Approximations for the likelihood ratio statistic for hypothesis testing between two Beta distributions

Filipe Marques, *NOVA University of Lisbon, Lisbon, Portugal.*
Email: fjm@fct.unl.pt
Frank Coolen, *Durham University, Durham, UK.*
Email: frank.coolen@durham.ac.uk
Tahani Coolen-Maturi, *Durham University, Durham, UK.*
Email: tahani.maturi@durham.ac.uk

Abstract

In this paper the likelihood ratio to test between two Beta distributions is addressed. The exact distribution of the likelihood ratio statistic, for simple hypotheses, is obtained in terms of Gamma or Generalized Integer Gamma distributions, when the first or the second of the two parameters of the Beta distributions are equal and integers. In the remaining cases addressed, near-exact or asymptotic approximations, are developed for the likelihood ratio statistic. Both the exact, asymptotic or near-exact representations are obtained using a logarithm transformation of the likelihood ratio statistic and by working with the corresponding characteristic function. The numerical studies illustrate the precision of the approximations developed. Simulations are developed to analyse the power and the reproducibility probability of the tests.

AMS Subject Classification: 62F03,62E15,62E20,62G99

Keywords: Likelihood ratio tests, Generalized Integer Gamma distribution, Generalized Near-Integer Gamma distribution, Mixtures, Reproducibility probability, Nonparametric predictive inference

---

1 Introduction

In this work we consider the likelihood ratio to test between two completely specified Beta distributions. The Beta distribution is an important tool for many statistical problems with applications in different areas. As examples we point out the following features; i) the i-th order statistic of a sample of size n, extracted from a continuous uniform distribution, has a Beta distribution, ii) in Bayesian statistics it is commonly used as conjugate prior for binomial and geometric random variables, iii) it is the distribution of Wilks's lambda in some particular cases and iv) in the general case, the Wilks's lambda distribution can be related with the product of independent Beta random variables. The test addressed in this work, is a useful procedure in the decision making between two Beta distributions, thus may be helpful in problems arising from the previous examples, such as in the Bayesian framework in the selection, between two Beta models, of the conjugate prior distribution of the probability of success of a Binomial or Geometric distribution. Some possible applications of this Bayesian approch may be found in biological assays, medicine (Gupta and Nadarajah, 2004; Griffiths, 1973; Pham et al., 2010), in social sciences (Wiley et al., 2015) and in financial problems (Rachev et al., 2008). We say that a random variable X has a Beta distribution with parameters $a > 0$ and $b > 0$, and we denote this fact by $X \sim \mathrm{Beta}(a, b)$, if its density function is given by

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, \text{ with } 0 \leq x \leq 1$$

where $B(.,.)$ denotes the usual beta function. For a sample of size n, $X_1, \ldots, X_n$, we consider the following simple hypotheses

$$H_0 : X_i \sim \mathrm{Beta}(a,b) \text{ vs } H_1 : X_i \sim \mathrm{Beta}(c,d) \tag{1}$$

and we study the following cases: I) $b = d = 1$ or $a = c = 1$, II) $b = d = \alpha$ or $a = c = \alpha$ with $\alpha \in \mathbb{N}$, III) $a = c = r$ or $b = d = r$ with $r \in \mathbb{R}\backslash\mathbb{N}$, IV) $a - c \in \mathbb{N}$ and $b - d \in \mathbb{N}$, and finally V) the general case with no restrictions on the parameters. In cases I and II the exact distribution is presented in terms of Gamma or Generalized Integer Gamma (GIG) distributions (Coelho, 1998). In case III, near-exact approximations (Coelho, 2004) are developed for the likelihood ratio statistic. These near-exact approximations are developed by working with the characteristic function of the logarithm of the likelihood ratio statistic. More precisely, first an adequate factorization of the expression of the characteristic function is obtained, and then one of the factors is approximated in such way that the resulting characteristic function corresponds to a known and manageable distribution. In cases IV and V, asymptotic approximations, based on shifted mixtures of (positive or negative) Gamma distributions are obtained for the likelihood ratio statistic. These approximations allow the computation of p-values and quantiles in a fast and precise way. Thus, using the approximations developed in Subsections 2.3-2.5, other more complex and time consuming techniques may be avoided to determine quantiles, such as numerical inversion formulas together with bisection methods or simulations. We should also point out that Wilks theorem (Wilks, 1983) which states that the distribution of the logarithm of likelihood ratio test statistics used to test composite hypotheses can be approximated by a χ^2 distribution can not be used when both hypotheses are completely specified, since in this case the number of degrees of freedom under the alternative and null hypotheses is equal to zero. This reinforces the importance of having approximations which may allow to perform these tests with the appropriate accuracy.

This work is organized as follows. In Section 2, five cases are considered; four with conditions on the parameters of the Beta distributions involved, and one with no restrictions on the parameters. In Section 3, the power and reproducibility of the test are illustrated with simulation studies. Section 4 is dedicated to the concluding remarks.

2 The likelihood ratio statistic and its distribution

For a random sample of size n, $X_1, \ldots, X_n$, we are interested in a test which may allow two decide between two Beta distributions with the parameters completely specified. The hypotheses of interest are defined in (1).

The likelihood ratio statistic is given by

$$\Lambda = \prod_{i=1}^{n} \frac{f_0(x_i)}{f_1(x_i)}$$

where $f_0(.)$ and $f_1(.)$ are the probability density functions of two Beta random variables with distributions $\mathrm{Beta}(a,b)$ and $\mathrm{Beta}(c,d)$. Thus, the likelihood ratio may be written as

$$\Lambda = \prod_{i=1}^{n} \frac{B(c,d)x_i^{a-1}(1-x_i)^{b-1}}{B(a,b)x_i^{c-1}(1-x_i)^{d-1}} = \left(\frac{B(c,d)}{B(a,b)}\right)^n \prod_{i=1}^{n} x_i^{a-c}(1-x_i)^{b-d}\,.$$

To study the null distribution of Λ, under H_0 in (1), we are essentially interested in the distribution of $\prod_{i=1}^{n} X_i^{a-c}(1-X_i)^{b-d}$, for $X_1, \ldots, X_n$ independent and identically distributed as $X_i \sim \mathrm{Beta}(a,b)$.

2.1 Case I: $b = d = 1$ or $a = c = 1$

For $b = d = 1$ we are interested in testing

$$H_0 : X_i \sim \mathrm{Beta}(a,1) \;\text{ vs }\; H_1 : X_i \sim \mathrm{Beta}(c,1)\,.$$

The expression of Λ, for an observed sample of size n, is given by

$$\Lambda = \prod_{i=1}^{n} \frac{a}{c}\, x_i^{a-c} = \left(\frac{a}{c}\right)^n \prod_{i=1}^{n} x_i^{a-c}\,.$$

This case was already addressed in Marques et al. (2018), however for completeness of this work we present it here with more detail, for example, now we specify the cases $a - c > 0$ and $c - a > 0$.

Theorem 2.1 *If $X_1, \ldots, X_n$ are independent and identically distributed with $X_i \sim$ $\mathrm{Beta}(a,1)$ then the cumulative distribution function of*

$$\Lambda = \left(\frac{a}{c}\right)^n \prod_{i=1}^{n} X_i^{a-c}\,,$$

with $a, c > 0$, is
i) for $a - c > 0$

$$1 - F_{\Gamma\left(n, \frac{a}{a-c}\right)}\left(-\log\left(\frac{x}{(a/c)^n}\right)\right) \tag{2}$$

ii) for $a - c < 0$

$$F_{\Gamma\left(n, -\frac{a}{a-c}\right)}\left(\log\left(\frac{x}{(a/c)^n}\right)\right) \tag{3}$$

where $F_{\Gamma(r,\lambda)}$ is the cumulative distribution function of a Gamma distribution with shape parameter $r > 0$ and rate parameter $\lambda > 0$.

Proof Let us consider $X_1, \ldots, X_n$ independent and identically distributed with $X_i \sim \mathrm{Beta}(a, 1)$ and the random variable $W = -\log(\prod_{i=1}^n X_i^{a-c}) = \sum_{i=1}^n -(a-c)\log(X_i)$. Since we know that $-\log(X_i)$ has an exponential distribution with parameter a, the h-th moment of X_i^{a-c} is given by

$$E\left[X_i^{(a-c)h}\right] = \frac{a}{a + h(a-c)}$$

and, given the relation $E\left[e^{\mathrm{i}tW}\right] = E\left[\Lambda^{-\mathrm{i}t}\right]$, the expression of the characteristic function of $-(a-c)\log(X_i)$ is given by

$$\Phi_{-(a-c)\log(X_i)}(t) = \frac{a}{a - \mathrm{i}t(a-c)}\,.$$

If $a - c > 0$, then, we may say that the characteristic function of W is given by

$$\Phi_W(t) = \left(\frac{\frac{a}{a-c}}{\frac{a}{a-c} - \mathrm{i}t}\right)^n\,.$$

This is the characteristic function of a Gamma distribution with shape parameter n and rate parameter $\frac{a}{a-c}$. Therefore it is easy to show, with the necessary transformations, that the cumulative distribution function of Λ when $a - c > 0$ is given by

$$1 - F_{\Gamma\left(n, \frac{a}{a-c}\right)}\left(-\log\left(\frac{x}{(a/c)^n}\right)\right)\,.$$

Following a similar procedure it is possible to obtain the result stated for $a - c < 0$, we just have to note that, for $a - c < 0$

$$\Phi_W(t) = \left(\frac{\frac{a}{c-a}}{\frac{a}{c-a} + \mathrm{i}t}\right)^n\,.$$

This is the characteristic function of a negative Gamma random variable. Again, after simple transformations we obtain the expression

$$F_{\Gamma\left(n, -\frac{a}{a-c}\right)}\left(\log\left(\frac{x}{(a/c)^n}\right)\right)\,.$$

$\square$

Please note that:

1. if one considers the case $a = c = 1$, and $b - d > 0$ or $d - b > 0$

$$H_0 : X_i \sim \mathrm{Beta}(1, b) \text{ vs } H_1 : X_i \sim \mathrm{Beta}(1, d)$$

   the expression of Λ, for a sample of size n, is given by

$$\Lambda = \prod_{i=1}^{n} \frac{b}{d} (1 - x_i)^{b-d},$$

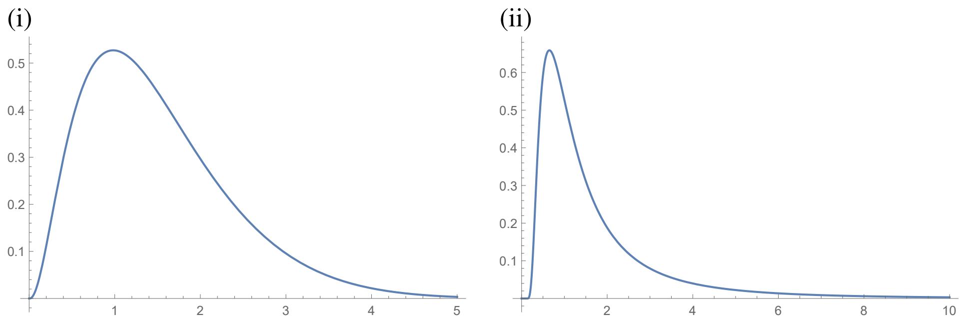
   and using the mirror property of the beta distribution if one takes $X_1, \ldots, X_n$ independent and identically distributed from $\mathrm{Beta}(1, b)$ we know that $1 - X_i \sim \mathrm{Beta}(b, 1)$, and thus this case is the same as the previous one;

2. clearly, the result in Theorem 2.1, is obtained under the null hypothesis, however the distribution of the likelihood ratio statistic under the alternative hypothesis is obtained following the same procedure.

These last notes also apply to the following cases considered.

Just as an illustration we present, in Figure 1, plots of the density functions corresponding to the distributions derived in Theorem 2.1, for two different scenarios: (i) $a = \frac{1}{4}$, $c = \frac{1}{5}$ and $b = d = 1$ and (ii) $a = 4$, $c = 5$ and $b = d = 1$.

Figure 1: Plots of the probability density functions, for the first case, in the following scenarios: (i) $a = \frac{1}{4}$ and $c = \frac{1}{5}$, (ii) $a = 4$ and $c = 5$



2.2 Case II: $b = d = \alpha$ or $a = c = \alpha$ with $\alpha \in \mathbb{N}$

For $b = d = \alpha$, with $\alpha \in \mathbb{N}$, we consider the hypotheses

$$H_0 : X_i \sim \mathrm{Beta}(a, \alpha) \text{ vs } H_1 : X_i \sim \mathrm{Beta}(c, \alpha).$$

The expression of Λ, for an observed sample of size n, is given by

$$\Lambda = \left(\frac{B(c,\alpha)}{B(a,\alpha)}\right)^n \prod_{i=1}^{n} x_i^{a-c} \,.$$

Although the expression of the likelihood ratio statistic is similar to the one in Subsection 2.1, since the underlying populations may be different (if $\alpha > 1$), the distribution of Λ, under the null hypothesis, will also be different.

Theorem 2.2 *If $X_1, \ldots, X_n$ are independent and identically distributed with $X_i \sim \mathrm{Beta}(a,\alpha)$, with $a > 0$ and $\alpha \in \mathbb{N}$, the cumulative distribution function of*

$$\Lambda = \left(\frac{B(c,\alpha)}{B(a,\alpha)}\right)^n \prod_{i=1}^{n} X_i^{a-c}$$

with $c > 0$, is (using the notation in Appendix 1 of Marques et al. (2015) for the GIG distribution)
i) for $a - c > 0$

$$1 - F_{\mathrm{GIG}}\left(-\log\left(\frac{x}{\left(\frac{B(c,\alpha)}{B(a,\alpha)}\right)^n}\right); \underline{n}, \underline{v}, \alpha\right)$$

with

$$\underline{n} = \{n, \ldots, n\}_{1\times\alpha}\,,\ \underline{v} = \left\{\frac{a+0}{a-c}, \ldots, \frac{a+\alpha-1}{a-c}\right\}_{1\times\alpha} \tag{4}$$

ii) for $a - c < 0$

$$F_{\mathrm{GIG}}\left(\log\left(\frac{x}{\left(\frac{B(c,\alpha)}{B(a,\alpha)}\right)^n}\right); \underline{n}, -\underline{v}, \alpha\right)$$

where $F_{\mathrm{GIG}}(.)$ denotes the cumulative distribution function of a Generalized Integer Gamma (GIG) distribution (Coelho, 1998) with integer shape parameters $\underline{n}$ and rate parameters $\underline{v}$ given in (4).

Proof Similar to the proof of Theorem 2.1, we consider $X_1, \ldots, X_n$ independent and identically distributed random variables with $X_i \sim Beta(a,\alpha)$ and the random variable $W = -\log(\prod_{i=1}^{n} X_i^{a-c}) = \sum_{i=1}^{n} -(a-c)\log(X_i)$. It is known that the characteristic function of W is given by

$$\Phi_W(t) = \prod_{j=1}^{n} \frac{\Gamma(a+\alpha)}{\Gamma(a)} \frac{\Gamma(a-(a-c)\mathrm{i}t)}{\Gamma(a+\alpha-(a-c)\mathrm{i}t)} = \left(\frac{\Gamma(a+\alpha)}{\Gamma(a)} \frac{\Gamma(a-(a-c)\mathrm{i}t)}{\Gamma(a+\alpha-(a-c)\mathrm{i}t)}\right)^n \,.$$

Since $\alpha \in \mathbb{N}$ and using the following equality, for $z \in \mathbb{C}$

$$\frac{\Gamma(z+\alpha)}{\Gamma(z)} = \prod_{k=0}^{\alpha-1} z + k \tag{5}$$

we may write

$$\Phi_W(t) = \left(\prod_{k=0}^{\alpha-1}(a+k)(a+k-(a-c)\mathrm{i}t)^{-1}\right)^n = \prod_{k=0}^{\alpha-1}\left(\frac{\frac{a+k}{a-c}}{\frac{a+k}{a-c}-\mathrm{i}t}\right)^n . \qquad (6)$$

When $a-c>0$, the characteristic function in expression (6) corresponds to the sum of α independent Gamma distributions, all with integer shape parameters, that is a GIG distribution with shape parameters given by $\underline{n}=\{n,...,n\}_{1\times\alpha}$ and rate parameters $\underline{v}=\left\{\frac{a+0}{a-c},\ldots,\frac{a+\alpha-1}{a-c}\right\}_{1\times\alpha}$. After the necessary transformations the cumulative distribution function of Λ is given by

$$1-F_{\mathrm{GIG}}\left(-\log\left(\frac{x}{\left(\frac{B(c,\alpha)}{B(a,\alpha)}\right)^n}\right);\underline{n},\underline{v},\alpha\right) .$$

When $a-c<0$, expression (6) can be written as

$$\Phi_W(t) = \prod_{k=0}^{\alpha-1}\left(\frac{\frac{a+k}{-a+c}}{\frac{a+k}{-a+c}+\mathrm{i}t}\right)^n$$

which is the characteristic function of a random variable Y such that $-Y$ has a GIG distribution with shape parameters given by $\underline{n}=\{n,...,n\}_{1\times\alpha}$ and rate parameters $-\underline{v}=\left\{\frac{a+0}{-a+c},\ldots,\frac{a+\alpha-1}{-a+c}\right\}_{1\times\alpha}$. In this case the cumulative distribution function of Λ is given by

$$F_{\mathrm{GIG}}\left(\log\left(\frac{x}{\left(\frac{B(c,\alpha)}{B(a,\alpha)}\right)^n}\right);\underline{n},-\underline{v},\alpha\right) .$$

□

In Figure 2 we present the probability density functions corresponding to the distributions in Theorem 2.2 for two scenarios: (i) $a=2$, $c=3/2$, $b=d=3$, (ii) $a=7/5$, $c=3$, $b=d=3$.

Results for the cases $a=c=\alpha$, $b-d>0$ or $b-d<0$ may be obtained in a similar way.

2.3 Case III: $b=d=r$ or $a=c=r$ with $r\in\mathbb{R}\backslash\mathbb{N}$

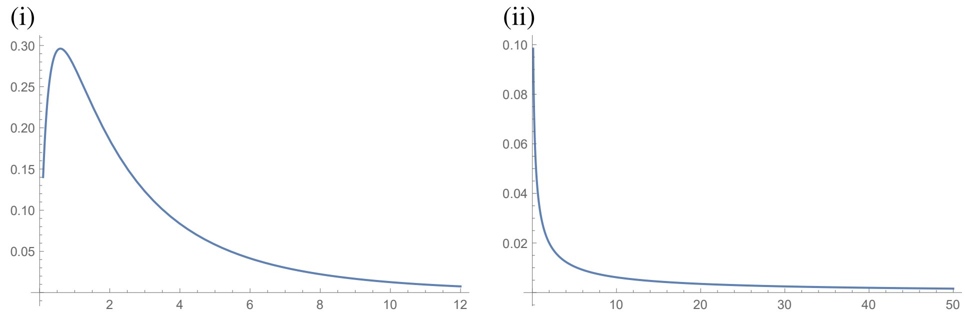
Similar to what was done in the previous section, we will just address one of the cases. For $b=d=r$ with $r\in\mathbb{R}\backslash\mathbb{N}$, we consider

$$H_0: X_i\sim \mathrm{Beta}(a,r) \;\text{ vs }\; H_1: X_i\sim\mathrm{Beta}(c,r)$$

the expression of Λ, for an observed sample of size n, is given by

$$\Lambda = \left(\frac{B(c,r)}{B(a,r)}\right)^n \prod_{i=1}^n x_i^{a-c} .$$

Figure 2: Plots of the probability density functions, for the second case, in the following scenarios: (i) $a = 2$, $c = 3/2$, $b = d = 3$, (ii) $a = 7/5$, $c = 3$, $b = d = 3$



In this case we do not have the exact cumulative distribution function of Λ in a manageable expression but we will show how it is possible to obtain precise approximations. We consider independent and identically distributed random variables $X_1, \ldots, X_n$ with $X_i \sim \mathrm{Beta}(a, r)$ and the random variable

$$W = -\log\left(\prod_{i=1}^{n} X_i^{a-c}\right) = \sum_{i=1}^{n} -(a-c)\log(X_i)\,.$$

The characteristic function of W is given by

$$\Phi_W(t) = \prod_{j=1}^{n} \frac{\Gamma(a+r)}{\Gamma(a)} \frac{\Gamma(a-(a-c)\mathrm{i}t)}{\Gamma(a+r-(a-c)\mathrm{i}t)}\,. \tag{7}$$

If $r > 1$ we may develop near-exact approximations for the distribution of the likelihood ratio statistic. These approximations, introduced by Coelho (2004), have already been used in several works involving the study of the distribution of likelihood ratio statistics used to test the structure of covariance matrices in the multivariate setting (Coelho et al., 2010; Coelho and Marques, 2012; Marques et al., 2017). The process may be illustrated as follows. The characteristic function in (7) may be factorized as follows

$$\Phi_W(t) = \left(\frac{\Gamma(a+r^\star)}{\Gamma(a)} \frac{\Gamma(a-(a-c)\mathrm{i}t)}{\Gamma(a+r^\star-(a-c)\mathrm{i}t)}\right)^n \left(\frac{\Gamma(a+r)}{\Gamma(a+r^\star)} \frac{\Gamma(a+r^\star-(a-c)\mathrm{i}t)}{\Gamma(a+r-(a-c)\mathrm{i}t)}\right)^n \tag{8}$$

with integer $r^\star = \lfloor r \rfloor$. Using the equality in (5) we may write

$$\begin{aligned}
\Phi_W(t) &= \left(\prod_{k=0}^{r^\star-1} \frac{a+k}{a+k-(a-c)\mathrm{i}t}\right)^n \left(\frac{\Gamma(a+r)}{\Gamma(a+r^\star)}\,\frac{\Gamma(a+r^\star-(a-c)\mathrm{i}t)}{\Gamma(a+r-(a-c)\mathrm{i}t)}\right)^n \\
&= \underbrace{\prod_{k=0}^{r^\star-1}\left(\frac{\frac{a+k}{a-c}}{\frac{a+k}{a-c}-\mathrm{i}t}\right)^n}_{\Phi_{W_1}(t)} \underbrace{\left(\frac{\Gamma(a+r)}{\Gamma(a+r^\star)}\,\frac{\Gamma(a+r^\star-(a-c)\mathrm{i}t)}{\Gamma(a+r-(a-c)\mathrm{i}t)}\right)^n}_{\Phi_{W_2}(t)} . \qquad (9)
\end{aligned}$$

The characteristic function Φ_{W_1} in (9) corresponds to the characteristic function of a GIG distribution with shape parameters given by $\underline{n} = \{n, ..., n\}_{1\times r^\star}$ and rate parameters $\underline{v} = \left\{\frac{a+0}{a-c}, \ldots, \frac{a+r^\star-1}{a-c}\right\}_{1\times r^\star}$. The characteristic function Φ_{W_2} in (9) corresponds to the sum of n independent Logbeta random variables, multiplied by $a-c$, with parameters $a+r^\star$ and $r-r^\star$. As a basis for the development of the near-exact approximations we consider the expansion for the ratio of Gamma functions given in expressions (11)–(14) of Tricomi and Erdélyi (1951) or in expression (12) of Luke (1969) which may be used to show that a Logbeta distribution may be represented as an infinite mixture of Gamma distributions. Thus, we propose as an approximation for the characteristic function Φ_{W_2} in (9) the characteristic function of a mixture of $m^\star+1$ Gamma distributions, all with rate parameter λ and with shape parameters $r+j$, $j = 0, \ldots, m^\star$ given by

$$\Phi_{W_2^\star}(t) = \sum_{j=0}^{m^\star} \pi_j\, \lambda^{s+j}(\lambda-\mathrm{i}t)^{-(s+j)} . \qquad (10)$$

Following the results in Coelho et al. (2010) we define s equal to the sum of the second parameters of the Logbeta distributions involved in the characteristic function of Φ_{W_2} in (9)

$$s = n(r-r^\star) . \qquad (11)$$

Then, the process has two main steps, first the parameter λ is determined as the rate parameter of a mixture of two Gamma distributions which equates the first 4 moments of the exact distribution of W_2, and second, assuming a fixed value for λ, the weights, π_j, are determined ensuring that the approximating distribution equates the first $m^\star$ exact moments, being thus the weights, π_j ($j = 0, \ldots, m^\star-1$), obtained as a solution of the system

$$\left.\frac{\partial^h}{\partial t^h}\,\Phi_{W_2}(t)\right|_{t=0} = \left.\frac{\partial^h}{\partial t^h}\,\Phi_{W_2^\star}(t)\right|_{t=0} , h = 1, \ldots, m^\star , \qquad (12)$$

with $\pi_{m^\star} = 1 - \sum_{j=0}^{m^\star-1}$. The resulting approximating characteristic function is given by

$$\Phi_{W^\star}(t) = \Phi_{W_1}(t) \times \Phi_{W_2^\star}(t)$$

and corresponds to a mixture Generalized Near-Integer Gamma (GNIG) distribution (Coelho, 2004) with weights π_j and with the GNIG parameters given by

$$\begin{aligned} \underline{n}_j &= \{n, ..., n, s+j\}_{1\times(r^\star+1)} \text{ with } j=0,\ldots,m^\star \text{ and } s \text{ in (11)} && (13) \\ \underline{v} &= \left\{\frac{a+0}{a-c}, \ldots, \frac{a+r^\star-1}{a-c}, \lambda\right\}_{1\times(r^\star+1)} . && (14) \end{aligned}$$

Using the notation in Appendix 1 of Marques et al. (2015), the corresponding approximating cumulative distribution function for Λ may be represented as

$$1-\sum_{j=0}^{m^\star} \pi_j F_{\text{GNIG}}\left(-\log\left(\frac{x}{\left(\frac{B(c,r)}{B(a,r)}\right)^n}\right), \underline{n}_j, \underline{v}, r^\star+1\right) . \tag{15}$$

If $a-c<0$, then the procedure to develop near-exact approximations for the distribution of the likelihood ratio statistic is similar to the previous one, but the approximating distributions will correspond to mixtures of negative GNIG distributions with weights π_j determined as solutions of the system of equations in (12) and the negative GNIG distributions with parameters given by $\underline{n}_j$ in (13) and $-\underline{v}$ with $\underline{v}$ given in (14).

The previous results may be summarized in the following theorem.

Theorem 2.3 *If $X_1, \ldots, X_n$ are independent and identically distributed with $X_i \sim \mathrm{Beta}(a,r)$, $a>0$, $r \in \mathbb{R}\backslash\mathbb{N}$ and $r>1$, and by approximating Φ_{W_2} in (9) by $\Phi_{W_2^\star}$ in (10) we obtain for*

$$\Lambda = \left(\frac{B(c,r)}{B(a,r)}\right)^n \prod_{i=1}^n X_i^{a-c}$$

with $c>0$, near-exact cumulative distribution functions given by
i) for $a-c>0$

$$1-\sum_{j=0}^{m^\star} \pi_j F_{\text{GNIG}}\left(-\log\left(\frac{x}{\left(\frac{B(c,r)}{B(a,r)}\right)^n}\right), \underline{n}_j, \underline{v}, r^\star+1\right)$$

and, ii) for $a-c<0$

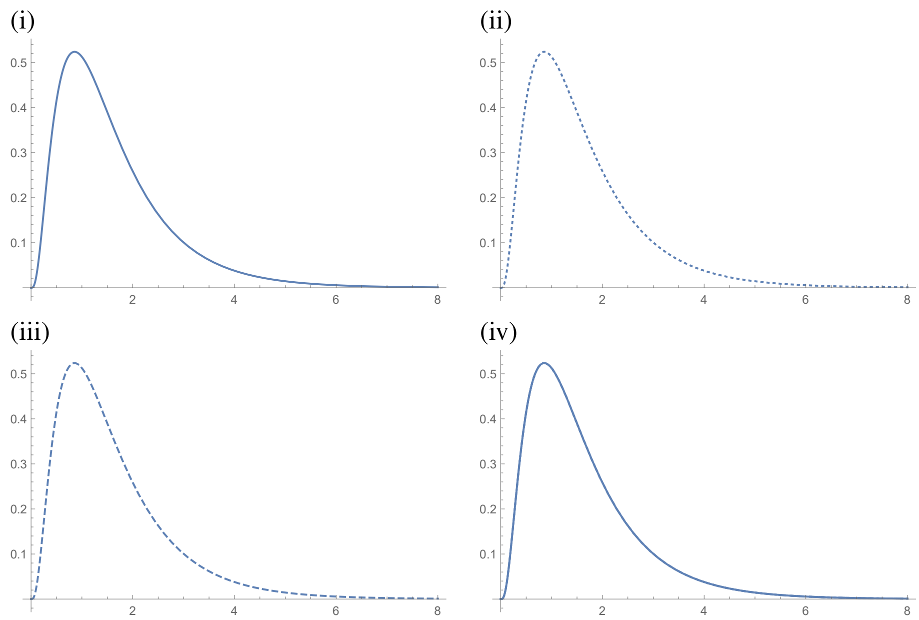
$$\sum_{j=0}^{m^\star} \pi_j F_{\text{GNIG}}\left(\log\left(\frac{x}{\left(\frac{B(c,r)}{B(a,r)}\right)^n}\right), \underline{n}_j, -\underline{v}, r^\star+1\right)$$

where $F_{\text{GNIG}}(.)$ denotes the cumulative distribution of a GNIG distribution (Coelho, 2004) with shape parameters $\underline{n}_j$ in (13) and rate parameters $\underline{v}$ in (14). The weights π_j are obtained as solution of the system of equations in (12).

If $r<1$, then the approximation is obtained using a similar approach but for the characteristic function Φ_W in (8), making $s=rn$ and λ as the rate parameter of a mixture of two Gamma distributions which equates the first 4 moments of the exact distribution of W and the weights π_j as the solution of the system of equations given in (12) replacing Φ_{W_2} by Φ_W. The resulting approximation is a simple mixture of Gamma distributions.

To illustrate the precision of these approximations, in Figure 3 we present for scenario $a = 17/5$, $r = 9/4$, $c = 3$ and $n = 20$ the plots for: (i) the exact probability density function, obtained using the inversion formulas in Gil-Pelaez (1951), (ii) the near-exact probability density function obtained for $m^\star = 2$, (iii) the near-exact probability density function obtained for $m^\star = 4$ and (iv) the representation in the same plot of (i), (ii) and (iii). The use of the inversion formulas in Gil-Pelaez (1951) is somehow limited. For example, if one wants to determine the exact quantiles of Λ we have to use these formulas together with the bisection method and this process may require a high computing time. Moreover, in the following subsections with more complex scenarios we were not able to plot the exact densities with the inversion formulas in Gil-Pelaez (1951). In Figure 3, the differences between the exact and approximating densities are indistinguishable even for small values of $m^\star$ such as 2 and 4.

Figure 3: Plots of the probability density functions, for Case III, scenario $a = 17/5$, $r = 9/4$, $c = 3$ and $n = 20$, of the (i) the exact probability density function, obtained using the inversion formulas in Gil-Pelaez (1951), (ii) the near-exact probability density function obtained for $m^\star = 2$, (iii) the near-exact probability density function obtained for $m^\star = 4$ and (iv) the representation in the same plot of (i), (ii) and (iii).



2.4 Case IV: $a - c \in \mathbb{N}$ and $b - d \in \mathbb{N}$

For $a - c \in \mathbb{N}$ and $b - d \in \mathbb{N}$, we consider the hypotheses

$$H_0 : X_i \sim \mathrm{Beta}(a, b) \text{ vs } H_1 : X_i \sim \mathrm{Beta}(c, d)$$

and the expression of Λ, for an observed sample of size n, is now given by

$$\Lambda = \prod_{i=1}^{n} \frac{B(c,d)}{B(a,b)}\, x_i^{a-c}(1 - x_i)^{b-d}\,.$$

This is a more complex case, for which we will only be able to derive asymptotic approximations for the distribution of Λ. The procedure is as follows. Let us consider independent and identically distributed $X_1, \ldots, X_n$ with $X_i \sim \mathrm{Beta}(a, b)$ and $a, b > 0$.

Just as a side note we point out that $X_i/(1 - X_i)$ has a beta prime distribution, so if $b - d < 0$ such that $b - d = -(a - c)$ we have as particular case

$$\begin{aligned} \Lambda &= \left(\frac{B(c,d)}{B(a,b)}\right)^n \prod_{i=1}^{n} X_i^{a-c}(1 - X_i)^{-(a-c)} \\ &= \left(\frac{B(c,d)}{B(a,b)}\right)^n \prod_{i=1}^{n} \left(\frac{X_i}{1 - X_i}\right)^{a-c} \end{aligned}$$

where we may identify the product of beta prime independent random variables to the power $a - c$.

Considering again the case addressed in this subsection, we have

$$\Lambda = \left(\frac{B(c,d)}{B(a,b)}\right)^n \prod_{i=1}^{n} X_i^{a-c}(1 - X_i)^{b-d}\,, \tag{16}$$

with $a - c \in \mathbb{N}$, $b - d \in \mathbb{N}$ and a, b, c and d positive real numbers. The characteristic function of $-\log\{X_i^{a-c}(1 - X_i)^{b-d}\}$ is given by

$$\frac{\Gamma(a - (a - c)\mathrm{i}t)\Gamma(b - (b - d)\mathrm{i}t)}{B(a, b)\Gamma(a + b - (a - c)\mathrm{i}t - (b - d)\mathrm{i}t)}$$

and as such the characteristic function of

$$W = -\log\left(\prod_{i=1}^{n} X_i^{a-c}(1 - X_i)^{b-d}\right) = \sum_{i=1}^{n} -\log\{X_i^{a-c}(1 - X_i)^{b-d}\}$$

is, in the general case, given by

$$\Phi_W(t) = \left(\frac{\Gamma(a - (a - c)\mathrm{i}t)\Gamma(b - (b - d)\mathrm{i}t)}{B(a, b)\Gamma(a + b - (a - c)\mathrm{i}t - (b - d)\mathrm{i}t)}\right)^n\,. \tag{17}$$

Given that in the particular case considered one has $a-c \in \mathbb{N}$ and $b-d \in \mathbb{N}$, after some technical developments of this last expression and using the Gauss multiplication formula which, for a positive integer η, is given by

$$\prod_{k=0}^{\eta-1} \Gamma\left(\frac{k}{\eta}+z\right) = (2\pi)^{\frac{\eta-1}{2}} \eta^{\frac{1}{2}-\eta z} \Gamma(\eta z)$$

we may write the characteristic function of W as

$$\Phi_W(t) = K_1 \mathrm{e}^{-\mathrm{i}tK_2} \frac{\left(\left\{\prod_{k=0}^{a-c-1} \Gamma\left(\frac{a}{a-c}+\frac{k}{a-c}-\mathrm{i}t\right)\right\}\left\{\prod_{k=0}^{b-d-1} \Gamma\left(\frac{b}{b-d}+\frac{k}{b-d}-\mathrm{i}t\right)\right\}\right)^n}{\left(\prod_{k=0}^{a+b-c-d-1} \Gamma\left(\frac{a+b}{a+b-c-d}+\frac{k}{a+b-c-d}-\mathrm{i}t\right)\right)^n} \tag{18}$$

with the constants K_1 and K_2 given by

$$K_1 = \left(\frac{\sqrt{2\pi}(a-c)^{a-\frac{1}{2}}(b-d)^{b-\frac{1}{2}}(a+b-c-d)^{-a-b+\frac{1}{2}}}{B(a,b)}\right)^n$$

and

$$K_2 = n(-(a+b-c-d)\log(a+b-c-d)+(a-c)\log(a-c)+(b-d)\log(b-d)) .$$

The characteristic function in (18) may be written as

$$\begin{aligned}
\Phi_W(t) &= \mathrm{e}^{-\mathrm{i}tK_2} \\
&\quad \underbrace{\times K_1\left(\prod_{k=0}^{a-c-1} \frac{\Gamma\left(\frac{a}{a-c}+\frac{k}{a-c}-\mathrm{i}t\right)}{\Gamma\left(\frac{a+b}{a+b-c-d}+\frac{k}{a+b-c-d}-\mathrm{i}t\right)}\right)^n \times\left(\prod_{k=0}^{b-d-1} \frac{\Gamma\left(\frac{b}{b-d}+\frac{k}{b-d}-\mathrm{i}t\right)}{\Gamma\left(\frac{a+b}{a+b-c-d}+\frac{k+a-c}{a+b-c-d}-\mathrm{i}t\right)}\right)^n}_{\Phi_{W_1}(t)} \\
&= \mathrm{e}^{-\mathrm{i}tK_2}\Phi_{W_1}(t) .
\end{aligned} \tag{19}$$

We should note that, in expression (19), K_2 corresponds to a shift in the main distribution. In order to obtain approximations for the distribution of the likelihood ratio one will use a similar procedure to the one given in Coelho and Alberto (2012) and Marques et al. (2017). More precisely, we will approximate the characteristic function of W_1 in (19) by a simple mixture of Gamma distributions, all with rate parameter λ and with shape parameters $r+j$, $j=0,\dots,m^\star$. Thus, we obtain as an approximating characteristic function of Φ_W in (17) the characteristic function

$$\Phi_{W^\star}(t) = \mathrm{e}^{-\mathrm{i}tK_2} \sum_{j=0}^{m^\star} \pi_j\, \lambda^{r+j}(\lambda-\mathrm{i}t)^{-(r+j)} \tag{20}$$

where λ and the weights π_j are determined using a matching moments technique in two steps. First λ is determined as the solution of the following system of equations

$$\frac{\partial^h}{\partial t^h}\,\Phi_{W_1}(t)\bigg|_{t=0} = \frac{\partial^h}{\partial t^h}\,\left\{(\lambda)^{r_1}(\lambda-\mathrm{i}t)^{-r_1}\right\}\bigg|_{t=0} \tag{21}$$

for $h = 1, 2$, that is, λ is the rate parameter of a Gamma distribution which matches the first 2 moments of the exact distribution of W_1 and, following a same procedure as the one used in Subsection 2.3, r is defined as

$$\begin{aligned} r = n\Bigg(&\left\{\sum_{k=0}^{b-d-1}\frac{k}{b-d}+\frac{b}{b-d}-\left(\frac{a-c+k}{a-c+b-d}+\frac{a+b}{a-c+b-d}\right)\right\} \\ &+\left\{\sum_{k=0}^{a-c-1}\frac{k}{a-c}+\frac{a}{a-c}-\left(\frac{k}{a-c+b-d}+\frac{a+b}{a-c+b-d}\right)\right\}\Bigg)\,. \end{aligned} \tag{22}$$

Given the complexity of the distribution in this case and in order to improve the quality of the approximations, in some cases one will consider $r = r_1$ given as solution of the system in (21). Finally, assuming fixed values for λ and r, the weights, π_j, are determined as a solution of the system

$$\frac{\partial^h}{\partial t^h}\,\Phi_{W_1}(t)\bigg|_{t=0} = \frac{\partial^h}{\partial t^h}\,\sum_{j=0}^{m^\star}\pi_j\,\lambda^{r+j}(\lambda-\mathrm{i}t)^{-(r+j)}\Bigg|_{t=0}\,, h=1,\ldots,m^\star \tag{23}$$

with

$$\pi_{m^\star} = 1-\sum_{j=0}^{m^\star-1}\pi_j\,.$$

Thus, we have as approximating distributions of W, mixtures of shifted Gamma distributions which, by simple transformation, give rise to the following cumulative distribution function

$$1-\sum_{j=0}^{m^\star}\pi_j F_{\Gamma(r+j,\lambda)}\left(-\log\left(\frac{x}{\left(\frac{B(c,d)}{B(a,b)}\right)^n}\right)-K_2\right)\,. \tag{24}$$

This procedure is summarized in the following theorem.

Theorem 2.4 *If $X_1,\ldots,X_n$ are independent and identically distributed with $X_i \sim \mathrm{Beta}(a,b)$, $a,b>0$, by approximating Φ_W in (17) by $\Phi_{W^\star}$ in (20) we obtain for*

$$\Lambda = \left(\frac{B(c,d)}{B(a,b)}\right)^n\prod_{i=1}^n X_i^{a-c}(1-X_i)^{b-d}\,,$$

with $a-c\in\mathbb{N}$, $b-d\in\mathbb{N}$ and a, b, c and d positive real numbers, the following approximating cumulative distribution function

$$1-\sum_{j=0}^{m^\star}\pi_j F_{\Gamma(r+j,\lambda)}\left(-\log\left(\frac{x}{\left(\frac{B(c,d)}{B(a,b)}\right)^n}\right)-K_2\right)$$

where $F_{\Gamma(r+j,\lambda)}(.)$ denotes the cumulative distribution function of a Gamma distribution with shape parameter $r+j$ and rate parameter λ. The parameter λ is obtained as solution of the system in (21) and r is defined as in (22) or is set equal to r_1 which is obtained as solution of the system in (21). The weights π_j are obtained as solution of the system of equations in (23).

In Case IV, we were not able to plot the exact probability density functions using the inversion formulas in Gil-Pelaez (1951). Therefore, in order to illustrate the precision of these approximations we present in Table 1, the exact, 0.01, 0.05 and 0.1, quantiles of Λ computed using the inversion formulas in Gil-Pelaez (1951) and the bisection method, and the equal decimal places of the approximating quantiles, obtained using expression (24). In Table 1 we considered the following scenarios: $a = 17/5$, $b = 16/3$, $c = 12/5$, $d = 10/3$, $n = 10, 50, 100$ and $m^\star = 2, 6, 10$. We would like to point out that the computing time needed for the proposed approximations is nearly zero.

Table 1: Comparison between exact and approximating quantiles, Case IV

	0.025	0.05	0.1
exact quantile $n = 10$	0.2199569845289	0.33488111698353	0.5241324001102
$m^\star = 2$	0.2199	0.3348	0.5241
$m^\star = 6$	0.21995698	0.334881116	0.52413240
$m^\star = 10$	0.219956984	0.3348811170	0.5241324001
exact quantile $n = 50$	0.2042221741896	0.4348764209495	1.00458042559089
$m^\star = 2$	0.204	0.4348	1.004
$m^\star = 6$	0.204222174	0.43487642	1.004580425
$m^\star = 10$	0.204222174189	0.434876420949	1.00458042559
exact quantile $n = 100$	0.5410521524667	1.4815871653844	4.57678796920922
$m^\star = 2$	0.54	1.481	4.5767
$m^\star = 6$	0.541052152	1.481587165	4.576787969
$m^\star = 10$	0.541052152466	1.481587165384	4.576787969209

The cases where $c - a \in \mathbb{N}$ and $d - c \in \mathbb{N}$, or other possible combinations, may also be addressed using similar procedures, but in these cases we may have to consider mixtures of shifted negative gamma distributions.

2.5 General case

Finally, having as basis the procedure described in Subsection 2.4, we propose as an approximation for Λ, in the general case with no restrictions on the parameters, mixtures of shifted (positive or negative) Gamma distributions. One will approximate the characteristic function in (17) of $W = -\log(\Lambda)$ with Λ given in (16) with no restrictions on the parameters $a, b, c, d > 0$, by the characteristic function of a mixture of shifted (positive or negative) Gamma distributions with shape parameters $r + j$ ($j = 0, \ldots, m^\star$), rate parameter λ and shift parameter w, given by

$$\Phi_{W^\star}(t) = \sum_{j=0}^{m^\star} \pi_j \, \lambda^{r+j} (\lambda - \mathrm{i}t)^{-(r+j)} \mathrm{e}^{\mathrm{i}tw} \, . \tag{25}$$

The parameters r, λ and w will be determined as solutions of the system of equations

$$\frac{\partial^h}{\partial t^h}\,\Phi_W(t)\bigg|_{t=0} = \frac{\partial^h}{\partial t^h}\,\left\{(\lambda)^r(\lambda-\mathrm{i}t)^{-r}\mathrm{e}^{\mathrm{i}tw}\right\}\bigg|_{t=0} \tag{26}$$

for $h = 1, 2, 3$. One should note that, when $a - c < 0$ or $b - d < 0$ one will obtain a negative rate parameter, that is $\lambda < 0$, which means that in these cases one will have instead of a mixture of gamma distributions, $Y_j \sim \Gamma(r + j, \lambda)$, with a shift parameter equal to w, a mixture of $-Y_j \sim \Gamma(r + j, -\lambda)$ distributions with the same shift parameter. Through the system of equations in (26) we define the rate, shape and shift parameters of the gamma distributions involved, and then we move forward to determine the weights. The weights are determined for fixed values of r, λ and w, by matching a given number, let us say $m^\star$, of exact moments, that is by solving the system

$$\frac{\partial^h}{\partial t^h}\,\Phi_W(t)\bigg|_{t=0} = \frac{\partial^h}{\partial t^h}\,\Phi_{W^\star}(t)\bigg|_{t=0}\;, h = 1, \ldots, m^\star \tag{27}$$

with

$$\pi_{m^\star} = 1 - \sum_{j=0}^{m^\star-1} \pi_j$$

with $\Phi_{W^\star}$ in (25).

Theorem 2.5 *If $X_1, \ldots, X_n$ are independent and identically distributed with $X_i \sim \mathrm{Beta}(a, b)$, $a, b > 0$, by approximating the characteristic function in (17) of $W = -\log(\Lambda)$ by $\Phi_{W^\star}$ in (25) we obtain for*

$$\Lambda = \left(\frac{B(c,d)}{B(a,b)}\right)^n \prod_{i=1}^n X_i^{a-c}(1 - X_i)^{b-d}\,,$$

the following approximating cumulative distribution functions
i) $a - c > 0$ and $b - d > 0$

$$1 - \sum_{j=0}^{m} \pi_j F_{\Gamma(r+j,\lambda)}\left(-\log\left(\frac{x}{\left(\frac{B(c,d)}{B(a,b)}\right)^n}\right) - w\right) \tag{28}$$

and when $a - c < 0$ or $b - d < 0$

$$\sum_{j=0}^{m} \pi_j F_{\Gamma(r+j,-\lambda)}\left(\log\left(\frac{x}{\left(\frac{B(c,d)}{B(a,b)}\right)^n}\right) + w\right) \tag{29}$$

where $F_{\Gamma(r+j,\lambda)}(.)$ denotes the cumulative distribution function of a Gamma distribution with shape parameter $r + j$ and rate parameter λ. The parameters λ, r and w are obtained as solutions of the system in (26). The weights π_j are obtained as solution of the system of equations in (27).

In this case, as also happened in Subsection 2.4, we were not able to plot the exact probability density functions. In Table 2 we present the exact, 0.01, 0.05 and 0.1, quantiles of Λ computed using the inversion formulas in Gil-Pelaez (1951), and the equal decimal places of the approximating quantiles computed using expression (28). In this table we consider the following scenarios: $a = 22/5$, $b = 3$, $c = 10/3$, $d = 7/4$, $n = 10, 25, 50$ and $m^\star = 2, 6, 10$.

Table 2: Comparison between exact and approximating quantiles, Case V

	0.025	0.05	0.1
exact quantile $n = 10$	0.1856432083191	0.3166371783289	0.5623492446855
$m^\star = 2$	0.185	0.31	0.56
$m^\star = 6$	0.185	0.316	0.56
$m^\star = 10$	0.185	0.316	0.562
exact quantile $n = 50$	0.5214269462216	1.3885795976733	4.1295434205751
$m^\star = 2$	0.521	1.38	4.1
$m^\star = 6$	0.521	1.388	4.129
$m^\star = 10$	0.5214	1.3885	4.1295
exact quantile $n = 100$	7.7459658168541	28.799851109827	125.90314598279
$m^\star = 2$	7.74	28.8	126.
$m^\star = 6$	7.746	28.80	125.90
$m^\star = 10$	7.74596	28.79985	125.9031

The results in Table 2, when compared with the ones in Table 1, show that the approximation developed for the general case is not as precise as the one developed for Case IV. Even so, it is a very reasonable approximation equating, in most cases, two decimal places of the exact quantile in the scenario under consideration.

3 Numerical studies and simulations

In this section, the power and the reproducibility properties of the test are illustrated through simulations.

3.1 Power study

To illustrate the power of these tests we consider the same scenarios addressed in Subsections 2.1 and 2.3 which are

Case I - i) $a = 1/4$, $c = 1/5$ and $b = d = 1$ $(a > c)$

$$H_0 : X_i \sim \text{Beta}(1/4, 1) \text{ vs } H_1 : X_i \sim \text{Beta}(1/5, 1)$$

Case I - ii) $a = 4$, $c = 5$ and $b = d = 1$ $(a < c)$

$$H_0 : X_i \sim \text{Beta}(4, 1) \text{ vs } H_1 : X_i \sim \text{Beta}(5, 1)$$

Case III - $a = 17/5$, $c = 3$ and $b = d = r = 9/4$

$$H_0 : X_i \sim \text{Beta}(17/5, 9/4) \text{ vs } H_1 : X_i \sim \text{Beta}(3, 9/4)\,.$$

The $\alpha = 0.05$ quantile was computed using the expressions of the exact cumulative distribution functions in (2) or (3) for Case I, and the near-exact cumulative distribution function in (15) for Case III. To compute the empirical power we considered 100 000 replications of samples of size $n = 50, 100, 200$ and 500 from the following distributions

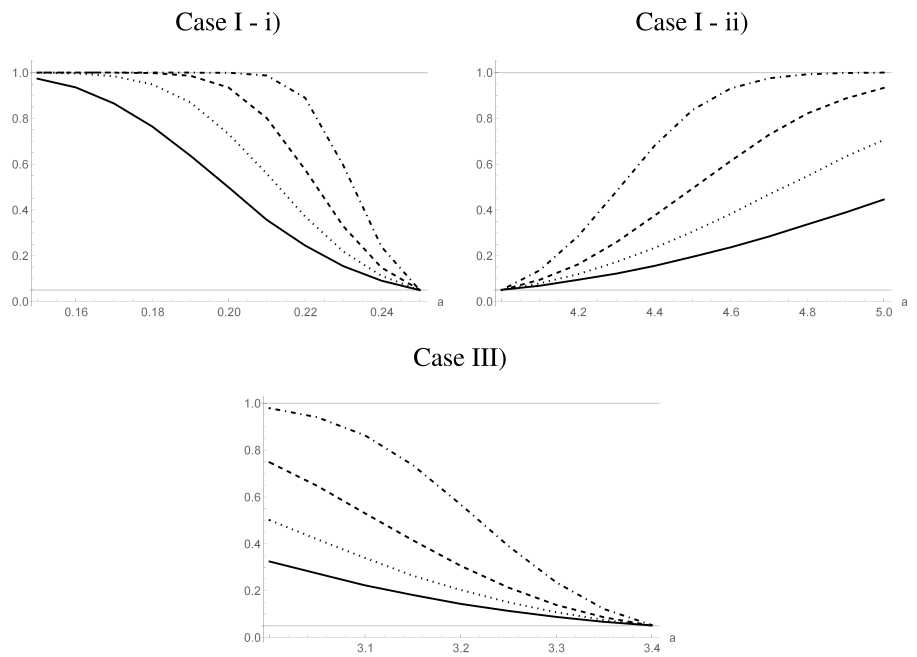
Case I - i) $X_i \sim \mathrm{Beta}(a, 1)$ with a from 0.15 to 0.25 step 0.01

Case II - ii) $X_i \sim \mathrm{Beta}(a, 1)$ with a from 4.0 to 5.0 step 0.1

Case III - $X_i \sim \mathrm{Beta}(a, 9/4)$ with a from 3.0 to 3.4 step 0.05 .

In Figure 4 it is possible to observe, as expected, the convergence of the power to 1 when a moves away from the value considered under the null hypothesis and also as a function of the sample size. These are known properties of likelihood ratio tests. In addition, we may say that the simulations point to an unbiased test since, under H_0, the simulated power is 0.05.

Figure 4: Power plots for different sample sizes. $n = 50$ (solid line), $n = 100$ (dotted line), $n = 200$ (dashed line) and $n = 500$ (dotted-dashed line)



3.2 Reproducibility probability

In this subsection we illustrate the reproducibility property of these likelihood ratio tests. The reproducibility probability (RP) of a test is the probability of making the same decision if a test were repeated under the same circumstances. This problem was first addressed by Goodman (1992) and has received, recently, increasing attention. The nonparametric predictive inference (NPI) for RP was first presented in Coolen and Bin Himd (2014) for two basic nonparametric tests, the one-sample sign test and the one-sample signed-rank test, and in Coolen and Alqifari (2018) RPs were computed for the quantile test and for a precedence test. We consider the NPI method to compute the lower and upper RPs of likelihood ratio tests for simple hypotheses introduced in Marques et al. (2018). The method can be summarized as follows. For n data observations $x_1 < x_2 < \cdots < x_n$ we may consider $n+1$ intervals (x_{i-1}, x_i), $i = 1, \ldots, n+1$. The values x_0 and x_{n+1} may be defined, for a distribution with support (0,1), as $x_0 = x_1/2$ and $x_{n+1} = (x_n + 1)/2$. We consider Hill's assumption (Hill, 1968; Arts et al, 2004) which assigns for a future real-valued observation, given the n data observations, probability $1/(n+1)$ to each open interval between consecutive data observations. Thus, for the m future observations, the $\binom{n+m}{m}$ different orderings of all these observations are all equally likely. For each ordering, we may count the number of future observations in each interval and compute, for the likelihood ratio, the minimum possible value, $\underline{LR}$, and the maximum possible value, $\overline{LR}$, and finally compute the NPI lower and upper RPs (for more details please see Marques et al. (2018)). In Marques et al. (2018), Section 4, the case considered in Section 2.1 of the present work was already addressed. Thus, in this subsection we consider more general set-ups, the cases IV and V in Sections 2.4 and 2.5. One considers the hypotheses

$$H_0 : X_i \sim \text{Beta}(a, b) \text{ vs } H_1 : X_i \sim \text{Beta}(c, d)$$

and the scenarios considered in Sections 2.4 and 2.5.

Case IV) $a = 17/5$, $b = 16/3$, $c = 12/5$ and $d = 10/3$

$$H_0 : X_i \sim \text{Beta}(17/5, 16/3) \text{ vs } H_1 : X_i \sim \text{Beta}(12/5, 10/3)$$

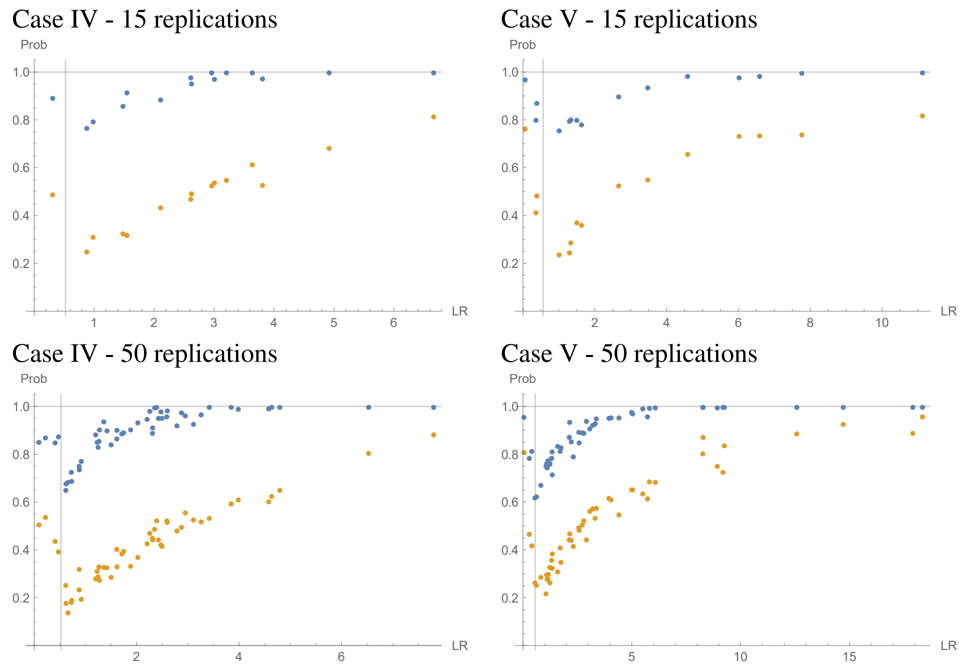
Case V) $a = 22/5$, $b = 3$, $c = 10/3$ and $d = 7/4$

$$H_0 : X_i \sim \text{Beta}(22/5, 3) \text{ vs } H_1 : X_i \sim \text{Beta}(10/3, 7/4)\,.$$

Using the results in Marques et al. (2018) and the 0.1 quantiles in Tables 1 and 2, we computed NPI lower and upper RPs. For Cases IV and V, we consider $n = 10$ and $m = n$ future observations, then we consider 15 and 50 replications simulated under H_0. In Figure 5, the blue dots are the upper RPs and the yellow dots are the lower RPs evaluated for each simulated value of the likelihood ratio statistic.

In Figure 5 the vertical line marks the value of the exact quantile and the horizontal line marks the value 1. From Figure 5 we may observe the same features already described in Marques et al. (2018) for Case I, which are: the upper and lower RPs tend to increase and to be closer to each other when the simulated value of the likelihood

Figure 5: Lower (yellow dots) and upper (blue dots) reproducibility probabilities



ratio statistic moves away from the quantile considered. When close to the quantile considered the lower RP is quite small which is a feature present in comparisons of two groups (Coolen and Bin Himd, 2014). As already mentioned the RP is a measure of how likely is to make the same decision if we repeat a test under the same circumstances. When the lower RP is high it indicates that we may have a reasonable security that if the test were repeated we would end up with the same decision with regard to rejection of the null hypothesis, thus ensuring the reproducibility of the test results. From Figure 5 we may observe that the lower RP only reaches values close, or equal, to 0.8 for quite distant values of the likelihood ratio statistics from the 0.1 quantile. This may suggest that in these cases the reproducibility of the test results is only guarantee for large values of the likelihood ratio.

4 Concluding remarks

In this paper we have studied the distribution of the likelihood ratio test statistic used to test between two Beta distributions. When two of the corresponding parameters of the two Beta distributions are equal and integers, representations of the exact distribution of the likelihood ratio statistic were obtained as transformations of a Gamma or of a GIG distribution. For the other three cases considered, near-exact or asymptotic approximations, were developed. Using the exact distributions or the approximations developed, quantiles and p-values can be computed in a fast and precise way. This way, other more complex and time consuming methods, based on numerical methods or simulations, may be avoided. Similar results for the distribution of the likelihood ratio under the alternative hypothesis may be easily obtained using similar procedures. The power of the test increases with the sample size and when the values of the parameters are considerably different from the ones assumed in the null hypothesis, these are the already expected behaviours for likelihood ratio tests. The lower and upper RPs show that only for distant values of the likelihood ratio from the fixed quantile we may ensure the reproducibility of the test results. The authors aim, in the future, to address the case where the decision making involves more than two Beta distributions, and also tests between other types of distributions such as two Beta type II distributions or two Kumaraswamy distributions.

Acknowledgements

The authors would like thank the two Reviewers for their careful reading and constructive comments. This work was partially supported by the Fundação para a Ciência e a Tecnologia (Portuguese Foundation for Science and Technology) through the project UID/MAT/00297/2013 (Centro de Matemática e Aplicações).

References

Arts G.R.J., Coolen F.P.A., van der Laan P., 2004. Nonparametric predictive inference in statistical process control. *Quality Technology and Quantitative Management*, 1, 201-216.

Coelho C.A., 1998. The Generalized Integer Gamma Distribution—A Basis for Distributions in Multivariate Statistics. *Journal of Multivariate Analysis* 64, 86–102.

Coelho C.A., 2004. The Generalized Near-Integer Gamma Distribution: A Basis for 'Near-Exact' Approximations to the Distribution of Statistics which are the Product of an Odd Number of Independent Beta Random Variables. *Journal of Multivariate Analysis*, 89, 191–218.

Coelho C.A. and Alberto R.P., 2012. On the Distribution of the Product of Independent Beta Random Variables Applications. *Technical Report, CMA*, 12.

Coelho C.A. and Marques F.J., 2012. Near-exact distributions for the likelihood ratio test statistic to test equality of several variance-covariance matrices in elliptically contoured distributions. *Computational Statistics*, 27, 627–659.

Coelho C.A., Arnold B.C. and Marques F.J., 2010. Near-exact distributions for certain likelihood ratio test statistics. *Journal of Statistical Theory and Practice*, 4, 711–725.

Coolen F.P.A., Alqifari H.N., 2018. Nonparametric predictive inference for reproducibility of two basic tests based on order statistics. *REVSTAT - Statistical Journal*, 16, 167-185

Coolen F.P.A., Bin Himd S., 2014. Nonparametric predictive inference for reproducibility of basic nonparametric tests. *Journal of Statistical Theory and Practice*, 8, 591-618.

Gil-Pelaez J., (1951). Note on the inversion theorem. *Biometrika*, 38, 481–482.

Goodman S.N., (1992). A comment on replication, p-values and evidence. *Statistics in Medicine*, 11, 875–879.

Griffiths, D. (1973). Maximum Likelihood Estimation for the Beta-Binomial Distribution and an Application to the Household Distribution of the Total Number of Cases of a Disease. *Biometrics*, 29, 637–648.

Gupta, A.K. and Nadarajah, S., (2004). *Handbook of Beta Distribution and Its Applications*. Statistics: A Series of Textbooks and Monographs, Taylor & Francis.

Hill B.M., 1968. Posterior distribution of percentiles: Bayes' theorem for sampling from a population. *Journal of the American Statistical Association*, 63, 677–691.

Luke Y.L., (1969). *The special functions and their approximations*. Academic Press, Inc., London.

Marques F.J., Coelho C.A. and de Carvalho M. 2015. On the distribution of linear combinations of independent Gumbel random variables. *Stat Comput*, 25, 683–701.

Marques, F.J., Coelho, C.A. and Rodrigues, P.C., 2017. Testing the equality of several linear regression models. *Computational Statistics*, in press

Marques F.J., Coolen F.P.A. and Coolen-Maturi T., 2018. Introducing nonparametric predictive inference methods for reproducibility of likelihood ratio tests. *Journal of Statistical Theory and Practice*, submitted for publication.

Pham, T.V., Piersma, S.R., Warmoes, M. and Jimenez, C.R. (2010). On the beta-binomial model for analysis of spectral count data in label-free tandem mass spectrometry-based proteomics. *Bioinformatics*, 26, 363-369.

Rachev, S.T., Hsu, J.S.J., Bagasheva, B.S. and Fabozzi, F.J. (2008). *Bayesian Methods in Finance*, Frank J. Fabozzi Series, John Wiley & Sons

Tricomi F.G. and Erdélyi A., 1951. The asymptotic expansion of a ratio of Gamma functions. *Pacific Journal of Mathematics* 1, 133–142.

Wiley, J.A., Martin, J.L., Herschkorn, S.J., and Bond, J. (2015). A New Extension of the Binomial Error Model for Responses to Items of Varying Difficulty in Educational Testing and Attitude Surveys. *PLoS ONE*, 10, e0141981.

Wilks S.S., 1983. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *The Annals of Mathematical Statistics*, 9, 60-62.