A Slew of Mixture Relationships Involving Discrete and Continuous Generalized Hypergeometric Distributions and Their Special Cases

by M.C. Jones[a] & N. Balakrishnan[b]

[b]*School of Mathematics and Statistics, The Open University, Walton Hall, Milton Keynes, MK7 6AA, United Kingdom*

[a]*Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, L8S 4K1, Canada*

Abstract

Our starting point is recognition of some mixture relationships involving the (continuous) Gauss hypergeometric distribution. Our main emphasis is then to generalize these relationships to ones involving (discrete) generalized hypergeometric distributions and their rarely considered continuous counterparts. Two such sets of relationships are derived, one involving beta distributions, the other gamma distributions. A wide variety of interesting special cases arise along the way: Poisson, binomial, negative binomial, logarithmic, Conway-Maxwell-Poisson and Libby-Novick distributions all appear. There are also comments on the wider context within which the relationships of interest in this article arise.

Keywords: Beta distribution; Conway-Maxwell-Poisson distribution; Gamma distribution; Gauss hypergeometric distribution; Libby-Novick distribution.

1. Introduction and Motivating Results

Kemp (1968) introduced the discrete generalized hypergeometric distributions; see also Johnson, Kemp & Kotz (2005, Section 2.4.1). These are distributions for $J = 0, 1, ...,$ with probability mass function (p.m.f.)

$$ {}_pp_q(j; \underline{\gamma}_p, \underline{\delta}_q, \lambda) \equiv \frac{(\gamma_1)_j \cdots (\gamma_p)_j \, \lambda^j}{(\delta_1)_j \cdots (\delta_q)_j \, j! \, {}_pF_q(\underline{\gamma}_p, \underline{\delta}_q, \lambda)}. \tag{1} $$

Here, $(\cdot)_j \equiv \Gamma(\cdot + j)/\Gamma(\cdot)$ where $\Gamma(\cdot)$ is the gamma function, $\underline{\gamma}_p \equiv \{\gamma_1, ..., \gamma_p\}$, $\underline{\delta}_q \equiv \{\gamma_1, ..., \gamma_q\}$, p and q are non-negative integers such that $p \leq q + 1$, and

$$ {}_pF_q(\underline{\gamma}_p, \underline{\delta}_p, \lambda) = \sum_{k \geq 0} \frac{(\gamma_1)_k \cdots (\gamma_p)_k \, \lambda^k}{(\delta_1)_k \cdots (\delta_q)_k \, k!} $$

is the generalized hypergeometric function; there are constraints on the values of the parameters $\underline{\gamma}_p, \underline{\delta}_q$ and λ to make these functions converge, and hence for the

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/sta4.550

distributions to be well-defined, which will be returned to later. Among the many special cases of these distributions are the Poisson, binomial, negative binomial and logarithmic (or logseries) distributions. We will write $J \sim {}_p\mathrm{HP}_q(\underline{\gamma}_p, \underline{\delta}_q, \lambda)$ to denote J following such a distribution.

When $p = 2$, $q = 1$, ${}_2F_1(\gamma_1, \gamma_2, \delta_1, \lambda)$ is the Gauss hypergeometric function, defined when all the parameters are positive and also $\lambda < 1$. In a Bayesian context, Armero & Bayarri (1994) first defined the continuous Gauss hypergeometric distribution for $0 < X < 1$ to have probability density function (p.d.f.)

$$ {}_2f_1^{\beta}(x; \gamma, \alpha, \beta, \theta) \equiv \frac{x^{\alpha-1}(1-x)^{\beta-1}\{1-(1-\theta)x\}^{-\gamma}}{B(\alpha,\beta)\,{}_2F_1(\gamma,\alpha,\alpha+\beta,1-\theta)}. \tag{2} $$

Here, $B(\cdot,\cdot)$ is the beta function. This definition holds for $\alpha, \beta, \gamma > 0$ and $0 < \theta < 1$ but is extended to all $\theta > 0$ by observing that, in an obvious notation, if $X \sim \mathrm{GH}(\gamma, \alpha, \beta, \theta)$ then $1 - X \sim \mathrm{GH}(\gamma, \beta, \alpha, 1/\theta)$. Special cases of this distribution include the beta and Libby-Novick (Libby & Novick, 1982) distributions.

There are nice links between the continuous Gauss hypergeometric distribution (2) and the discrete Gauss hypergeometric distribution (the discrete generalized hypergeometric distribution (1) with $p = 2$, $q = 1$). By series expansion of $\{1-(1-\theta)x\}^{-\gamma}$ in (1) when $0 < \theta < 1$, it is straightforward to show that the continuous Gauss hypergeometric distribution can be expressed as a mixture of beta distributions with a discrete Gauss hypergeometric mixing distribution:

$$ \text{if } X|J=j \sim \mathrm{Beta}(\alpha+j,\beta) \text{ and } J \sim {}_2\mathrm{HP}_1(\gamma,\alpha,\alpha+\beta,1-\theta), \text{ then } X \sim \mathrm{GH}(\gamma,\alpha,\beta,\theta). \tag{3} $$

Relationship (3) is dual to the following relationship expressing the discrete Gauss hypergeometric distribution as a mixture of negative binomial distributions with a continuous Gauss hypergeometric mixing distribution:

$$ \begin{aligned} &\text{if } J|X=x \sim \mathrm{NegBin}(\gamma, 1-(1-\theta)x) \text{ and } X \sim \mathrm{GH}(\gamma,\alpha,\beta,\theta), \\ &\qquad \text{then } J \sim {}_2\mathrm{HP}_1(\gamma,\alpha,\alpha+\beta,1-\theta). \end{aligned} \tag{4} $$

Relationships (3) and (4) are dual in the sense that they comprise the full set of marginal and conditional distributions of a certain bivariate distribution for X and J (of which, more later). Collectively, for $0 < \theta < 1$, we have:

$$ X|J=j \sim \mathrm{Beta}(\alpha+j,\beta), \tag{5} $$

$$ X \sim \mathrm{GH}(\gamma,\alpha,\beta,\theta), \tag{6} $$

$$ J \sim {}_2\mathrm{HP}_1(\gamma,\alpha,\alpha+\beta,1-\theta), \tag{7} $$

and

$$ J|X=x \sim \mathrm{NegBin}(\gamma,1-(1-\theta)x), \tag{8} $$

(For more on this particular case, see Section 2.2.) A similar set of results is obtained for $\theta > 1$ by consideration of the distribution of $1 - X$ in place of that of X.

In addition to observing the above, the purposes of this article are:

1. to set these particular relationships in the wider context of general p and q;
2. thereby, to explore links between what will newly be called beta-generated continuous generalized hypergeometric distributions and discrete generalized hypergeometric distributions;
3. also, to consider relationships involving the gamma distribution in place of the beta distribution;
4. and to note a number of interesting special cases.

The general forms of relationship (4) are given briefly in Johnson et al. (2005, Section 8.4), on which this article seeks to elaborate considerably. Section 5 embeds our specific results within an even wider context.

2. General Beta-Based Results and More Special Cases

Consider the bivariate distribution for $J = 0, 1, ...,$ and $0 < X < 1$ with p.m.f./p.d.f.

$$\begin{aligned} f(j,x) &= \frac{x^j B(\alpha,\beta)}{B(\alpha+j,\beta)} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} \\ &\quad \times \frac{(\gamma_1)_j \cdots (\gamma_p)_j (\alpha)_j \lambda^j}{(\delta_1)_j \cdots (\delta_q)_j (\alpha+\beta)_j \, j! \; {}_{p+1}F_{q+1}(\underline{\gamma}_p, \alpha, \underline{\delta}_q, \alpha+\beta, \lambda)} \\ &= \frac{x^{\alpha+j-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} \frac{(\gamma_1)_j \cdots (\gamma_p)_j \lambda^j}{(\delta_1)_j \cdots (\delta_q)_j \, j! \; {}_{p+1}F_{q+1}(\underline{\gamma}_p, \alpha, \underline{\delta}_q, \alpha+\beta, \lambda)}. \end{aligned} \quad (9)$$

The marginal and conditional distributions of this bivariate distribution are given by $X|J = j \sim \text{Beta}(\alpha + j, \beta)$ as at (5) together with

$$X \sim {}_{p+1}\text{GH}^{\beta}_{q+1}(\underline{\gamma}_p, \underline{\delta}_q, \alpha, \beta, \lambda), \quad (10)$$

$$J \sim {}_{p+1}\text{HP}_{q+1}(\underline{\gamma}_p, \alpha, \underline{\delta}_q, \alpha+\beta, \lambda) \quad (11)$$

and

$$J|X = x \sim {}_p\text{HP}_q(\underline{\gamma}_p, \underline{\delta}_q, \lambda x). \quad (12)$$

Here, the beta-generated continuous generalized hypergeometric distribution ${}_{p+1}\text{GH}^{\beta}_{q+1}(\underline{\gamma}_p, \underline{\delta}_q, \alpha, \beta, \lambda)$ is defined to have density

$${}_{p+1}f^{\beta}_{q+1}(x; \underline{\gamma}_p, \underline{\delta}_q, \alpha, \beta, \lambda) \equiv \frac{x^{\alpha-1}(1-x)^{\beta-1} \; {}_pF_q(\underline{\gamma}_p, \underline{\delta}_q, \lambda x)}{B(\alpha,\beta) \; {}_{p+1}F_{q+1}(\underline{\gamma}_p, \alpha, \underline{\delta}_q, \alpha+\beta, \lambda)}, \qquad 0 < x < 1. \quad (13)$$

(This reduces to (2) when $p = 1$, $q = 0$, where λ has been replaced by $1 - \theta$ in (2) to follow the usual convention. Of course, (10) and (11) reduce to (6) and (7) for the same choices of p and q.)

So, the beta-generated continuous generalized hypergeometric distribution at (13) is naturally connected with the corresponding discrete generalized hypergeometric distribution in two ways:

(i) the beta-generated continuous generalized hypergeometric distribution can be expressed as a mixture of beta distributions with a discrete generalized hypergeometric mixing distribution via (5), (10) and (11);

(ii) the discrete generalized hypergeometric distribution can be expressed as a mixture of discrete generalized hypergeometric distributions with each of p and q reduced by 1 with a beta-generated continuous generalized hypergeometric mixing distribution via (10), (11) and (12) (as in Johnson et al., 2005, Section 8.4).

We need, however, to pay attention to the conditions under which both generalized hypergeometric functions featured above converge and are positive, so that the distributions above are all well-defined. Let n be a positive integer and define $(-n)_j \equiv (-1)^j j! \binom{n}{j}$. Then, using Johnson et al. (2005, Sections 1.1.8 and 2.4.1), three sets of appropriate conditions are:

(a) positive parameters when $p \leq q$;

(b) positive parameters and $0 < \lambda < 1$ when $p = q + 1$;

(c) positive parameters aside from $\gamma_1 = -n$, $\lambda \equiv -\psi < 0$ when $p \geq 1$.

In case (c), the corresponding discrete distributions have support $0, 1, ..., n$ (and there are further ways of acquiring such finite support discrete distributions that will not be utilized here).

Because the beta-generated continuous generalized hypergeometric distribution at (13) is not well known, it is worth mentioning the straightforward formula it possesses for its moments. If X follows this distribution, as at (10), then

$$E(X^r) = \frac{(\alpha)_r}{(\alpha+\beta)_r} \frac{{}_{p+1}F_{q+1}(\underline{\gamma}_p, \alpha + r, \underline{\delta}_q, \alpha + \beta + r, \lambda)}{{}_{p+1}F_{q+1}(\underline{\gamma}_p, \alpha, \underline{\delta}_q, \alpha + \beta, \lambda)}.$$

2.1 Special Case $p = q = 0$

When $p = q = 0$, $X|J = j \sim \text{Beta}(\alpha + j, \beta)$ and (10)–(12) reduce to

$$X \sim {}_1\text{GH}_1^{\beta}(\alpha, \beta, \lambda) \tag{14}$$

$$J \sim {}_1\mathrm{HP}_1(\alpha, \alpha+\beta, \lambda) \tag{15}$$

and

$$J|X = x \sim \text{Poisson}(\lambda x). \tag{16}$$

Here, ${}_1F_1$ is the confluent hypergeometric function and ${}_1\mathrm{HP}_1$ is the confluent hypergeometric distribution mentioned in an abstract of Hall (1956) and discussed in Bhattacharya (1966) (see Johnson et al., 2005, Section 4.12.4). Moreover, ${}_1\mathrm{GH}_1^{\beta}(\alpha, \beta, \lambda)$, which has density

$$\frac{x^{\alpha-1}(1-x)^{\beta-1}e^{\lambda x}}{B(\alpha,\beta)\ {}_1F_1(\alpha, \alpha+\beta, \lambda)}, \qquad 0 < x < 1,$$

is the Kummer beta distribution of Ng & Kotz (1995) and Gordy (1998). The mixture construction of the Kummer beta distribution (15) via (5) and (14) is given by Ng & Kotz (1995); the dual link between the Kummer beta distribution and the confluent hypergeometric distribution via the Poisson mixture relationship (14)–(16) is not.

When $\alpha = \beta = 1$, intriguing relationships between simple distributions ensue. In this case, the conditional distribution of X is that of $U^{1/(j+1)}$ where $U \sim$ Uniform$(0,1)$ and its marginal distribution is the distribution with p.d.f. $f_{IE}(x) \equiv \lambda e^{\lambda x}/(e^{\lambda}-1)$ on $0 < x < 1$ (a truncated increasing exponential function). Also, the marginal distribution of J is that of $P-1$ where P follows the zero-truncated Poisson distribution, which has p.m.f. $\lambda^j/\{(e^{\lambda}-1)j!\}$, $j = 1, 2, \ldots$. We therefore learn from (5), (14) and (15) a construction of the distribution with density f_{IE}: it is the distribution of $U^{1/P}$, where U and P are independent. In addition, from (14), (15) and (16), we find that mixing a Poisson distribution with the f_{IE} distribution results in a zero-truncated Poisson distribution minus one.

It is also worth noting that when $\alpha = 1$, $\beta \neq 1$, the ${}_1\mathrm{HP}_1(\alpha, \alpha+\beta, \lambda)$ distribution reduces to the hyper-Poisson distribution of Bardwell & Crow (1964) with parameter $\beta+1 > 1$; relations involving this distribution therefore result but are not spelled out here.

2.2 Special Case $p = 1$, $q = 0$, $0 < \lambda < 1$

This special case is precisely that given in (5)–(8). Here are some further aspects of it.

We start with its special case with $\gamma = \alpha = \beta = 1$ and recall that $\lambda = 1-\theta$. Again, $X|J = j$ is distributed as $U^{1/(j+1)}$, $U \sim$ Uniform$(0,1)$, while, marginally, X follows the particular ${}_2\mathrm{GH}_1^{\beta}$ distribution with p.d.f. $f_{IRL}(x) \equiv (1-\theta)/[(-\log\theta)\{1-(1-\theta)x\}]$, $0 < x < 1$ (a truncated increasing reciprocal linear function). In this case, $J \equiv L-1$ where L follows the logarithmic distribution with parameter $1-\theta$ on $\ell = 1, 2, \ldots$, where the logarithmic distribution with parameter $0 < p < 1$ has p.m.f. $p^{\ell}/\{-\log(1-p)\,\ell\}$. Thus, f_{IRL} arises as the distribution of $U^{1/L}$, L independent of U. Additionally, (6)–(8) reduce to saying that if $J|Y = y$ follows the geometric distribution on $1, 2, \ldots$, with parameter $\theta < y < 1$, and Y follows the distribution on $(\theta, 1)$ with density $1/\{(-\log\theta)y\}$, then J follows the logarithmic distribution with parameter $1-\theta$.

Returning to general choice of λ, α and β, if $X \sim \text{Beta}(\alpha, \beta)$, then

$$L = \theta X/\{1-(1-\theta)X\} \tag{17}$$

follows the Libby-Novick distribution with density

$$f_L(\ell; \alpha, \beta, \theta) \equiv \frac{\theta^\alpha \ell^{\alpha-1}(1-\ell)^{\beta-1}}{B(\alpha,\beta)\{1-(1-\theta)\ell\}^{\alpha+\beta}}, \qquad 0 < \ell < 1,$$

(Libby & Novick, 1982, Chen & Novick, 1984). This is the GH distribution (2) when $\gamma = \alpha + \beta$. It follows from (5)–(7) that, for $0 < \theta < 1$,

$$\text{if } L|J = j \sim \text{Beta}(\alpha + j, \beta) \text{ and } J \sim \text{NegBin}(\alpha, \theta), \text{ then } L \sim \text{LibNov}(\alpha, \beta, \theta). \tag{18}$$

This mixture relationship between beta and Libby-Novick distributions complements the transformation relationship between them. It can first be found in Chabot (2016); see Jones & Marchand (2021). The following novel result also arises, from (6)–(8):

$$\text{if } J|L = \ell \sim \text{NegBin}(\alpha+\beta, 1-(1-\theta)\ell) \text{ and } L \sim \text{LibNov}(\alpha, \beta, \theta),$$
$$\text{then } J \sim \text{NegBin}(\alpha, \theta).$$

It is also straightforward to show that if $X \sim \text{GH}(\gamma, \alpha, \beta, \theta)$ then $L \sim \text{GH}(\alpha+\beta-\gamma, \alpha, \beta, 1/\theta)$. Applying this transformation to X in (5) and (6) shows that the Gauss hypergeometric distribution can also be written as a mixture of Libby-Novick distributions. For $0 < \theta < 1$, we have:

$$\text{if } X|J = j \sim \text{LibNov}(\alpha + j, \beta, 1/\theta) \text{ and } J \sim {}_2\text{HP}_1(\gamma, \alpha, \alpha+\beta, 1-\theta),$$
$$\text{then } X \sim \text{GH}(\alpha+\beta-\gamma, \alpha, \beta, 1/\theta).$$

This includes, when $\gamma = \alpha + \beta$, the relationship

$$\text{if } X|J = j \sim \text{LibNov}(\alpha + j, \beta, 1/\theta) \text{ and } J \sim \text{NegBin}(\alpha, \theta), \text{ then } X \sim \text{Beta}(\alpha, \beta),$$

something of a 'reverse' of (18). The results in this section can be extended to $\theta > 1$ by consideration of the distributions of $1 - L$ and $1 - X$.

2.3 Special Case $p = 1$, $q = 0$, $\gamma_1 = -n$, $\lambda < 0$

The choice of negative integer γ_1 and negative $\lambda \equiv -\psi$ means that the discrete distributions in (11) and (12) have finite support; indeed that of $J|X = x$ reduces to a binomial distribution. All told, we have, again, that $X|J = j \sim \text{Beta}(\alpha + j, \beta)$ and

$$X \sim {}_2\text{GH}_1^\beta(-n, \alpha, \beta, -\psi),$$

$$J \sim \ {}_2\mathrm{HP}_1(-n, \alpha, \alpha+\beta, -\psi)$$

and

$$J|X = x \sim \ \text{Binomial}\left(n, \tfrac{\psi x}{1+\psi x}\right). \tag{19}$$

Here, the distribution of X has the following form on $0 < x < 1$, as a particular finite mixture of beta distributions:

$$f_J(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}\ (1+\psi x)^n}{\sum_{j=0}^{n} \binom{n}{j} B(\alpha+j,\beta)\psi^j}.$$

When $\alpha = \beta = 1$, the distribution of J reduces to that of $B-1$, say, where B follows the zero-truncated Binomial$(n+1, \psi/(1+\psi))$ distribution. The distribution of X in this case has (mixture of power-law) p.d.f. $f_{MP}(x) = \psi(n+1)(1+\psi x)^n/\{(1+\psi)^{n+1}-1\}$ on $0 < x < 1$. This distribution therefore serves as a mixture distribution taking the binomial distribution at (19) to the zero-truncated binomial minus one distribution just mentioned; alternatively, f_{MP} arises as the distribution of $U^{1/B}$ where $U \sim$ Uniform$(0,1)$ and B are independent.

3. General Gamma-Based Results and More Special Cases

Consider the rescaled version of (9) corresponding to βX (on $(0,\beta)$) and J, and allow $\beta, \lambda \to \infty$ in such a way that λ/β tends to a constant. Reusing X and λ for their limiting scaled versions, we have the distribution for $J = 0, 1, \ldots$, and $X > 0$ given by

$$f(j,x) = \frac{x^{\alpha+j-1}e^{-x}}{\Gamma(\alpha)} \frac{(\gamma_1)_j \cdots (\gamma_p)_j \lambda^j}{(\delta_1)_j \cdots (\delta_q)_j\, j!\ {}_{p+1}F_q(\underline{\gamma}_p, \alpha, \underline{\delta}_q, \lambda)}. \tag{20}$$

For this distribution, we have

$$X|J = j \sim \text{Gamma}(\alpha + j), \tag{21}$$

$$X \sim \ {}_{p+1}\mathrm{GH}_q^{\gamma}(\underline{\gamma}_p, \underline{\delta}_q, \alpha, \lambda), \tag{22}$$

$$J \sim \ {}_{p+1}\mathrm{HP}_q(\underline{\gamma}_p, \alpha, \underline{\delta}_q, \lambda) \tag{23}$$

and $J|X = x \sim \ {}_p\mathrm{HP}_q(\underline{\gamma}_p, \underline{\delta}_q, \lambda x)$, that is, (12). Here, the gamma-generated continuous generalized hypergeometric distribution ${}_{p+1}\mathrm{GH}_q^{\gamma}$ has density

$${}_{p+1}f_q^{\gamma}(x; \underline{\gamma}_p, \underline{\delta}_q, \alpha, \lambda) = \frac{x^{\alpha-1}e^{-x}\ {}_pF_q(\underline{\gamma}_p, \underline{\delta}_q, \lambda x)}{\Gamma(\alpha)\ {}_{p+1}F_q(\underline{\gamma}_p, \alpha, \underline{\delta}_q, \lambda)}, \qquad x > 0. \tag{24}$$

Relationships (21)–(23) and (12) constitute a gamma-based analogue of the beta-based relationships (5) and (10)–(12):

(I) the gamma-generated continuous generalized hypergeometric distribution can be expressed as a mixture of gamma distributions with a discrete generalized hypergeometric mixing distribution via (21), (22) and (23);

(II) the discrete generalized hypergeometric distribution can be expressed as a mixture of discrete generalized hypergeometric distributions with just p reduced by 1 with a gamma-generated continuous generalized hypergeometric mixing distribution via (22), (23) and (12) (as also in Johnson et al., 2005, Section 8.4).

The conditions under which the above distributions are all well-defined differ a little from the beta-generated case in Section 2. Conditions (a) and (b) are replaced by conditions (d) and (e) below; condition (c) concerning finite support discrete distributions remains the same. Conditions (d) and (e) are:

(d) positive parameters when $p+1 \leq q$;

(e) positive parameters and $0 < \lambda < 1$ when $p = q$.

The gamma-generated continuous generalized hypergeometric distribution at (24) has moments

$$E(X^r) = \frac{\Gamma(\alpha+r)}{\Gamma(\alpha)} \frac{{}_{p+1}F_q(\underline{\gamma}_p, \alpha+r, \underline{\delta}_q, \lambda)}{{}_{p+1}F_q(\underline{\gamma}_p, \alpha, \underline{\delta}_q, \lambda)}.$$

Note that all positive moments exist whenever the corresponding distribution exists.

3.1 Special Case $p = q = 0$, $0 < \lambda < 1$

When $p = q = 0$ and $0 < \lambda < 1$, (22) and (23) become

$$X \sim \text{Gamma}(\alpha)/\theta, \tag{25}$$

$$J \sim \text{NegBin}(\alpha, \theta) \tag{26}$$

while also $X|J = j \sim \text{Gamma}(\alpha + j)$ as at (21) and $J|X = x \sim \text{Poisson}(\lambda x)$ as at (16). Here, we have put $0 < \theta = 1-\lambda < 1$. Together, (21), (25) and (26) constitute the well known result that a scaled chi-squared distribution is a negative binomial mixture of chi-squared distributions when the scaling factor is bigger than one (Robbins & Pitman, 1949, Neuts & Zacks, 1967). In addition, (25), (26) and (16) show the standard construction of the negative binomial distribution as a gamma mixture of Poissons. The duality of these relationships was noted in Jones & Marchand (2021).

3.2 Special Case $p = q = 1$, $0 < \lambda < 1$

Simply setting $p = q = 1$ in (21), (22) and (12) is, in itself, of little interest. Its further special case with $\gamma_1 = \alpha = 1$ and $\delta_1 = 2$ may be, however. When we do so, we get another non-standard but relatively simple (reciprocal times exponential difference) distribution for X, namely one with p.d.f. $f_{RTED}(x) \equiv e^{-x}(e^{\lambda x}-1)/[x\{-\log(1-\lambda)\}]$, $x>0$. Recalling that the unit gamma distribution with parameter n is the distribution of $\sum_{i=1}^{n} E_i$ where the independent E_i's are each (unit) exponentials, then (21)–(23) tell us that f_{RTED} is the distribution of $\sum_{i=1}^{L} E_i$ where L has the logarithmic distribution with parameter λ, independently of $E_1, E_2, \ldots$. The dual of this, from (22), (23) and (12), is that mixing a truncated Poisson distribution with parameter λx with the distribution with p.d.f. f_{RTED} for X results in a logarithmic distribution with parameter λ.

3.3 Special Case $p=1$, $q=0$, $\gamma_1=-n$ *and* $\lambda<0$

It is immediate to write down the following in addition to (21), $X|J=j \sim \text{Gamma}(\alpha+j)$:

$$X \sim {}_2\text{GH}_0^{\gamma}(-n,\alpha,-\psi),$$

$$J \sim {}_2\text{HP}_0(-n,\alpha,-\psi)$$

and

$$J|X=x \sim \text{Binomial}\left(n, \tfrac{\psi x}{1+\psi x}\right).$$

where $\psi=-\lambda$. Note that

$${}_2f_0^{\gamma}(-n,\alpha,-\psi)=\frac{x^{\alpha-1}e^{-x}\,(1+\psi x)^n}{\sum_{j=0}^{n}\binom{n}{j}\Gamma(\alpha+j)\psi^j}$$

is a special case of the Kummer-gamma distribution of Armero & Bayarri (1997) and Ng & Kotz (1995), the general case of which allows a real-valued power of $1+\psi x$. The special case $\alpha=1$ is of a little interest but will not be spelled out here.

3.4 Special Case $p=0$, $q\in\mathbb{N}$, $\alpha=1$ *and* $\delta_1=\cdots=\delta_q=1$

The Conway-Maxwell-Poisson (COM-Poisson) distribution, $\text{CMP}(\lambda,\nu)$ (Conway & Maxwell, 1962, Shmueli et al., 2005), is a two-parameter generalization of the Poisson distribution with p.m.f. $\lambda^j/\{(j!)^\nu Z(\lambda,\nu)\}$, $j=0,1,\ldots$, where $Z(\lambda,\nu)=\sum_{k=0}^{\infty}\lambda^k/(k!)^\nu$. For integer $\nu=1,2,\ldots$, the COM-Poisson distribution is the ${}_0\text{HP}_{\nu-1}(\underline{1}_{\nu-1},\lambda)$ distribution, where $\underline{1}_{\nu-1}$ has $\nu-1$ elements all of which have value 1. Its normalizing constant is $Z(\lambda,\nu)={}_0\text{F}_{\nu-1}(\underline{1}_{\nu-1},\lambda)$. An interesting set of mixture relationships then arises by setting $p=0$, $q=\nu$, $\alpha=1$ and $\delta_i=1$, $i=1,\ldots,\nu$ in (21)–(24) and (12). We get

$$X|J=j \sim \text{Gamma}(1+j), \tag{27}$$

$$X \sim \text{ContCMP}(\lambda,\nu), \tag{28}$$

$$J \sim \text{CMP}(\lambda,\nu) \tag{29}$$

and

$$J|X = x \sim \text{ CMP}(\lambda x, \nu + 1). \tag{30}$$

Here, ContCMP(λ, ν) is the "continuous COM-Poisson distribution" newly defined to have density

$$\frac{e^{-x}\ {}_0F_\nu(\underline{1}_\nu, \lambda x)}{{}_0F_{\nu-1}(\underline{1}_{\nu-1}, \lambda)} = \frac{e^{-x}\sum_{j=0}^\infty (\lambda x)^j/(j!)^{\nu+1}}{\sum_{j=0}^\infty \lambda^j/(j!)^\nu} = \frac{e^{-x} Z(\lambda x, \nu+1)}{Z(\lambda, \nu)}, \qquad x > 0.$$

As well as its own mixture derivation through (27)–(29), note how the continous COM-Poisson distribution acts as the mixture distribution of X in arriving at the COM-Poisson distribution (29) from the COM-Poisson distribution (30). Note too that relationships (27)–(30), although derived for integer ν, continue to hold for real $\nu > 0$.

4. A Further Connection Between Beta-Based and Gamma-Based Results

At the beginning of Section 3, we pointed out that our gamma-based results are related to our beta-based results through a limiting argument. Another link between beta and gamma distributions also generalizes to the current results. Recall that the Beta(α, β) distribution is that of $G_\alpha/(G_\alpha + G_\beta)$ where G_α and G_β are independent random variables, each following a (unit-scale) gamma distribution with the specified parameter. It follows that the ${}_{p+1}\text{GH}^\beta_{q+1}(\underline{\gamma}_p, \underline{\delta}_q, \alpha, \beta, \lambda)$ distribution at (10) is that of $R \equiv H/(H + G_\beta)$ where H is independent of G_β, $H|J = j \sim \text{Gamma}(\alpha + j)$ and $J \sim {}_{p+1}\text{HP}_{q+1}(\underline{\gamma}_p, \alpha, \underline{\delta}_q, \alpha + \beta, \lambda)$ from (11). It follows from (21)–(23) with q increased to $q + 1$ that $H \sim {}_{p+1}\text{GH}^\gamma_{q+1}(\underline{\gamma}_p, \underline{\delta}_q, \alpha + \beta, \alpha, \lambda)$ distribution.

5. Comments on the Wider Context

The key to obtaining the results of this paper is to explore the marginal and conditional distributions of a well-chosen bivariate distribution, interpreting the links between them as mixture relationships. In particular, bivariate distributions (9) and (20) are examples of a joint distribution of continuous X and discrete J which has p.d.f./p.m.f of the form $f(x, j) = f_j(x)p(j)$ so that $X|J = j \sim f_j$, the distribution with p.d.f. f_j, $X \sim f^p_{\text{mix}}$, which has the mixture density $\sum_k f_k(x)p(k)$, $J \sim p$ and $J|X = x$ has the distribution p^f_j say, with p.m.f. $f_j(x)p(j)/\sum_k f_k(x)p(k)$. So, if $X|J = j \sim f_j$ and $J \sim p$ then $X \sim f^p_{\text{mix}}$, dual to the other mixture result that if $J|X = x \sim p^f_j$ and $X \sim f^p_{\text{mix}}$ then $J \sim p$.

The challenge has been to identify interesting and interpretable (and sometimes novel and/or surprising) examples of these relationships. To this end, let $X > 0$ and

$J = 0, 1, ...$, be continuous and discrete random variables on the given supports, respectively, and consider the joint distribution with p.d.f/p.m.f. given by

$$f(x,j) = \frac{x^j f(x)}{\mu_j} \frac{a_j \lambda^j}{\sum_{k=0}^{\infty} a_k \lambda^k}, \qquad x > 0, \quad j = 0, 1, \dots . \tag{31}$$

Here, f is a density with all moments $\mu_j = \mathbb{E}_f(X^j)$ finite, $j = 0, 1, ...$, while $\underline{a} \equiv \{a_0, a_1, ...\}$, with $a_j > 0$, $j = 0, 1, ...$, and $\lambda > 0$ are the parameters of a power series distribution, PSD$(\underline{a}, \lambda)$ say. For such a distribution, it is easy to see that, conditionally, $X|J = j$ follows the j'th power-weighted version of f with density $x^j f(x)/\mu_j$ and, marginally, X follows the mixture of power-weighted versions of f having density $\sum_{j=0}^{\infty}(a_j/\mu_j)(\lambda x)^j f(x)/\sum_{k=0}^{\infty} a_k \lambda^k$. Correspondingly, $J \sim$ PSD$(\underline{a}, \lambda)$ and $J|X = x \sim$ PSD$(\underline{a/\mu}, \lambda x)$, where $\underline{\mu} = \{1, \mu_1, \mu_2, ...\}$ and the division is performed elementwise.

In the spirit of X and $X|J = j$ following weighted versions of f, it might be of interest to note that the PSD distributions above can be viewed as weighted Poisson distributions. For instance, rewrite (31) as

$$f(x,j) = \frac{x^j f(x)}{\mu_j} \frac{a_j^* e^{-\lambda} \lambda^j / j!}{\sum_{k=0}^{\infty} a_k^* e^{-\lambda} \lambda^k / k!}, \qquad x > 0, \quad j = 0, 1, \dots$$

where $a_j^* = j! a_j$, $j = 0, 1, ...$, and interpret PSD$(\underline{a}, \lambda)$ as WPD$(\underline{a}^*, \lambda)$, the weighted Poisson$(\lambda)$ distribution with weights a_j^*, $j = 0, 1, \dots$. This might be done, for example, to take advantage of results of Kokonendji et al. (2008) which relate concavity properties of the weight sequence to dispersion properties of the resulting distribution.

6. Conclusion

Despite all the intriguing generality in Section 5, we have found the choices made in this article, namely gamma and beta distributions for f and their interactions with generalized hypergeometric distributions, to work out most nicely – (5), (10), (11) and (12) in the beta case, (21), (22), (23) and (12) in the gamma case – and to form, via their many special cases, a most interesting slew of related relations.

References

Armero, C. & Bayarri, M.J. (1994) Prior assessments for prediction in queues. *J. Roy. Statist. Soc. Ser. D*, 43, 139–153.

Armero, C. & Bayarri, M.J. (1997) A Bayesian analysis of a queuing system with unlimited service. *J. Statist. Planning Inference*, 58, 241–261.

Bardwell, G.E. & Crow, E.L. (1964) A two-parameter family of hyper-Poisson distributions. *J. Amer. Statist. Assoc.*, 59, 133–141.

Bhattacharya, S.K. (1966) Confluent hypergeometric distributions of discrete and continuous type with applications to accident proneness. *Calcutta Statist. Assoc. Bull.*, 15, 20–31.

Chabot, M. (2016) Sur l'estimation du rapport de deux paramètres d'intensité poissonniens et l'estimation par fonctions de masse prédictives. M.Sc. thesis, Faculté de Science, Université de Sherbrooke, Canada.

Chen, J.J. & Novick, M.R. (1984) Bayesian analysis for binomial models with generalized beta prior distributions. *J. Educ. Statist.,* 9, 163–175.

Conway, R.W. & Maxwell, W.L. (1962) A queuing model with state dependent service rates. *J. Indust. Eng.,* 12, 132–136.

Gordy, M.B. (1998) Computationally convenient distributional assumptions for common-value actions. *Comput. Econ.*, 12, 61–78.

Hall, W.J. (1956) Some hypergeometric series distributions occurring in birth-and-death processes at equilibrium, (preliminary report). Abstract in *Ann. Math. Statist.*, 27, 221.

Johnson, N.L., Kemp, A.W. & Kotz, S. (2005) *Univariate Discrete Distributions*, 3rd ed. Hoboken, NJ: Wiley.

Jones, M.C. & Marchand, É. (2021) A (non-central) chi-squared mixture of non-central chi-squareds is (non-central) chi-squared, and related results, corollaries and applications. *Stat,* 10, e398.

Kemp, A.W. (1968) A wide class of discrete distributions and the associated differential equations. *Sankhyā Ser. A*, 30, 401–410.

Kokonendji, C.C., Mizère, D. & Balakrishnan, N. (2008) Connections of the Poisson weight function to overdispersion and underdispersion. *J. Statist. Planning Inference,* 138, 1287–1296.

Libby, D.L. & Novick, M.R. (1982) Multivariate generalized beta distributions with applications to utility assessment. *J. Educ. Statist.,* 7, 271–294.

Neuts, M.F. & Zacks, S. (1967) On mixtures of χ^2- and F-distributions which yield distributions of the same family. *Ann. Inst. Statist. Math.*, **19**, 527–536.

Ng, K.W. & Kotz, S. (1995) *Kummer-Gamma and Kummer-Beta univariate and multivariate distributions.* Research Report No. 84, Department of Statistics, University of Hong Kong.

Robbins, H. & Pitman, E.J.G. (1949) Application of the method of mixtures to quadratic forms in normal variates. *Ann. Math. Statist.,* 20, 552–560.

Shmueli, G., Minka, T., Kadane, J.B., Borle, S. & Boatwright, P. (2005) A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. *J. Roy. Statist. Soc. Ser. C,* 54, 127–142.