Anais da Academia Brasileira de Ciências 2019 91(2): e20180955
(Annals of the Brazilian Academy of Sciences)
Printed version ISSN 0001-3765 / Online version ISSN 1678-2690
http://dx.doi.org/10.1590/0001-3765201920180955
www.scielo.br/aabc | www.fb.com/aabcjournal

Inflated Kumaraswamy distributions

FRANCISCO CRIBARI-NETO and JÉSSICA SANTOS

Departamento de Estatística, Universidade Federal de Pernambuco,
Av. Prof Luiz Freire, s/n, Cidade Universitária, 50670-901 Recife, PE, Brazil

Manuscript received on September 12, 2018; accepted for publication on February 14, 2019

How to cite: CRIBARI-NETO F AND SANTOS J. 2019. Inflated Kumaraswamy distributions. An Acad Bras Cienc 91: e20180955. DOI 10.1590/10.1590/0001-3765201920180955.

Abstract: The Kumaraswamy distribution is useful for modeling variables whose support is the standard unit interval, i.e., $(0,1)$. It is not uncommon, however, for the data to contain zeros and/or ones. When that happens, the interest shifts to modeling variables that assume values in $[0,1)$, $(0,1]$ or $[0,1]$. Our goal in this paper is to introduce inflated Kumaraswamy distributions that can be used to that end. We consider inflation at one of the extremes of the standard unit interval and also the more challenging case in which inflation takes place at both interval endpoints. We introduce inflated Kumaraswamy distributions, discuss their main properties, show how to estimate their parameters (point and interval estimation) and explain how testing inferences can be performed. We also present Monte Carlo evidence on the finite sample performances of point estimation, confidence intervals and hypothesis tests. An empirical application is presented and discussed.

Key words: Inflated distribution, Kumaraswamy distribution, likelihood ratio test, maximum likelihood estimation, score test, Wald test.

INTRODUCTION

Oftentimes practitioners need to model variables that assume values in the standard unit interval, $(0,1)$, such as rates, proportions and concentration indices. The beta distribution is the most commonly used model in such applications, since its density can assume a wide range of shapes depending on the parameter values. Nonetheless, it was noted by Kumaraswamy (1976) that the beta law may fail to fit well hydrological data, especially when the data are hydrological observations of small frequency. He then proposed a new distribution, which can be considered as an alternative to the well known beta model. That distribution is now known as the *Kumaraswamy distribution*. We say that the random variable Y is Kumaraswamy-distributed with shape parameters $\alpha > 0$ and $\beta > 0$, denoted by $Y \sim \mathrm{Kum}(\alpha,\beta)$, if its probability density function (pdf) is given by

$$g(y;\alpha,\beta) = \alpha\beta y^{\alpha-1}(1-y^{\alpha})^{\beta-1}, \quad y \in (0,1), \tag{1}$$

Correspondence to: Francisco Cribari-Neto
E-mail: cribari@gmail.com
ORCid: https://orcid.org/0000-0002-5909-6698

the corresponding cumulative distribution function (cdf) being $G(y;\alpha,\beta) = 1-(1-y^{\alpha})^{\beta}$. We note that if $Y \sim \text{Kum}(\alpha,1)$, then $1-Y \sim \text{Kum}(1,\alpha)$ and $-\ln(Y)$ is exponentially distributed with parameter α; likewise, if $Y \sim \text{Kum}(1,\beta)$, then $1-Y \sim \text{Kum}(\beta,1)$ and $-\ln(1-Y)$ is exponentially distributed with parameter β.

The Kumaraswamy model has received considerable attention in the recent literature. Carrasco et al. (2010) proposed a new five-parameter distribution that generalizes the beta and Kumaraswamy distributions. Lemonte (2011) obtained nearly unbiased estimators for the parameters that index the Kumaraswamy law. A method for distinguishing between the Kumaraswamy and beta models was proposed by Silva and Barreto-Souza (2014). Barreto-Souza and Lemonte (2013) introduced a bivariate Kumaraswamy distribution for which the marginal distributions are univariate Kumaraswamy laws.

According to Mitnik and Baek (2013), the Kumaraswamy distribution has an advantage relative to beta model: its distribution and quantile functions can be expressed in closed form. That renders, for instance, random number generation based on the inversion method an easy task; see Jones (2009). It is thus, for instance, very easy to generate sequences of pseudo-random numbers from that law using the inversion method. To that end, one only needs to generate a sequence of pseudo-random standard uniform numbers and evaluate the Kumaraswamy quantile function at each value. In contrast, beta random number generation requires the use of acceptance-rejection algorithms, which are more computationally intensive. Additionally, the Kumaraswamy density can assume many different shapes depending on the parameter values, which makes the corresponding law quite flexible for representing rates and proportions. Finally, Wang et al. (2017) note that the Kumaraswamy distribution is particularly useful for modeling variables that describe natural and biological phenomena that are restricted to the standard unit interval.

It is not uncommon, however, for the data to contain zeros and/or ones. When that happens, the interest shifts to modeling variables that assume values in $[0,1)$, $(0,1]$ or $[0,1]$. The Kumaraswamy distribution cannot be used in such cases since, like the beta law, its support is $(0,1)$. Ospina and Ferrari (2010) introduced the class of inflated beta distributions, which allows for the presence of extreme values in the data. In this paper we develop alternative laws: we introduce the class of inflated Kumaraswamy distributions. We consider inflation at one of the endpoints of the standard unit interval and also the more challenging case where inflation takes place at both zero and one, that is, we first consider variables whose support are $[0,1)$ and $(0,1]$ and then we consider the double inflation case, i.e., variables that assume values in $[0,1]$. Such distributions are obtained by combining the Kumaraswamy distribution (continuous component) with a degenerate or with a couple of degenerate distributions (discrete component).

The paper unfolds as follows.The next section presents the zero or one inflated Kumaraswamy distribution (single inflation). Point and interval estimation are also discussed. Notice that inflation only takes place at a single point. In the following section, we go further and introduce the zero and one inflated Kumaraswamy distribution (double inflation). We also show how to perform point and interval estimation. Next, we focus on hypothesis testing inference. Finally, we present and discuss: (i) Monte Carlo simulation evidence and (ii) an empirical application.

THE ZERO OR ONE INFLATED KUMARASWAMY DISTRIBUTION

Data on rates and proportions may contain zeros and/or ones. When that happens the underlying data generating process contains a discrete component that causes a given value or a couple of specific values to be observed with positive probability. It is thus necessary to combine continuous and discrete data generating mechanisms into a more general law. In what follows, we shall focus on random variables that assume values in $(0,1)$ but that can also equal c with positive probability, where $c = 0$ or $c = 1$. We say there is data inflation at one of the standard unit interval endpoints.

We introduce the inflated Kumaraswamy distribution in c (IK_c), whose cdf is given by

$$IK_c(y;\lambda,\alpha,\beta) = \lambda \mathbb{1}_{[c,1]}(y) + (1-\lambda)G(y;\alpha,\beta), \tag{2}$$

where $\mathbb{1}_A(y)$ is an indicator function that equals 1 when $y \in A$ and 0 when $y \notin A$ and $0 < \lambda < 1$ is the mixture parameter. Notice that, with probability $1-\lambda$, Y follows the Kumaraswamy distribution with parameters (α,β) and, with probability λ, it follows a degenerate distribution at c.

Let Y be a random variable with cdf given by (2), denoted by $Y \sim IK_c(\lambda,\alpha,\beta)$. Its pdf is given by

$$ik_c(y;\lambda,\alpha,\beta) = \begin{cases} \lambda, & \text{if} \quad y = c, \\ (1-\lambda)g(y;\alpha,\beta), & \text{if} \quad y \in (0,1), \end{cases} \tag{3}$$

where $0 < \lambda < 1$, $\alpha > 0$ and $\beta > 0$ are the parameters that index the Kumaraswamy distribution and $g(y;\alpha,\beta)$ is the density given in (1). Note that $\lambda = \Pr(Y = 0)$ or $\lambda = \Pr(Y = 1)$.

Figure 1 shows different Kumaraswamy densities inflated at $c = 0$ and at $c = 1$, for different values of α and β, with $\lambda = 0.5$ (recall that λ is the mixture parameter). Note that the probability density function of the inflated Kumaraswamy distribution at c given in (3) may assume a wide variety of shapes; e.g., it can be U-shaped, increasing, decreasing, asymmetric to the left, asymmetric to the right, bell-shaped, and even constant.

The rth moment of Y is

$$\mathbb{E}(Y^r) = \lambda c + (1-\lambda)\mu_r, \quad r = 1,2,\ldots,$$

where $\mu_r = [\beta\Gamma(1+r/\alpha)\Gamma(\beta)]/[\Gamma\left(1+r/\alpha+\beta\right)]$ is the rth moment of the Kumaraswamy distribution, $\Gamma(\cdot)$ denoting the gamma function. In particular, the mean and variance of Y are

$$\begin{aligned} \mathbb{E}(Y) &= \lambda c + (1-\lambda)\mu_1 = \lambda c + \beta(1-\lambda)B\left(1+\frac{1}{\alpha},\beta\right) \quad \text{and} \\ \mathrm{Var}(Y) &= \lambda c + (1-\lambda)\mu_2 - [\lambda c + (1-\lambda)\mu_1]^2 \\ &= \lambda c + \beta(1-\lambda)B\left(1+\frac{2}{\alpha},\beta\right) - \left[\lambda c + \beta(1-\lambda)B\left(1+\frac{1}{\alpha},\beta\right)\right]^2 \\ &= \lambda c(1-\lambda c) + (1-\lambda)\beta\left\{B\left(1+\frac{2}{\alpha},\beta\right)\right. \\ &\quad \left. - B\left(1+\frac{1}{\alpha},\beta\right)\left[2\lambda c + \beta(1-\lambda)B\left(1+\frac{1}{\alpha},\beta\right)\right]\right\}, \end{aligned}$$

respectively, where $B(\cdot,\cdot)$ is the beta function.

It is noteworthy that the density function presented in (3) can be written as

$$ik_c(y;\lambda,\alpha,\beta) = \left[\lambda^{\mathbb{1}_{\{c\}}(y)}(1-\lambda)^{1-\mathbb{1}_{\{c\}}(y)}\right] \times \left[g(y;\alpha,\beta)^{1-\mathbb{1}_{\{c\}}(y)}\right]. \tag{4}$$

The density in (4) is expressed as the product of two terms: the first term only depends on λ whereas the second term only involves α and β.

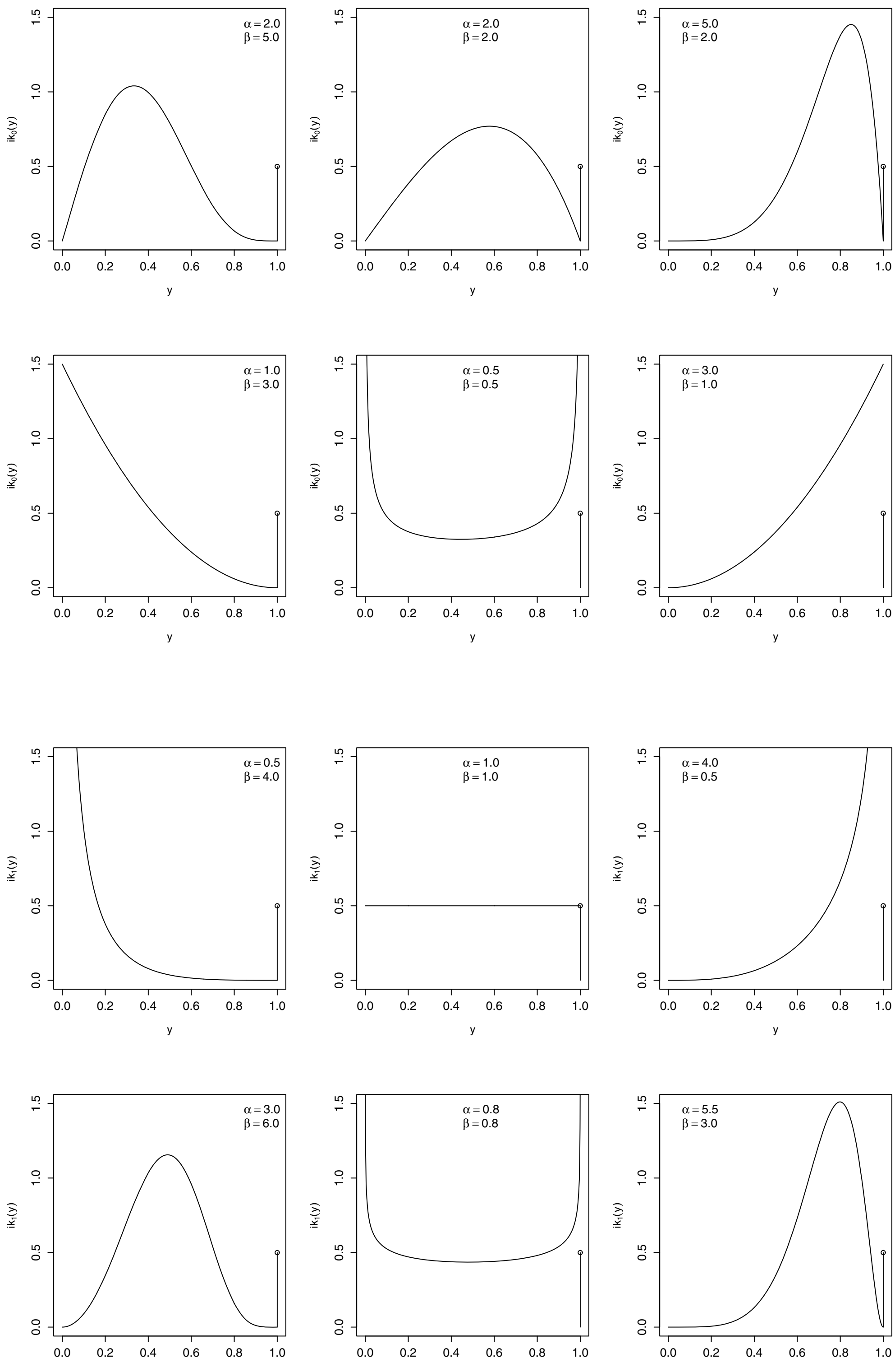

Figure 1 - Inflated Kumaraswamy densities at c = 0 and c = 1, λ = 0.5.

The likelihood function for $\vartheta = (\lambda,\alpha,\beta)'$ based on $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$, a $\mathrm{IK_c}$ random sample, is

$$L(\vartheta;\mathbf{y}) = \prod_{i=1}^{n} \mathrm{ik_c}(y_i;\lambda,\alpha,\beta) = L_1(\lambda;\mathbf{y}) \times L_2(\alpha,\beta;\mathbf{y}),$$

where

$$L_1(\lambda;\mathbf{y}) = \prod_{i=1}^{n} \lambda^{\mathbb{1}_{\{c\}}(y_i)}(1-\lambda)^{1-\mathbb{1}_{\{c\}}(y_i)} = \lambda^{\sum_{i=1}^{n}\mathbb{1}_{\{c\}}(y_i)}(1-\lambda)^{n-\sum_{i=1}^{n}\mathbb{1}_{\{c\}}(y_i)} \quad \text{and}$$

$$L_2(\alpha,\beta;\mathbf{y}) = \prod_{\substack{i=1 \\ y_i\in(0,1)}} g(y_i;\alpha,\beta).$$

The zero or one inflated Kumaraswamy log-likelihood function is then given by

$$\ell(\vartheta;\mathbf{y}) = \ell_1(\lambda;\mathbf{y}) + \ell_2(\alpha,\beta;\mathbf{y}),$$

where

$$\ell_1(\lambda;\mathbf{y}) = \ln(\lambda)\sum_{i=1}^{n}\mathbb{1}_{\{c\}}(y_i) + \ln(1-\lambda)\left[n - \sum_{i=1}^{n}\mathbb{1}_{\{c\}}(y_i)\right] \quad \text{and}$$

$$\ell_2(\alpha,\beta;\mathbf{y}) = \sum_{\substack{i=1 \\ y_i\in(0,1)}} \ln(\alpha\beta) + (\alpha-1)\sum_{\substack{i=1 \\ y_i\in(0,1)}} \ln(y_i) + (\beta-1)\sum_{\substack{i=1 \\ y_i\in(0,1)}} \ln(1-y_i^{\alpha}).$$

The score function, which is obtained by differentiating the log-likelihood function, is denoted by $U(\vartheta) = \left[U_\lambda(\lambda), U_\alpha(\alpha,\beta), U_\beta(\alpha,\beta)\right]$, where

$$U_\lambda(\lambda) = \frac{\partial \ell_1(\lambda;\mathbf{y})}{\partial\lambda} = \frac{1}{\lambda}\sum_{i=1}^{n}\mathbb{1}_{\{c\}}(y_i) - \frac{1}{1-\lambda}\left[n - \sum_{i=1}^{n}\mathbb{1}_{\{c\}}(y_i)\right],$$

$$U_\alpha(\alpha,\beta) = \frac{\partial \ell_2(\alpha,\beta;\mathbf{y})}{\partial\alpha} = \frac{1}{\alpha}\left[n - \sum_{i=1}^{n}\mathbb{1}_{\{c\}}(y_i)\right] + \sum_{\substack{i=1 \\ y_i\in(0,1)}} \ln(y_i) + (\beta-1)\sum_{\substack{i=1 \\ y_i\in(0,1)}} \left(\frac{y_i^{\alpha}}{y_i^{\alpha}-1}\right)\ln(y_i) \quad \text{and}$$

$$U_\beta(\alpha,\beta) = \frac{\partial \ell_2(\alpha,\beta;\mathbf{y})}{\partial\beta} = \frac{1}{\beta}\left[n - \sum_{i=1}^{n}\mathbb{1}_{\{c\}}(y_i)\right] + \sum_{\substack{i=1 \\ y_i\in(0,1)}} \ln(1-y_i^{\alpha}).$$

The maximum likelihood estimator (mle) of λ is $\hat{\lambda} = n^{-1}\sum_{i=1}^{n}\mathbb{1}_{\{c\}}(y_i)$, i.e., it is given by the proportion of sample values that equal c. The maximum likelihood estimators of α and β cannot be expressed in closed-form. They can be obtained, however, by numerically maximizing the log-likelihood function using a nonlinear optimization method, such as a Newton or quasi-Newton method. The BFGS quasi-Newton method is commonly used for numerically maximizing log-likelihood functions; for details on such a method, see Nocedal and Wright (2006) and Press et al. (1992).

The Fisher information matrix for the zero or one inflated Kumaraswamy law is

$$K(\vartheta) = \begin{pmatrix} k_{\lambda\lambda} & 0 & 0 \\ 0 & k_{\alpha\alpha} & k_{\alpha\beta} \\ 0 & k_{\beta\alpha} & k_{\beta\beta} \end{pmatrix}, \tag{5}$$

where

$$k_{\lambda\lambda} = \frac{n}{\lambda(1-\lambda)}, \quad k_{\alpha\alpha} = \frac{n(1-\lambda)}{\alpha^2} + \frac{n\beta(1-\lambda)}{\alpha^2(\beta-2)}\left\{\left[\psi(\beta)-\psi(2)\right]^2 - \left[\psi'(\beta)-\psi'(2)\right]\right\},$$
$$k_{\alpha\beta} = k_{\beta\alpha} = -\frac{n(1-\lambda)}{\alpha(\beta-1)}\left\{\left[\psi(\beta+1)-\psi(2)\right]\right\}, \quad k_{\beta\beta} = \frac{n(1-\lambda)}{\beta^2}.$$

Here, $\psi(z) = \partial \ln\Gamma(z)/\partial z$ is the digamma function and $\psi'(z) = \partial\psi(z)/\partial z$ is the trigamma function.

Let $\hat{\vartheta} = (\hat{\lambda}, \hat{\alpha}, \hat{\beta})'$ denote the mle of ϑ. In large samples $\hat{\vartheta}$ is expected to be approximately normally distributed: $\hat{\vartheta} \overset{a}{\sim} \mathcal{N}_3(\vartheta, K(\vartheta)^{-1})$, where $K(\vartheta)$ is the information matrix given in (5) and $\overset{a}{\sim}$ denotes approximately distributed. Using such a result, it is possible to construct approximate confidence intervals for the model parameters. Let $\delta \in (0, 0.5)$. It follows that $(1-\delta)\times 100\%$ asymptotic confidence intervals for λ, α and β are given, respectively, by $\hat{\lambda} \pm z_{(1-\delta/2)}\mathrm{se}(\hat{\lambda})$, $\hat{\alpha} \pm z_{(1-\delta/2)}\mathrm{se}(\hat{\alpha})$ and $\hat{\beta} \pm z_{(1-\delta/2)}\mathrm{se}(\hat{\beta})$, where $\mathrm{se}(\cdot)$ denotes standard error and $z_{(1-\delta/2)}$ is the $1-\delta/2$ standard normal quantile. The standard errors are obtained as square roots of the diagonal elements of the inverse of Fisher's information matrix after the unknown parameters are replaced with the corresponding maximum likelihood estimates.

ZERO AND ONE INFLATED KUMARASWAMY DISTRIBUTION

The distribution introduced in the previous section is not suitable for modeling fractional data that contain both zeros and ones, i.e., when data inflation occurs at both ends of the standard unit interval. In what follows we shall introduce a distribution that can be used to model variables that have support in $[0,1]$. We shall now introduce the appropriate law for that case. We say that the random variable Y follows the zero and one inflated Kumaraswamy distribution, denoted by $Y \sim \mathrm{ZOIK}(y;\lambda,p,\alpha,\beta)$, if its cdf is given by

$$\mathrm{ZOIK}(y;\lambda,p,\alpha,\beta) = \lambda \mathrm{Ber}(y;p) + (1-\lambda)G(y;\alpha,\beta),$$

with $y \in [0,1]$, where $\lambda \in (0,1)$ is the mixture parameter and $\mathrm{Ber}(y;p)$ denotes the cumulative distribution function of a Bernoulli random variable with parameter $p = \Pr(Y = 1)$.

It follows that the pdf of Y is

$$\mathrm{zoik}(y;\lambda,p,\alpha,\beta) = \begin{cases} \lambda p, & \text{if} \quad y = 1, \\ \lambda(1-p), & \text{if} \quad y = 0, \\ (1-\lambda)g(y;\alpha,\beta), & \text{if} \quad y \in (0,1). \end{cases} \quad (6)$$

Note that $\lambda p = \Pr(Y = 1)$ and $\lambda(1-p) = \Pr(Y = 0)$. For $y \in (0,1)$ and $0 < a < b < 1$, $\Pr(Y \in (a,b)) = (1-\lambda)\int_a^b g(y;\alpha,\beta)dy$.

Figure 2 presents several ZOIK densities for $\lambda = 0.2$ and $p = 0.5$. Notice the many different shapes that the density can assume. The distribution is thus a very flexible law for variables that assume values in the standard unit interval with inflation at both interval limits.

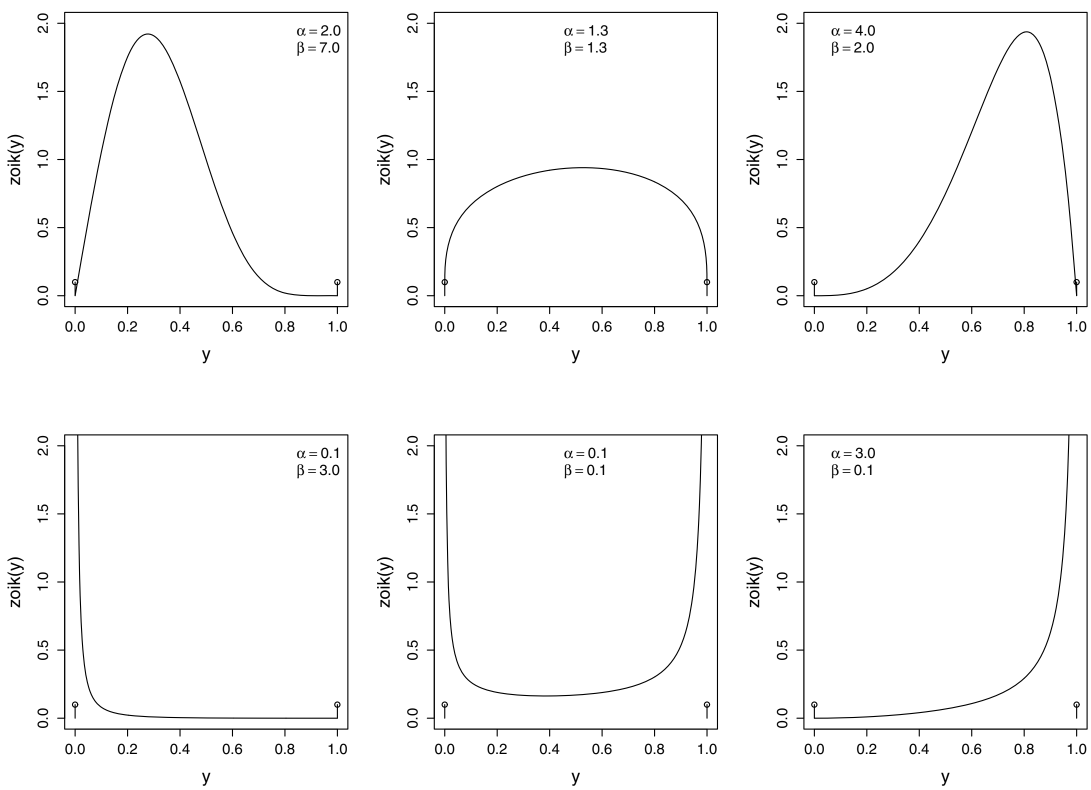

Figure 2 - Zero and one inflated Kumaraswamy densities, $\lambda = 0.2$ and $p = 0.5$.

Let Y be a zero and one inflated Kumaraswamy random variable. Its rth moment is $\mathbb{E}(Y^r) = \lambda p + (1-\lambda)\mu_r$, $r = 1, 2, \ldots$. Hence,

$$\mathbb{E}(Y) = \lambda p + (1-\lambda)\mu_1 = \lambda p + \beta(1-\lambda)B\left(1+\frac{1}{\alpha},\beta\right) \quad \text{and}$$
$$\begin{aligned}\mathrm{Var}(Y) &= \lambda p + (1-\lambda)\mu_2 - [\lambda p + (1-\lambda)\mu_1]^2 \\ &= \lambda p + \beta(1-\lambda)B\left(1+\frac{2}{\alpha},\beta\right) - \left[\lambda p + \beta(1-\lambda)B\left(1+\frac{1}{\alpha},\beta\right)\right]^2 \\ &= \lambda p(1-\lambda p) + (1-\lambda)\beta\left\{B\left(1+\frac{2}{\alpha},\beta\right) - B\left(1+\frac{1}{\alpha},\beta\right)\left[2\lambda p + \beta(1-\lambda)B\left(1+\frac{1}{\alpha},\beta\right)\right]\right\},\end{aligned}$$

where μ_1 and μ_2 are the first and second Kumaraswamy moments, respectively.

Consider the zero and one inflated Kumaraswamy density given in (6). It is possible to write it as

$$\begin{aligned}\mathrm{zoik}(y;\lambda,p,\alpha,\beta) &= \left[\lambda p^y(1-p)^{1-y}\right]^{\mathbb{1}_{\{0,1\}}(y)} \times \left[(1-\lambda)g(y;\alpha,\beta)\right]^{1-\mathbb{1}_{\{0,1\}}(y)} \\ &= \left[\lambda^{\mathbb{1}_{\{0,1\}}(y)}(1-\lambda)^{1-I_{\{0,1\}}(y)}\right]\left[p^y(1-p)^{1-y}\right]^{\mathbb{1}_{\{0,1\}}(y)}\left[g(y;\alpha,\beta)^{1-\mathbb{1}_{\{0,1\}}(y)}\right], \qquad (7)\end{aligned}$$

where now $\mathbb{1}_{\{0,1\}}(y)$ is the indicator function that equals one if $y \in \{0,1\}$ and equals zero if $y \notin \{0,1\}$. The pdf in (7) factors into three terms: the first term only depends on λ, the second term only depends on p and the third term involves α and β.

The likelihood function for $\vartheta = (\lambda, p, \alpha, \beta)'$ based on the random sample $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$ is

$$L(\vartheta;\mathbf{y}) = \prod_{i=1}^{n} \text{zoik}(y_i;\lambda,p,\alpha,\beta) = L_1(\lambda;\mathbf{y}) \times L_2(p;\mathbf{y}) \times L_3(\alpha,\beta;\mathbf{y}),$$

where

$$L_1(\lambda;\mathbf{y}) = \prod_{i=1}^{n} \lambda^{\mathbb{1}_{\{0,1\}}(y_i)}(1-\lambda)^{1-\mathbb{1}_{\{0,1\}}(y_i)} = \lambda^{\sum_{i=1}^{n}\mathbb{1}_{\{0,1\}}(y_i)}(1-\lambda)^{n-\sum_{i=1}^{n}\mathbb{1}_{\{0,1\}}(y_i)},$$

$$L_2(p;\mathbf{y}) = \prod_{i=1}^{n}\left[p^{y_i}(1-p)^{1-y_i}\right]^{\mathbb{1}_{\{0,1\}}(y_i)} = p^{\sum_{i=1}^{n} y_i \mathbb{1}_{\{0,1\}}(y_i)}(1-p)^{\sum_{i=1}^{n}(1-y_i)\mathbb{1}_{\{0,1\}}(y_i)}$$

$$= p^{\sum_{i=1}^{n}\mathbb{1}_{\{1\}}(y_i)}(1-p)^{\left[\sum_{i=1}^{n}\mathbb{1}_{\{0,1\}}(y_i)-\sum_{i=1}^{n}\mathbb{1}_{\{1\}}(y_i)\right]} \quad \text{and}$$

$$L_3(\alpha,\beta;\mathbf{y}) = \prod_{\substack{i=1\\ y_i\in(0,1)}} g(y_i;\alpha,\beta) = \prod_{\substack{i=1\\ y_i\in(0,1)}} (\alpha\beta)y_i^{(\alpha-1)}(1-y_i^{\alpha})^{(\beta-1)}.$$

The corresponding log-likelihood function can be expressed as

$$\ell(\vartheta;\mathbf{y}) = \ell_1(\lambda;\mathbf{y}) + \ell_2(p;\mathbf{y}) + \ell_3(\alpha,\beta;\mathbf{y}),$$

where

$$\ell_1(\lambda;\mathbf{y}) = \ln(\lambda)\sum_{i=1}^{n}\mathbb{1}_{\{0,1\}}(y_i) + \ln(1-\lambda)\left[n - \sum_{i=1}^{n}\mathbb{1}_{\{0,1\}}(y_i)\right],$$

$$\ell_2(p;\mathbf{y}) = \ln(p)\sum_{i=1}^{n}\mathbb{1}_{\{1\}}(y_i) + \ln(1-p)\left[\sum_{i=1}^{n}\mathbb{1}_{\{0,1\}}(y_i) - \sum_{i=1}^{n}\mathbb{1}_{\{1\}}(y_i)\right] \quad \text{and}$$

$$\ell_3(\alpha,\beta;\mathbf{y}) = \sum_{\substack{i=1\\ y_i\in(0,1)}} \ln(\alpha\beta) + (\alpha-1)\sum_{\substack{i=1\\ y_i\in(0,1)}} \ln(y_i) + (\beta-1)\sum_{\substack{i=1\\ y_i\in(0,1)}} \ln(1-y_i^{\alpha}).$$

The score function is given by $U(\vartheta) = \left[U_\lambda(\lambda), U_p(p), U_\alpha(\alpha,\beta), U_\beta(\alpha,\beta)\right]$, where

$$U_\lambda(\lambda) = \frac{\partial \ell_1(\lambda;\mathbf{y})}{\partial\lambda} = \frac{1}{\lambda}\sum_{i=1}^{n}\mathbb{1}_{\{0,1\}}(y_i) - \frac{1}{1-\lambda}\left[n - \sum_{i=1}^{n}\mathbb{1}_{\{0,1\}}(y_i)\right],$$

$$U_p(p) = \frac{\partial \ell_2(p;\mathbf{y})}{\partial p} = \frac{1}{p}\sum_{i=1}^{n}\mathbb{1}_{\{1\}}(y_i) - \frac{1}{1-p}\left[\sum_{i=1}^{n}\mathbb{1}_{\{0,1\}}(y_i) - \sum_{i=1}^{n}\mathbb{1}_{\{1\}}(y_i)\right],$$

$$U_\alpha(\alpha,\beta) = \frac{\partial \ell_3(\alpha,\beta;\mathbf{y})}{\partial\alpha} = \frac{1}{\alpha}\left[n - \sum_{i=1}^{n}\mathbb{1}_{\{0,1\}}(y_i)\right] + \sum_{\substack{i=1\\ y_i\in(0,1)}} \ln(y_i) + (\beta-1)\sum_{\substack{i=1\\ y_i\in(0,1)}}\left(\frac{y_i^{\alpha}}{y_i^{\alpha}-1}\right)\ln(y_i) \quad \text{and}$$

$$U_\beta(\alpha,\beta) = \frac{\partial \ell_3(\alpha,\beta;\mathbf{y})}{\partial\beta} = \frac{1}{\beta}\left[n - \sum_{i=1}^{n}\mathbb{1}_{\{0,1\}}(y_i)\right] + \sum_{\substack{i=1\\ y_i\in(0,1)}} \ln(1-y_i^{\alpha}).$$

The maximum likelihood estimators of λ and p are, respectively, $\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{0,1\}}(y_i)$, which is the proportion of discrete values in the sample, and $\hat{p} = \sum_{i=1}^{n} \mathbb{1}_{\{1\}}(y_i)/\sum_{i=1}^{n} \mathbb{1}_{\{0,1\}}(y_i)$, which is the proportion of degenerate values that equal one.

The Fisher information matrix for the zero and one inflated Kumaraswamy distribution is

$$K(\vartheta) = \begin{pmatrix} k_{\lambda\lambda} & 0 & 0 & 0 \\ 0 & k_{pp} & 0 & 0 \\ 0 & 0 & k_{\alpha\alpha} & k_{\alpha\beta} \\ 0 & 0 & k_{\beta\alpha} & k_{\beta\beta} \end{pmatrix}, \quad (8)$$

where

$$k_{\lambda\lambda} = \frac{n}{\lambda(1-\lambda)}, \quad k_{pp} = \frac{n\lambda}{p(1-p)}, \quad k_{\alpha\alpha} = \frac{n(1-\lambda)}{\alpha^2} + \frac{n\beta(1-\lambda)}{\alpha^2(\beta-2)}\left\{\left[\psi(\beta)-\psi(2)\right]^2 - \left[\psi'(\beta)-\psi'(2)\right]\right\},$$
$$k_{\alpha\beta} = k_{\beta\alpha} = -\frac{n(1-\lambda)}{\alpha(\beta-1)}\left\{\left[\psi(\beta+1)-\psi(2)\right]\right\}, \quad k_{\beta\beta} = \frac{n(1-\lambda)}{\beta^2}.$$

As before, approximate confidence intervals can be constructed based on the asymptotic normality of $\hat{\vartheta}$, the mle of ϑ. In large samples, it is expected that $\hat{\vartheta} \overset{a}{\sim} \mathcal{N}_4(\vartheta, K(\vartheta)^{-1})$, where $K(\vartheta)$ is the information matrix given in (8). Using such a limiting distribution, it is possible to construct asymptotic confidence intervals for λ, p, α and β. For $\delta \in (0, 0.5)$, the $(1-\delta)\times 100\%$ asymptotic confidence intervals for such parameters are given, respectively, by $\hat{\lambda} \pm z_{(1-\delta/2)}\mathrm{se}(\hat{\lambda})$, $\hat{p} \pm z_{(1-\delta/2)}\mathrm{se}(\hat{p})$, $\hat{\alpha} \pm z_{(1-\delta/2)}\mathrm{se}(\hat{\alpha})$ and $\hat{\beta} \pm z_{(1-\delta/2)}\mathrm{se}(\hat{\beta})$.

HYPOTHESIS TESTING INFERENCE

The asymptotic normality of $\hat{\vartheta}$ can also be used to construct hypothesis tests. Suppose the interest lies in making testing inference on a subset of parameters. Let $\vartheta = (\vartheta_1', \vartheta_2')'$, where ϑ_1 is an $r \times 1$ vector of parameters of interest and ϑ_2 is an $(m-r)\times 1$ vector of nuisance parameters. We wish to test the null hypothesis $\mathscr{H}_0 : \vartheta_1 = \vartheta_1^{(0)}$ against the alternative hypothesis $\mathscr{H}_1 : \vartheta_1 \neq \vartheta_1^{(0)}$. The inference can be based on the following criteria: likelihood ratio (LR), Wald (W) and score (S). For details on these tests, see Buse (1992), Cox and Hinkley (1979, Chapter 9) and Welsh (1996, Section 4.5).

Let $\hat{\vartheta}$ be the unrestricted maximum likelihood estimator of ϑ and let $\tilde{\vartheta} = \left(\vartheta_1^{(0)\prime}, \tilde{\vartheta}_2'\right)'$ be the restricted maximum likelihood estimator of ϑ which is obtained by imposing $\mathscr{H}_0$. The likehood ratio test statistic is given by

$$LR = 2[\ell(\hat{\vartheta}) - \ell(\tilde{\vartheta})],$$

the Wald test statistic can be written as

$$W = \left(\hat{\vartheta}_1 - \vartheta_1^{(0)}\right)' \left[K^{rr}(\hat{\vartheta})\right]^{-1} \left(\hat{\vartheta}_1 - \vartheta_1^{(0)}\right)$$

and the score test statistic is

$$S = U_r(\tilde{\vartheta})' K^{rr}(\tilde{\vartheta}) U_r(\tilde{\vartheta}),$$

where $K^{rr}(\hat{\vartheta})$ is the $r \times r$ block of Fisher's information matrix inverse that corresponds to ϑ_1 evaluated at $\hat{\vartheta}$, $U_r(\tilde{\vartheta})$ denotes the $r \times 1$ vector that contains the r elements of the score function corresponding to the parameters of interest and $K^{rr}(\tilde{\vartheta})$ is the $r \times r$ block of Fisher's information matrix inverse that corresponds to ϑ_1 evaluated at $\tilde{\vartheta}$.

Notice that in order to compute LR one needs to obtain $\hat{\vartheta}$ and $\tilde{\vartheta}$, i.e., it is necessary to perform both unrestricted and restricted parameter estimation. In contrast, in order to compute W one only needs to perform unrestricted estimation and in order to compute S one only needs to carry out restricted estimation.

Under $\mathscr{H}_0$ and under some regularity conditions outlined by Serfling (1980), $\text{LR} \xrightarrow{d} \chi^2_r$, $\text{W} \xrightarrow{d} \chi^2_r$ and $\text{S} \xrightarrow{d} \chi^2_r$, where $\xrightarrow{d}$ denotes convergence in distribution. The three test statistics thus share the same asymptotic null distribution. The tests are typically carried out using asymptotic (i.e., approximate) critical values. The null hypothesis $\mathscr{H}_0$ is rejected at significance level $\delta \in (0,1)$ if the selected criterion exceeds $\chi^2_{r;1-\delta}$, the $1-\delta$ χ^2_r upper quantile.

NUMERICAL EVALUATION

In what follows we shall report results of Monte Carlo simulations that were carried out to evaluate the finite sample performances of point estimators, confidence intervals and hypothesis tests. We consider inflation at one and also inflation at both zero and one. The reported results are based on 10,000 replications and were obtained using the Ox matrix programming language; see Cribari-Neto and Zarkos (2003) and Doornik (2009). Log-likelihood maximization was performed using the quasi-Newton BFGS method with analytical first derivatives, which is typically regarded as the best performing method; see Mittelhammer et al. (2000, Section 8.13). The initial values used in the BFGS iterative scheme were arbitrarily selected, being different from the true parameter values. We varied such initial values and noticed that they had little impact on the results.

At the outset we focus on point estimation. Tables I and II contain the variances, relative biases and mean squared errors (MSEs) of the maximum likelihood estimators of the parameters that index the Kumaraswamy distribution with inflation at one and with inflation at zero and one, respectively. Relative bias is computed as the difference between the mean estimate and the true parameter valued divided by the latter. We report results for different sample sizes. The mixture parameter (λ) assumes two values: 0.05 and 0.50. The results show that the relative biases, variances and mean squared errors decay as the sample size increases. The results in Table I show that point estimation of λ is less accurate when the true parameter value is small. Consider, e.g., $n = 50$. The relative bias of $\hat{\lambda}$ equals 8.60% when $\lambda = 0.05$ and –0.02% when $\lambda = 0.50$. It is noteworthy that point estimation of β is less accurate than that of α and λ, especially when the value of λ is large (0.50). This seems to be a characteristic Kumaraswamy maximum likelihood point estimation that is carried over to the new class of inflated distributions. Consider, for instance, the numerical evidence reported by Lemonte (2011). Except when the value of β is quite small, the numerical evidence in his paper shows that the maximum likelihood estimator of β is considerably less accurate than that of α both in terms of bias and mean squared error.

Next, we evaluate the accuracy of interval estimation in finite samples. The confidence intervals empirical coverages and non-coverages are presented in Tables III (single inflation) and IV (double inflation); entries are percentages. The results show that the empirical coverages approach the nominal ones as the sample size increases. The non-coverages also become better balanced as number of data points is increased. Consider, e.g., $n = 100$, $\lambda = 0.50$ and $1-\delta = 0.95$. Under single inflation, the empirical coverage rates for λ, α and β are, respectively, 94.62%, 94.53% and 96.32%. Under double inflation, the corresponding

TABLE I
Relative biases, variances and MSEs of the maximum likelihood estimators of the parameters that index the IK_1 distribution; $\alpha = 1.5, \beta = 3.0$.

λ	Measure	Estimator	n = 50	100	200	500
0.05	Rel. Bias	$\hat{\lambda}$	0.0860	0.0040	0.0020	0.0000
		$\hat{\alpha}$	0.0375	0.0193	0.0095	0.0043
		$\hat{\beta}$	0.0933	0.0449	0.0223	0.0099
	Variance	$\hat{\lambda}$	0.0008	0.0005	0.0002	0.0001
		$\hat{\alpha}$	0.0579	0.0266	0.0130	0.0051
		$\hat{\beta}$	0.8120	0.3045	0.1358	0.0524
	MSE	$\hat{\lambda}$	0.0008	0.0005	0.0002	0.0001
		$\hat{\alpha}$	0.0610	0.0274	0.0132	0.0052
		$\hat{\beta}$	0.8903	0.3226	0.1403	0.0533
0.50	Rel. Bias	$\hat{\lambda}$	–0.0002	–0.0004	–0.0004	–0.0002
		$\hat{\alpha}$	0.0803	0.0381	0.0174	0.0079
		$\hat{\beta}$	0.2044	0.0875	0.0411	0.0172
	Variance	$\hat{\lambda}$	0.0049	0.0024	0.0012	0.0005
		$\hat{\alpha}$	0.1282	0.0557	0.0258	0.0099
		$\hat{\beta}$	2.6576	0.7277	0.2943	0.1030
	MSE	$\hat{\lambda}$	0.0049	0.0024	0.0012	0.0005
		$\hat{\alpha}$	0.1427	0.0590	0.0264	0.0101
		$\hat{\beta}$	3.0335	0.7965	0.3095	0.1057

coverage rates for λ, α, β and p are 94.27%, 94.77%, 96.50% and 96.63%. Overall, the confidence intervals display reasonably accurate coverages except the confidence interval for λ when the true parameter value is very small ($\lambda = 0.05$, Table III). For instance, when $n = 100$ and $1 - \delta = 90\%$, the exact interval coverage was slightly below 86%. For α and β, the corresponding coverage figures were 89.76% and 91.28%.

We also carried out simulations to evaluate the finite performances of testing inferences based on the LR, W and S asymptotic chi-squared criteria. The interest lies in testing $\mathscr{H}_0 : \lambda = \lambda_0 \times \mathscr{H}_1 : \lambda \neq \lambda_0$ for the IK_1 law. For the ZOIK law, we test $\mathscr{H}_0 : \lambda = \lambda_0 \times \mathscr{H}_1 : \lambda \neq \lambda_0$ and also $\mathscr{H}_0 : p = p_0 \times \mathscr{H}_1 : p \neq p_0$.

In the former case, $\alpha = 1.5$ and $\beta = 3.0$; in the latter case, for the test on λ we generated data using $p = 0.5, \alpha = 1.5, \beta = 3.0$ and for the test on p we performed data generation using $\lambda = 0.2, \alpha = 1.5, \beta = 3.0$. Data generation was performed under the null hypothesis. The significance levels are 5% and 10%. The tests null rejection rates are presented in Tables V (test on λ, IK_1 law), VI (test on λ, ZOIK, law) and VII (test on p, ZOIK law). Notice that the empirical null rejection rates converge to the corresponding nominal significance levels as the sample size increases. Overall, the likelihood ratio test is the best performing

TABLE II
Relative biases, variances and MSEs of the maximum likelihood estimators of the parameters that index the ZOIK distribution; $p = 0.5, \alpha = 1.5, \beta = 3.0$.

λ	Measure	Estimator	n = 50	100	200	500
0.05	Rel. Bias	$\hat{\lambda}$	0.3920	0.0880	0.008	0.0020
		$\hat{p}$	–0.0028	–0.0026	–0.0014	–0.0006
		$\hat{\alpha}$	0.0390	0.0185	0.0084	0.0038
		$\hat{\beta}$	0.0964	0.0439	0.0209	0.0089
	Variance	$\hat{\lambda}$	0.0007	0.0004	0.0002	0.0001
		$\hat{p}$	0.0243	0.0310	0.0255	0.0104
		$\hat{\alpha}$	0.0593	0.0267	0.0130	0.0051
		$\hat{\beta}$	0.8137	0.3104	0.1381	0.0521
	MSE	$\hat{\lambda}$	0.0011	0.0004	0.0002	0.0001
		$\hat{p}$	0.0243	0.0310	0.0255	0.0104
		$\hat{\alpha}$	0.0627	0.0275	0.0132	0.0052
		$\hat{\beta}$	0.8973	0.3278	0.1420	0.0528
0.50	Rel. Bias	$\hat{\lambda}$	–0.0016	–0.0016	–0.0006	–0.0004
		$\hat{p}$	0.0024	0.0012	0.0008	0.0002
		$\hat{\alpha}$	0.0779	0.0359	0.0163	0.0071
		$\hat{\beta}$	0.2033	0.0868	0.0395	0.0166
	Variance	$\hat{\lambda}$	0.0049	0.0025	0.0012	0.0005
		$\hat{p}$	0.0101	0.0051	0.0025	0.0010
		$\hat{\alpha}$	0.1283	0.0552	0.0253	0.0095
		$\hat{\beta}$	2.6944	0.7409	0.2873	0.0996
	MSE	$\hat{\lambda}$	0.0049	0.0025	0.0012	0.0005
		$\hat{p}$	0.0102	0.0051	0.0025	0.0010
		$\hat{\alpha}$	0.1419	0.0581	0.0259	0.0096
		$\hat{\beta}$	3.0663	0.8087	0.3013	0.1020

test, i.e., it is typically the least size-distorted test. For example, when $n = 100$, $\lambda = 0.10$ ($\lambda = 0.50$) and at the 5% significance level in Table V, the likelihood ratio null rejection rate is 4.44% (5.38%) under single inflation. The corresponding figures for the score and Wald tests are, respectively, 6.35% (5.38%) and 7.05 (5.38%). The null rejection rates of the three tests coincide when $\lambda_0 = 0.50$ (IK_1 and ZOIK), even though the test statistics values are slightly different in each replication. The tests become less accurate when they are used to make inference on p (Table VII), especially when the value of p_0 is small. The tests become more

TABLE III
Confidence intervals empirical coverages and noncoverages (to the left; to the right) rates (%), IK_1 distribution; $\alpha = 1.5$ and $\beta = 3.0$.

λ	$1-\delta$	Parameter	n 50	100	200	500
0.05	95%	λ	99.63	88.01	92.60	93.32
			(0.37; 0.00)	(0.52; 11.47)	(1.31; 6.09)	(1.25; 5.43)
		α	94.91	95.15	95.00	95.02
			((2.84; 2.25)	(2.57; 2.28)	(2.65; 2.35)	(2.77; 2.21)
		β	96.07	95.85	95.59	95.23
			(0.49; 3.44)	(1.10; 3.05)	(1.54; 2.87)	(2.02; 2.75)
	90%	λ	98.69	85.58	84.97	88.61
			(1.31; 0.00)	(2.95; 11.47)	(2.59; 12.44)	(2.83; 8.56)
		α	89.64	89.76	89.89	90.24
			(6.05; 4.31)	(5.78; 4.46)	(5.48; 4.63)	(5.30; 4.46)
		β	91.75	91.28	90.82	90.11
			(2.50; 5.75)	(3.39; 5.33)	(3.91; 5.27)	(4.68; 5.21)
0.50	95%	λ	93.72	94.62	94.43	94.54
			(3.29; 2.99)	(2.71; 2.67)	(2.87; 2.70)	(2.83; 2.63)
		α	94.52	94.53	94.93	94.80
			(3.19; 2.29)	(3.05; 2.42)	(2.61; 2.46)	(2.72; 2.48)
		β	96.53	96.32	95.67	95.33
			(0.00; 3.47)	(0.49; 3.19)	(1.20; 3.13)	(1.67; 3.00)
	90%	λ	88.42	91.66	89.99	90.28
			(6.00; 5.58)	(4.11; 4.23)	(5.20; 4.81)	(4.79; 4.93)
		α	89.51	89.46	89.58	89.86
			(6.70; 3.79)	(6.16; 4.38)	(5.90; 4.52)	(5.40; 4.74)
		β	94.08	92.23	90.55	90.38
			(0.56; 5.36)	(2.50; 5.27)	(3.74; 5.71)	(4.43; 5.19)

accurate when $n \geq 200$. Consider, for example, $p_0 = 0.10$, $\delta = 10\%$ and $n = 200$. The null rejection rates of the likelihood ratio, score and Wald tests are 9.38%, 8.14% and 12.40%.

We have also carried out power simulation, i.e., simulations in which data generation was performed under the alternative hypothesis. For brevity, we shall only report results for the test on λ in the ZOIK law. Data generation was carried using $\lambda = 0.20$ and $\lambda = 0.40$ when $\lambda_0 = 0.10$ and $\lambda_0 = 0.50$, respectively. Since no test is very liberal, the tests are performed using asymptotic (χ^2) critical values. The tests nonnull rejection

TABLE IV
Confidence intervals empirical coverages and noncoverages (to the left; to the right) rates (%), ZOIK distribution; $p = 0.5, \alpha = 1.5$ and $\beta = 3.0$.

λ	$1-\delta$	Parameter	n = 50	n = 100	n = 200	n = 500
0.05	95%	λ	99.39	94.79	93.49	93.47
			(0.61; 0.00)	(0.42; 4.79)	(1.18; 5.33)	(1.23; 5.30)
		p	98.84	94.20	89.53	93.24
			(0.59; 0.57)	(2.88; 2.92)	(5.25; 5.22)	(3.29; 3.47)
		α	94.68	94.93	95.07	95.06
			(2.88; 2.44)	(2.63; 2.44)	(2.55; 2.38)	(2.57; 2.37)
		β	96.30	95.70	95.57	95.26
			(0.43; 3.27)	(1.22; 3.08)	(1.58; 2.85)	(1.93; 2.81)
	90%	λ	97.48	92.07	86.11	88.64
			(2.52; 0.00)	(3.14; 4.79)	(2.29; 11.60)	(2.96; 8.40)
		p	94.49	87.46	85.95	87.61
			(2.62; 2.89)	(6.13; 6.41)	(7.14; 6.91)	(6.03; 6.36)
		α	89.85	90.07	90.27	89.84
			(5.84; 4.31)	(5.47; 4.46)	(4.95; 4.78)	(5.25; 4.91)
		β	92.11	91.31	90.64	90.34
			(2.47; 5.42)	(3.65; 5.04)	(3.84; 5.52)	(4.55; 5.11)
0.50	95%	λ	93.87	94.27	94.40	94.45
			(3.08; 3.05)	(2.77; 2.96)	(2.79; 2.81)	(2.74; 2.81)
		p	93.38	93.63	94.82	94.63
			(3.32; 3.30)	(3.23; 3.14)	(2.55; 2.63)	(2.57; 2.80)
		α	94.72	94.77	94.88	95.53
			(3.13; 2.15)	(2.96; 2.27)	(2.66; 2.46)	(2.19; 2.28)
		β	96.59	96.50	95.66	95.78
			(0.00; 3.41)	(0.43; 3.07)	(1.27; 3.07)	(1.40; 2.82)
	90%	λ	88.48	90.98	89.78	90.15
			(5.69; 5.83)	(4.34; 4.68)	(5.04; 5.18)	(4.86; 4.99)
		p	87.95	88.71	89.81	89.51
			(6.22; 5.83)	(5.64; 5.65)	(5.05; 5.14)	(4.97; 5.52)
		α	89.42	89.72	90.05	90.51
			(6.74; 3.84)	(5.61; 4.67)	(5.49; 4.46)	(5.02; 4.47)
		β	94.12	91.88	91.04	90.88
			(0.45; 5.43)	(2.69; 5.43)	(3.40; 5.56)	(3.97; 5.15)

TABLE V
Null rejection rates (%), IK_1 distribution,
$\mathscr{H}_0 : \lambda = \lambda_0 \times \mathscr{H}_1 : \lambda \neq \lambda_0$.

λ_0	δ	Test	n			
			50	100	200	500
0.10	5%	LR	5.25	4.44	5.87	5.21
		W	12.20	7.05	6.88	5.75
		S	2.39	6.35	4.39	4.37
	10%	LR	8.64	9.87	9.68	10.31
		W	13.70	13.90	11.88	11.22
		S	8.64	13.29	12.24	8.59
0.50	5%	LR	6.28	5.38	5.57	5.46
		W	6.28	5.38	5.57	5.46
		S	6.28	5.38	5.57	5.46
	10%	LR	11.58	8.34	10.01	9.72
		W	11.58	8.34	10.01	9.72
		S	11.58	8.34	10.01	9.72

TABLE VI
Null rejection rates (%), ZOIK distribution,
$\mathscr{H}_0 : \lambda = \lambda_0 \times \mathscr{H}_1 : \lambda \neq \lambda_0$.

λ_0	δ	Test	n			
			50	100	200	500
0.10	5%	LR	2.90	3.91	6.07	5.44
		W	5.61	6.03	7.31	5.72
		S	2.90	5.80	4.33	4.61
	10%	LR	7.00	8.96	10.03	10.10
		W	7.45	12.59	12.04	10.83
		S	7.00	12.36	12.30	8.67
0.50	5%	LR	6.13	5.73	5.60	5.55
		W	6.13	5.73	5.60	5.55
		S	6.13	5.73	5.60	5.55
	10%	LR	11.52	9.02	10.22	9.85
		W	11.52	9.02	10.22	9.85
		S	11.52	9.02	10.22	9.85

TABLE VII
Null rejection rates (%), ZOIK distribution,
$\mathscr{H}_0 : p = p_0 \times \mathscr{H}_1 : p \neq p_0$.

p_0	δ	Test	n			
			50	100	200	500
0.10	5%	LR	3.20	2.37	3.51	5.20
		W	0.67	0.79	8.35	7.05
		S	6.32	4.65	3.48	4.90
	10%	LR	6.32	5.11	9.38	10.11
		W	2.02	3.01	12.40	11.81
		S	10.29	7.49	8.14	9.69
0.75	5%	LR	2.39	4.84	5.43	5.55
		W	5.30	8.61	7.08	6.30
		S	4.08	4.09	4.80	5.29
	10%	LR	6.30	10.06	10.33	10.51
		W	10.00	14.00	12.16	11.11
		S	5.96	9.20	10.07	10.63

TABLE VIII
Nonnull rejection rates (%), ZOIK distribution,
$\mathscr{H}_0 : \lambda = \lambda_0 \times \mathscr{H}_1 : \lambda \neq \lambda_0$.

λ_0	δ	Test	n			
			50	100	200	500
0.10	5%	LR	56.36	80.54	97.89	100.00
		W	42.53	72.68	96.77	99.99
		S	56.36	86.90	97.89	100.00
	10%	LR	70.11	86.90	98.55	100.00
		W	56.42	80.56	97.89	100.00
		S	70.11	91.89	99.15	100.00
0.50	5%	RV	34.39	55.55	82.55	99.53
		W	34.39	55.55	82.55	99.53
		S	34.39	55.55	82.55	99.53
	10%	RV	45.67	62.92	89.12	99.82
		W	45.67	62.92	89.12	99.82
		S	45.67	62.92	89.12	99.82

rates are presented in Table VIII. It is noteworthy that the tests are less powerful when the value of λ_0 is large. Consider, e.g., n = 200 and $\delta = 5\%$. The estimated powers of the likelihood ratio, score and Wald tests are around 98% whereas for $\lambda_0 = 0.10$ they are around 82%. We also note that the powers of the tests coincide when $\lambda_0 = 0.50$.

EMPIRICAL APPLICATION

In what follows we shall present an empirical application of the IK_1 distribution. The variable of interest assumes values in $(0,1]$. It is the proportion of inhabitants in each of the 5,566 Brazilian municipalities that lived in homes with at least one bathroom and piped water in 2010. The data source is the 2013 edition of the Brazilian Atlas of Human Development; http://www.atlasbrasil.org.br/2013. The data contain 73 observations that equal one. Table IX displays some descriptive statistics on the variable of interest. Notice that 75% of the data points exceed 0.6778 and that there if left-skewness.

TABLE IX
Descriptive statistics.

minimum	1st quartile	median	mean	3rd quartile	maximum
0.0326	0.6778	0.9124	0.8087	0.9800	1.0000

We fitted the inflated Kumaraswamy (IK_1) and beta distributions (BEOI), both with inflation at one. The maximum likelihood estimates of the parameters that index that IK_1 distribution (standard errors in parentheses) are $\hat{\lambda} = 0.0131$ (0.0015), $\hat{\alpha} = 2.3513$ (0.0503) and $\hat{\beta} = 0.5292$ (0.0085). The parameter estimates we obtained for the BEOI law are $\hat{\lambda} = 0.0131$ (0.0015), $\hat{\mu} = 0.8026$ (0.0027) and $\hat{\varphi} = 2.7160$ (0.0518). Again, log-likelihood maximization was performed using the BFGS quasi-Newton method and the Ox matrix programming language. Figure 3 contains the data histogram and the fitted IK_1 density. The fitted BEOI density is not included in the plot because it is very similar to the fitted IK_1 density, as expected given the large sample size.

We performed the Kolmogorov-Smirnov test; for details on such a test, see Pestman (1998, Section 7.4). The interest lies in determining whether the sample at hand came from the the postulated distribution. The test was performed for each of the two inflated laws. For the inflated Kumaraswamy and beta laws, the test statistics are, respectively, 0.1896 and 0.1938. Even though the null hypothesis is not rejected for both distributions, the fact that the test statistic is smaller for the inflated Kumaraswamy law indicates there is more evidence in favor of the inflated Kumaraswamy distribution relative to the alternative law.

Using the maximum likelihood estimate of λ (IK_1 law), we constructed the asymptotic 95% confidence interval for such a parameter. The lower interval limit is 0.0102 and the upper limit equals 0.0160.

Finally, we tested the following null hypotheses against the corresponding two sided alternative hypotheses (IK_1 law): (i) $\mathscr{H}_0 : \lambda = 0.010$, (ii) $\mathscr{H}_0 : \lambda = 0.015$ and (iii) $\mathscr{H}_0 : \lambda = 0.015$, the respective likelihood ratio test statistics (p-values in parentheses) being 4.9695 (0.0258), 1.3970 (0.2372) and 15.3039 (0.0001). The corresponding score [Wald] figures are 5.4566 (0.0195) [4.1736 (0.0411)], 1.3381 (0.2474) [1.5274 (0.2165)] and 13.4602 (0.0002) [20.3827 (< 0.0001)]. It is thus clear that the second null hypothesis is not rejected at the usual nominal levels, and one can safely take the value of λ to be 0.015.

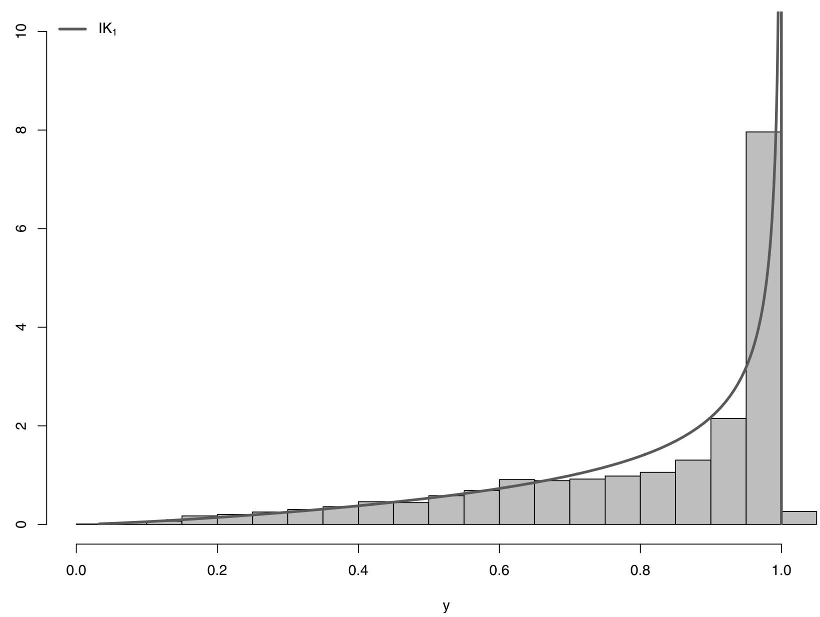

Figure 3 - Data histogram and fitted inflated Kumaraswamy density.

CONCLUSIONS

Applied statisticians oftentimes need to model variables that assume values in the standard unit interval, $(0,1)$; e.g., rates, proportions, income inequality indices, etc. The beta and Kumaraswamy distributions are commonly used with such variables. There are instances, however, when the variable of interest may display inflation, i.e., it may equal zero and/or one with positive probability. Put differently, it assumes values in $[0,1)$ (inflation at zero), $(0,1]$ (inflation at one) or $[0,1]$ (inflation at both interval limits). In this paper, we introduced inflated Kumaraswamy distributions that can be used as underlying laws for variables that assume values in those intervals. We considered two separate cases, namely: (i) inflation at zero *or* one and (ii) inflation at zero *and* one. For both cases, we introduced the appropriate law and also discussed point estimation, interval estimation and hypothesis testing inference. We presented Monte Carlo simulation evidence on the finite sample performances of point estimates, confidence intervals and hypothesis tests. Finally, an empirical application was presented and discussed.

ACKNOWLEDGMENTS

This work was supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) under Grant 301651/2017-5. We are also thankful to two anonymous referees whose comments and suggestions led to a much improved manuscript.

AUTHOR CONTRIBUTIONS

The current paper was jointly developed by the two authors. The first author proposed the research topic. The analytical results were derived by the second author and checked by the first author. The Monte Carlo simulations were performed by the second author. The empirical application was carried out jointly by the two authors. The manuscript was written by the first author.

REFERENCES

BARRETO-SOUZA W AND LEMONTE AJ. 2013. Bivariate Kumaraswamy distribution: properties and a new method to generate bivariate classes. Statistics 47(6): 1321-1342.

BUSE A. 1992. The likelihood ratio, Wald and Lagrange Multiplier tests: an expository note. Amer Statist 36(3): 153-157.

CARRASCO JMF, FERRARI SLP AND CORDEIRO GM. 2010. A new generalized Kumaraswamy distribution. Technical Report arXiv: 1004.0911v1 [stat.ME]. URL https://arxiv.org/abs/1004.0911v1.

COX DR AND HINKLEY DV. 1979. Theoretical statistics. New York: Chapman & Hall/CRC.

CRIBARI-NETO F AND ZARKOS SG. 2003. Econometric and statistical computing using Ox. Comput Econ 21(3): 277-295.

DOORNIK JA. 2009. An object-oriented matrix programming language Ox 6. London: Timberlake Consultants Press.

JONES MC. 2009. Kumaraswamy's distribution: a beta-type distribution with some tractability advantages. Stat Methodol 6(1): 70-81.

KUMARASWAMY P. 1976. Sinepower probability density function. J Hydrol 31(1-2): 181-184.

LEMONTE AJ. 2011. Improved point estimation for the Kumaraswamy distribution. J Stat Comput Simul 81(12): 1971-1982.

MITNIK PA AND BAEK S. 2013. The Kumaraswamy distribution: median-dispersion re-parameterizations for regression modeling and simulation-based estimation. Statist Papers 54(1): 177-192.

MITTELHAMMER RC, JUDGE GG AND MILLER DJ. 2000. Econometric foundations. New York: Cambridge University Press.

NOCEDAL J AND WRIGHT SJ. 2006. Numerical optimization. New York: Springer. 2^{nd} ed.

OSPINA R AND FERRARI SLP. 2010. Inflated beta distributions. Statist Papers 51(1): 111-126.

PESTMAN WR. 1998. Mathematical statistics. New York: Walter de Gruyter.

PRESS WH, TEUKOLSKY SA, VETTERLING WT AND FLANNERY BP. 1992. Numerical recipes in C: the art of scientific computing. New York: Cambridge University Press. 2^{nd} ed.

SERFLING RJ. 1980. Approximation theorems of mathematical statistics. New York: John Wiley & Sons.

SILVA RB AND BARRETO-SOUZA W. 2014. Beta and Kumaraswamy distributions as non-nested hypotheses in the modeling of continuous bounded data. Technical Report arXiv:1406.1941 [stat.ME]. URL https://arxiv.org/abs/1406.1941.

WANG BX, WANG XK AND YU K. 2017. Inference on the Kumaraswamy distribution. Comm Statist Theory Methods 46(5): 2079-2090.

WELSH AH. 1996. Aspects of statistical inference. New York: John Wiley & Sons.