On Non-central Beta distributions

Sulle distribuzioni Beta non centrali

Andrea Ongaro and Carlo Orsi

Abstract A new Non-central Beta distribution is defined. Several properties are derived (including various representations and moments expressions) both for the new and the standard Non-central Beta distribution, showing in many respects a greater tractability of the former.
Abstract *Nel presente articolo viene introdotta una nuova distribuzione Beta non centrale. Vengono derivate varie proprietà (in particolare varie rappresentazioni ed espressioni dei momenti) sia per la nuova Beta non centrale sia per quella standard, mostrando come sotto molti aspetti la prima appaia maggiormente trattabile.*

Key words: Beta distribution, non-centrality, moments, coniugacy, mixture representations, hypergeometric functions

1 Introduction and definitions

In this paper we shall introduce a new distribution (distr.) on $(0,1)$ which generalizes the Beta distr.. It can be viewed as an analogue of the Doubly Non-central Beta distr. (see [1]). We shall perform a preliminary investigation of the properties of such distr.. We shall also provide a novel parallel analysis of the Doubly Non-central Beta distr., which is based on the same techniques employed in the study of the new one. Such parallel analysis will allow a better understanding of the similarities and differences between the two distr.s.

Let us first introduce some notations and recall some useful facts.

Andrea Ongaro
Department of Economics, Management and Statistics, University of Milan-Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milan, e-mail: andrea.ongaro@unimib.it

Carlo Orsi
Department of Economics, Management and Statistics, University of Milan-Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milan, e-mail: c.orsi@campus.unimib.it

It is well known that if Y_i, $i=1,2$, are independent (ind.) χ^2 random variables (r.v.s) with $2\alpha_i > 0$ degrees of freedom (d.f.), denoted by $\chi^2_{2\alpha_i}$, then the r.v.:

$$X = \frac{Y_1}{Y_1+Y_2} \tag{1}$$

has a Beta distr. with shape parameters (s.p.s) α_i, denoted by Beta(α_1,α_2). Its density is:

$$\mathrm{Beta}\left(x;\alpha_1,\alpha_2\right) = \frac{x^{\alpha_1-1}\left(1-x\right)^{\alpha_2-1}}{B\left(\alpha_1,\alpha_2\right)}, \quad 0<x<1\ . \tag{2}$$

The Beta distr. can also be obtained as conditional distr. of X given the sum $Y^+ = Y_1+Y_2$, as a consequence of the independence of X and Y^+. The latter is a characterizing property of (ind.) Gamma r.v.s.

Now, let $\chi'^2_g(\lambda)$ be a Non-central χ^2 distr. with $g>0$ d.f. and non-centrality (n.c.) parameter $\lambda \geq 0$. In the following, we shall rely on the representation of the $\chi'^2_g(\lambda)$ distr. as mixture of central χ^2 distr.s, which can be easily derived from its characteristic function. More specifically, let M be a Poisson r.v. with mean $\lambda/2$, $\lambda \geq 0$ (the case $\lambda=0$ corresponding to a r.v. degenerate at zero). Then Y' has a $\chi'^2_g(\lambda)$ distr., if Y', conditionally on M, has a χ^2_{g+2M} distr..

A r.v. is said to have a Doubly Non-central Beta distr. with s.p.s α_1,α_2 and n.c.p.s λ_1,λ_2, denoted by DNcB$\left(\alpha_1,\alpha_2,\lambda_1,\lambda_2\right)$, if it is distributed as $X' = Y'_1/(Y'_1+Y'_2)$ where Y'_i, $i=1,2$, are ind. $\chi'^2_{2\alpha_i}(\lambda_i)$ r.v.s (see [1]).

Our basic idea to introduce a new parametric family of distr.s on $(0,1)$ which extends the Beta one is to consider the conditional distr. of X' given $Y'_1+Y'_2$. Because of the above mentioned characterizing property of the Gamma r.v., such distr. must depend on $Y'_1+Y'_2$ and therefore must be different from the DNcB distr., unless $\lambda_1=\lambda_2=0$. The latter case correspond to the Beta distr..

The density of the new distr. (as well as of the standard one) can be derived by using the above mentioned mixture representation of the Non-central χ^2 distr.. Specifically, let M_i, $i=1,2$, be ind. Poisson r.v.s with mean $\lambda_i/2$. Then the ind. Non-central χ^2 r.v.s (Y'_1,Y'_2) can be obtained by asking that, conditionally on (M_1,M_2), Y'_i be ind. with distr. $\chi^2_{2\alpha_i+2M_i}$, $i=1,2$. By noting that, conditionally on (M_1,M_2), $X' = Y'_1/(Y'_1+Y'_2)$ and $Y'_1+Y'_2$ are ind., one has that X' given $(M_1,M_2,Y'_1+Y'_2)$ has a Beta$(\alpha_1+M_1,\alpha_2+M_2)$ distr.. A direct application of Bayes theorem then shows that, conditionally on $Y'_1+Y'_2=y$, (M_1,M_2) has a probability function given by:

$$\Pr\left(M_1=i,M_2=j\,|\,Y'_1+Y'_2=y\right) = \frac{1}{{}_0F_1\left(\alpha^+,y\lambda^+/4\right)}\,\frac{\left(\frac{y\lambda_1}{4}\right)^i\left(\frac{y\lambda_2}{4}\right)^j}{i!\,j!\,(\alpha^+)_{i+j}},\ i,j\in\mathrm{N}\cup\{0\}\ , \tag{3}$$

where $(a)_0=1$, $(a)_i = a\,(a+1)\ldots(a+i-1)$, $\forall i\in\mathrm{N}$, $\alpha_1+\alpha_2=\alpha^+$, $\lambda_1+\lambda_2=\lambda^+$ and:

$${}_0F_1\left(a;x\right) = \sum_{i=0}^{+\infty}\frac{1}{(a)_i}\frac{x^i}{i!}, \qquad a>0,\ x\geq 0 \tag{4}$$

is the generalized hypergeometric function ${}_pF_q$ with $p = 0$ and $q = 1$ coefficients respectively at numerator and denominator (see [2]).

As the conditioning value y in (3) can be incorporated in the n.c.p.s λ_i's, we are then led to the following definition of the new distr..

Definition 1. A r.v. is said to have a Conditional Doubly Non-central Beta distr. with s.p.s α_1, α_2 and n.c.p.s λ_1, λ_2, denoted by $\mathrm{CDNcB}(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$, if it is distributed as $\frac{Y_1'}{Y_1'+Y_2'}\,|Y_1' + Y_2' = 1$ where Y_i', $i = 1,2$, are ind. $\chi'^2_{2\alpha_i}(\lambda_i)$ r.v.s.

2 Representations and density plots

The above discussion directly leads to the following mixture representations.

Property 1 (Mixture representation). Let X' have a $\mathrm{DNcB}(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$ distr. and X'' have a $\mathrm{CDNcB}(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$ distr.. Then X' and X'' admit the following representation:

$$X' \,|\, (M_1 = i, M_2 = j) \sim X'' \,|\, (N_1 = i, N_2 = j) \sim \mathrm{Beta}(\alpha_1 + i, \alpha_2 + j) \ , \tag{5}$$

where M_i, $i = 1,2$, are ind. Poisson r.v.s with mean $\lambda_i/2$ and the probability function of (N_1, N_2) is given by (3) with $y = 1$.

Although the distr. of (N_1, N_2) is slightly more complicated than the one of (M_1, M_2), it can be easily handled and understood in terms of the r.v.s (N_1, N^+), with $N^+ = N_1 + N_2$. In fact, it is easy to see that $N_1 \,|\, N^+ \sim \mathrm{Binomial}(N^+, \lambda_1/\lambda^+)$. Interestingly, $N_1 \,|\, N^+$ thus shares the same distr. of $M_1 | M^+$, with $M^+ = M_1 + M_2$. Furthermore, N^+ has the following probability function:

$$\Pr\left(N^+ = i\right) = \frac{1}{{}_0F_1\left(\alpha^+; \lambda^+/4\right)\,(\alpha^+)_i\; i!}\left(\frac{\lambda^+}{4}\right)^i, \quad i \in \mathrm{N} \cup \{0\} \ . \tag{6}$$

This allows us to derive expressions for the moments of X' and X'' (see Sect. 3) and for the covariance between N_1 and N_2 as follows:

$$\begin{aligned} \mathrm{Cov}(N_1, N_2) &= \mathrm{E}_{N^+}\left\{\mathrm{Cov}\left[(N_1, N_2)\,|N^+\right]\right\} + \mathrm{Cov}\left[\mathrm{E}\left(N_1|N^+\right), \mathrm{E}\left(N_2|N^+\right)\right] = \\ &= \frac{\lambda_1\,\lambda_2}{(\lambda^+)^2}\left[\mathrm{Var}\left(N^+\right) - \mathrm{E}\left(N^+\right)\right] \ , \end{aligned} \tag{7}$$

which can be easily expressed in terms of the hypergeometric function ${}_0F_1$. An extensive numerical investigation shows that such covariance is negative.

It is also simple to check that the CDNcB admits an unconditional "ratio" type representation.

Property 2 (Ratio representation). Let (N_1, N_2) be distributed as in Property 1. Furthermore, let (Z_1, Z_2) be conditionally ind. r.v.s given (N_1, N_2), with $Z_i \,|\, (N_1, N_2) \sim \chi^2_{2(\alpha_i + N_i)}$, $i = 1,2$. Then $\frac{Z_1}{Z_1+Z_2} \sim \mathrm{CDNcB}(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$.

Notice that (Z_1, Z_2) are negatively correlated as $\mathrm{Cov}(Z_1, Z_2) = 4\,\mathrm{Cov}(N_1, N_2)$.

To fully understand the behavior of the Non-central Beta distr.s it is instructive to realize how they perturbate the Beta distr..

Property 3 (Perturbation representation). Let $X' \sim \mathrm{DNcB}(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$ and $X'' \sim \mathrm{CDNcB}(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$. Then the densities $f_{X'}$ of X' and $f_{X''}$ of X'' can be written as:

$$f_{X'}(x) = \mathrm{Beta}(x; \alpha_1, \alpha_2)\, \mathrm{e}^{-\frac{\lambda^+}{2}}\, \Psi_2 \left[\alpha^+; \alpha_1, \alpha_2; \frac{\lambda_1 x}{2}, \frac{\lambda_2 (1-x)}{2}\right], \tag{8}$$

Ψ_2 being the Humbert's (two variables, double series) confluent hypergeometric function (see [2]) and

$$f_{X''}(x) = \mathrm{Beta}(x; \alpha_1, \alpha_2) \frac{{}_0F_1\left(\alpha_1; \frac{\lambda_1 x}{4}\right) {}_0F_1\left[\alpha_2; \frac{\lambda_2(1-x)}{4}\right]}{{}_0F_1\left(\alpha^+; \frac{\lambda^+}{4}\right)}. \tag{9}$$

Such representations highlight the greater interpretability and tractability of the CDNcB distr. with respect to the DNcB. The perturbing factor of the Beta density in the CDNcB distr., unlike the DNcB one, has a product form involving only the ${}_0F_1$ function (and therefore a single series) in a completely symmetric fashion. Furthermore, ${}_0F_1(a; x)$ has a very simple behavior and is implemented in common statistical packages. For any fixed $a > 0$, it is increasing (with values in $[1, +\infty)$) and convex. For fixed $x \geq 0$, ${}_0F_1(a; x)$ and its first derivative with respect to x are decreasing in a. See plots in Fig. 1.

It follows that ${}_0F_1\left(\alpha_1; \frac{\lambda_1 x}{4}\right)$ in (9) has the effect of giving more weight to the right tail of the Beta density through the scale p. λ_1 and the s.p. α_1 determining the rate of increase of the function. Perfectly symmetric considerations hold for the other component ${}_0F_1\left[\alpha_2; \frac{\lambda_2(1-x)}{4}\right]$ of the perturbing factor.

Figures 2, 3 display the CDNcB density for selected values of the parameters. They show that it can assume a large variety of shapes far beyond the Beta model ones. The effect of the perturbing factor can be clearly seen when $\alpha_1 = \alpha_2 = 1$, because in this case the Beta density reduces to the Uniform one (see Fig. 3). In particular notice that the CDNcB density can take on arbitrary finite and positive limits at 0 and 1. On the contrary, Beta densities cannot display such behavior.

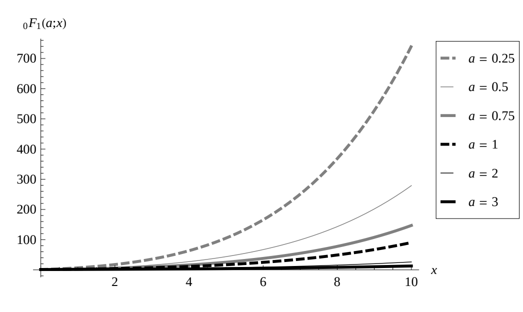

Fig. 1 Plots of ${}_0F_1(a; x)$ for $x \geq 0$ and selected values of $a > 0$

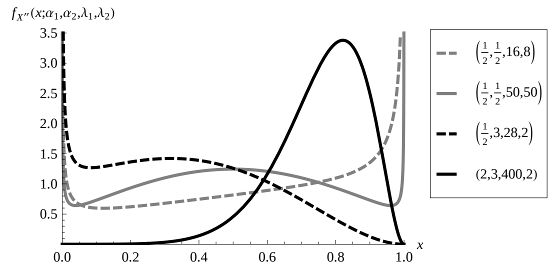

Fig. 2 Plots of the density of $X'' \sim \mathrm{CDNcB}\,(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$ for selected values of $\alpha_1, \alpha_2, \lambda_1, \lambda_2$

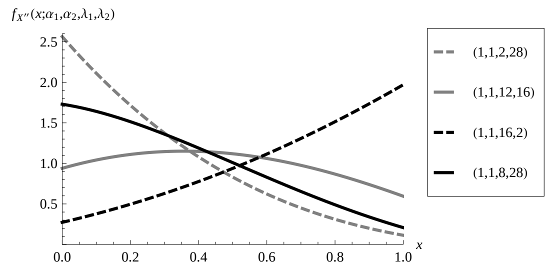

Fig. 3 Plots of the density of $X'' \sim \mathrm{CDNcB}\,(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$ for $\alpha_1 = \alpha_2 = 1$ and selected values of λ_1, λ_2

3 Moments

By exploiting the mixture representation in Property 1, it is possible to derive new general formulas for the moments of both Non-central Beta distr.s. Specifically, in the notation of Property 1, the moments of the CDNcB distr. can be written as:

$$\mathrm{E}\left[\left(X''\right)^r | N^+\right] = \mathrm{E}\left\{\mathrm{E}\left[\left(X''\right)^r \middle| N_1, N^+\right] \middle| N^+\right\} , \tag{10}$$

which then leads to

$$\mathrm{E}\left[\left(X''\right)^r | N^+\right] = \frac{1}{\left(\alpha^+ + N^+\right)_r} \sum_{i=0}^{N^+} \left(\alpha_1 + i\right)_r \binom{N^+}{i} \theta_1^i \left(1 - \theta_1\right)^{N^+ - i} , \tag{11}$$

where $\theta_1 = \lambda_1 / \lambda^+$. This reduces the computation of moments from a double series to a single one. In particular, by setting $r = 1$ and $r = 2$ in (11), we have:

$$\mathrm{E}\left(X''\right) = \mathrm{E}\left[\frac{\alpha_1 + \theta_1 N^+}{\alpha^+ + N^+}\right] , \tag{12}$$

$$\mathrm{E}\left[\left(X''\right)^2\right] = \mathrm{E}\left[\frac{\alpha_1\left(\alpha_1 + 1\right) + \left(2\alpha_1 + 2 - \theta_1\right)\theta_1 N^+ + \theta_1^2 N^{+2}}{\left(\alpha^+ + N^+\right)\left(\alpha^+ + 1 + N^+\right)}\right] . \tag{13}$$

Moments formulas (11), (12) and (13) hold for the DNcB as well, provided we simply replace N^+ by M^+.

Quite remarkably, after some manipulations, the first two moments of the CD-NcB distr. can be written in terms of the easily tractable ${}_0F_1$ function as follows:

$$\mathrm{E}\left(X''\right) = \frac{\alpha_1}{\alpha^+} \frac{{}_0F_1\left(\alpha^+ + 1; \frac{\lambda^+}{4}\right) + \frac{\theta_1 \frac{\lambda^+}{4} \times {}_0F_1\left(\alpha^+ + 2; \frac{\lambda^+}{4}\right)}{\alpha_1 (\alpha^+ + 1)}}{{}_0F_1\left(\alpha^+; \frac{\lambda^+}{4}\right)} , \tag{14}$$

$$\mathrm{E}\left[(X'')^2\right] = \frac{\alpha_1 (\alpha_1 + 1)}{\alpha^+ (\alpha^+ + 1) \; {}_0F_1\left(\alpha^+; \frac{\lambda^+}{4}\right)} \left[{}_0F_1\left(\alpha^+ + 2; \frac{\lambda^+}{4}\right) + \right.$$

$$\left. + \frac{2\theta_1 \frac{\lambda^+}{4} \times {}_0F_1\left(\alpha^+ + 3; \frac{\lambda^+}{4}\right)}{\alpha_1 (\alpha^+ + 2)} + \frac{\theta_1^2 \left(\frac{\lambda^+}{4}\right)^2 \times {}_0F_1\left(\alpha^+ + 4; \frac{\lambda^+}{4}\right)}{\alpha_1 (\alpha_1 + 1)(\alpha^+ + 2)(\alpha^+ + 3)} \right] . \tag{15}$$

4 Concluding remarks

We conclude by pointing to two possible developments relative to the new and in some respects more tractable Non-central Beta distr. here introduced.

In Bayesian analysis, the Beta distr. is often used as a (conjugate) prior for the Binomial model. Let us shortly analyze the behavior of the CDNcB in such context. If $X|P \sim \mathrm{Binomial}\,(n,P)$ and $P \sim \mathrm{CDNcB}\,(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$, then it is easy to see that the posterior density of P is proportional to:

$$\mathrm{Beta}(p; \gamma_1, \gamma_2) \; {}_0F_1\left(\alpha_1; \frac{\lambda_1 p}{4}\right) \; {}_0F_1\left[\alpha_2; \frac{\lambda_2 (1-p)}{4}\right] , \tag{16}$$

where $\gamma_1 = \alpha_1 + x$ and $\gamma_2 = \alpha_2 + n - x$. It follows that the six parameters density defined by (16) is a conjugate class for the binomial model. Although it is an extension of the CDNcB distr., it maintains the simple structure of a Beta density perturbed by the same type of product factor. In fact, the only difference with respect to the CDNcB distr. is the introduction of distinct parameters for the s.p.s of the Beta and the perturbing part. Therefore the extended distr. keeps a simple interpretation and it seems worthwhile to further investigate it.

A second promising line of research appears to be the extension of the CDNcB distr. to the multidimensional setting to obtain a more tractable Non-central Dirichlet distr.. Indeed, the method employed to construct the CDNcB distr. naturally extends to such setting.

References

1. Johnson, N.L., Kotz, S., Balakrishnan, N.: Continuous Univariate Distributions Vol. 2, 2nd edition, pp. 502–503. John Wiley & Sons, New York (1995)
2. Srivastava, H.M., Karlsson, W.: Multiple Gaussian Hypergeometric Series, pp. 19–20, 25–26. Ellis Horwood, Chichester (1985)