1 + \delta, a_2 + \delta, \ldots, a_p + \delta$$

and

$$\underline{b} + \underline{\delta}_q = b_1 + \delta, b_2 + \delta, \ldots, b_q + \delta.$$

Result 3. *Let Y be a random variable distributed as EGHRD with PGF (2.2). Then the mean and variance of EGHRD are*

$$E(Y) = D_0 T_0^{-1} T_1(\theta_1 + m\theta_2),$$
$$\mathrm{Var}(Y) = D_0 T_0^{-2}(D_1 T_0 T_2 - D_0 T_1^2)(\theta_1 + m\theta_2)^2 + D_0 T_0^{-1} T_1(\theta_1 + m^2\theta_2),$$

where for $i = 0, 1, 2, \ldots,$

$$D_i = \frac{\prod_{r=1}^{p}(a_r + i)}{\prod_{s=1}^{q}(b_s + i)}$$

and

$$T_i = {}_pF_q(\underline{a} + \underline{i}_p; \underline{b} + \underline{i}_q; \theta_1 + \theta_2 + \xi). \tag{2.5}$$

The proof is straightforward and hence omitted.

3 Recurrence relations for probabilities, raw moments and factorial moments

Result 4. *The following is a simple recurrence relation for probabilities $Q_n(\underline{a}; \underline{b})$ of EGHRD for $n \geq 0$, in which $Q_{-r}(\underline{a}; \underline{b}) = 0$, for $r > 0$.*

$$\begin{aligned}(n+1)Q_{n+1}(\underline{a}; \underline{b}) = D_0 T_0^{-1} T_1\{&\theta_1 Q_n(\underline{a} + \underline{1}_p; \underline{b} + \underline{1}_q) \\ &+ m\theta_2 Q_{n-m+1}(\underline{a} + \underline{1}_p; \underline{b} + \underline{1}_q)\},\end{aligned} \tag{3.1}$$

where $\underline{a} + \underline{1}_p$ and $\underline{b} + \underline{1}_q$ are as given in Result 2 with $\delta = 1$.

Proof. On differentiating (2.3) with respect to t, we get

$$\begin{aligned}&\sum_{n=0}^{\infty}(n+1)Q_{n+1}(\underline{a}; \underline{b})t^n \\ &\quad = D_0 T_0^{-1}(\theta_1 + m\theta_2 t^{m-1})\, {}_pF_q(\underline{a} + \underline{1}_p; \underline{b} + \underline{1}_q; \theta_1 t + \theta_2 t^m + \xi).\end{aligned} \tag{3.2}$$

Replace $\underline{a}$, $\underline{b}$ by $\underline{a} + \underline{1}_p$, $\underline{b} + \underline{1}_q$ in (2.3) to obtain the following:

$${}_pF_q(\underline{a} + \underline{1}_p; \underline{b} + \underline{1}_q; \theta_1 t + \theta_2 t^m + \xi) = T_1 \sum_{n=0}^{\infty} Q_n(\underline{a} + \underline{1}_p; \underline{b} + \underline{1}_q)t^n. \tag{3.3}$$

By using (3.3) in (3.2) we have

$$\sum_{n=0}^{\infty}(n+1)Q_{n+1}(\underline{a}; \underline{b})t^n = D_0 T_0^{-1} T_1(\theta_1 + m\theta_2 t^{m-1}) \sum_{n=0}^{\infty} Q_n(\underline{a} + \underline{1}_p; \underline{b} + \underline{1}_q)t^n.$$

On equating coefficients of t^n on both sides we get the relation (3.1). □

Result 5. *The following is a recurrence relation for raw moments of EGHRD for* $r \geq 0$:

$$\mu_{r+1}(\underline{a}; \underline{b}) = D_0 T_0^{-1} T_1 \sum_{j=0}^{r} \binom{r}{j} (\theta_1 + m^{j+1}\theta_2)\mu_{r-j}(\underline{a} + \underline{1}_p; \underline{b} + \underline{1}_q). \tag{3.4}$$

Proof. In Section 2 we proved that the rth raw moment μ_r of EGHRD exists finitely. Therefore the characteristic function $\phi_Y(t)$ of EGHRD with PGF (2.2) has the following series representation, in which $\mu_r(\underline{a}; \underline{b}) = \mu_r$. For t in R,

$$\begin{aligned}\phi_Y(t) &= G(e^{it}) \\ &= T_0^{-1} {}_pF_q(\underline{a}; \underline{b}; \theta_1 e^{it} + \theta_2 e^{mit} + \xi) \\ &= \sum_{r=0}^{\infty} \mu_r(\underline{a}; \underline{b}) \frac{(it)^r}{r!}.\end{aligned} \tag{3.5}$$

Now the proof of relation (3.4) follows on differentiating both sides of the above equation with respect to t and equating the coefficients of $\frac{(it)^r}{r!}$ on both sides, in the light of the arguments similar to those in the proof of Result 4. □

Result 6. *The following is a simple recurrence relation for factorial moments* $\mu_{[n]}(\underline{a}; \underline{b})$ *of EGHRD for* $n \geq m - 1$, *in which* $\mu_{[0]}(\underline{a}; \underline{b}) = 1$:

$$\begin{aligned}&\mu_{[n+1]}(\underline{a}; \underline{b}) \\ &= D_0 T_0^{-1} T_1 \Bigg[\theta_1 \mu_{[n]}(\underline{a} + \underline{1}_p; \underline{b} + \underline{1}_q) \\ &\qquad + m\theta_2 \sum_{j=0}^{m-1} \binom{m-1}{j} \frac{n!}{(n-j)!} \mu_{[n-j]}(\underline{a} + \underline{1}_p; \underline{b} + \underline{1}_q)\Bigg].\end{aligned} \tag{3.6}$$

Proof. The factorial moment generating function $F_Y(t)$ of EGHRD with PGF (2.2) is

$$\begin{aligned}F_Y(t) &= G(1+t) \\ &= T_0^{-1} {}_pF_q[\underline{a}; \underline{b}; \theta_1(1+t) + \theta_2(1+t)^m + \xi] \\ &= \sum_{n=0}^{\infty} \mu_{[n]}(\underline{a}; \underline{b}) \frac{t^n}{n!}.\end{aligned} \tag{3.7}$$

The relation (3.6) follows on differentiating both sides of (3.7) with respect to t and equating the coefficients of $\frac{t^n}{n!}$ on both sides, in the light of the arguments similar to those in the proof of Result 4. □

4 Mixtures of EGHRD

In Theorem 1 given below it is shown that certain mixtures of EGHRD are also EGHRD.

Theorem 1. *For $j = 1, 2$ let Y_j be a non-negative integer valued random variable and let λ_1 and λ_2 be independent random variables having the following probability density functions, in which $c > 0$, $d > 0$ and $\beta > 0$.*

$$f_1(\lambda_1) = \frac{\lambda_1^{c-1}(1-\lambda_1)^{d-1}}{B(c,d)} \frac{{}_pF_q[\underline{a};\underline{b};(\theta_1+\theta_2+\xi)\lambda_1]}{{}_{p+1}F_{q+1}(\underline{a},c;\underline{b},c+d;\theta_1+\theta_2+\xi)}, \quad 0 < \lambda_1 < 1, \tag{4.1}$$

$$f_2(\lambda_2) = \frac{\lambda_2^{\beta-1}e^{-\lambda_2}}{\Gamma(\beta)} \frac{{}_pF_q[\underline{a};\underline{b};(\theta_1+\theta_2+\xi)\lambda_2]}{{}_{p+1}F_q(\underline{a},\beta;\underline{b};\theta_1+\theta_2+\xi)}, \qquad \lambda_2 > 0. \tag{4.2}$$

Assume that the conditional distribution of Y_1 given λ_1 and Y_2 given λ_2 are EGHRD and let their PGF's be

$$\frac{{}_pF_q[\underline{a};\underline{b};\lambda_1(\theta_1 t+\theta_2 t^m+\xi)]}{{}_pF_q[\underline{a};\underline{b};\lambda_1(\theta_1+\theta_2+\xi)]}$$

and

$$\frac{{}_pF_q[\underline{a};\underline{b};\lambda_2(\theta_1 t+\theta_2 t^m+\xi)]}{{}_pF_q[\underline{a};\underline{b};\lambda_2(\theta_1+\theta_2+\xi)]}$$

(*where $p \le q$*) *respectively. Then Y_j has EGHRD for $j = 1, 2$.*

Proof. Let $G_j^*(t)$ denote the PGF of Y_j for $j = 1, 2$. Then from the well-known properties of integrals of a power series, we obtain the following:

$$\begin{aligned} G_1^*(t) &= E(t^{Y_1}) \\ &= E_{\lambda_1}\{E(t^{Y_1}|\lambda_1)\} \\ &= \int_0^1 \frac{{}_pF_q[\underline{a};\underline{b};\lambda_1(\theta_1 t+\theta_2 t^m+\xi)]}{{}_pF_q[\underline{a};\underline{b};\lambda_1(\theta_1+\theta_2+\xi)]} f_1(\lambda_1)\,d\lambda_1 \\ &= \frac{{}_{p+1}F_{q+1}(\underline{a},c;\underline{b},c+d;\theta_1 t+\theta_2 t^m+\xi)}{{}_{p+1}F_{q+1}(\underline{a},c;\underline{b},c+d;\theta_1+\theta_2+\xi)}, \end{aligned} \tag{4.3}$$

$$\begin{aligned} G_2^*(t) &= E(t^{Y_2}) \\ &= \int_0^\infty \frac{{}_pF_q[\underline{a};\underline{b};\lambda_2(\theta_1 t+\theta_2 t^m+\xi)]}{{}_pF_q[\underline{a};\underline{b};\lambda_2(\theta_1+\theta_2+\xi)]} f_2(\lambda_2)\,d\lambda_2 \\ &= \frac{{}_{p+1}F_q(\underline{a},\beta;\underline{b};\theta_1 t+\theta_2 t^m+\xi)}{{}_{p+1}F_q(\underline{a},\beta;\underline{b};\theta_1+\theta_2+\xi)}. \end{aligned} \tag{4.4}$$

□

As a special case of Theorem 1 with $\xi = -(\theta_1 + \theta_2)$, we establish that the gamma mixture of Hermite distribution is the Gegenbour distribution and the gamma mixture of generalized Hermite distribution is the generalized Gegenbour distribution.

5 Limiting distributions

The following is a standard result for generalized hypergeometric functions.

Lemma 2. *The generalized hypergeometric function ${}_pF_q(\underline{a}; \underline{b}; \theta)$ with $p \leq q$ can be obtained as the limiting form of the following function as $v \to \infty$.*

$$F = {}_{p+1}F_q(\underline{a}, v; \underline{b}; v^{-1}\theta), \qquad v > |\theta|.$$

The proof easily follows by observing that

$$\lim_{v \to \infty} \frac{(v)_r}{v^r} = 1.$$

A direct application of Lemma 2 leads to the following result regarding EGHRD.

Result 7. *The EGHRD with PGF (2.2) is the limiting form as $d \to \infty$ of the distribution with the following PGF, in which $d > 0$:*

$$P(t) = \frac{{}_{p+1}F_q[\underline{a}, d; \underline{b}; d^{-1}(\theta_1 t + \theta_2 t^m + \xi)]}{{}_{p+1}F_q[\underline{a}, d; \underline{b}; d^{-1}(\theta_1 + \theta_2 + \xi)]}.$$

As a special case of Result 7 with $\xi = 0$ or $\xi = -(\theta_1 + \theta_2)$, we obtain that Hermite distribution is the limit of the Gegenbour distribution and that the generalized Hermite distribution is the limit of the generalized Gegenbour distribution.

Acknowledgments

The author is highly grateful to the Editor-in-Chief and both the anonymous referees for their valuable comments that greatly improved the presentation of the paper.

References

Gupta, R. P. and Jain, G. C. (1974). A generalized Hermite distribution and its properties. *SIAM Journal of Applied Mathematics* **27** 359–363. MR0350087

Johnson, N. L., Kemp, A. W. and Kotz, S. (2005). *Univariate Discrete Distributions*. Wiley, New York. MR2163227

Kemp, A. W. (1968). A wide class of discrete distributions and the associated differential equations. *Sankhya, Series A* **30** 401–410. Zbl 0186.53004.

Kemp, A. W. and Kemp, C. D. (1974). A family of discrete distributions defined via their factorial moments. *Communications in Statistics: Theory Methods* **3** 1187–1196. MR0362620

Kemp, C. D. and Kemp, A. W. (1965). Some properties of the Hermite distribution. *Biometrika* **52** 381–394. MR0208717

Kemp, C. D. and Kemp, A. W. (1969). Some distributions arising from an inventory decision problem. *Bulletin of the International Statistical Institute* **43** 336–338.

Kumar, C. S. (2002). Extended generalized hypergeometric probability distributions. *Statistics and Probability Letters* **59** 1–7. MR1925281

Kumar, C. S. (2008). A unified approach to bivariate discrete distributions. *Metrika* **67** 113–123. MR2353798

Mathai, A. M. and Saxena, R. K. (1973). *Generalized Hypergeometric Functions with Applications in Statistics and Physical Sciences*. *Lecture Notes* **348**. Springer, Heidelberg. MR0463524

Medhi, J. and Borah, M. (1984). On generalized Gegenbauer polynomials and associated probabilities. *Sankhya, Series B* **46** 157–165. MR0783031

Plunkett, I. G. and Jain, G. C. (1975). Three generalized negative binomial distributions. *Biometrische Zeitschrift* **17** 286–302. MR0405662

Slater, L. J. (1966). *Generalized Hypergeometric Functions*. Cambridge Univ. Press, Cambridge. MR0201688

Department of Statistics
University of Kerala
Trivandrum–695 581
India
E-mail: drcsatheeshkumar@gmail.com