

# A (non-central) chi-squared mixture of non-central chi-squareds is (non-central) chi-squared and related results, corollaries and applications

M. C. Jones<sup>1</sup>  | Éric Marchand<sup>2</sup>

<sup>1</sup>School of Mathematics and Statistics, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK

<sup>2</sup>Département de mathématiques, Université de Sherbrooke, Sherbrooke, Quebec, J1K 2R1, Canada

## Correspondence

M. C. Jones, School of Mathematics and Statistics, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK.  
Email: m.c.jones@open.ac.uk

Our main, novel, result is that a certain non-central chi-squared mixture of non-central chi-squared distributions is itself a scaled non-central chi-squared distribution. From this and a link to a known result on a mixture representation for a scaled central chi-squared distribution, numerous further mixture results, both old and new, ensue. These include mixture results for central  $F$ , non-central  $F$  and Libby–Novick distributions. The main result involves distributions all with the same degrees of freedom; it is also extended to the case where the mixing non-central chi-squared distribution has degrees of freedom an even number larger than that of the conditional non-central chi-squared distribution, with further consequences pursued.

## KEYWORDS

beta distribution, binomial mixture,  $F$  distribution, gamma distribution, Libby–Novick distribution, negative binomial mixture

## 1 | INTRODUCTION

We seek to provide a unified presentation of a number of mixture results, both old and new, concerning central and non-central chi-squared distributions, as well as various other distributions. Our findings link various results from the literature. This is achieved from the base of our novel main result (Section 2), which is that a certain non-central chi-squared mixture of non-central chi-squared random variables is itself distributed as a scaled non-central chi-squared. Our presentation highlights the relative ease with which we can derive further mixture results from results such as this. For instance, in Section 3, we link the main result to a known result on a mixture representation for a scaled (central) chi-squared random variable (e.g., Neuts & Zacks, 1967) and pursue several corollaries thereof. The main result of Section 2 is then extended to a more general choice of non-central chi-squared mixing distribution in Section 4, with further corollaries and consequences. As the article unfolds, specific applications to, and links with, results in Bayesian estimation of the chi-squared non-centrality parameter, balanced variance component models, M/M/1 queues, central and non-central  $F$  distributions, the Libby–Novick distribution, and the distribution of the sample coefficient of determination are described. Generally, non-central distributions, such as chi-squared,  $F$  and beta, arise in various situations, typically related to quadratic forms in normal linear models and to the distribution of various non-null test statistics (e.g., Johnson et al. 1995)

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Stat* published by John Wiley & Sons Ltd.

## 2 | MAIN RESULT AND CONSEQUENCES THEREOF

The main result of this article is that a non-central chi-squared mixture of non-central chi-squared distributions – where the degrees of freedom of both non-central chi-squareds is the same – is itself a scaled non-central chi-squared distribution with the same degrees of freedom. The non-central chi-squared distribution with  $\nu$  degrees of freedom,  $\nu \in \mathbb{R}$ , and non-centrality parameter  $\delta \geq 0$ , is written as  $\chi_\nu^2(\delta)$  or – the central case – just  $\chi_\nu^2$  when  $\delta = 0$ . In what follows,  $\sim$  denotes ‘is distributed as’.

**Result 1.** Let  $h, \delta \geq 0, \nu > 0$ . If  $X|Y = y \sim \chi_\nu^2(hy)$  and  $Y \sim \chi_\nu^2(\delta)$ , then  $\frac{X}{1+h} \sim \chi_\nu^2\left(\frac{h\delta}{1+h}\right)$ .

Trivially, a special case of this result is that a central chi-squared mixture of non-central chi-squared distributions is distributed as central chi-squared.

**Corollary 1.** Let  $h \geq 0, \nu > 0$ . If  $X|Y = y \sim \chi_\nu^2(hy)$  and  $Y \sim \chi_\nu^2$ , then  $\frac{X}{1+h} \sim \chi_\nu^2$ .

*Proof of Result 1.* Let  $M_Y(t) = (1-2t)^{-\nu/2} e^{\delta t/(1-2t)}$ ,  $t < 1/2$ , be the moment generating function (m.g.f.) of the  $\chi_\nu^2(\delta)$  distribution. Then,

$$M_X(t) = \mathbb{E}_Y\{M_{X|Y}(t)\} = (1-2t)^{-\nu/2} M_Y\{ht/(1-2t)\} = \{1-2(1+h)t\}^{-\nu/2} e^{\delta ht/\{1-2(1+h)t\}}$$

so that

$$M_{X/(1+h)}(t) = M_X\left(\frac{t}{1+h}\right) = (1-2t)^{-\nu/2} e^{\delta ht/\{(1+h)(1-2t)\}}, \quad t < 1/2.$$

In the important special case that  $\nu = m$ , say, is a positive integer, Result 1 arises from consideration of properties of underlying multivariate normal distributions. Let  $N_m(\mu, \Sigma)$  denote the  $m$ -dimensional normal distribution with mean  $\mu$  and variance  $\Sigma$  and let  $I_m$  be the  $m$ -dimensional identity matrix.

*Alternative Proof of Result 1 for Positive Integer Degrees of Freedom.* Consider

$$Z|W = w \sim N_m(\sqrt{hw}, I_m), \quad W \sim N_m(\sqrt{\delta}, I_m) \quad (1)$$

so that  $\sqrt{h}W \sim N_m(\sqrt{h\delta}, hI_m)$ . We infer Result 1 by setting  $\|Z\|^2 = X$  and  $\|W\|^2 = Y$ . A familiar argument to recover the marginal distribution of  $Z$ , and hence that of  $X$ , is as follows. From (1), we obtain

$$\begin{aligned} Z - \sqrt{h}W|W = w \sim N_m(0, I_m) &\Rightarrow Z - \sqrt{h}W \sim N_m(0, I_m) \text{ independently of } W \\ &\Rightarrow \{(Z - \sqrt{h}W) + \sqrt{h}W\} \sim N_m(\sqrt{h\delta}, (1+h)I_m) \\ &\Rightarrow \frac{Z}{\sqrt{1+h}} \sim N_m\left(\sqrt{\frac{h\delta}{1+h}}, I_m\right) \end{aligned}$$

from which the claimed distribution of  $X/(1+h)$  follows. □

For the multivariate normal distribution defined by (1), it is also easy to verify that

$$W|Z = z \sim N_m\left(\frac{\sqrt{hz} + \sqrt{\delta}}{1+h}, \frac{I_m}{1+h}\right)$$

and hence

$$(1+h)Y|Z = z \sim \chi_m^2\left(\frac{\|\sqrt{hz} + \sqrt{\delta}\|^2}{1+h}\right).$$

When  $\delta = 0$ , the above reduces to  $(1+h)Y|Z=z \sim \chi_m^2\left(\frac{h\|z\|^2}{(1+h)}\right)$  and, since this conditional distribution depends on  $z$  only through  $\|z\|^2 = x$ , we infer that  $(1+h)Y|X=x \sim \chi_m^2(hx/(1+h))$ . The combination of the conditional and marginal distributions of  $Y$  corresponds to the following equivalent reformulation of Corollary 1: if  $(1+h)Y|X=x \sim \chi_m^2(hx/(1+h))$  and  $X/(1+h) \sim \chi_m^2$ , then  $Y \sim \chi_m^2$ .

*Remark 1.* Denote by  $\sqrt{h}w = \mu$  the mean of the multivariate normal distribution of  $Z|W=w$  in (1) and write  $\theta = \|\mu\|^2$  for the non-centrality parameter of the corresponding  $\chi_m^2(\theta)$  distribution of  $X|\sqrt{h}W = \mu$ . Then, if  $\mu \sim N_m(0, I_m)$  is the prior distribution for  $\mu$ , the posterior distribution for  $\theta|X=x$  is given by  $(1+h)\theta/h|X=x \sim \chi_m^2(hx/(1+h))$ . This version of Corollary 1 was given by Perlman and Rasmussen (1975). See also L'Moudden and Marchand (2021, Example 3.2).

**Application 1.** Corollary 1 was independently observed by M. Bilodeau who, in Bilodeau (2021), has updated his result to a version of Result 1. Bilodeau (2021) gives a clear exposition of how Result 1 allows for the ready derivation of distributional results for balanced variance component models, that is, (normal-based) ANOVA models incorporating (normally distributed) random effects.

Now, define the doubly non-central  $F(p, q, \delta_1, \delta_2)$  distribution as the distribution of  $F = qX_1/(pX_2)$  where  $X_1 \sim \chi_p^2(\delta_1)$  and  $X_2 \sim \chi_p^2(\delta_2)$ , independently,  $p, q > 0, \delta_1, \delta_2 \geq 0$ . By applying Result 1 to numerator and denominator, it extends to the non-central  $F$  distribution as follows.

**Corollary 2.** If  $F|Y_1 = y_1, Y_2 = y_2 \sim F(p, q, h_1y_1, h_2y_2)$  and  $Y_1 \sim \chi_p^2(\delta_1)$  and  $Y_2 \sim \chi_q^2(\delta_2)$ , independently, then  $(1+h_2)F/(1+h_1) \sim F\left(p, q, \frac{h_1\delta_1}{1+h_1}, \frac{h_2\delta_2}{1+h_2}\right)$ .

Not only, however, can any doubly non-central  $F$  distribution be obtained by mixing another non-central  $F$  distribution with the same degrees of freedom over its non-centrality parameters in this way, but since  $\delta_1$  and/or  $\delta_2$  can be zero, the doubly non-central  $F$  can give rise to singly non-central and central  $F$  distributions, and singly non-central to central. Similar results pertain for non-central beta and  $t$  distributions.

### 3 | SCALED CHI-SQUARED AS A MIXTURE OF CHI-SQUAREDS AND CONSEQUENCES THEREOF

Corollary 1 can be decomposed using the well-known representation of the non-central chi-squared distribution as that of a Poisson mixture of chi-squared distributions, namely, if  $X|J=j \sim \chi_{\nu+2j}^2$  and  $J \sim \text{Poisson}(\delta/2)$ , then  $X \sim \chi_{\nu}^2(\delta)$ . This leads to the following result which was first observed in the mid-twentieth century (Neuts & Zacks, 1967; Robbins & Pitman, 1949): providing the scaling factor is greater than 1, a scaled chi-squared distribution is a negative binomial mixture of chi-squared distributions. We write  $\text{NegBin}(\lambda, p)$  to mean that version of the negative binomial distribution with probability mass function (p.m.f.)  $\frac{\Gamma(\lambda+j)}{\Gamma(\lambda)!} p^\lambda (1-p)^j$  on  $j = 0, 1, \dots$

**Result 2.** Let  $\theta \geq 1, \nu > 0$ . If  $X|J=j \sim \chi_{\nu+2j}^2$  and  $J \sim \text{NegBin}(\nu/2, 1/\theta)$ , then  $X/\theta \sim \chi_{\nu}^2$ .

*Proof.* Expand Corollary 1 in the following way: if  $X|J=j, Y = y \sim \chi_{\nu+2j}^2, J|Y = y \sim \text{Poisson}((\theta-1)y/2)$  and  $Y \sim \chi_{\nu}^2$ , then  $X/\theta \sim \chi_{\nu}^2$ . Then, make the standard calculation to obtain the marginal distribution of  $J$ . □

Write  $G_\alpha$  for a gamma random variable with shape parameter  $\alpha$  and unit scale parameter, writing  $\text{Gamma}(\alpha)$  for its distribution, and note that  $X_\nu/2 \stackrel{d}{=} G_{\nu/2}$  when  $X_\nu \sim \chi_{\nu}^2$ . Here,  $\stackrel{d}{=}$  denotes ‘has the same distribution as’. Corollary 3 (Teicher, 1960) is immediate.

**Corollary 3.** Let  $\theta \geq 1, \alpha > 0$ . If  $X|J=j \sim \text{Gamma}(\alpha+j)$  and  $J \sim \text{NegBin}(\alpha, 1/\theta)$ , then  $X/\theta \sim \text{Gamma}(\alpha)$ .

That is, a gamma distribution can be represented as a negative binomial shape mixture of gamma distributions. In particular, when  $\alpha = 1$ , an exponential distribution with mean  $\theta > 1$  arises from choosing from the set of unit scale Erlang (gamma with integer shape parameter) distributions according to a geometric distribution with probability parameter  $1/\theta$ .

**Application 2** (of exponential special case of Corollary 3). Consider a classical M/M/1 queueing process with arrivals following a Poisson process with intensity  $\lambda$  and service times occurring independently and identically distributed as exponential with rate  $\mu$  (i.e., mean  $1/\mu$ ). Consider the number of individuals,  $L$ , in the system as well as  $W$ , the total time spent by an individual in the system. In the steady state, whenever  $\mu > \lambda$ , it can be shown that  $L \sim \text{Geometric}((\mu - \lambda)/\mu)$  and  $(\mu - \lambda)W \sim \text{Gamma}(1)$  (e.g., Karlin &

Taylor, 1975). By the properties of the Poisson process, we have that  $\mu W|L = \ell \sim \text{Gamma}(1 + \ell)$ . The exponential distribution of  $W$  corresponds to the  $\alpha = 1$  version of Corollary 3 by setting  $X = \mu W$  and  $\theta = \mu/(\mu - \lambda)$ .

A new observation concerning Result 2/Corollary 3 is that they are ‘dual’ to versions of the representation of a negative binomial distribution as a gamma mixture of Poisson distributions which was used in the proof of Result 2. In particular, if  $X$  and  $J$  are distributed as in Corollary 3, then it is easy to see that  $J|X = x \sim \text{Poisson}(x(\theta - 1)\theta)$  which combines with  $X/\theta \sim \text{Gamma}(\alpha)$  to produce  $J \sim \text{NegBin}(\alpha, 1/\theta)$ .

Now, introduce the beta distribution of the second kind,  $\text{Beta}_2(\alpha, \beta)$ , as the distribution of  $G_\alpha/G_\beta$  where  $G_\alpha$  and  $G_\beta$  are independent,  $\alpha, \beta > 0$ , and write  $F(p, q)$  for the central  $F$  distribution. By applying Corollary 3 to numerator and denominator and accounting for  $p$  and  $q$  or  $p/q$  taking values above or below unity gives the following.

**Corollary 4.** *If any of the following four mixtures pertains, then  $F \sim F(p, q)$ :*

- (i) for  $p, q > 1$ ,  $F|J_1 = j_1, J_2 = j_2 \sim \text{Beta}_2((p/2) + j_1, (q/2) + j_2)$  where  $J_1 \sim \text{NegBin}(p/2, 1/q)$  and  $J_2 \sim \text{NegBin}(q/2, 1/p)$ , independently;
- (ii) for  $p, q < 1$ ,  $F|J_1 = j_1, J_2 = j_2 \sim \text{Beta}_2((p/2) + j_1, (q/2) + j_2)$  where  $J_1 \sim \text{NegBin}(p/2, p)$  and  $J_2 \sim \text{NegBin}(q/2, q)$ , independently;
- (iii) for  $q > p$ ,  $F|J = j \sim \text{Beta}_2((p/2) + j, q/2)$  and  $J \sim \text{NegBin}(p/2, p/q)$ ;
- (iv) for  $q < p$ ,  $F|J = j \sim \text{Beta}_2(p/2, (q/2) + j)$  and  $J \sim \text{NegBin}(q/2, q/p)$ .

Corollary 4(i) was given by Neuts and Zacks (1967) for  $p = q (> 1)$ .

The usual beta distribution (of the first kind),  $\text{Beta}(\alpha, \beta)$ , is, of course, the distribution of  $G_\alpha/(G_\alpha + G_\beta)$  where  $G_\alpha$  and  $G_\beta$  are independent,  $\alpha, \beta > 0$ . If the gamma random variables are allowed to have different scales, an almost equally simple – but less well known – distribution arises which is known as the Libby–Novick distribution (Libby & Novick, 1982): for  $\alpha, \beta, \mu > 0$ ,  $\text{LN}(\alpha, \beta, \mu)$  is the distribution of  $G_\alpha/(G_\alpha + \mu G_\beta)$ . Corollary 3 can now be applied to  $\mu G_\beta$  in  $G_\alpha/(G_\alpha + \mu G_\beta)$  when  $\mu > 1$  and to  $G_\alpha/\mu$  in  $(G_\alpha/\mu)/\{(G_\alpha/\mu) + G_\beta\}$  when  $\mu < 1$  to obtain the following:

**Corollary 5.** *If either of the following two mixtures pertains, then  $L \sim \text{LN}(\alpha, \beta, \mu)$ :*

- (i) for  $\mu < 1$ ,  $L|I = i \sim \text{Beta}(\alpha + i, \beta)$  and  $I \sim \text{NegBin}(\alpha, \mu)$ ;
- (ii) for  $\mu > 1$ ,  $L|I = i \sim \text{Beta}(\alpha, \beta + i)$  and  $I \sim \text{NegBin}(\beta, 1/\mu)$ .

This is a construction of the Libby–Novick distribution that seems previously to have been given only in Chabot (2016).

Let  $B_{\alpha, \beta} \sim \text{Beta}(\alpha, \beta)$ . Then, a standard relationship is that  $G_{\alpha+\beta} B_{\alpha, \beta} \stackrel{d}{=} G_\alpha$  when  $G_{\alpha+\beta}$  and  $B_{\alpha, \beta}$  are independent. Writing this in an equivalent manner, we have

$$G_\gamma B_{\alpha, \gamma-\alpha} \stackrel{d}{=} G_\alpha, \quad 0 < \alpha < \gamma, \quad (2)$$

$G_\gamma$  and  $B_{\alpha, \gamma-\alpha}$  being independent. Multiplying both sides of (2) by  $\psi$  where  $0 < \psi < 1$ , Corollary 3 shows that

$$G_{\alpha+1} \text{ with } I \sim \text{NegBin}(\alpha, \psi) \stackrel{d}{=} G_{\gamma+J} B_{\alpha, \gamma-\alpha} \text{ with } J \sim \text{NegBin}(\gamma, \psi). \quad (3)$$

Standard manipulations given in Appendix A yield the possibly novel fact that, for fixed  $j = 0, 1, \dots$ , and independent  $G_{\gamma+j}$  and  $B_{\alpha, \gamma-\alpha}$

$$G_{\gamma+j} B_{\alpha, \gamma-\alpha} \stackrel{d}{=} G_{\alpha+M} \text{ with } M \sim \text{BetaBin}(j, \alpha, \gamma - \alpha). \quad (4)$$

Here,  $\text{BetaBin}(n, \alpha, \beta)$  is the beta-binomial distribution on  $0, \dots, n$  which is the distribution of  $N$  when  $N|P = p \sim \text{Binomial}(n, p)$  and  $P \sim \text{Beta}(\alpha, \beta)$ ; see, e.g., Johnson et al. (2005, section 6.2.2). Noting that  $I$  on the left-hand side of (3) has the same distribution as  $M$  when  $M|J = j$  is distributed as on the right-hand side of (4) and  $J$  is distributed as on the right-hand side of (3) yields the attractive, and possibly unknown, relationship in Corollary 6.

**Corollary 6.** *Let  $0 < \alpha < \gamma$ . If  $M|N = n, P = p \sim \text{Binomial}(n, p)$  where  $N \sim \text{NegBin}(\gamma, \theta)$  and, independently,  $P \sim \text{Beta}(\alpha, \gamma - \alpha)$ , then  $M \sim \text{NegBin}(\alpha, \theta)$ .*

## 4 | EXTENSION OF MAIN RESULT AND CONSEQUENCES THEREOF

In this section, we give an extension of Result 1 to the case where the mixing non-central chi-squared distribution has degrees of freedom an even number,  $2k$ , larger than that of the conditional non-central chi-squared distribution.

**Result 3.** Let  $h, \delta \geq 0, \nu > 0, k = 0, 1, \dots$ . If  $X|Y = y \sim \chi_{\nu}^2(hy)$  with  $Y \sim \chi_{\nu+2k}^2(\delta)$ , and  $\frac{X'}{1+h}|J=j \sim \chi_{\nu+2j}^2\left(\frac{h\delta}{1+h}\right)$  with  $J \sim \text{Binomial}\left(k, \frac{h}{1+h}\right)$ , then  $X$  and  $X'$  have the same distribution.

*Proof.* Adapting the m.g.f.-based proof of Result 1, for  $t < 1/2$ , we get to

$$M_X(t) = \{1 - 2(1+h)t\}^{-\nu/2} \left(1 - \frac{2ht}{1-2t}\right)^{-k} e^{\delta ht / \{1-2(1+h)t\}} = \{1 - 2(1+h)t\}^{-\nu/2-k} (1-2t)^k e^{\delta ht / \{1-2(1+h)t\}}.$$

Now, for integer  $k$ , we can write

$$(1-2t)^k = \left\{ \frac{h+1-2(1+h)t}{1+h} \right\}^k = \frac{1}{(1+h)^k} \sum_{j=0}^k \binom{k}{j} h^j \{1-2(1+h)t\}^{k-j}$$

which gives  $M_X(t) = M_{X'}(t)$  because the binomial p.m.f. of  $J$  is  $\binom{k}{j} h^j / (1+h)^k, j = 0, \dots, k$ . □

Result 3 with  $k = 0$  reduces to Result 1 since then  $J \equiv 0$ .

There is a host of corollaries which flow from Result 3 of which just a few of the most interesting are explicitly provided here. For avoidance of doubt, a direct proof of Corollary 7 is provided in Appendix B.

**Corollary 7.** Let  $0 < \psi < 1, \alpha > 0, k = 0, 1, \dots$ . If  $Z|I=i \sim \text{Gamma}(\alpha+i)$  with  $I \sim \text{NegBin}(\alpha+k, \psi)$ , and  $\psi Z'|J=j \sim \text{Gamma}(\alpha+j)$  with  $J \sim \text{Binomial}(k, 1-\psi)$ , then  $Z$  and  $Z'$  have the same distribution.

Corollaries 8 and 9 follow from Corollary 7 in much the same ways as Corollaries 5 and 2 follow from Corollaries 3 and 1, respectively.

**Corollary 8.** Let  $\alpha, \beta > 0, k = 0, 1, \dots$ . In each part below,  $L$  and  $L'$  have the same distribution.

- (i) Let  $\mu < 1$ . If  $L|I=i \sim \text{Beta}(\alpha+i, \beta)$  with  $I \sim \text{NegBin}(\alpha+k, \mu)$ , and  $L'|J=j \sim \text{LN}(\alpha+j, \beta, \mu)$  with  $J \sim \text{Binomial}(k, 1-\mu)$ .
- (ii) Let  $\mu > 1$ . If  $L|I=i \sim \text{Beta}(\alpha, \beta+i)$  with  $I \sim \text{NegBin}(\beta+k, 1/\mu)$ , and  $L'|J=j \sim \text{LN}(\alpha, \beta+j, \mu)$  with  $J \sim \text{Binomial}(k, 1-1/\mu)$ .

**Corollary 9.** Let  $h_i, \delta_i \geq 0, k_i = 0, 1, \dots, i = 1, 2$  and  $p, q > 0$ . If  $F|Y_1 = y_1, Y_2 = y_2 \sim F(p, q, h_1 y_1, h_2 y_2)$  with  $Y_1 \sim \chi_{p+2k_1}^2(\delta_1)$  and, independently,  $Y_2 \sim \chi_{q+2k_2}^2(\delta_2)$ , and

$$\frac{(1+h_2)p(q+2J_2)}{(1+h_1)q(p+2J_1)} F'|J_1 = j_1, J_2 = j_2 \sim F\left(p+2j_1, q+2j_2, \frac{h_1 \delta_1}{1+h_1}, \frac{h_2 \delta_2}{1+h_2}\right),$$

with  $J_1 \sim \text{Binomial}\left(k_1, \frac{h_1}{1+h_1}\right)$  and, independently,  $J_2 \sim \text{Binomial}\left(k_2, \frac{h_2}{1+h_2}\right)$ , then  $F$  and  $F'$  have the same distribution.

**Application 3** (of Corollaries 8(i) and 9). The coefficient of determination is the fraction of variance explained by a multiple linear regression and is equal to the square of the multiple correlation coefficient. Let  $R^2$  be the sample coefficient of determination of a multivariate normal sample of size  $n$  and dimension  $p$ , with population coefficient of determination  $\rho^2$ . Marchand (1997, Section 4) notes that

$$R^2|I=i \sim \text{Beta}\left(\frac{p-1}{2}+i, \frac{n-p}{2}\right) \text{ and } I \sim \text{NegBin}\left(\frac{p-1}{2}+\frac{n-p}{2}, 1-\rho^2\right). \tag{5}$$

It follows anew from Corollary 8(i) that, when  $n - p$  is even, it is the case that  $R^2$  has the same distribution as  $R'^2$  when

$$R^2|J=j \sim \text{LN}\left(\frac{p-1}{2}+j, \frac{n-p}{2}, 1-\rho^2\right) \text{ and } J \sim \text{Bin}\left(\frac{n-p}{2}, \rho^2\right). \quad (6)$$

It then follows from a property of the Libby–Novick distribution that, when  $n - p$  is even,

$$\frac{R^2}{1-R^2}|J=j \sim \frac{1}{1-\rho^2} \text{Beta}_2(p-1+2j, n-p) \text{ and } J \sim \text{Bin}\left(\frac{n-p}{2}, \rho^2\right)$$

which is Theorem 2 of Gurland (1968). And Result 3 with  $\delta = 0$  tells us that this is equivalent to

$$\frac{R^2}{1-R^2}|Y=y \sim \text{Beta}_2\left(p-1, n-p, \frac{\rho^2 y}{1-\rho^2}\right) \text{ and } Y \sim \chi_{n-1}^2$$

which is given in Example 3 of Wijsman (1959). Here, the (singly) non-central  $\text{Beta}_2(p, q, \delta)$  distribution is the distribution of  $pF/q$  where  $F \sim F(p, q, \delta, 0)$ .

Benton and Krishnamoorthy (2003) are amongst those concerned with computing the distribution of  $R^2$  using the representation (5). Their main contribution is to recommend starting recursive computations from summands corresponding to a high value of the p.m.f. of the mixing distribution rather than from zero, thereby alleviating computational inaccuracies due to computer under-flow. Benton and Krishnamoorthy's algorithm also requires a stopping rule for the number of summands in the infinite sum associated with the negative binomial mixing distribution. It follows from (6) that, when  $n - p$  is even, this second difficulty can be circumvented by employing the finite sum associated with the binomial mixing distribution.

## 5 | A CLOSING REMARK

*Remark 2.* The non-central chi-squared distribution with zero degrees of freedom (and non-zero non-centrality parameter) is defined as the distribution of  $X$  where  $X|Y=y \sim \chi_{2y}^2$ , and  $Y \sim \text{Poisson}(\delta)$  (Siegel, 1979), while its central version is identically zero. Results 1 and 3 continue to hold for  $\nu = 0$  but Result 2 becomes vacuous. We have chosen not to complicate corollaries of Results 1 and 3 by laying out all the instances where  $\nu$  could be set to zero in them; the reader should find them obvious.

## DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## ORCID

M. C. Jones  <https://orcid.org/0000-0001-6914-0675>

## REFERENCES

- Benton, C., & Krishnamoorthy, K. (2003). Computing discrete mixtures of continuous distributions: Noncentral chisquare, noncentral t and the distribution of the square of the sample multiple correlation coefficient. *Computational Statistics & Data Analysis*, 43, 249–267.
- Bilodeau, M. (2021). ANOVA of balanced variance component models, arXiv:2103.13202v2.
- 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 des Sciences, Université de Sherbrooke, Canada.
- Gurland, J. (1968). A relatively simple form of the distribution of the multiple correlation coefficient. *Journal of the Royal Statistical Society Series B*, 30, 276–283.
- Johnson, N. L., Kemp, A. W., & Kotz, S. (2005). *Univariate Discrete Distributions*. (3rd ed.). Hoboken, NJ: Wiley.
- Johnson, N. L., Kotz, S., & Balakrishnan, N. (1995). *Continuous Univariate Distributions*. (Vol. 2, 2nd ed.). New York: Wiley.
- Karlin, S., & Taylor, H. M. (1975). *A First Course in Stochastic Processes*. (2nd ed.). San Diego, CA: Academic Press.
- Libby, D. L., & Novick, M. R. (1982). Multivariate generalized beta distributions with applications to utility assessment. *Journal of Educational Statistics*, 7, 271–294.
- L'Moudden, A., & Marchand, É. (2021). Bayesian estimation and prediction for certain type of mixtures. *Communications in Statistics - Theory and Methods*. <https://doi.org/10.1080/03610926.2021.1913185>
- Marchand, É. (1997). On moments of beta mixtures, the noncentral beta distribution, and the coefficient of determination. *Journal of Statistical Computation and Simulation*, 59, 161–178.
- Neuts, M. F., & Zacks, S. (1967). On mixtures of  $\chi^2$ - and  $F$ -distributions which yield distributions of the same family. *Annals of the Institute of Statistical Mathematics*, 19, 527–536.

Perlman, M. D., & Rasmussen, U. A. (1975). Some remarks on estimating a noncentrality parameter. *Communications in Statistics*, 4, 455–468.  
 Robbins, H., & Pitman, E. J. G. (1949). Application of the method of mixtures to quadratic forms in normal variates. *Annals of Mathematical Statistics*, 20, 552–560.  
 Siegel, A. F. (1979). Non-central chi-squared distribution with zero-degrees of freedom and testing for uniformity. *Biometrika*, 66, 381–386.  
 Teicher, H. (1960). On the mixtures of distributions. *Annals of Mathematical Statistics*, 31, 55–73.  
 Wijsman, R. A. (1959). Applications of a certain representation of the Wishart matrix. *Annals of Mathematical Statistics*, 30, 597–606.

**How to cite this article:** 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(1), e398. <https://doi.org/10.1002/sta4.398>

**APPENDIX A: PROOF OF (4)**

The joint density of independent  $G_{\gamma+j} > 0$  and  $0 < B_{\alpha, \gamma-\alpha} < 1$  is proportional to  $g^{\gamma+j-1} e^{-g} b^{\alpha-1} (1-b)^{\gamma-\alpha-1}$  so the density, at  $x > 0$ , of their product is proportional to

$$\begin{aligned} x^{\alpha-1} \int_x^\infty y^j (y-x)^{\gamma-\alpha-1} e^{-y} dy &= x^{\alpha-1} e^{-x} \int_0^\infty (w+x)^j w^{\gamma-\alpha-1} e^{-w} dw \\ &= x^{\alpha-1} e^{-x} \sum_{m=0}^j \binom{j}{m} x^m \int_0^\infty w^{\gamma-\alpha+j-m-1} e^{-w} dw \\ &= \sum_{m=0}^j \frac{x^{\alpha+m-1} e^{-x}}{\Gamma(\alpha+m)} \binom{j}{m} \Gamma(\alpha+m) \Gamma(\gamma-\alpha+j-m) \end{aligned}$$

which yields the required result since the p.m.f. of the specified beta-binomial distribution is proportional to  $\binom{j}{m} \Gamma(\alpha+m) \Gamma(\gamma-\alpha+j-m)$ .

**APPENDIX B: DIRECT PROOF OF COROLLARY 7**

The density, at  $x > 0$ , of  $Z$  when  $Z|I=i \sim \text{Gamma}(\alpha+i)$  and  $I \sim \text{NegBin}(\alpha+k, \psi)$  is given by

$$\frac{x^{\alpha-1} e^{-x/2}}{2^\alpha} \sum_{i=0}^\infty \frac{1}{\Gamma(\alpha+i)} \binom{x}{2}^i \frac{\Gamma(\alpha+k+i)}{\Gamma(\alpha+k)!} \psi^{\alpha+k} (1-\psi)^i = \frac{\psi^{\alpha+k} x^{\alpha-1} e^{-x/2}}{2^\alpha \Gamma(\alpha)} \sum_{i=0}^\infty \frac{(\alpha+k)_i z^i}{(\alpha)_i i!} = \frac{\psi^{\alpha+k} x^{\alpha-1} e^{-x/2}}{2^\alpha \Gamma(\alpha)} {}_1F_1(\alpha+k, \alpha, z).$$

Here,  $(\alpha)_i = \Gamma(\alpha+i)/\Gamma(\alpha)$ ,  $z = x(1-\psi)/2$  and  ${}_1F_1$  is the confluent hypergeometric function. On the other hand, the density, at  $x > 0$ , of  $Z'$  when  $\psi Z'|J=j \sim \text{Gamma}(\alpha+j)$  and  $J \sim \text{Bin}(k, 1-\psi)$  is given by

$$\left(\frac{\psi}{2}\right)^\alpha x^{\alpha-1} e^{-\psi x/2} \sum_{j=0}^k \frac{x^j}{\Gamma(\alpha+j)} \left(\frac{\psi}{2}\right)^j \binom{k}{j} (1-\psi)^j \psi^{k-j} = \frac{\psi^{\alpha+k} x^{\alpha-1} e^{-x/2}}{2^\alpha \Gamma(\alpha)} \sum_{j=0}^k \binom{k}{j} \frac{z^j}{(\alpha)_j}.$$

These two representations are the same because, for  $k = 0, 1, \dots$ ,

$${}_1F_1(\alpha+k, \alpha, z) = e^z \sum_{j=0}^k \binom{k}{j} \frac{z^j}{(\alpha)_j},$$

a fact that can be proved by induction using the standard contiguity relationship

$${}_1F_1(\alpha, \gamma, z) = {}_1F_1(\alpha-1, \gamma, z) + \frac{z}{\gamma} {}_1F_1(\alpha, \gamma+1, z).$$

;